From: Analytic Approaches to Product and Marketing Planning A.D. Shucker, ed. Marketing Science Institute, Cambridge, MA

EFFECTIVE MARKETING RESEARCH: AN EMPIRICAL COMPARISON OF TECHNIQUES TO MODEL CONSUMERS' PERCEPTIONS AND PREFERENCES

by John R. Hauser and Frank S. Koppelman

INTRODUCTION AND SUMMARY OF RESULTS

The primary goal of marketing analysis is to provide managers with the consumer response information necessary to develop, refine, and evaluate alternative strategies. For example, a new product manager might want to know what dimensions consumers use to evaluate laundry detergents, or he or she might want to know the relative importance of those dimensions. Similarly, a retailer might want to identify the basic dimensions that consumers consider when choosing among shopping locations. Knowing these dimensions and their relative importance helps managers select optimal store locations.

The marketing literature abounds with techniques to help managers identify dimensions and importances. For example, the most commonly used approach for identifying perceived dimensions is nonmetric scaling and its variations (Green and Rao, 1972; Green and Wind, 1973), but many researchers also use discriminant analysis (Johnson, 1971; Johnson, 1970; Pessemier, 1976) and factor analysis (Urban, 1975). The most common approach to identify the importances of those perceived dimensions is expectancy values (Fishbein, 1967; Rosenberg, 1956; Wilkie and Pessemier, 1973), but recent advances in econometrics have allowed researchers to use statistical techniques such as preference regression (Hauser and Urban, 1977) and preference logit (McFadden, 1972; McFadden, 1975) for estimating the importance weights.

Each combination of perception model and preference model provides managerial insight. But which combination is best? If a manager or analyst is to undertake marketing research to develop, refine, or evaluate a marketing strategy, which combination of models should he select? To answer the question, this paper empirically evaluates the perception and preference models relative to the criteria of (1) ability to provide managerial interpretation, (2) ability to accurately predict consumer preference, (3) ease of use, and (4) cost.

The outcomes of those tests are clear, albeit surprising. Of the perception models, factor analysis is superior to nonmetric scaling on all criteria and is superior to discriminant analysis on interpretability and predictability. Of the preference models, logit analysis predicts better but

---

Author's Note: We wish to thank Bruce Bagamery, Yosy Prashker, and Steve Shugan for their careful computer analysis, constant checking, and insightful questions.

This study was funded by university research grant DOT-OS-40001, from the U.S. Department of Transportation, Peter Stopher and Peter Watson, principal investigators.

costs more than preference regression. Both logit analysis and preference regression yield similar managerial interpretations and predict better than unit weights with perception models, but when the analysis is restricted to fundamental attributes, multicollinearity undermines the predictability of statistical models and both expectancy value and unit weights are better. Overall, the best combined model is logit analysis used with factor scores followed by expectancy value with fundamental attributes. All comparisons are confirmed with saved data and sampling tests.

The structure of this paper is to review the theory underlying each perception or preference model; discuss the empirical setting and experimental design; present the tests of interpretability, predictability, ease of use, and cost; and provide the confirmatory saved data and sampling tests. A final section suggests further research and testing.

REVIEW OF THE UNDERLYING THEORIES

This section is meant to compare the basic model assumptions rather than completely review the theory; the references in the introduction contain complete discussions. The basic characteristics of the perceptions and preference models are summarized in Tables 1 and 2, respectively.

Perception Models

Focus groups, open-ended surveys, and other qualitative measurement can identify a large number of fundamental attributes which consumers might use to describe a particular product category. But to understand the true perceptual process, to gain managerial insight, and to enhance creative strategy development, market researchers must use models to identify the few basic perceptual dimensions (Bruner et al., 1956) consumers use to reduce the cognitive strain in evaluating products or services. These perception models either reduce the set of fundamental attributes through correlation (factor analysis) or discriminant ability (discriminant analysis), or they independently uncover the dimensions based on measures of dissimilarity (nonmetric scaling).

Nonmetric scaling asks consumers to consider all products in a category and to indicate the relative similarity or dissimilarity between products. Based on the assumption that judged dissimilarity between products is proportional to the distance between stimuli in a 2, 3, or 4 dimensional space, nonmetric scaling, MDSCAL (Kruskal, 1964) selects the "positioning" of products in this perceptual space to best recover judged dissimilarity. Individual differences are modeled by differentially weighting the common dimensions, INDSCAL (Carroll and Chang, 1970). Identification of the names of the dimensions is aided by a regression-like procedure, PROFIT (Carroll and Chang, 1971) which projects the dimensions on the fundamental attributes. The appeal of nonmetric scaling is that it makes very few assumptions about how individuals process information and, at least in the "positioning" step, it is not sensitive to the selection of the fundamental attributes. Its drawbacks are that the number of dimensions is limited by the stimuli [at least 7 or 8 stimuli are needed for 2 or 3 dimensions (Klahr, 1969; Koppleman et al., 1977)] and that the individual difference scaling implicitly assumes that, relative to each dimension, all individuals perceive the stimuli in the same rank order. For example, INDSCAL would not allow one individual to perceive that Pepsi is

Table 1. Theoretical Constructs Behind the Four Models of Consumer Perceptions

NONMETRIC SCALING	FACTOR ANALYSIS
• dissimilarity distance	• searches for common component of scale rating
• "positions" stimuli to best recover distance	• rating = common + specific + error
• "fits" attribute ratings to explain stimuli postions	• correlations identify dimensions
DISCRIMINANT ANALYSIS	**FUNDAMENTAL ATTRIBUTES**
• can identify stimuli by its ratings	• either no reduction possible
• search for dimensions that discriminate best	• or reduction sacrifices information
• discriminant weights identify dimensions	

Table 2. Theoretical Constructs Behind the Four Models of Consumer Preference

PREFERENCE REGRESSION	1st PREFERENCE LOGIT
• weighted attributes • rank order preference	• weighted attributes probability of 1st preference
• statistically estimate "average" weights	• maximum likelihood estimate of "average" weights
UNIT WEIGHTS	**EXPECTANCY VALUE**
• either cannot distinguish differential weighting	• consumers can "self-explicate" importance weights
• or all attributes have equal weight	• individual specific importance weights

sweeter than Coke while another individual perceived that Coke is sweeter than Pepsi.

