Simple and Weighted Unfolding Threshold Models for the Spatial Representation of Binary Choice Data

Wayne S. DeSarbo
University of Pennsylvania

Donna L. Hoffman
Columbia University

This paper describes the development of an unfolding methodology designed to analyze "pick any" or "pick any/n" binary choice data (e.g., decisions to buy or not to buy various products). Maximum likelihood estimation procedures are used to obtain a joint space representation of both persons and objects. A review of the relevant literature concerning the spatial treatment of such binary choice data is presented. The nonlinear logistic model type is described, as well as the alternating maximum likelihood algorithm used to estimate the parameter values. The results of an application of the spatial choice model to a synthetic data set in a monte carlo analysis are presented. An application concerning consumer (intended) choices for nine competitive brands of sports cars is discussed. Future research may provide a means of generalizing the model to accommodate three-way choice data.

Social scientists are often concerned with the problem of explaining and/or predicting individual behavior, especially individual choice behavior, in a wide variety of settings. Often the choices can be considered to be selections from a continuum of alternatives, as illustrated by the conventional theory of the household (or firm) that deals with "how much" to consume (or produce) of a certain commodity. Economic theory provides various models of individual behavior in these situations, and standard procedures allow for statistical inference about typical population behavior, given a random sample of data from a population of individuals (see Judge, Griffiths, Hill, Lutkepohl, & Lee, 1985). Increasingly, however, as microdata (e.g., panel data) have become more widely available, researchers are faced with situations in which choice alternatives are limited in number (i.e., the alternatives are discrete). Examples of discrete choice alternatives include choices arising from product selection, retail store selection, or job choice.

The most common source of data associated with choice behavior is the situation in which people select a preferred object or objects from a fixed set of n alternatives. Such data are known as "pick k/n" or "pick any/n" data (Coombs, 1964). For example, in product concept testing in marketing (Wind, 1982), people are often asked to select the best product(s) from a set of 10 or so fixed items presented on a questionnaire. Or, people may be asked to select the products from the set of 10 that they would actually buy if offered in the marketplace. Still another example of "pick any/n" data occurs with management's attempt to evaluate competitive positioning through a systematic examination of advertising and package claims, using a products × benefits matrix (Wind, 1982).

Another type of choice data is that in which people choose preferred objects from an unconstrained, person-specific set of alternatives. For example, people may choose cars they like, persons who are friends, or organizations of which they are a member. The identifying characteristic of these data is that the alternatives from which the preferred object is selected may be assumed to vary from person to person—for example, friends are chosen from among persons met. Coombs (1964) has suggested the phrase "pick any" to describe this type of choice data. An example is where vacationers may initially choose a number of preliminary locations for a possible vacation, but later reduce this set to one. Another example is sociometric choice data where a person can select any number of friends in a social network to associate with, request money from, etc.

The unfolding methodology proposed in this paper has been designed to analyze "pick any" or "pick any/n" individual choice data. Its purpose is to represent choice behavior, defined by dichotomous variables that indicate whether a particular object was chosen, in terms of a joint space of persons and objects.

Scaling Models for Binary Data

A number of multidimensional scaling (MDS) methodologies have been suggested in the psychometric literature to accommodate the spatial analysis of "binary" choice data. Torgerson (1958) provided deterministic and stochastic methodologies to analyze such data spatially. He developed deterministic conjunctive and disjunctive models of choice and corresponding spatial methods that operate on binary choice data. Torgerson also generalized Thurstone's (1929) method of similar reactions in obtaining a matrix of "directed distances," which is converted into scalar products and then factor analyzed.

Nonmetric MDS methods (e.g., Lingoes, 1972) estimate a joint space of persons and objects, in which choices correspond to short distances between points while non-choices correspond to long distances. Here, the data would show a massive number of "ties", and it is not clear what effect this would have on resulting solutions. Nonmetric factor analysis (Kruskal & Shepard, 1974) and binary factor analysis (Christoffersson, 1975; Muthén, 1981) are methodologies which extend classical factor analysis to the analysis of binary data.

Correspondence analysis (Benzecri, 1973) is another type of scaling procedure which typically analyzes aggregate choice data in the form of a frequency matrix and derives a joint space of row and column objects based on an eigenstructure analysis of a normalized frequency matrix. It can also be used to analyze raw binary data. Greenacre (1984) and Lebart, Morineau, and Warwick (1984) discussed extensions of the procedure to multiple correspondence analysis. Hoffman and Franke (1986) have discussed the various uses of correspondence analysis in marketing.

Two "optimal scaling" approaches have also been suggested to analyze such choice data. Dual scaling methods (Nishisato, 1980) and homogeneity analysis (De Leeuw, 1984; Gifi, 1981a, 1981b; Heiser, 1981) attempt to derive best fitting continuous weights to discrete categorical data values. Many of these approaches handle correspondence analysis as a special case. Gifi (1981a) and Nishisato (1980) presented a more detailed literature review of dual scaling approaches.

Levine (1979) proposed an eigenstructure-based technique to provide a joint space to "pick any" data. According to Levine, a non-chosen alternative may be either (1) unacceptable; (2) acceptable but not considered; or (3) acceptable and considered, but not chosen. He proposed a scaling method where persons are represented by points whose coordinates are proportional to the centroids of the points representing their choices. Similarly, choice alternatives are represented as points whose coordinates are proportional to the centroids of the points representing persons who have chosen them. Green and DeSarbo (1980) and Holbrook, Moore, and Winer (1980, 1982) have presented applications of this model in marketing. This approach can be viewed as a special case of correspondence analysis where a different type of normalization is employed.

Takane (1983) has recently presented an item response model for the analysis of "pick any/n" data. Such data are viewed as a special type of successive categories data in which there are only two response categories. Here, each item (stimulus), rather than each category of an item, may be represented as a point, and persons are assumed to choose (or not to choose) the item according to its closeness to their respective ideal points. Takane developed an EM algorithm to maximize a marginal likelihood in order to estimate the desired set of parameters. His model is somewhat similar to a special case of one of the unfolding models developed below.

