Available online at www.sciencedirect.com
ScienceDirect
Journal of Mathematical Psychology 51 (2007) 305–318
Journal of Mathematical Psychology
www.elsevier.com/locate/jmp

Internal multidimensional unfolding about a single-ideal—A probabilistic solution

David B. MacKay*

Kelley School of Business, Indiana University, 1309 East Tenth Street, Bloomington, IN 47405-1709, USA

Received 4 August 2006; received in revised form 10 April 2007
Available online 2 July 2007

Abstract

A solution is presented for an internal multidimensional unfolding problem in which all the judgments of a rectangular proximity matrix are a function of a single-ideal object. The solution is obtained by showing that when real and ideal objects are represented by normal distributions in a multidimensional Euclidean space, a vector of distances among a single-ideal and multiple real objects follows a multivariate quadratic form in normal variables distribution. An approximation to the vector's probability density function (PDF) is developed which allows maximum likelihood (ML) solutions to be estimated. Under dependent sampling, the likelihood function contains information about the parametric distances among real object pairs, permitting the estimation of single-ideal solutions and leading to more robust multiple-ideal solutions. Tests for single- vs. multiple-ideal solutions and dependent vs. independent sampling are given. Properties of the proposed model and parameter recovery are explored. Empirical illustrations are also provided.

Keywords: Probabilistic scaling; Thurstonian models; Multidimensional scaling; PROSCAL

1. Introduction

Internal unfolding models are commonly assumed to require the estimation of parameters for multiple-ideal objects. In this paper, it is shown that this assumption is incorrect and that multidimensional solutions about a single-ideal object may be derived when the data are characterized by dependent sampling.

Marketing and sensory analysis are two application areas in which the ability to estimate single-ideal solutions is essential. A common question in both areas is whether consumers or subjects should be modeled as being unique, as being members of a relatively small number of segments, or as being elements of one undifferentiated market or set. Marketing and product development decisions flow from the answer to this segmentation question (MacKay, 2005, 2006; MacKay, Easley, & Zinnes, 1995). If internal unfolding analysis is used to answer this question, a single-ideal solution must be able to be estimated to determine if consumers or subjects should be modeled as being elements of one undifferentiated market or set. Lacking an ability to estimate single-ideal solutions from liking rating or liking ranking data (the most common data used in internal unfolding), analysts usually bring in similarity or dissimilarity data and do what is termed an external analysis. The problem posed by external analysis is that dissimilarity judgments may be based on attributes (for example, product color and smell) that are different from those used in product liking (such as brand identification and taste).

First proposed by Coombs (1950), unfolding models were initially used to represent a rectangular matrix of preferences by p subjects for n objects as distances between p ideal objects and n real objects by estimating the coordinates of the objects in a latent space. Rank order data, n judgments for each subject, formed the initial set of experimental data but other types of preference data, such as liking ratings and paired comparisons, were soon investigated as well. A large distance between a real and an ideal object indicated that the real object had a high disutility. (To relate spatial model parameters to data it is convenient to refer to disutility instead of utility, liking or preference.) In 1958, Coombs made use of Thurstone's

*Fax: +1 812 855 6440.
E-mail address: mackay@indiana.edu

0022-2496/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2007.05.002

(1927) comparative judgment model when he conceptualized a unidimensional probabilistic version of the unfolding model in which the real and ideal objects were represented not by points but by distributions. Zinnes and Griggs (1974) proposed a multidimensional version of Coombs' probabilistic unfolding model. Making use of Hefner's (1958) realization that squared distances in a Thurstonian framework follow a noncentral chi-square distribution, the Zinnes and Griggs model was estimated from binary choice probabilities, required equal variances for all objects on all dimensions, and assumed independent sampling. For binary choice probabilities, independent sampling assumed that when a subject evaluated a pair of objects, the subject drew two samples from an ideal distribution and used one sample to compute the distance to the first real object and the other sample to compute the distance to the second real object. (See Böckenholt, 2006 for a recent review of Thurstonian models.)