Discriminant analysis assumes that it is possible to identify a particular product or service by knowing how consumers rate that product or service on the fundamental attributes. The model then searches for dimensions (combined attributes) that best discriminate between products. That is, the dependent variable is "product rated" and the independent variables are the attribute scores. The discriminant weights, relative weightings of the attributes in the discriminant function, identify the name of each dimension. Some researchers rotate the discriminant dimensions with varimax rotation (Pessemier, 1976) to obtain better interpretability of the dimensions. The appeal of discriminant analysis is that the dimensions are specifically chosen to discriminate among products or services. Its drawbacks are that it implicitly assumes that the attribute ratings are interval-scaled and that it confounds differences between products based on a cognitive reduction with physical differences between products in the marketplace. The limitation on dimensions, one less than the number of stimuli, is not as severe as in nonmetric scaling.

Factor analysis assumes that there are some underlying and cognitive dimensions, and that when a consumer rates an attribute his rating has a common component, an attribute-specific component, and some measurement error. The comon components can be found by "factor analyzing" (Rummel, 1970) the attribute ratings across products and consumers. The dimensions are named by examining the correlations, called factor scores, between each dimension and the attributes. The appeal of factor analysis is that it is directly based on assumptions about the measurement task. Its drawback is the assumption that the attribute ratings are interval-scaled. Note that, like MDSCAL, a common space is obtained, but, unlike INDSCAL, individuals are free to reverse orderings on the underlying dimensions.

Fundamental attributes, i.e., the raw standardized attribute ratings, are included in the empirical tests as a check of the hypothesis that underlying cognitive dimensions exist and describe consumer response effectively.

Preference Models

It is important to managers to identify the underlying cognitive dimensions and to know product or service positions relative to these dimensions. But, if they want to develop effective marketing strategies, managers must also know the relative importances of the dimensions. For example, it is important for a health services manager to know that service is perceived relative to quality, personalness, convenience, and value, but to develop strategy he or she must know whether to stress quality (i.e., high quality, premium cost), value (i.e., inexpensive but adequate), or another dimension. Preference models determine these relative importances by measuring or estimating a linear compensatory model[1]:

---

[1] There are also nonlinear models such as the disjunctive, conjunctive, additive conjoint, interaction conjoint, and von Neumann-Morgenstern utility models (Green and Devita, 1975; Green and Wind, 1975; Hauser and Urban, 1976; Johnson, 1974; Kotler, 1976).

$$p_{ij} \backsim \sum_k w_{ik} \cdot d_{ijk} \qquad (1)$$

that states that consumer i's preference, P_{ij}, for product j is determined by differentially weighting his or her perceptions, d_{ijk}, of product j relative to cognitive dimension k.

Preference regression statistically estimates the importance weights using rank order preference as the dependent variable and the consumers' perceptions as independent variables. The statistical techniques are either monotonic regression (Johnson, 1974) or ordinary least squares (OLS) regression. Because recent simulation (Carmone et al., 1976; Cattin and Wittink, 1976) and empirical tests (Hauser and Urban, 1977) show that OLS performs as well as the more complex and expensive montonic regression, the empirical tests discussed in this paper use OLS. The appeal of preference regression is that full rank order information is used. Its drawbacks are the metric assumption and the necessity of estimating average importance weights, w_k's, to gain sufficient degrees of freedom.

Preference logit assumes that the true preference, p^T_{ij}, is composed of an observable part, p_{ij} as in equation 1, plus an error term, e_{ij}, i.e.:

$$p^T_{ij} = P_{ij} + e_{ij}. \qquad (2)$$

Assuming a probability distribution for the error term[2] makes it possible to derive a functional form for the probability, L_{ij}, that consumer i ranks product j as his or her first preference. This probability, given by

$$L_{ij} = \exp(p_{ij})/\sum_m \exp(p_{im}), \qquad (3)$$

where the sum is over all products, m, is called the first preference logit model. The importance weights are estimated by maximum likelihood techniques (McFadden, 1970; McFadden and Wills, no date).

The appeal of the logit model is that it explicitly models stochastic behavior (Bass, 1974) and it makes no metric assumptions about preference rankings. Its drawbacks are that it uses only first preference information and that it estimates average importance weights to gain degrees of freedom.

Expectancy value does not statistically estimate importance weights but rather asks consumers to state their own relative importances. These self-explicated weights are then used in equation 1. The appeal of the expectancy value model is individual specific weights. Its drawbacks are the scaling problems inherent in using self-explicated importances and the often questioned ability of consumers to accurately provide these weights. Furthermore, because the self-explicated weights must be measured in the original survey, expectancy value can only be used with fundamental attributes.

---

2 The error terms are independent and identically distributed Weibull random variables (McFadden, 1970).

Unit weights, i.e., the assumption that all dimensions are of equal weight (Einhorn, no date), are included in the empirical tests as a check of the hypothesis that market research can identify relative importance.