Objectives

For the analysis of such binary choice data, a new multidimensional unfolding based choice model is proposed which would eliminate some of the limitations that exist in current psychometric and econometric approaches. The model, which estimates a joint space of persons and objects from two-mode, two-way binary choice data, can be characterized as follows:

1. *Stochastic model*—The underlying choice process of a particular person choosing a particular object is modeled as a stochastic process.
2. *An unfolding representation*—The proposed model represents both persons (ideal points) and objects as points in the derived joint space.
3. *Dimensionality identification*—In many other spatial models, the appropriate dimensionality is determined in part by examining a plot of the goodness-of-fit measure versus the number of dimensions (i.e., the scree plot) as well as subsequent interpretation of the dimensions. In addition to these criteria, the proposed model, because of its stochastic nature, allows for an asymptotic statistical test for identifying the appropriate dimensionality.
4. *Type of binary data*—The proposed model will be able to accommodate "pick any" and "pick any/n" binary choice data.
5. *Effect of object features*—The proposed model features a number of nested models that can be estimated. One available option involves reparameterizing estimated object coordinates to be specified functions of measurable features, attributes, etc. Given this option, once a function is specified for such a reparameterization (discussed below) and coefficients estimated, the user can then directly link such features to locations in the derived space. The estimated coefficients would indicate the importance or impact of the feature variable on a dimension.
6. *Effect of person information*—The proposed model has options which also allow for person coordinates to be specified functions of selected demographic, psychographic, and/or other relevant individual background data. As above, reparameterization coefficients are estimated which indicate the importance or impact of a particular background variable on a particular dimension. This aids in interpreting the obtained solution in terms of person characteristics, which is useful for classification.
7. *Prediction of person and/or object locations*—Current MDS methods are limited with respect to ability to draw inferences concerning persons and/or objects not included in the input data. Reparameterized models can be used to make predictions on the spatial locations for persons and/or objects not included in the analysis, with suitable assumptions. In addition, such reparameterizations can also be used to examine the impact of changed feature variables on a particular object's coordinates.
8. *Utility theory based*—The proposed methodology is based on an indirect utility function (McFadden, 1976) which posits choice as a function of levels of a latent unobservable variable, which in turn can be a function of object and person attributes. This will be demonstrated below.
9. *Nested models and associated tests*—The proposed methodology allows for the estimation of simple or weighted unfolding models, with or without coordinate reparameterization options. As such, a variety of nested models can be tested by means of asymptotic statistical significance tests to identify the "best" model.

The Spatial Choice Model

Let: $t = 1, ..., T$ dimensions (extracted in an MDS context),

$i = 1, ..., I$ persons,

$j = 1, ..., J$ objects,

$k = 1, ..., K$ object features,

$l = 1, ..., L$ person characteristics,

$y_{ij} = \begin{cases} 1 \text{ if person } i \text{ chooses object } j, \\ 0 \text{ otherwise,} \end{cases}$

$x_{jk} =$ the value of feature k for object j,

$z_{il} =$ the value of characteristic l for person i,

$p_{ij} =$ the probability person i chooses object j,

$a_{it} =$ the tth coordinate for person i,

$w_{it} =$ the salience of dimension t for person i,

$c_i =$ an additive constant for person i, and

$b_{jt} =$ the tth coordinate for object j.

It is assumed, as in most logit models, that y_{ij} is binomial $(1, p_{ij})$ and that there is independence across persons and objects. This is tantamount to a stochastic choice model which can be roughly categorized in Thurstone's (1927) Random Utility Model. It is also assumed that the choice process of person i choosing object j is Bernoulli, with probability of choice given by p_{ij}. Unlike the conditional logit (McFadden, 1976) and conditional probit (Hausman & Wise, 1978), there need not be the constraint that $\sum_j p_{ij} = 1$, because the choices for objects are modeled as being independent of each other. (By not requiring this constraint, the model can be used to accommodate choice situations where, for example, complementary or multiple-purchased objects/brands are bought, with correspondingly high probabilities. For example, large families may consistently buy the same two or more brands of breakfast cereals on the same purchase occasion.) The independence assumption across persons and objects presupposes that each choice is an independent ''coin toss'' with given probability p_{ij}. As such, the model is not restricted by considerations of dependence from irrelevant alternatives. Each choice is assumed to be made independently.

In many cases, the independence assumption across objects may be questionable. For example, in questionnaires where the order of presentation for objects is not randomized across persons, it can be legitimately argued that the choices are systematically not independent. Here, there can be order effects where the probability of a person choosing the kth object presented may be a function of the previous $k - 1$ objects presented and of the respective choice decisions made. (The effects of violations of this independence assumption are examined in the next section.)

The likelihood function is

$$L = \prod_{i=1}^{I} \prod_{j=1}^{J} p_{ij}^{y_{ij}} (1 - p_{ij})^{1-y_{ij}} \quad . \tag{1}$$

Converting to logs, the expression becomes

$$\Phi = \ln L = \sum_{i=1}^{I} \sum_{j=1}^{J} \left[y_{ij} \ln \left(\frac{p_{ij}}{1-p_{ij}} \right) + \ln (1 - p_{ij}) \right] \quad . \tag{2}$$

The spatial model is expressed as

$$\ln \left(\frac{1-p_{ij}}{p_{ij}} \right) = \sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i \quad , \tag{3}$$

a nonlinear logistic function with the ln of the odds on the left and the weighted unfolding relationship (Carroll, 1972) on the right,

where a_{it} is the ideal point for person i,

b_{jt} is the coordinate of object j, and

w_{it} is the salience or importance of person i for dimension t.

The proximity of an object to a person's ideal point indicates some degree of the magnitude of probability of choice. Thus, the closer an object is to an ideal point, the higher the probability of that person choosing/ buying that particular object.

There is more to this spatial choice model than these geometrical notions of proximity and choice. A latent, unobservable variable D_{ij} can be defined as

$$D_{ij} = \sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i + \mu_{ij} \quad , \tag{4}$$

where μ_{ij} is a stochastic error component. D_{ij} can be viewed as a latent disutility index. The latent variable D_{ij} is defined such that if $D_{ij} \leq d_i^*$ (some individual threshold value), then $y_{ij} = 1$; for $D_{ij} > d_i^*$, $y_{ij} = 0$. Thus, object j is selected by person i if it is not very far from that person's ideal notion of an object. Note that this threshold value varies to allow for individual differences between persons. Therefore,

$$P(y_{ij} = 1) = P[D_{ij} \leq d_i^*] = P\left[\sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i + \mu_{ij} \leq d_i^* \right]$$

$$= P\left[\mu_{ij} \leq -\sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 - c_i^* \right] = P(\mu_{ij} \leq -f_{ij}) \quad , \tag{5}$$

where

$$c_i^* = c_i - d_i \tag{6}$$

and

$$f_{ij} = \sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i^* \quad . \tag{7}$$

Thus, the estimated c_i^* term, assuming $\mu_{ij} = 0$, is an estimate of the threshold value (actually its negative). (Takane, 1983, presented a similar simple unfolding model where additive constants were estimated by object.) Similarly,

$$P(y_{ij} = 0) = P(D_{ij} > d_i^*) = 1 - P(\mu_{ij} \leq -f_{ij}) \quad . \tag{8}$$

Assuming that the stochastic error term has a logistic cumulative distribution function (see McFadden, 1976, for a discussion of this assumption in logit type models), then

$$P(y_{ij} = 1) = \frac{1}{1 + \exp(f_{ij})} \tag{9}$$

and

$$P(y_{ij} = 0) = \frac{\exp(f_{ij})}{1 + \exp(f_{ij})} \quad . \tag{10}$$

The likelihood function is then of the general form

$$L = \prod_{y_{ij} = 0} P(y_{ij} = 0) \prod_{y_{ij} = 1} P(y_{ij} = 1) \quad , \tag{11}$$

which can be shown to be equivalent to Expression 1. Expression 3 follows directly from a simple algebraic manipulation of Expressions 9 and 10.