An unfolding model based upon preference ratio data, in which subjects indicated the degree to which an object in a stimulus pair was preferred, was proposed by MacKay and Zinnes (1995). It relaxed the Case V (equal variances on all objects) and isotropic space (equal variances on all dimensions) assumptions of the Zinnes and Griggs model and permitted the estimation of a single-ideal object solution in a multidimensional space (MacKay et al., 1995) without requiring the collection of similarity or dissimilarity data, as was necessary with external ideal point models (Carroll, 1972). The key to estimating a single-ideal solution was dependent sampling which assumed that when a subject provided a preference ratio judgment, the ideal distribution was sampled only once and those multidimensional coordinates were then compared to both of the real objects in the preference ratio. Dependent sampling also permitted the unique estimation of variances for ideal and real objects. Multidimensional dependent sampling models for discrimination, identification, and preferential choice have been reported by Ennis (1993).

While preference ratio judgments provide the promise of permitting the estimation of single-ideal unfolding models, applications involving preference ratios are few. Monadic judgments of n objects, the problem originally addressed by Coombs, are much more common but the estimation of single-ideal unfolding models from monadic judgments encounters severe indeterminacy problems. Schönemann and Wang (1972), describing their metric unfolding model, state that in an r-dimensional space a minimum of $r+1$ ideal points are needed to obtain a unique solution. Coxon (1982), referring to his experience with nonmetric unfolding models, recommends at least 30 ideal points for a two-dimensional solution. Anderson (1981) states that one must have "ideal points distributed across the stimulus range … a [single-ideal point] analysis is not possible." The SPSS implementation of the ALSCAL algorithm (Takane, Young, & De Leeuw, 1977) requires a minimum of four ideal objects to perform an unfolding analysis.

This paper will propose a model for the dependent sampling of monadic disutility judgments that permits the estimation of single- or multiple-ideal unfolding models. Covariances of distances between ideal and real objects provide the information needed to estimate single-ideal solutions. The covariances depend not just upon the means and variances of the distances between real and ideal objects, but upon the relationship of the real objects to each other. This is illustrated in Fig. 1, Panel A, where distributions for three real and one ideal object, whose means lie on a line in two-dimensional space, are portrayed. Standard deviational ellipses are represented by the circles about the centroids of the objects. The means and variances of distances d_{BI} and $d_{B'I}$ are the same but the covariance of d_{AI} and d_{BI} will be positive while the