EMPIRICAL SETTING AND EXPERIMENTAL DESIGN

The empirical problem is to model consumers' perceptions and preferences relative to the attractiveness of shopping locations. Although distance and other measures of accessibility influence a consumer's choice of where to shop, most of today's consumers, especially in large metropolitan areas, are faced with a myriad of shopping locations, all within easy driving distances of their homes. Thus other attributes of shopping locations, such as variety of merchandise, prestige, "specials", and reasonable price, are becoming important determinants of shopping behavior. To begin to understand this construct of shopping location attractiveness it is necessary to determine the cognitive dimensions of attractiveness and their relative importances.[3]

The study develops models based on measures of seven shopping locations including downtown Chicago and six suburban shopping centers of widely differing characteristics (see Table 3). The locations were chosen to represent the types of shopping opportunities available to residents in the suburbs north of Chicago. The data were obtained by sampling individual shoppers at four of these locations. The models reported use choice based adjustments to eliminate sampling bias in the estimation of importance weights (Lerman et al., 1976; Manski and Lerman, no date; Wallace and Hassain, 1969). Since choice based sampling theory is relatively new, the study inluded tests to ensure that there was no estimation bias. This sampling reliability was checked by performing two distinct streams of analyses, one stream for the full set of seven locations and one stream for the set of four sampled locations. The results, explained in a later section, confirm the choice-based sampling theory, and indicate that the resulting models and comparisons are unbiased. This result alone holds promise for market researchers because it indicates that random samples may be replaced by more efficient ways to collect data.

Table 3. Description of Shopping Locations

1.	Chicago Loop	-	Downtown Chicago Central Shopping District.
2.	Woodfield	-	One of the largest shopping centers in the Midwest
3.	Plaza del Lago	-	An exclusive shopping center characterized by Spanish architecture and specialty shops
4.	Korvette City	-	A small discount shopping center

[3] In a companion study, shopping center choice is modeled as determined by accessibility and attractiveness. The perception/preference study is necessary to develop those models. All comparative results hold true in the expanded model (Hauser and Koppelman, 1977).

5. Old Orchard - Relatively large suburban shopping center

6. Edens Plaza - Moderate size shopping center on major highways

7. Golf Mill - Moderate size shopping center on major highways

The data were collected by self-administered questionnaire. The data used in this analysis include rank order preference for the attractiveness of each shopping location, similarity judgments for all pairs of shopping locations, direct ratings of each shopping location for sixteen attributes, and self-explicated importances of those attributes. (See Figure 1.) The attributes chosen to describe the general characteristics of shopping locations and the questionnaire were developed through extensive literature review, preliminary surveys, and analysis of developmental questionnaires (Stopher, et al., 1974).

After a series of pretests, 37,500 mail-back questionnaires were distributed. Of these, 6,000 consumers returned complete questionnaires and 1,600 of these consumers reported familiarity with all seven shopping locations. Five hundred of these respondents were randomly selected for analysis and an additional 500 respondents were selected for saved data testing.

The experimental design is a full factorial for all feasible cells. Preference regression, preference logit, and unit weighting models are estimated for each perception model (nonmetric scaling, discriminant analysis, factor analysis). Preference regression, unit weights, and expectancy value are estimated for the fundamental attributes.[4] In addition, two base models are used for comparison. These are equally likely preference and preference proportional to market share. All models are compared with respect to (1) interpretability, (2) predictability, (3) ease of use, and (4) cost.

MANAGERIAL INTERPRETABILITY

Good models provide insight to analysts and managers by helping them understand consumer response. Thus our first model comparison is managerial interpretability. This criterion, which serves as a first screen, is primarily a test of face validity but is also a test of intermodel consistency and robustness. The next section will test accuracy by using multiple measures to test predictability. (Predictability and interpretability serve as surrogates for explanatory power since the models are all of the same basic structure.)

Perception Models

The first test of interpretability is based on perceptual structure. Although the analyses differ, each perception model identifies the underlying cognitive dimensions with structure matrices which relate the cognitive dimensions to the sixteen fundamental attributes (Table 4). Despite strong

---

[4] Preference logit is too costly to run with the 22 variables (16 attributes and 6 sampling parameters) for fundamental attribute model. Without a resurvey, expectancy value is only good for the fundamental attributes.

If all the following shopping centers were equally easy to get to, which of them would you prefer to shop at for the goods you came to buy?

Please indicate your order of preference by placing a number beside each center. Start with number 1 for the most preferred shopping center, number 2 for the second most preferred, and so on down to the least preferred shopping center.

Please rank all the shopping centers.

Chicago Loop	[ ]
Edens Plaza (Wilmette)	[ ]
Golf Mill Shopping Center	[ ]
Korvette City (Dempster & Waukegan)	[ ]
Plaza del Lago	[ ]
Old Orchard	[ ]
Woodfield	[ ]

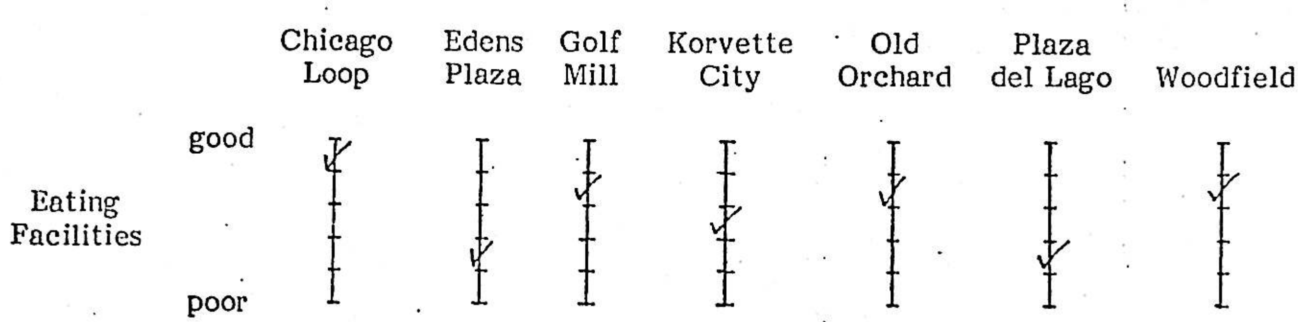
(a) Rank order preferences for attractiveness

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.