In order to express the log likelihood as a function of the parameters a_{it} and b_{jt}, Equation 3 can be rewritten according to Expression 9 as

$$p_{ij} = \frac{1}{1 + \exp\left[\sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i^*\right]} \quad , \tag{12}$$

and the log likelihood (Equation 2) can be reexpressed as

$$\Phi = \ln L = \sum_{i=1}^{I} \sum_{j=1}^{J} \{(1 - y_{ij}) f_{ij} - \ln [1 + \exp(f_{ij})]\} \quad . \tag{13}$$

Estimates of the a_{it}s, w_{it}s, c_is, and b_{jt}s, given $Y = \|y_{ij}\|$ and T, are obtained by maximum likelihood methods that are discussed in the next section and the appendix.

The model defined in Equation 3 can be generalized to incorporate additional data in the form of object attributes and/or person background variables. The coordinates for objects or persons are reparameterized as linear functions of background variables (e.g., object features and person demographic variables). If object attribute data are available, b_{jt} can be reparameterized as $\sum_{k=1}^{K} x_{jk}\gamma_{kt}$, where x_{jk} is the value of feature k for object j, and γ_{kt} is the impact of feature k on dimension t. As in CANDELINC (Carroll, Pruzansky, & Kruskal, 1980), three-way multivariate conjoint analysis (DeSarbo, Carroll, Lehmann, & O'Shaughnessy, 1982), and GENFOLD2 (DeSarbo & Rao, 1984, 1986), the locations of objects can be modeled to be linear functions of their respective features. Thus, the x_{jk} are objectively quantified features which are related to subjective attributes (see Lancaster, 1966, 1971, 1979, for a discussion of his consumer demand theory based on demand for attributes). Similarly, when person background data are available, a_{it} can be reparameterized as $\sum_{l=1}^{L} z_{il}\alpha_{lt}$, where z_{il} is the value of characteristic l for person i and α_{lt} is the coefficient of the lth person characteristic on dimension t. When both object attribute and person background data are available, both person and object coordinates can be reparameterized.

Note that the simple unfolding model (Coombs, 1964) can be expressed as a special case of the weighted unfolding model by constraining $w_{it} = 1$, for all i,t. As Carroll (1972, 1980) demonstrated, the simple unfolding model can be viewed as nested within the weighted unfolding model.

It is also important to note the particular indeterminacies involved in each of the models discussed. Let:

$\mathbf{a}_i$ = the $T \times 1$ vector of coordinates for the ith person;
$\mathbf{W}^i$ = the $I \times T \times T$ diagonal array where the ith slice is a $T \times T$ diagonal matrix whose main diagonal elements are w_{it};
$\mathbf{b}_j$ = the $T \times 1$ vector of coordinates for the jth object.

Then,

$$f_{ij} = \sum_{t=1}^{T} w_{it}(a_{it} - b_{jt})^2 + c_i^* = (\mathbf{a}_i - \mathbf{b}_j)'\mathbf{W}^i(\mathbf{a}_i - \mathbf{b}_j) + c_i^* \tag{14}$$

defines the weighted unfolding model, and

$$f_{ij} = \sum_{t=1}^{T} (a_{it} - b_{jt})^2 + c_i^* = (\mathbf{a}_i - \mathbf{b}_j)'(\mathbf{a}_i - \mathbf{b}_j) + c_i^* \tag{15}$$

defines the simple unfolding model. For the weighted unfolding model, including all reparameterization options, there are T scalar indeterminacies, because it is possible to multiply the tth dimension of $\mathbf{A}$ and

B by some scalar and mulitply the tth dimension of $\mathbf{W}^i$ by the reciprocal of that constant squared and still obtain the same f_{ij}. In addition, for the non-reparameterized version, there are T indeterminacies, because it is possible to jointly center

$$\mathbf{L} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \tag{16}$$

by dimension and not alter the distances between the two sets of points (here, $\mathbf{A} = \|a_{it}\|$ and $\mathbf{B} = \|b_{jt}\|$). Thus, there are $2T$ total indeterminacies in the non-reparameterized weighted unfolding model.

The simple unfolding model suffers from even more indeterminacies. In addition to the T indeterminacies concerning centering $\mathbf{L}$ (for the non-reparameterized version) mentioned above, each set of coordinates in Expression 15 can be premultiplied by some orthogonal matrix $\mathbf{R}$ to obtain the same f_{ij}. Here, the likelihood function value is not altered by applying an orthogonal transformation (e.g., rotation) to both sets of coordinates. For example, it might be possible to improve interpretation by rotating the coordinates in $\mathbf{A}$ (or $\mathbf{B}$) using some orthogonal matrix $\mathbf{R}$.

The degrees of freedom of the model are defined as the effective number of model parameters. The rotational indeterminacy inherent in the simple unfolding model necessitates an adjustment to the degrees of freedom by subtracting T^2, the number of elements in $\mathbf{R}$ (actually, if $\mathbf{R}$ is orthogonal as discussed, only the upper triangle half of $\mathbf{R}$ need be counted as estimated parameters—subtract $T \times [(T-1)/2]$ from degrees of freedom), minus T more (in the non-reparameterized version) for the centering indeterminacies. In the weighted unfolding model, T must be subtracted from the degrees of freedom because of the scale indeterminacies, minus T more (in the non-reparameterized version) for the centering indeterminacies.

Model Options

Simple Versus Weighted Unfolding

The spatial choice methodology presented here allows the user to estimate either a simple unfolding (Coombs, 1964) or weighted unfolding (Carroll, 1972, 1980) model. The simple unfolding model assumes that a given distance on a dimension makes as much difference to one person as to another. The "iso-choice" contours, or contours of equiprobable choice, are circles (for $T = 2$), spheres (for $T = 3$), or hyperspheres (for $T > 3$).

In the weighted unfolding model, persons are allowed to weight the dimensions differently, and w_{it} can be thought of as the "salience" or "importance" of the tth dimension for person i. Here the iso-choice contours are ellipses (for $T = 2$), ellipsoids (for $T = 3$), or hyperellipsoids (for $T > 3$).

There is controversy in the literature over the desirability of constraining w_{it}s in the weighted unfolding model to be positive. Carroll (1972) claimed that, in the weighted unfolding model, a negative w_{it} has a clear interpretation: If w_{it} is negative with respect to dimension t, the ideal point for individual i indicates the least preferred, rather than the most preferred value, and the farther an object is along that dimension from the ideal point, the more highly preferred is the object. He thus argued for not constraining the w_{it} to be positive. Other authors, such as Srinivasan and Shocker (1973) and Davison (1976), disputed the utility of unconstrained analyses; they claimed it may lead to unreasonable predictions. The present spatial choice model allows for either unconstrained weights or nonnegative constrained weights. Note, however, that if the weights are constrained to be nonnegative, it would be necessary to reduce the degrees of freedom for the model if any weight parameters are on the boundary of the feasible region (i.e., when some $w_{it} = 0$).