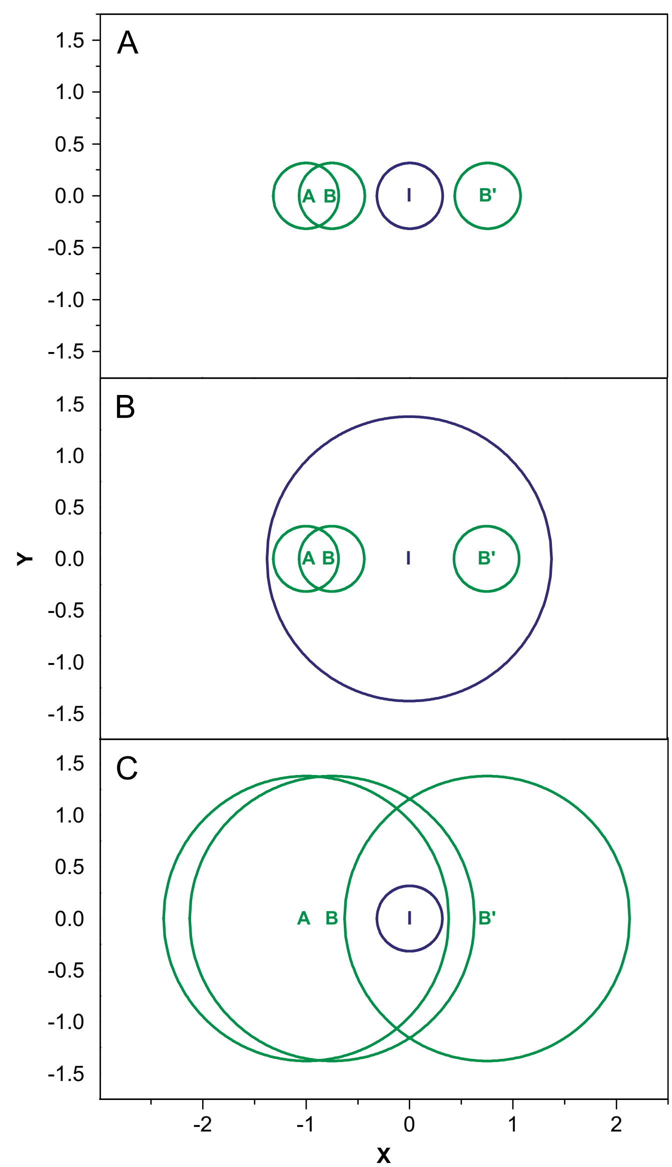

Fig. 1. Parameters used for laterality illustrations. In Panel A, the real objects (A, B, B') and the ideal object (I) have dimensional variances (0.1, 0.1), in Panel B real and ideal object variances are (0.1, 1.9), and in Panel C they are (1.9, 0.1).

covariance of d_{AI} and $d_{B'I}$ will be negative. This change in sign parallels Coombs' (1958) observation that transformations of inconsistency measures into psychological distance measures differed when a real object pair was unilateral, on the same side of an ideal object, or bilateral, on different sides of an ideal object. Formulae for calculating moments of the distances and the probability density function (PDF) of the distances are given in Section 2, where the dependent sampling model is defined and an approach for finding maximum likelihood (ML) estimates is presented. Section 3 investigates some of the properties of the proposed model. Section 4 outlines the estimation process. Two empirical examples, one involving rating data and the other involving ranking data, are given in Section 5. Section 6 concludes with a brief discussion of the results.

The proposed model differs from the wandering ideal point model (De Soete, Carroll, & DeSarbo, 1986) which assumes that only the ideal objects are probabilistic; the stimulus objects remain deterministic. This may be an appropriate assumption for some data sets, but for many data sets involving complex multidimensional stimuli, our expectation is that the variances of the stimuli may be larger than those of the ideal objects.

Bossuyt (1990) refers to the probabilistic unfolding models of the type mentioned above as random coordinate models because it is the random variable nature of the coordinates that makes the judgments random variables. In contrast, random distance models, such as those of Ramsay (1980) and DeSarbo, De Soete, and Eliashberg (1987), make no explicit assumption about the distribution of the coordinates and instead model the distances directly as random variables. However, in both types of models, the values of the distances at the moment of choice determine which object a subject will prefer. Bechtel (1968), Schönemann and Wang (1972), and others have followed a random response theory approach which does not make any assumption concerning momentary fluctuations due to the randomness of the coordinates or distances, but instead treats choice probabilities as a direct function of unfolding model parameters, such as the disutility and dissimilarity of real objects. Some authors have also defined probabilistic versions of the unfolding model that are aspatial, viz. the unfolding tree model of Carroll and De Soete (1990).

2. Multivariate probabilistic unfolding

If independent sampling assumptions hold for all disutility judgments then univariate unfolding models may be constructed. If independent sampling assumptions do not hold, then multivariate unfolding models are required. This section starts by considering the distribution of univariate distances obtained under the assumption of independent sampling. A multivariate framework for a vector of independently sampled disutility judgments for a single subject is then presented and generalized to dependently sampled judgments. An approximation for the multivariate PDF is described which provides the basis for ML estimation, results of which are given in the following sections.