	Completely similar (identical)						Completely different
	1	2	3	4	5	6	7
Woodfield and Chicago Loop	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Edens Plaza and Gold Mill	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
Woodfield and Plaza del Lago	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

(b) 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:



(c) Ratings of the attributes of shopping locations

Figure 1. Examples of the Survey Measurement

Table 4. Structure Matrices for the Four Models of Consumer Perception

	VARIETY	QUALITY vs. VALUE	PARKING and SATISFACTION
1. Layout of store	.217	.497	.849
2. Return and service	.318	.122	.940
3. Prestige of store	.297	.804	.515
4. Variety of merchandise	.929	.360	.084
5. Quality of merchandise	.295	.811	-.505
6. Availability of credit	.880	-.085	-.468
7. Reasonable price	.485	-.853	.192
8. "Specials"	.786	-.594	.173
9. Free parking	-.294	-.550	.782
10. Center layout	-.447	.036	.894
11. Store atmosphere	-.199	.452	.869
12. Parking available	.463	-.478	.747
13. Center atmosphere	-.099	.480	.872
14. Sales assistants	-.052	.411	.910
15. Store availability	.872	.429	-.236
16. Variety of stores	.921	.385	-.054

(a) Nonmetric Scaling: directional cosines

	VARIETY vs. SATISFACTION	QUALITY vs. VALUE
1. Layout of store	.016	-.010
2. Return and service	.088	.040
3. Prestige of store	.249	.409
4 Variety of merchandise	.440	-.071
5. Quality of merchandise	-.536	.553
6. Availability of credit	.154	-.047
7. Reasonable price	.259	-.334
8. Specials"	.060	-.307
9. Free parking	.829	.631
10. Center layout	-.124	.004
11. Store atmosphere	-.180	-.008
12. Parking available	.088	-.089
13. Center atmosphere	-.053	.314
14. Sales assistants	-.233	-.005
15. Store availability	.064	-.000
16. Variety of stores	.902	.571

(b) Discriminant analysis: discriminant coefficients

Table 4. Structure Matrices for the Four Models of Consumer Perception

Cont'd

	VARIETY, QUALITY, and SATISFACTION	PARKING	VALUE		VARIETY	QUALITY AND SATIS-FACTION	PARKING	VALUE
1. Layout of store	.619	.256	.168	1.	.267	.583	.200	.156
2. Return and service	.468	.361	.289	2.	.095	.528	.255	.343
3. Prestige of store	.878	.042	-.001	3.	.388	.822	-.058	-.001
4. Variety of merchandise	.613	-.270	.455	4.	.665	.327	-.185	.309
5. Quality of merchandise	.847	.038	.026	5.	.307	.810	-.074	.037
6. Availability of credit	.341	.121	.454	6.	.159	.337	.049	.487
7. Reasonable price	-.057	.121	.596	7.	.067	-.063	.113	.599
8. "Specials"	.140	.022	.747	8.	.223	.074	.008	.739
9. Free parking	-.061	.800	.028	9.	-.150	.068	.811	.043
10. Center layout	.246	.585	.071	10.	.030	.308	.560	.074
11. Store atmosphere	.583	.506	-.009	11.	.080	.658	.400	.034
12. Parking available	-.033	.845	.088	12.	.145	.105	.841	.108
13. Center atmosphere	.702	.460	-.027	13.	.244	.694	.404	-.040
14. Sales assistants	.545	.389	.137	14.	.173	.560	.319	.147
15. Store availability	.573	-.064	.350	15.	.619	.320	.034	.204
16. Variety of stores	.652	-.311	.363	16.	.829	.288	.173	.160

(c) Factor analysis: factor loadings for three factors

(d) Factor analysis: factor loadings for four factors superficial similarities, the different models demonstrate striking differences in interpretation. First, examine the three three-dimensional perception models. The nonmetric scaling model and the factor analysis model generally have strong loadings on a single dimension indicating strong relationships within groups of attributes. The discriminant model has some attributes related strongly to two or three dimensions and six attributes which are not strongly related to any of the dimensions identified.[5] The nonmetric scaling and discriminant models include mixed signs for some of the major loadings; that is, some attributes load positively and others negatively on the same dimensions. These mixed loadings prevent the manager or analyst from identifying any natural direction of goodness along the affected dimensions. The factor analysis model does not include mixed signs for any of the major loadings.

The effect of increasing the number of dimensions was examined by developing four-dimensional discriminant and factor analyses (the small number of shopping locations do not allow the development of a reliable four-dimensional scaling analysis). The four-dimensional discriminant analysis continued to include both unclear and mixed loadings and was not analyzed further.[6] The four-dimensional factor analysis produced clearer dimensions by separating variety from quality and satisfaction (Table 4, part d).

The second test of interpretability is based on the perceptual maps (Figure 2). These maps help managers identify how each product, service, or shopping location is "positioned" in the market place. This way a manager can know the relative strengths and weaknesses of each shopping location and can identify opportunities in the market.

Careful inspection of Figure 2 reveals some consistency among the reduced maps even though the measures of dimensions come from different models. 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. But there are important differences. The factor analysis map gives slightly better insight because the four dimensions do not mix quality with value nor parking with satisfaction. All of the reduced perceptual maps are easier to work with and understand than the fundamental attribute map, which presents too much information to readily internalize for strategy development.

Thus all the perception models provide useful insight and some overall consistency of interpretation. However, factor analysis which has strong face validity is superior because of the clearer loadings, absence of mixed loadings, and the ability to identify four managerially important dimensions. Final judgement on the importance of these differences must await results of the predictive ability tests.

---

5 Varimax rotation of the discriminant coefficients did not improve interpretability and was dropped from further analysis.

6 The four-dimensional discriminant also provided very little improvement in explaining differences in perception among individuals and shopping centers.