Full or Reparameterized Models

The proposed spatial choice model allows reparameterization of either **A** or **B** (or both) to be linear functions of some known set of preselected descriptor variables (**Z** for persons, **X** for objects):

$$\mathbf{A} = \mathbf{Z}\alpha \quad , \tag{17}$$

$$\mathbf{B} = \mathbf{X}\gamma \quad . \tag{18}$$

As in CANDELINC (Carroll et al., 1980), GENFOLD2 (DeSarbo & Rao, 1984, 1986), and three-way multivariate conjoint analysis (DeSarbo et al., 1982), these reparameterizations can aid in the interpretation of the dimensions derived (see Bentler & Weeks, 1978; Bloxom, 1978; de Leeuw & Heiser, 1980; Noma & Johnson, 1977) and can replace the post-analyses property-fitting methods often used to interpret the solution obtained.

It should be noted that when a linear function replaces an object or person coordinate, the number of background variables in the linear function cannot exceed the number of entities that exist for those variables. For example, if J objects have K attributes, $J \geq K$, because only JT coordinates can possibly be identified (excluding rotational indeterminacy). Similarly, if I persons have L background variables, $I \geq L$, because only IT coordinates can possibly be identified (excluding rotational indeterminacy). Thus, in most applications, such reparameterizations actually improve the degrees of freedom (i.e., by reducing the number of parameters to be estimated).

Internal Versus External Analyses

The proposed methodology allows the user to specify coordinate values for objects (**B**) and/or persons (**A**) obtained from, say, a previous analysis or some theory (external analysis in the terminology of Carroll, 1972). Alternatively, the methodology allows for the estimation of both these sets of points (i.e., internal analysis).

Centroid Restrictions on the Matrices A or B

Because of potential degenerate solutions common to unfolding, options exist to constrain objects (ideal points) to be at the centroid of the locations of the ideal points (objects) that selected (were selected by) them. Heiser (1981) developed such a set of restrictions in his two-way preference analysis for only one set of points (objects). The present methodology allows such restrictions on either set of coordinates.

Table 1 depicts some of the major types of models that can be estimated using the proposed spatial choice methodology. Many of these various models are nested within others. Because maximum likelihood procedures are used to estimate parameters, nested and full models can theoretically be statistically tested (asymptotically) against each other in order to obtain a more parsimonious model.

The Algorithm

Maximum likelihood methods are used to maximize Φ (or minimize $-\Phi$) in Equation 2 with respect to the given set of unknown parameters specified in one of the various models selected by the user. Because Equation 2 is nonlinear, the method of conjugate gradients is used to estimate the desired set of parameters in an alternating fashion. The computational details of the algorithm are presented in the Appendix. Because Φ in Equation 2 has an upper bound of zero, and because each estimating stage of the maximum likelihood can be shown to conditionally increase Φ, a limiting sums argument (Courant, 1965) can be used to prove convergence to at least a locally optimum solution.

Table 1
Different Unfolding Model
Specification Options

Unfolding Model
1. Simple
2. Weighted
a. w_{it} unconstrained
b. w_{it} constrained

Type of Analysis
1. Internal
2. External
a. $\underset{\sim}{A}$ (or $\underset{\sim}{\alpha}$) given
b. $\underset{\sim}{\tilde{B}}$ (or $\underset{\sim}{\tilde{\gamma}}$) given
c. $\underset{\sim}{\tilde{A}}$ (or $\underset{\sim}{\alpha}$) and $\underset{\sim}{B}$ (or γ) given

Reparameterization Option
1. None
2. $\underset{\sim}{A} = Z\underset{\sim}{\alpha}$
3. $\underset{\sim}{\tilde{B}} = \tilde{X}\tilde{\gamma}$
4. $\underset{\sim}{\tilde{A}} = \underset{\sim}{\tilde{Z}}\underset{\sim}{\alpha}$ and $\underset{\sim}{B} = \underset{\sim}{X}\gamma$

Four goodness-of-fit measures are computed for the spatial choice model:

1. The ln likelihood function Φ;
2. Deviance measure (McCullagh & Nelder, 1983; Nelder & Wedderburn, 1972):

$$D = -2\left[\sum_{i=1}^{I}\sum_{j=1}^{J}y_{ij}\ln(\hat{p}_{ij}) + (1 - y_{ij})\ln(1 - \hat{p}_{ij})\right] = -2\Phi \quad , \tag{19}$$

where $\hat{p}_{ij}$, the estimated probability, is expressed as in Equation 9 using estimated values for a_{it}, w_{it}, and b_{jt}. Note that nested models can *theoretically* be tested as the difference between respective deviance measures, which is asymptotically χ^2 distributed with the difference in model degrees of freedom as the appropriate χ^2 test degrees of freedom. This is appropriate in testing dimensionality and various models, because of the obvious nesting. One question that arises with the use of maximum likelihood procedures here concerns the validity of much of the statistical properties of estimators and associated tests. This question is of particular importance with respect to the spatial choice model presented, due to the presence of incidental parameters (e.g., a_i, b_j, c_i^*) whose number varies according to the order of Y. According to Andersen (1980), maximum likelihood estimators in such cases may not be consistent. Andersen suggested the use of conditional maximum likelihood estimators in such cases. Potential problems arising from use of the χ^2 test for various nested models are investigated below.

3. Sums of squares:

$$\text{SSQ} = \sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij} - \hat{p}_{ij})^2 \quad ; \tag{20}$$

4. Point-biserial correlation between Y and $\hat{P}$: r_{pb} $(Y,\hat{P})$, where $\hat{P} = \|\hat{p}_{ij}\|$. A point-biserial correlation between Y and arcsin $(\hat{P})$ might also be considered. Note that the last three measures are also given for row and column objects to examine goodness of fit for each person and object.

Synthetic Data Results

A synthetic data set was employed to illustrate the relationship between a set of choice data and its spatial representation. The number of objects was limited to six, with the expectation that the solution space would be of sufficiently low dimensionality to facilitate visual examination and interpretation. The choice patterns were repeated in order to check certain properties of the algorithm. The spatial interrelationship between these patterns could, therefore, be examined for consistency and interpretability.

Synthetic data were constructed for $I = 30$ persons evaluating $J = 6$ objects by repeating the patterns in Table 2 five times. Each (hypothetical) person's choices can be represented by one of these 6 patterns, where 1 indicates that an object was chosen by that person and 0 that it was not chosen. Note that the patterns for A, B, and C are the same as for D, E, and F, respectively. It was desired to examine the performance of the algorithm in (1) determining the dimensionality; (2) producing a reasonable spatial representation of such data; (3) examining whether persons with the same choice pattern had the same coordinates (local minimum solutions are possible with such an algorithm, which might render different coordinates for persons with the same choice values); (4) examining whether objects always selected together had the same coordinates; (5) examining the relationship of persons and the objects they selected; and (6) investigating what happens when the independence assumption is violated (the patterns for A, B, and C are the same for D, E, and F respectively, and each group of six persons has the same choice patterns).