2.1. Univariate independent distances

Unfolding models represent disutilities by the distances between unobserved coordinates of real and ideal objects in a latent r-dimensional space. It is assumed that the coordinates x_{ik} of object i on dimension k are normally distributed with mean μ_{ik} and variance σ^2_{ik}. Temporarily assume that the covariances of the coordinates are zero and define d_{ij}, a Euclidean distance random variable, as

$$d_{ij} = \left(\sum_{k=1}^{r}(x_{ik} - x_{jk})^2\right)^{1/2}, \quad 1 \leqslant i \leqslant m, \; 1 \leqslant j \leqslant n.$$

Object i belongs to the set of ideal objects and object j belongs to the set of real objects. This formulation generalizes the early work on unfolding by allowing multiple subjects to be defined by the same ideal object. Thus, $p \geqslant m$ with no requirement that p be a multiple of m. The value of m and the assignment of subjects to ideal objects is decided by the modeler. (See Section 4 below for a discussion of how subjects may be assigned to ideal objects.)

Assume for the moment an identity relationship between d_{ij} and the disutility judgment. (Section 4 also discusses how measurement transformations may be used to relax the identity relationship assumption.)

Let $d^2_{ij} = \sum_{k=1}^{r} d^2_{ijk}$ where $d_{ijk} = x_{ik} - x_{jk}$. Then,

$$d_{ijk} \sim N(\delta_{ijk}, \sigma^2_{ijk}),$$

where

$$\delta_{ijk} = \mu_{ik} - \mu_{jk},$$

$$\sigma^2_{ijk} = \sigma^2_{ik} + \sigma^2_{jk}.$$

To find the distribution of d^2_{ij}, let

$$z_{ijk} = d^2_{ijk}/\sigma^2_{ijk}.$$

It is well known (Suppes & Zinnes, 1963) that z_{ijk} is distributed as a noncentral chi-square χ'^2 distribution with one degree of freedom v and noncentrality parameter

$$\lambda_{ijk} = \mu^2_{ijk}/\sigma^2_{ijk}.$$

Therefore, d^2_{ij} will be distributed as a weighted sum of independent noncentral chi-square variates

$$f_{D^2_{ij}}(d^2_{ij}) = \sum_k \sigma^2_{ijk} \chi'^2_{v=1,\lambda_{ijk}}(d^2_{ijk}/\sigma^2_{ijk}),$$

where

$$D^2_{ij} = \sum_{k=1}^{r} \delta^2_{ijk}.$$

The PDF of distance (disutility judgment) d_{ij} follows as $f_{D_{ij}}(d_{ij}) = 2d_{ij} f_{D^2_{ij}}(d^2_{ij})$. Given $f_{D_{ij}}(d_{ij})$, it is possible to use

ML methods to estimate the location and variance parameters of the real and ideal objects. If there are p_i subjects whose ideal object is i, the likelihood is calculated over all np, $p=\sum_i p_i$ univariate observations. Procedures for calculating PDFs of the noncentral chi-square distribution and the quadratic forms in normal variables distribution (see below) may be found in Mathai and Provost (1992).

A limitation of the noncentral chi-square approach is that the definition of σ_{ij}^2 does not allow for nonzero covariances on the coordinates. To overcome this limitation, express d_{ij}^2 as a quadratic form by letting $\mathbf{d}_{ij}$ be an r-dimensional column vector of coordinate differences. Then, $d_{ij}^2=\mathbf{d}_{ij}'\mathbf{C}\mathbf{d}_{ij}$ where $\mathbf{C}$ is an identity matrix. The theory of quadratic forms in normal variables gives us a means of expressing the PDF of the quadratic form representation of d_{ij}^2 as a function of an $r\times r$ covariance matrix that allows for dependent coordinates. Quadratic form representations of distances are also used in the probabilistic multidimensional scaling (MDS) analysis of dissimilarities, where i and j both belong to the same set of real objects (MacKay, 1989).