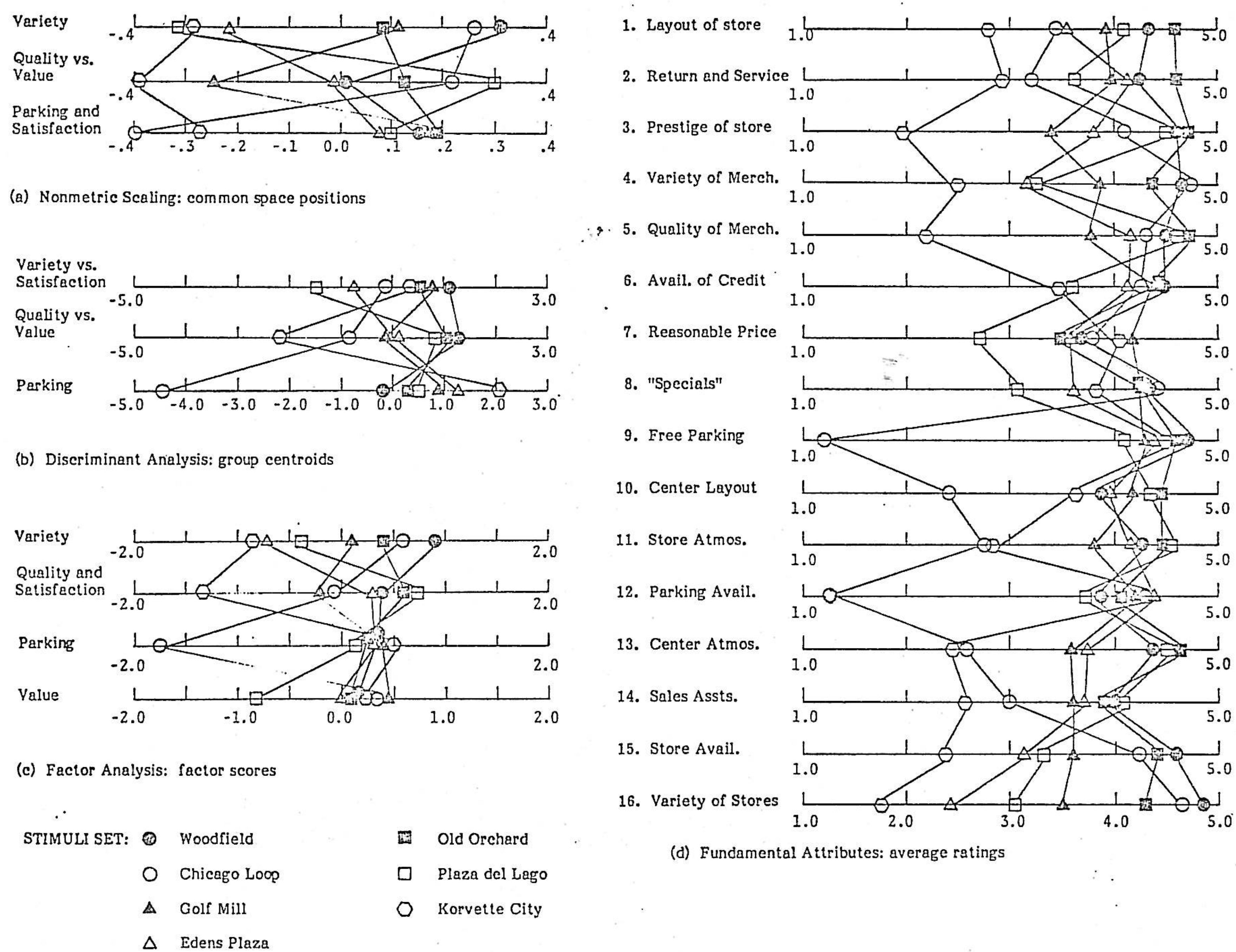


Figure 2. Perceptual Maps for the Four Models of Consumer Perceptions

Preference Models

The normalized estimated importance weights for each preference model for the three perception models are shown in Table 5. Although the exact values of the importances vary, their rank order is identical for both statistical models. This statistical robustness plus the fact that the estimated weights are definitely not equal suggests that the dimensions do have differential importances and that a unit weighting model probably neglects importance information. (The negative value for parking further indicts the discriminant analysis model.) Final judgement on robustness and significance relative to unit weights must await the predictability tests.

The expectancy value model incorporates individual importances for the fundamental attributes. These can be averaged to obtain an overall index of the importance of these attributes but cannot be used to obtain importance weights for the underlying cognitive dimensions.[7] Thus expectancy values provide a different type of importance information which may augment those obtained from the other preference models.

Table 5. Normalized Importance Weights for the Preference Models (* = not significant at the .05 level)

CONSUMER MODEL	NORMALIZED IMPORTANCE WEIGHTS			
Factor Scores	Variety	Quality/ Satisfaction	Value	Parking
Pref. Regress.	.38	.54	.07*	.01*
1st Pref. Logit	.30	.41	.23	.07*
Unit Weights	.25	.25	.25	.25
Nonmetric Scaling	Variety	Quality/ Value	Parking/ Satisfaction	
Pref. Regress.	.26	.43	.31	-
1st Pref. Logit	.26	.49	.26	-
Unit Weights	.33	.33	.33	-
Discriminant Analysis	Variety/ Satisfaction	Quality/ Value	Parking	
Pref. Regress.	.10	.70	-.20	-
1st Pref. Logit	.17	.63	-.20	-
Unit Weights	.33	.33	.33	-

[7] The structure matrices are not full rank and cannot be inverted.

PREDICTIVE ABILITY

A good model provides managerial insight and predicts well. Furthermore, predictive ability acts as a surrogate for explanatory power if the models have similar structure. Three distinct but related tests examine predictive ability. This use of multiple tests minimizes the problem that some models, such as preference logit, optimize one measure, i.e., first preference prediction, at the expense of other measures.

The importance weights estimated for the "calibration" sample of 500 consumers are used in equation 1 to predict consumers' preference rating among the seven shopping locations for the "calibration" sample, and for an additional "save data" sample of 500 consumers. The preference ratings are rank ordered to obtain individual preference ranks. These data are used to test (1) the percent of times each model correctly predicts first preference, (2) the percent of times the model correctly predicts the seven preferences, and (3) the mean of the absolute difference between the predicted and actual first preference market shares. Table 6 reports the results of those tests.