Table 3 represents the statistical summary for the simple unfolding model for $T = 1$ and $T = 2$. Note that the deviance measure for going from $T = 1$ to $T = 2$ is significant, indicating that two dimensions is the appropriate dimensionality. Even if this test statistic were ignored here, the other measures clearly verify that these data are appropriately represented by two dimensions. Figure 1 portrays the rotated two-dimensional joint space for such data, where the letters A–F are used to label objects and the numbers 1–30 are used to label person ideal points.

The configuration in Figure 1 is a parsimonious spatial representation of the synthetic choice data. Each pattern can be identified by a cluster of ideal points representing the persons who have that choice pattern. Persons with the same pattern have clustered very tightly together in the space. The positioning of the objects relative to the ideal point clusters clearly reflects the choices made by a person. Note how well the points are aligned so that the objects A–F are closest to those persons that solely chose A, B, C, D, E, and/or F respectively. Those persons that chose multiple objects are represented by ideal points located between the two objects they chose. For example, those persons that chose objects B and C have ideal points in the quadrant between objects B and C (indicating close proximity of B and C to their ideal points) and are far from those persons who did *not* choose B and C, but chose solely A. Similarly, objects

Table 2
Synthetic Data Choice Patterns

		Object					
		A	B	C	D	E	F
	1.	1	0	0	1	0	0
	2.	0	1	0	0	1	0
	3.	0	0	1	0	0	1
Pattern	4.	1	1	0	1	1	0
	5.	1	0	1	1	0	1
	6.	0	1	1	0	1	1

Table 3
Summary of the Analysis for the Simulated Data

| T	$\Phi = \ln L$	Deviance	SSQ	r_{pb}	Diff D
1	-124.64	249.27	44.93	0.06	--
2	-0.00	0.00	0.00	1.00	249.27***

*** $\alpha < .01$

with identical patterns of choice ([A,D], [B,E], [C,F]) are located quite close to each other. Thus, individual preferences can be assessed quickly and comprehensively.

The analysis for the synthetic data was also performed using a different random start for the parameter estimates. The goal was to examine the stability of the resulting solutions obtained from the algorithm and to check for potential locally optimum solutions. Canonical correlation was used as an approximate configuration matching technique to examine the similarity of the two joint space solutions. Substantial congruence was found between the two solutions, with canonical correlations of .99 and .98.

While further monte carlo work is required to fully investigate the appropriateness of the χ^2 test for nested models with finite samples, this example does provide some evidence that the test renders reliable results for this one data set. Clearly, more extensive monte carlo analyses are required to investigate the performance of the algorithm, as well as the appropriateness of these statistical distributions, especially in light of incidental parameters and no replications.

Figure 1
Rotated Joint Space Plot for Synthetic Data
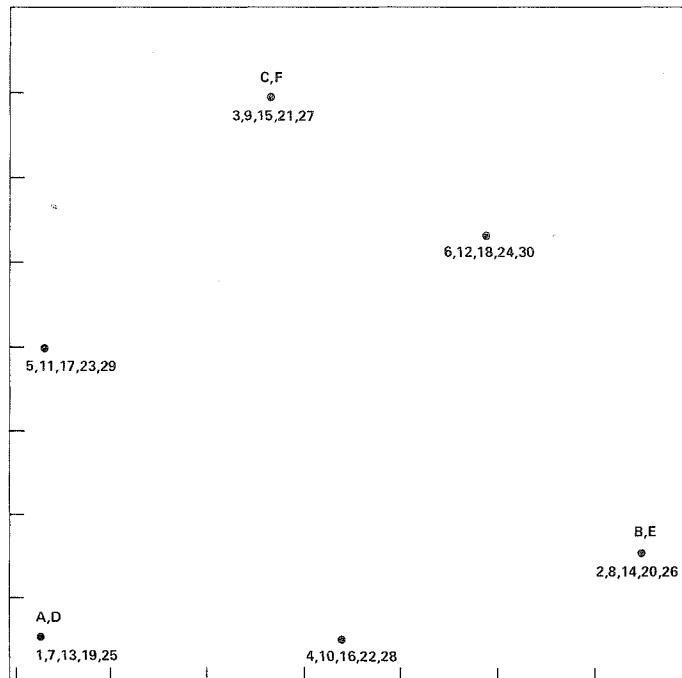
Application

Green and DeSarbo (1980) reported choice data collected from 35 MBA students at the University of Pennsylvania. Each student was asked to list all foreign sports cars that came to mind that he or she would like to own if money were not a major constraint. No specified number or list of choices was provided, so that this is "pick any" data. In addition, for each chosen car, the students were asked to pick as few or as many attributes as desired from a specified list of 22, shown in Table 4, that were most highly associated with that choice. Table 4 also lists the nine sports cars whose names appeared most frequently across the 35 students.

Green and DeSarbo applied Levine's (1979) model to the data, and the first three eigenvalues (dimensions) accounted for 55% of the variance. In addition to a constant vector always produced as the first dimension by this method, the two other dimensions identified were a "value" dimension separating cars of lower price (e.g., MG) from those of higher price (e.g., BMW), and an "appearance" dimension that separated highly elegant cars such as Ferrari and Jaguar from the most functional cars, such as MG and Porsche.

Spatial Choice Model Results

The 35×9 choice data were run in 1–4 dimensions using the simple unfolding model without degeneracy constraints. Table 5 presents the statistical results of the internal analyses which showed that three dimensions appear to be the appropriate dimensionality for fitting this model (this result was also

Table 4
Sports Car Names Mentioned Most Frequently
and Prespecified Car Attributes

Sports Cars	
	Plotting Codes
Mercedes 450SL	A
Triumph	B
Datsun 280ZX	C
BMW	D
Fiat	E
Ferrari	F
MGB	G
Jaguar SJ6	H
Porsche	I

| Attributes	
Economical to maintain	Plush interior
High acceleration	High reliability
High braking ability	Excellent cornering
Good for city driving	Sleek, racy lines
Easy-to-get parts	Conservative styling
Good gas mileage	Well engineered
High durability	Good service facilities nearby
Beautiful lines	Luxurious styling
High resale value	Not temperamental
Comfortable on long trips	Prestigious
High speed	Low cost

Table 5
Summary of the Analyses on
the Sports Car Data

T	$\Phi = \ln L$	Deviance	SSQ	r_{pb}	Diff D
1	−107.041	214.083	35.366	0.605	−
2	−76.451	152.901	24.636	0.746	61.182**
3	−10.051	20.101	2.496	0.978	132.800***
4	−1.194	2.389	0.044	0.999	17.712

*** $\alpha < .01$
** $\alpha < .025$

verified by using a different random start for the initial parameter estimates). Again, even if the χ^2 test were to be disregarded because of incidental parameters, the other measures presented in Table 5 clearly indicate that $T = 3$ dimensions is the most parsimonious solution. These three dimensions account for over 95% of the variance in $\mathbb{Y}$ as measured by $r_{pb}^2(\mathbb{Y}, \mathbb{P})$.