2.2. Multivariate independent disutilities

A judgment d_{ij} of the disutility of real object j for a subject whose ideal object is i may be expressed as a distance random variable following the development of the previous section. Given the disutilities, the unfolding problem is to estimate the means and variances of the objects which constitute the parameters of the latent space underlying the judgments. To do so, it will be convenient to find an expression for the PDF of the multivariate n-element vector of distances to the n real objects for ideal object i.

For a subject represented by ideal object i, consider the nr element vector of sample values for the n real objects in r dimensions $(x_{11},\ldots,x_{1r},x_{21},\ldots,x_{nr})$. Represent the subject's ideal object by n independent samples of r coordinates for ideal object i $(x_{i1}^1,\ldots,x_{ir}^1,x_{i1}^2,\ldots,x_{ir}^n)$. Then, the difference vector of sample values for ideal object i is

$$\mathbf{y}_i=(x_{i1}^1-x_{11},\ldots,x_{ir}^1-x_{1r},x_{i1}^2-x_{21},\ldots,x_{ir}^n-x_{nr}),\quad 1\leqslant i\leqslant m$$

and the $nr\times nr$ matrix of squared random variable differences is $\mathbf{W}_i=\mathbf{y}_i'\mathbf{y}_i$. Similarly, letting $\mathbf{m}_i$ equal the corresponding nr element difference vector of means, we can define $\mathbf{\Theta}_i=\mathbf{m}_i'\mathbf{m}_i$.

Partition $\mathbf{W}_i=[\mathbf{W}_{ijk}]$ with $\mathbf{W}_{ijk}$ being of order $r\times r$ and $1\leqslant j,k\leqslant n$. Then, $[\mathrm{tr}\,\mathbf{W}_{i11},\ldots,\mathrm{tr}\,\mathbf{W}_{inn}]$ is the vector of squared random variable distances (judgments) to the n real objects for ideal object i. In corresponding fashion, define $\mathbf{\Theta}_i=[\mathbf{\Theta}_{ijk}]$ where $\mathrm{tr}[\mathbf{\Theta}_{ijj}]$ is the squared parametric (inter-centroid) distance to real object j from ideal object i and define $\mathbf{\Sigma}_i=[\mathbf{\Sigma}_{ijk}]$ where the $\mathbf{\Sigma}_{ijj}$ are $(r\times r)$ diagonal matrices and the $\mathbf{\Sigma}_{ijk}$, $j\neq k$, are $(r\times r)$ zero matrices. Thus,

$$\mathbf{\Sigma}_{ijk}=\begin{cases}\begin{bmatrix}\sigma_{ijk1}^2 & & & \\ & \ddots & & \\ & & & \sigma_{ijkr}^2\end{bmatrix}, & j=k,\quad 1\leqslant i\leqslant m,\ 1\leqslant j,k\leqslant n,\ \sigma_{ijjl}^2>0.\\ 0, & j\neq k,\end{cases}\tag{1}$$

As was true for univariate independent distances, the diagonal element, σ_{ijjl}^2, $1\leqslant l\leqslant r$, is equal to the sum of the variances of real object j and ideal object i on dimension l. (In the following sections, we shall generally assume that the off diagonal elements of $\mathbf{\Sigma}_{ijj}$ are zero but the mathematics are unaffected if this assumption is relaxed and the axes of the standard deviational ellipses about the real and ideal objects are allowed to vary in their orientation.) Letting $u_{ij}=\mathrm{tr}\,\mathbf{W}_{ijj}$, then $\mathbf{u}_i=[u_{i1},\ldots,u_{in}]$ will follow a *multi*variate quadratic forms in normal variables distribution $f_{\mathbf{\Theta}_i}(\mathbf{u}_i)$. The PDF of the n-element distance vector $\mathbf{d}_i$ will be