Perception models

It is clear from Table 6 that factor analysis dominates both nonmetric scaling and discriminant analysis, and in some cases is superior to fundamental attributes. This difference in predictive ability cannot be explained solely by the use of four versus three cognitive dimensions. In fact, the nonmetric scaling models predicted that no one would prefer Edens Plaza, Golf Mill, or Korvette City, which account for 13% of the actual first preferences. This is a direct result of the theoretical problem mentioned earlier that INDSCAL assumes common rankings of stimuli along each cognitive dimension.

The predictive ability of the fundamental attributes is sensitive to the preference model used. The poor performance of preference regression on fundamental attributes is surprising. A priori, one would expect that fundamental attributes would contain more information than the reduced perception models, but in this case the data is collinear and it appears that the collinearity degrades prediction. This multicollinearity also produces unreliable importance weights which degrade managerial interpretability. On the other hand, fundamental attributes combined with expectancy values or unit weights predict alomst as well as the best factor analysis models.

Thus the predictive tests reinforce the interpretability analysis. Factor analysis is the superior perception model to uncover consumers' cognitive dimensions.

Preference models

The interpretability tests indicate that preference regression and preference logit are similar and that both are superior to unit weights with the reduced perception models. The predictive and saved data tests confirm this hypothesis for nonmetric scaling and discriminant analysis, but for factor analysis they indiate that logit is slightly superior. All preference models for factor analysis, discriminant analysis and fundamental attributes predict significantly better than both the "equally likely" base model and the market share proportional model. Expectancy value predicts well with fundamental

Table 6. Predictive and Saved Data Tests for the Combined Models

CONSUMER MODEL	PREDICTIVE TESTS 1st preference	All preferences	Mean Absolute Error	SAVED DATA TESTS 1st preference	All preferences	Mean Absolute Error
Base Models						
Equally likely	14.3	14.3	12.0	14.3	14.3	12.0
Market Share	26.8	-	-	-	-	-
Factor Scores						
Pref. Regress.	50.1	32.9	4.1	46.6	31.8	4.3
1st Pref. Logit	55.0	37.0	3.6	50.8	36.6	4.5
Unit Weights	48.7	33.0	5.8	44.0	31.4	5.1
Nonmetric Scaling						
Pref. Regress	36.6	25.1	9.0	19.0	14.4	22.6
1st Pref. Logit	34.8	24.4	10.8	19.0	13.5	23.1
Unit Weights	32.4	24.8	9.7	19.0	11.3	22.9
Discriminant Analysis						
Pref. Regress.	35.6	26.6	9.9	38.4	29.1	8.2
1st Pref. Logit	36.2	27.8	10.1	39.6	29.7	8.3
Unit Weights	35.2	23.8	9.0	35.6	23.2	7.9
Fundamental Attributes						
Pref. Regress.	39.5	30.6	7.8	41.4	30.9	6.2
Unit Weights	51.1	36.4	5.7	47.0	34.2	5.3
Expectancy Value	52.8	34.4	5.4	47.0	34.0	5.4

attributes but not as well as preference logit/factor analysis. Expectancy value cannot provide relative importances for the reduced dimensions.

Thus, based on both managerial interpretability and predictive ability, the best combination model is factor analysis with preference logit.

EASE OF USE AND COST

In multimillion dollar managerial decisions it is worthwhile spending substantial money, time, and effort on analysis. But not all managerial decisions are multimillion dollar decisions. It is important that a good model be tailored to the decision it supports. While a model may be the best predictor, it may be too difficult or costly to use in a real environment. Thus the final model comparisons are ease of use and cost.

Perception models

Factor analysis and discriminant analysis are readily available in most standard statistical packages, e.g., QUAIL, BMDP, SPSS (Berkman et al., 1976; Dixon, 1975; Nie et al., 1975), are simple to access and use, and provide easily interpretable output. Both models cost about $10-$20 to run, including checks at three, four, and five dimensions (CDC-6400 at $510 per cpu hour). Finally, the models are readily transferable to new data sets via factor score coefficients or unstandardized discriminant coefficients.

In contrast, the special programs for nonmetric scaling require many exploratory runs, special FORTRAN programs to handle data transfer, and a series of statistical manipulations and data handling to develop a common space, estimate individual weights, and compute the directional cosines. A single set of runs costs about $40, but because various starting configurations and dimensions must be checked, the effective cost is about $150. Furthermore, new individual weights must be estimated to transfer the model. Finally, there is an added survey cost because nonmetric scaling requires direct similarity measures[8] in addition to attribute measures.

Preference models

Preference regression is available in most statistical packages (Dixon, 1975; Nie et al., 1975) and preference logit is becoming readily available (McFadden, no date). Both are simple to access and use, and provide easily interpretable output. Of the two, preference regression is less expensive. A typical run of 20 or more variables costs under $10. Logit costs increase rapidly with the number of variables. For example, for our analyses an estimation on reduced perceptions which requires 9 or 10 variables (3 or 4 cognitive dimensions and 6 sampling variables) cost about $30, while an estimation on fundamental attributes which requires 22 variables (16 attributes and 6 sampling variables) would cost over $100. [Fortunately, forthcoming estimation packages promise to be much less expensive to use (Berkman, 1976)].

---

[8] There are techniques based on indirect similiarity measures, but these techniques introduce metric assumptions (Green and Rao, 1972). Furthermore, when checked on this data set they did not outperform the direct measures.

Unit weights and expectancy values have no estimation cost, but expectancy values incur a survey cost because self-explicated importances must be measured in addition to attribute ratings.

Thus the best model for interpretability and predictability, preference logit-factor analysis, is relatively easy to use, not very expensive, and requires no survey questions beyond the attribute ratings and first preference. The last section outlines a summary of these comparisons, but first the next section presents a short discussion on the choice based sampling checks.