Figure 2 presents the projections of the products (labeled by the letters A–I) and ideal points (labeled as "*") on each of the three dimensions (i.e., the product and person coordinates). By computing correlations between these three dimensions and the aggregated frequencies of the 20 attributes (reparameterizing the brands is not appropriate here, because $K > J$), some insight can be gained into the interpretation of these three dimensions, and tentative labels can be created for each of the three dimensions which attempt to summarize many of the items which display higher correlations with these derived dimensions.

Figure 2
Plot of the Three Dimensions
for the Sports Car and Ideal Point Coordinates
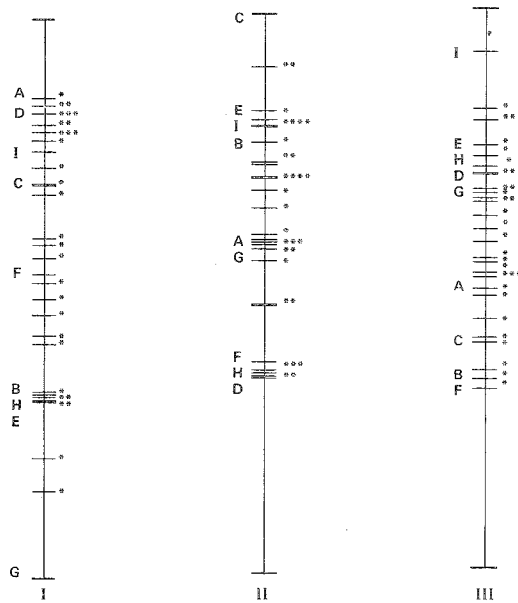
The first dimension correlates highly with *beautiful lines* (.694), *comfortable on long trips* (.853), *prestigious* (.591), *plush interior* (.519), and *well-engineered* (.613), and negatively with *low cost* ($-.497$). Mercedes, BMW, and Porsche load positively while MG, Triumph, and Fiat load negatively. This can be labeled the "quality" or "elegance" dimension. The one-dimensional plot indicates a heavy concentration of ideal points around brands occupying the positive part of this dimension, especially BMW. There are small groups of students near Triumph, Jaguar, and Fiat. MG is positioned quite far from the locus of ideal points.

The second dimension correlates highly with *good for city driving* (.587), *easy to get parts* (.725), *good service nearby* (.778), and *good gas mileage* (.446). Datsun, Fiat, Triumph, and Porsche load positively while BMW, Ferrari, and Jaguar load negatively. This can be labeled the "practicality" dimension. Ideal points on this dimension tend to be widely dispersed on the scale with some concentrations of points near Triumph, Mercedes, and MG, and the BMW-Ferrari-Jaguar group.

The third dimension correlates highly with *high acceleration* (.469), *high braking ability* (.517), *high speed* (.506), *high durability* (.573), *high resale value* (.560), *excellent cornering* (.587), *well-engineered* (.598), and *not temperamental* (.729). Porsche, Jaguar, and Fiat load positively while Triumph and Datsun load negatively. This can be labeled the "performance" dimension. On this dimension, ideal points appear to be most concentrated near the Jaguar-BMW-MG group, although many exist between MG and Mercedes where none of the nine car brands currently exist.

A final note of interest concerns an examination or comparison of students with the same choice patterns. Seven different choice patterns were shared by two or more students in the sample. An examination of the joint space plots (not shown) revealed considerable congruence for ideal points representing students with the same choice pattern. For example, Students 11, 27, and 31 shared the same choice pattern and also had ideal points in the joint space plots that were quite close to each other. Also, Students 4 and 32 shared the same choice pattern, as did Students 7 and 25, Students 14, 24, and 33, Students 15 and 30, Students 18 and 21, and Students 23 and 34. Again, there was substantial congruence in coordinates for students sharing the same choice patterns.

Discussion

The spatial choice model presented can be used in a variety of ways to analyze binary data. In addition to analyzing "pick any/n" or "pick any" data as described in the application, the spatial choice model could also provide a spatial representation of binary profile data (e.g., objects having or not having certain attributes or benefits). Another application would be to record items purchased in a grocery shopping trip by different people to investigate the issue of complementary goods. This model can also be used to investigate sociological networks where respondents denote choices for friends out of particular social groups. Another application concerns the use of the methodology to investigate potential job choices. Clearly, other behavioral applications exist involving choice data.

Perhaps the greatest potential for application concerns the use of the reparameterized models where object and/or person coordinates are constrained to be specified functions of associated background variables (e.g., for objects—features, attributes; for persons—demographics, psychographics). As in GENFOLD2 (DeSarbo & Rao, 1984, 1986), these reparameterizations can aid in interpreting the dimensions derived by examining how these background variables for both persons and objects load on them.

A number of interesting research directions have been suggested by the development of this procedure for spatially representing binary data. One extension would be to model the log of the odds as a vector model instead of a distance model. The ideal point model in Expression 3 may not be appropriate for many applications where "the more the better" type of utility functions affect individuals' underlying dimensions/bases of choice. Here, the vector model (Carroll, 1972) could be modeled to provide a different spatial representation. Such work has been investigated by DeSarbo and Keramidas (1984). Further monte

carlo work must be performed on all models under a diversity of conditions to test the algorithm for local minimum or other potential problems, such as other types of violations to the independence assumption. The statistical properties of the estimates must be evaluated by way of bootstrapping and using second derivative information. Also, the small sample properties of the χ^2 test must be more rigorously examined, as should the consistency of the estimated parameters. Finally, work is needed to generalize the methodology to handle three-way choice arrays where, for example, choices are obtained from persons for products purchased/chosen over time or for different situations.

Appendix

The Maximum Likelihood Algorithm for the Unfolding Spatial Choice Model

Phase I: Input Options

The user must specify T, the number of dimensions for the analysis; IT*, the maximum number of major iterations allowed; MIT*, the maximum number of minor iterations allowed; TOL, the convergence tolerance; the type of unfolding analysis requested (simple or weighted); restrictions on $\mathbf{W} = \|w_{it}\|$ if estimated (none or positivity constraints); whether $\mathbf{A} = \|a_{it}\|$ is reparameterized via $\mathbf{A} = \mathbf{Z}\alpha$; whether $\mathbf{B} = \|b_{jt}\|$ is reparameterized via $\mathbf{B} = \mathbf{X}\gamma$; whether centroid restrictions are requested on $\mathbf{A}$; whether centroid restrictions are required on $\mathbf{B}$; and the type of starting estimates. The matrices $\mathbf{Y} = \|y_{ij}\|$, $\mathbf{X} = \|x_{jk}\|$, and $\mathbf{Z} = \|z_{il}\|$ must also be input, depending upon which model is to be fit.