$$f_{\boldsymbol{\delta}_i}(\mathbf{d}_i)=\left|2^n\prod_{j=1}^n d_{ij}\right|f_{\mathbf{\Theta}_i}(\mathbf{u}_i),$$

where $d_{ij}=\sqrt{\mathrm{tr}[\mathbf{W}_{ijj}]}$, $\boldsymbol{\delta}_i=[\delta_{i1},\ldots,\delta_{in}]$, and $\delta_{ij}=\sqrt{\mathrm{tr}[\mathbf{\Theta}_{ijj}]}$. Due to the independent sampling of distances assumption, $f_{\boldsymbol{\delta}_i}(\mathbf{d}_i)=\prod_{j=1}^n f_{D_{ij}}(d_{ij})$. The log likelihood is calculated as the sum of the log likelihoods over p m-variate observation vectors.

2.3. Multivariate dependent disutilities

The advantage of the multivariate representation is that it can be generalized to a dependent sampling situation. (The independent quadratic forms representation allows for dependent sampling of coordinates. The multivariate dependent disutilities representation allows for dependent sampling of distances.) With dependent sampling for disutilities, it is assumed that for each set of n judgments, the ideal object is only sampled once. That single r-dimensional sample of coordinates is then compared to each of the n independently sampled sets of real object coordinates. The difference vector of sample values thus becomes

$$\mathbf{y}_i=(x_{i1}-x_{11},\ldots,x_{ir}-x_{1r},x_{i1}-x_{21},\ldots,x_{ir}-x_{nr}),\quad i=1,\ldots,m,$$

and the difference vector of means changes in a similar fashion. Defining $\mathbf{\Sigma}_i=[\mathbf{\Sigma}_{ijk}]$ as before, the $\mathbf{\Sigma}_{ijk}$, $j\neq k$, are now diagonal matrices with the r diagonal elements being equal to the variances of ideal object i on the r dimensions.

For $j = k$, $\mathbf{\Sigma}_{ijj}$ is unchanged. Thus,

$$\mathbf{\Sigma}_{ijk} = \begin{cases} \begin{bmatrix} \sigma^2_{ijkl} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \sigma^2_{ijkr} \end{bmatrix}, & j = k, \\ \begin{bmatrix} \sigma^2_{i1} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \sigma^2_{ir} \end{bmatrix}, & j \neq k, \end{cases} \quad \begin{array}{l} 1 \leqslant i \leqslant m, \\ 1 \leqslant j,k \leqslant n, \\ \sigma^2_{ijjl} > 0. \end{array} \tag{2}$$

2.4. Multivariate quadratic forms in normal variables PDF

The multivariate dependent disutilities model described above is a special case of quadratic forms in jointly Gaussian variates described by Jensen and Solomon (1994) who provide a closed form approximation for its PDF. The covariance matrix allowed by the general multivariate quadratic forms in normal variables distribution will, in most unfolding applications, be greatly simplified. If, for example, an isotropic Case V model is being estimated, the nonzero $\mathbf{\Sigma}_{ijk}$ will be constructed from just one variance σ^2. Closed form approximations are advisable since the distributions of quadratic forms depend upon expansions which may be slow to converge and are typically intractable in more than two dimensions.

In a manner similar to the Wilson and Hilferty (1931) approximation of the chi-square distribution, Jensen and Solomon use the moments of the multivariate quadratic forms in normal variables distribution to develop a Gaussian approximation that can be used to estimate the PDF. Using the unfolding model notation, they develop a Gaussian approximation for $[(u_{i1}/\theta_{i1})^{\alpha_{i1}}, \ldots, (u_{in}/\theta_{in})^{\alpha_{in}}]$ where u_{ij} is the squared observed disutility of real object j for a subject whose ideal object is i, $\theta_{ij} = \mathrm{tr}(\mathbf{\Sigma}_{ijj} + \mathbf{\Theta}_{ijj})$, $\alpha_{ij} = 1 - 2\theta_{ij}\omega_{ij3}/3\omega^2_{ij2}$, and $\omega_{ijs} = \mathrm{tr}\,\mathbf{\Sigma}^{s-1}_{ijj}(\mathbf{\Sigma}_{ijj} + s\mathbf{\Theta}_{ijj})$, $s = 1, 2, \ldots$. The mean and variance of the squared disutility of real object j for a subject whose ideal object is i are shown to be θ_{ij} and $2\omega_{ij2}$. The covariance of the squared disutilities d^2_{ij} and d^2_{ik} is