RELIABILITY AND EXTENDABILITY

Perceptions/preference models are important marketing tools, but standard procedures to collect data for these models can be difficult and expensive. For example, random samples collected at home locations require a survey frame, a sampling strategy, and either mail back or home interview surveys thoughout the metropolitan area. For the types of models tested in this paper these sampling costs would be high. Alternatively, choice-based samples can be collected. In a choice-based sample, consumers are sampled based on their choice of shopping location. To obtain unbiased estimates, samples should be collected in proportion to the number of consumers who would have chosen that location in a random sample. These sampling costs are less but still quite high. A third sampling strategy, made possible by recent developments in econometrics (Lerman, et al., 1976; Manski and Lerman, forthcoming; Wallace and Hassain, 1969), is to purposively collect samples at representative shopping locations and statistically correct the sampling bias. If this statistical analysis proves successful, market researchers with small or moderate budgets will be able to develop perception and preference models.

This study used choice-based sampling theory, and all empirical analyses use estimation procedures designed to eliminate sampling bias (Lerman, et al., 1976; Manski and Lerman, forthcoming; Wallace and Hassain, 1969). To test the reliability of the results, the models were estimated for the full seven stimuli and then again for only the four sampled stimuli. These empirical tests confirmed the predictions of the theory. For example, Table 7 shows that when factor analysis is the perceptual model, all estimated importances are statistically equivalent independent of the sampling strategy. Analyses when discriminant analysis and nonmetric scaling are the perception models provide similar results.

Table 7. Example of Comparison Between Full Stimuli Set and Choice-Based Sample Only.
(There were no statistically significant differences at the .05 level. The * indicates the estimated coefficient was not significant at the .05 level.)

	Variety	Quality/ Satisfaction	Value	Parking
Complete Set of 7 Stimuli				
Preference regression	.38	.54	*	*
1st preference logit	.30	.41	.23	*
Sampled Stimuli Only (4)				
Preference regression	.38	.52	*	*
1st preference logit	.27	.42	.24	*

The basic idea behind choice-based sampling theory is simple, although the formal proofs contained in the references are complex. Suppose that a sample is biased in the sense that it contains all the relevant population segments but in the wrong proportions. One such biased sample might result from drawing 50% of the sample from color TV users and 50% from nonusers when testing for differences among users and nonusers. If it were possible to identify the correct usage proportions in the population (suppose only 10% of the population were color TV users) it would be possible to differentially weigh each respondent and create a statistical sample more or less equivalent to a random sample. But in a choice-based sample this proportion bias may not be known. Choice-based sampling theory allows one to include sampling variables to simultaneously estimate both sample bias and importances (Lerman et al., 1976; Manski and Lerman, forthcoming; Wallace and Hassain, 1969). The resulting importances are unbiased.

For the analyses of this paper, the choice-based sampling theory and the empirical checks indicate that the results are extendable to other shopping locations within the Chicago area and perhaps to other metropolitan areas. At the very least, these sampling tests indicate that there is no systematic sampling bias confounding the intermodel comparisons.

SUMMARY

The model comparisons described previously are summarized in Table 8. Based on these comparisons, the best model for accuracy and insight appears to be preference logit/factor analysis. If a logit package is not available on the computer system, the next choice would be preference regression/factor analysis for interpretability and expectancy value for predictability.

Table 8. Summary Comparison of Perception and Preference Models

MODEL	CRITERIA			
PERCEPTION	INTERPRET-ABILITY	PREDICT-ABILITY	EASE OF USE	COST
FACTOR ANALYSIS	Excellent insight	Best	Standard statistical package	\$10-20
NONMETRIC SCALING	Good insight	Moderate	Special programs many analysis runs Requires similarity questions in survey	\$40-150
DISCRIMINANT ANALYSIS	Good insight	Moderate	Standard statistical package	\$10-20
FUNDAMENTAL ATTRIBUTES	Difficult to interpret	Good with non-statistical pref. models	No additional analysis	Expensive with Logit model
PREFERENCE				
PREFERENCE REGRESSION	Consistency	Good	Standard statistical package	\$10
	across			
PREFERENCE LOGIT	Methods	Best	Becoming available	\$10-100
EXPECTANCY VALUE	Only useful on fundamental attributes	Good	No additional analysis Requires importance questions in survey	No analysis cost
UNIT WEIGHTS	No insight	Good	No additional analysis	No analysis cost

When selecting preference models, both preference logit and preference regression provide consistent managerial insight and good predictions with relatively small cost and effort. Unit weight provides a good base model but does not help identify the relative importances of cognitive dimensions or attributes. Finally, expectancy value provides a basis for assessing relative importances for fundamental attributes but does not provide information about the relative importance of cognitive dimensions. However, it does provide good predictive ability and can add to managerial insight.

The analyses reported in this paper have been performed for a single product category and for a single measurement instrument. However, the product category, shopping location, is an important category for marketing research and the measurement instrument went through extensive and careful theoretical and empirical development. Nonetheless, the results obtained are subject to confirmation in other empirical or simulation tests. [Hauser and Urban (1976) report similar statistical consistency between preference logit and preference regression for models applied to the design of health services.]

Further interesting tests are possible to test the linear compensatory models relative to the extended expectancy value models (Ryan and Bonfield, 1975) and relative to nonlinear preference models (conjoint analysis, tradeoff analysis, direct utility assessment) which require followup personal interviews after the perceptual dimensions have been identified. This area of comparative model development is fruitful and deserves attention from marketing researchers.

REFERENCES

Bass, Frank, "The Theory of Stochastic Preference and Brand Switching," Journal of Marketing Research, 11 (February 1974), pp. 1-20.

Berkman, J., D. Brownstone, G.M. Duncan, and D. McFadden, "QUAIL User's Manual," Institute of Transportation Studies, University of California, Berkeley, 1976.

Bruner, J.S., J.J. Goodnaw, and G.R. Austin, A Study of Thinking. York: John Wiley & Sons, 1956.

Carmone, Frank J., Paul E. Green, and Arum K. Jain, "The Robustness of Analysis: Some Monte Carlo Results," Working Paper, University Pennsylvania, May 1976.