Phase II: Starting Estimates

The user has the choice of one or two methods to obtain starting estimates of $\mathbf{A}$ (or α), $\mathbf{B}$ (or γ), $\mathbf{W}$, and $\mathbf{c} = \|c_i^*\|$, either random starting estimates or given starting estimates. Options exist to perform external analyses where $\mathbf{A}$ and/or $\mathbf{B}$ is given and held fixed throughout the analyses. Set IT $= 0$.

Phase III: Estimation of the Matrices A (or α) and B (or γ)

Set IT $=$ IT $+ 1$. This phase of the algorithm conditionally estimates $\mathbf{A}$ (or α) and $\mathbf{B}$ (or γ), holding $\mathbf{W}$ and $\mathbf{c}$ set at their current values. Conditional estimates of $\mathbf{A}$ (or α) and $\mathbf{B}$ (or γ) are sought to maximize the log of the likelihood function in Expression 13. The method of conjugate gradients (Fletcher & Reeves, 1964) is used to solve this unconstrained optimization procedure. Assuming the likelihood function in Expression 13 obeys typical continuity and regularity conditions, partial derivatives of Expression 13 can be taken with respect to the desired set of parameters:

$$\frac{\partial \Phi}{\partial b_{jt}} = 2 \sum_{i=1}^{I} w_{it}(b_{jt} - a_{it}) \left[1 - y_{ij} - \frac{\exp(f_{ij})}{1 + \exp(f_{ij})} \right] , \tag{A1}$$

$$\frac{\partial \Phi}{\partial a_{it}} = -2 \sum_{j=1}^{J} w_{it}(b_{jt} - a_{it}) \left[1 - y_{ij} - \frac{\exp(f_{ij})}{1 + \exp(f_{ij})} \right] , \tag{A2}$$

$$\frac{\partial \Phi}{\partial \gamma_{kt}} = 2 \sum_{i=1}^{I} \sum_{j=1}^{J} w_{it}(b_{jt} - a_{it}) x_{jk} \left[1 - y_{ij} - \frac{\exp(f_{ij})}{1 + \exp(f_{ij})} \right] , \tag{A3}$$

$$\frac{\partial \Phi}{\partial \alpha_{lt}} = -2 \sum_{i=1}^{I} \sum_{j=1}^{J} w_{it}(b_{jt} - a_{it}) z_{il} \left[1 - y_{ij} - \frac{\exp(f_{ij})}{1 + \exp(f_{ij})} \right] . \tag{A4}$$

For external analyses, the relevant set of partial derivatives is set equal to zero. (See Heiser, 1981, for a method of imposing centroid constraints on a set of stimuli coordinates in a joint space analysis.)

For the sake of convenience, assume that the relevant parameters to be estimated are contained in the vector θ and that $\nabla\Phi$ is the vector of relevant partial derivatives for this desired set of parameters. Then the complete conjugate gradient procedure used to minimize $-\Phi$ can be summarized (see Rao, 1979, for further details) as follows:

1. Start with initial parameter estimate θ^1; set MIT $= 1$.
2. Set the first search direction $S^1 = -\nabla\Phi^{(1)}$.
3. Find $\theta^{(2)}$ according to the relation

$$\theta^{(2)} = \theta^{(1)} + \mu^{(1)}S^{(1)} \quad , \tag{A5}$$

    where $\mu^{(1)}$ is the optimal step length (obtained through quadratic interpolation) in the direction $S^{(1)}$. Set MIT $= 2$.
4. Calculate $\nabla\Phi^{(2)}$ and set

$$S^{(MIT)} = -\nabla\Phi^{(MIT)} + \frac{[\nabla\Phi^{(MIT)}]'[\nabla\Phi^{(MIT)}]}{[\nabla\Phi^{(MIT-1)}]'[\nabla\Phi^{(MIT-1)}]} S^{(MIT-1)} \quad . \tag{A6}$$

5. Compute the optimal step length $\mu^{(MIT)}$ in the direction $S^{(MIT)}$, and find

$$\theta^{(MIT+1)} = \theta^{(MIT)} + \mu^{(MIT)}S^{(MIT)} \quad . \tag{A7}$$

6. If $\theta^{(MIT+1)}$ is optimal and/or satisfies the minor iteration tests below, stop. Otherwise set MIT $=$ MIT $+ 1$ and go to step 4 above (i.e., undertake another minor iteration). A number of convergence tests are performed in this minor estimation cycle to test whether additional iterations of the conjugate gradient procedure are necessary or beneficial.

> If MIT $\geq$ MIT*, stop;
> If $\|\nabla\Phi^{(MIT)}\| \leq$ TOL, stop;
> If $|\Phi^{(MIT)} - \Phi^{(MIT-1)}| \leq$ TOL, stop.

It has been shown that conjugate gradient procedures can avoid the typical "cycling" often encountered with steepest descent algorithms. In addition, they demonstrate valuable quadratic terminations (Himmelblau, 1972) properties—that is, conjugate gradient procedures will typically find the globally optimum solution for a quadratic loss function in H steps, where H is the number of parameters to be solved for. This is relevant for the present application only if a Taylor expansion of $\ln L$ around the optimum is dominated by its second order term.

Phase IV: Estimation of the Matrices W and c

This phase of the algorithm conditionally estimates W (if a weighted unfolding model is stipulated) and c, holding A (or α) and B (or γ) at their current values estimated from the previous phase. The partial derivatives are

$$\frac{\partial\Phi}{\partial w_{it}} = \sum_{j=1}^{J}(b_{jt}-a_{it})^2\left[1-y_{ij} - \frac{\exp(f_{ij})}{1+\exp(f_{ij})}\right] \tag{A8}$$

and

$$\frac{\partial\Phi}{\partial c_i^*} = \sum_{j=1}^{J}\left[1-y_{ij} - \frac{\exp(f_{ij})}{1+\exp(f_{ij})}\right] \quad , \tag{A9}$$

and the same conjugate gradient procedure as in Phase III is used. (If positivity constraints are placed on w_{it}, then the suggestion by Gill, Murray, & Wright, 1981, can be used to estimate $\lambda_{it} = w_{it}^2$.)

Phase V: Major Iteration Tests

A number of convergence tests are performed after a complete major iteration cycle (completing phases I–IV) to test if additional computation is necessary:

If IT $\geq$ IT*, go to Phase VI;
If $|\Phi^{(IT)} - \Phi^{(IT-1)}| <$ TOL, go to Phase VI; otherwise, go to Phase III.

Phase VI: Output and Diagnostics

All parameter values with final IT value, Φ, deviance, sum of squares, and point-biserial correlation statistics are printed for the complete analysis, as well as by person and by product. Correlations between dimensions are output within coordinate sets for $\mathbf{A}$ (and/or $\boldsymbol{\alpha}$), $\mathbf{B}$ (and/or $\boldsymbol{\gamma}$), and $\mathbf{W}$.