Carroll, J. Douglas and Jih Jie Chang, "Analysis of Individual Differences Multidimensional Scaling via an N-way Generalization of Eckart-Decomposition," Psychometrika, 35 (1970), pp. 283-319.

Carroll, J.Douglas and Jih Jie Chang, "An Alternate Solution to the Metric Problem," Psychometric Society, St. Louis, Missouri, March 1971.

Cattin, Philippe and Dick R. Wittink, "A Comparison of Estimation Procedures Multiattribute Models," Presented at the Joint National Meeting the Operations Research Society of America and The Institute Management Sciences, November 3-5, 1976, Miami, Florida.

Dixon, W.J., BMDP: Biomedical Computer Programs. Los Angeles: University California Press, 1975.

Einhorn, Hillel E., "Unit Weighting," Organizational Behavior and Human 13, pp. 171-192.

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

Green, Paul. E., "On the Robustness of Multidimensional Scaling Techniques," Journal of Marketing Research, 12 (February 1975), pp. 73-81.

Green, Paul E. and Frank J. Carmone, "Multidimensional Scaling: An Introduction and Comparison of Nonmetric Unfolding Techniques," Journal of Marketing Research, 7 (August 1969), pp. 330-341.

Green, Paul E. and Michael Devita, "An Interaction Model of Consumer Utility," Journal of Consumer Research, 2 (September 1975).

Green, Paul E. and Vithala Rao, Applied Multidimensional Scaling. New York: Holt, Rinehart & Winston, Inc., 1972.

Green, Paul E. and Yoram Wind, Multiattribute Decisions in Marketing: A Measurement Approach. Hinsdale, Ill.: The Dryden Press, 1973.

Green, Paul E. and Yoram Wind, "New Way to Measure Consumer's Judgments," Harvard Business Review (July-August, 1975).

Hauser, John R. and Francis S. Koppelman, "An Empirical Model of Consumer Shopping Behavior," Transportation Center Working Paper, Northwestern University, Evanston, Illinois, January 1977.

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

Hauser, John R. and Glen L. Urban, "A Normative Methodology for Modeling Consumer Response to Innovation," Operations Research, 25 (July-August 1977), pp. 569-619.

Johnson, Richard M., "Market Segmentation: A Strategic Management Tool," Journal of Marketing Research, 8 (February 1971), pp. 13-18.

Johnson, Richard M., "Multiple Discriminant Analysis: Applications to Marketing Research," Market Facts, Inc., Chicago, Illinois, January 1970.

Johnson, Richard M., "Tradeoff Analysis of Consumer Values," Journal of Marketing Research, 11 (May 1974), pp. 121-127.

Johnson, Richard S., "Monotonic Regression," Working Paper, Market Facts, Inc., Chicago, Illinois, 1974.

Klahr, David, "A Monte Carlo Investigation of the Statistical Significance of Kruskal's Non-Metric Scaling Procedure," Psychometrika, 34:3 (September 1969), pp. 319-330.

Koppleman, Francis S., J.N. Prashker, and B. Bagemery, "Perceptual Maps of Destination Characteristics Based on Similarities Data," 56th Annual Meeting, Transportation Research Board, Washington, 1977.

Kotler, Philip, Marketing Management: Analysis, Planning and Control, 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1976.

Kruskal, Joseph B., "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika, 29 (March 1964) pp. 1-27.

Lerman, S.R., C.F. Manski, and T.J. Atherson, "Non-Random Sampling in the Calibration of Disaggregate Choice Models," Cambridge Systematics, Inc., Final Report No. PO-6-3-0021 for the U.S. Department of Transportation, Federal Highway Administration, February 1976.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zaremblea (ed.), Frontiers in Econometrics, New York: Academic Press, 1970, pp. 105-142.

McFadden, Daniel, "Quantal Choice Analysis: A Survey," NSF-NBER Conference on Decision Roles and Uncertainty, Department of Economics, University of California, Berkeley, March 22-23, 1974.

McFadden, Daniel, "The Revealed Preferences of a Government Bureaucracy: Theory," The Bell Journal of Economics, 6:2 (Autumn 1975).

McFadden, Daniel and H. Wills, "XLOGIT, a Program for Multinominal Logit Analysis," Travel Demand Forecasting Project, Institute for Transportation and Traffic Engineering, University of California, Berkeley.

Manski, C.F. and S.R. Lerman, "The Estimation of Choice Probabilities from Choice-Based Samples," Econometrika (forthcoming).

Nie, N.H., C. Hadlai Hall, J.G. Jenkins, K. Steinbrenner, and D.H. Bent, SPSS: Statistical Package for the Social Sciences, 2nd edition. New York: McGraw Hill, 1975.

Pessemier, Edgar A., "Market Structure Analysis of New Products, Market, and Communication Opportunities," Report No. 76-100, Marketing Science Institute, Cambridge, Massachusetts, June 1976.

Rosenburg, M.J., "Cognitive Structure and Attitude Effect," Journal of Abnormal and Social Psychology, 53 (1956), pp. 367-372.

Rummel, R.J., Applied Factor Analysis, Evanston, Ill.: Northwestern University Press, 1970.

Ryan, Michael J. and Edward H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2:2 (September 1975).

Stopher, P.R., P.L. Watson, and Jean Marie Blin, "A Method for Assessing Pricing and Structural Changes on Transport Mode Use," Transportation Center, Northwestern University, Interim Report to the Office of University Research, U.S. Department of Transportation, September 1974.

Urban, Glen L., "PERCEPTOR: A Model for Product Design," Management Science, 21:8 (April 1975), pp. 858-871.

von Neumann, J. and Oscar Morgenstern, The Theory of Games and Economic Behavior, 2nd ed. Princeton, N.J.: Princeton University Press, 1947.

Wallace, T.D. and Ashig Hassain, "The Use of Error Components Models in Combining Cross Section with Time Series Data," Econometrica, 37 (January 1969).

Wilkie, William L. and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November 1973), pp. 428-441.