References

Andersen, E. B. (1980). *Discrete statistical models with social science applications*. New York: North Holland.

Bentler, P. M., & Weeks, D. G. (1978). Restricted multidimensional scaling models. *Journal of Mathematical Psychology, 17*, 138–151.

Benzecri, J. P. (1973). *L'analyse des données. 1. La taxonomie, 2. L'analyse des correspondances*. Paris: Dunod.

Bloxom, B. (1978). Constrained multidimensional scaling in *n* spaces. *Psychometrika, 43*, 397–408.

Carroll, J. D. (1972). Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences, Vol. 1: Theory* (pp. 105–155). New York: Seminar Press.

Carroll, J. D. (1980). Model and methods for multidimensional analysis of preferential choice (or other dominance data). In E. D. Lantermann & H. Feger (Eds.), *Similarity and choice* (pp. 234–289). Bern: Hans Huber.

Carroll, J. D., Pruzansky, S., & Kruskal, J. B. (1980). CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika, 45*, 3–24.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika, 40*, 5–32.

Coombs, C. H. (1964). *A theory of data*. New York: Wiley.

Courant, R. (1965). *Differential and integral calculus, Vol. I* (2nd ed.). New York: Wiley.

Davison, M. L. (1976). Fitting and testing Carroll's weighted unfolding model for preferences. *Psychometrika, 41*, 233–247.

de Leeuw, J. (1984). *Canonical analysis of categorical data*. Leiden, the Netherlands: DWSO Press.

de Leeuw, J., & Heiser, W. (1980). Multidimensional scaling with restrictions on the configuration. In P. R. Krishnaiah (Ed.), *Multivariate analysis, Vol. 5* (pp. 501–522). New York: North Holland.

DeSarbo, W. S., Carroll, J. D., Lehmann, D. R., & O'Shaughnessy, J. (1982). Three-way multivariate conjoint analysis. *Marketing Science, 1*, 323–350.

DeSarbo, W. S., & Keramidas, E. M. (1984). *A new multidimensional scaling methodology for the spatial representation of binary choice data*. Unpublished manuscript, Bell Laboratories, Murray Hill NJ.

DeSarbo, W. S., & Rao, V. R. (1984). GENFOLD2: A set of models and algorithms for the GENeral unFOLDing analysis of preference/dominance data. *Journal of Classification, 1*, 147–186.

DeSarbo, W. S., & Rao, V. R. (1986). A constrained unfolding model for product positioning analysis. *Marketing Science, 5*, 1–19.

Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *Computer Journal, 7*, 149–154.

Gifi, A. (1981a). *Nonlinear multivariate analysis*. Department of Data Theory, University of Leiden, The Netherlands.

Gifi, A. (1981b). *Homogeneity analysis*. Department of Data Theory, University of Leiden, The Netherlands.

Gill, P. E., Murray, W., & Wright, M. H. (1981). *Practical optimization*. New York: Academic Press.

Green, P. E., & DeSarbo, W. S. (1980). Two models

for representing unrestricted choice data. In K. B. Monroe (Ed.), *Advances in consumer research proceedings*. Ann Arbor MI: Association for Consumer Research.

Greenacre, M. J. (1984). *Theory and applications of correspondence analysis*. London: Academic Press.

Hausman, J. A., & Wise, D. A. (1978). A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica, 46,* 403–426.

Heiser, W. J. (1981). *Unfolding analysis of proximity data*. Unpublished doctoral dissertation, University of Leiden, The Netherlands.

Himmelblau, D. M. (1972). *Applied nonlinear programming*. New York: McGraw-Hill.

Hoffman, D., & Franke, G. (1986). Correspondence analysis: Graphical representation of categorical data in marketing research. *Journal of Marketing Research, 23,* 213–227.

Holbrook, M. B., Moore, W. L., & Winer, R. (1980). Using "pick any" data to represent competitive positions. ORSA/TIMS *Marketing Measurement Proceedings* (pp. 129–134). Providence RI: Institute of Management Science.

Holbrook, M. B., Moore, W. L., & Winer, R. (1982). Constructing joint spaces from "pick any" data: A new tool for consumer analysis. *Journal of Consumer Research, 9,* 99–105.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T. (1985). *The theory and practice of econometrics* (2nd ed.). New York: Wiley.

Kruskal, J. B., & Shepard, R. N. (1974). A nonmetric variety of linear factor analyses. *Psychometrika, 39,* 123–157.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy, 74,* 132–157.

Lancaster, K. (1971). *Consumer demand: A new approach*. New York: Columbia University Press.

Lancaster, K. (1979). *Variety, equity, and efficiency*. New York: Columbia University Press.

Lebart, L., Morineau, A., & Warwick, K. M. (1984). *Multivariate descriptive statistical analysis*. New York: Wiley.

Levine, J. H. (1979). Joint-space analysis of "pick any" data: Analysis of choices from an unconstrained set of alternatives. *Psychometrika, 44,* 85–92.

Lingoes, J. C. (1972). A general survey of the Guttman-Lingoes nonmetric program series. In R. N. Shepard, A. K. Romney, & S. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences, Vol. 1: Theory* (pp. 52–68). New York: Seminar Press.

McCullagh, P., & Nelder, J. A. (1983). *Generalized linear models*. New York: Chapman and Hall.

McFadden, D. (1976). Quantal choice analysis: A survey. *Annals of Economic and Social Measurement, 5,* 363–390.

Muthén, B. (1981). Factor analysis of dichotomous variables: American attitudes toward abortion. In D. N. Jackson & E. F. Borgatta (Eds.), *Factor analysis and measurement in sociological research* (pp. 114–136). Beverly Hills CA: Sage Press.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, A, 135,* 370–384.

Nishisato, S. (1980). *Analysis of categorical data: Dual scaling and its applications*. Toronto: University of Toronto Press.

Noma, E., & Johnson, J. (1977). *Constraining nonmetric multidimensional scaling configurations* (Technical Report No. 60). Ann Arbor: University of Michigan, Human Performance Center.

Rao, S. S. (1979). *Optimization: Theory and applications*. New York: Wiley.

Srinivasan, V., & Shocker, A. D. (1973). Linear programming techniques for multidimensional analysis of preferences. *Psychometrika, 38,* 337–369.

Takane, Y. (1983). Choice model analysis of the "pick any/*n*" type of binary data. *Handout for presentation at the European meetings of the Psychometric and Classification Societies*. July, Jouy-en-Josas, France.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34,* 273–286.

Thurstone, L. L. (1929). Theory of attitude measurement. *Psychological Review, 36,* 222–241.

Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: Wiley.

Wind, J. (1982). *Product policy: Concepts, methods, and strategy*. Reading MA: Addison-Wesley.

Acknowledgments

Author's Address

Send requests for reprints or further information to Wayne S. DeSarbo, Marketing Department, Wharton School, University of Pennsylvania, Philadelphia PA 19104, U.S.A.