$$2\,\mathrm{tr}\,\mathbf{\Sigma}_{ijk}(\mathbf{\Sigma}_{ijk} + 2\mathbf{\Theta}_{ijk}). \tag{3}$$

Letting $v_{ij} = (u_{ij}/\theta_{ij})^{\alpha_{ij}}$, the mean and covariance matrix of $\mathbf{v}_i = [v_{i1}, \ldots, v_{in}]$ for a subject whose ideal object is i are $\boldsymbol{\mu}_i = [\mu_{i1}, \ldots, \mu_{in}]$ and $\mathbf{\Xi}_i = [\xi_{ijk}]$ where

$$\mu_{ij} = 1 + \omega_{ij2}\alpha_{ij}(\alpha_{ij} - 1)/\theta^2_{ij}, \quad 1 \leqslant j \leqslant n$$

and

$$\xi_{ijk} = 2\alpha_{ij}\alpha_{ik}(\mathrm{tr}\,\mathbf{\Sigma}_{ijk}(\mathbf{\Sigma}_{ikj} + 2\mathbf{\Theta}_{ikj}))/\theta_{ij}\theta_{ik}, \quad 1 \leqslant j,k \leqslant n.$$

To derive the PDF of the disutility judgments $\mathbf{d}_i = \mathbf{u}_i^{1/2}$, start with the multivariate normal distribution $\mathbf{v}_i \sim \mathbf{N}(\boldsymbol{\mu}_i, \mathbf{\Xi}_i)$ and take the Jacobian J_i by calculating the determinant of $\delta\mathbf{v}_i/\delta\mathbf{d}_i$:

$$J_i = \begin{vmatrix} \left(\frac{2d_{i1}\alpha_{i1}}{\theta_{i1}}\right)\left(\frac{u_{i1}}{\theta_{i1}}\right)^{\alpha_{i1}-1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \left(\frac{2d_{in}\alpha_{in}}{\theta_{in}}\right)\left(\frac{u_{in}}{\theta_{in}}\right)^{\alpha_{in}-1} \end{vmatrix}$$

which gives the PDF of the distance vector for a subject whose ideal object is i as

$$f_{\boldsymbol{\delta}_i}(\mathbf{d}_i) = |J_i|(2\pi)^{-r/2}|\mathbf{\Xi}_i|^{-1/2}\exp[-\tfrac{1}{2}(\boldsymbol{\delta}_i'\mathbf{\Xi}_i^{-1}\boldsymbol{\delta}_i)]. \tag{4}$$

Jensen and Solomon evaluate their approximation by comparing estimated cumulative distribution functions (CDFs) to known CDFs of bivariate central chi-squared distributions and multivariate central chi-squared distributions computed in earlier studies. The reported results are quite satisfactory.

CDF analysis focuses on the tail of the distribution. ML estimation involves the entire distribution. To visualize the differences in bilateral and unilateral pairs, and dependent and independent sampling, PDFs of bilateral and unilateral pairs under dependent and independent sampling conditions were estimated by the Jensen and Solomon approximation and compared to simulated distributions of 10 million distances. Objects A and B of Panel B, Fig. 1, formed the unilateral pair and objects A and B' formed the bilateral pair. The large variance on the ideal object I was chosen to highlight the difference between dependent sampling, where the variance of the ideal object forms the diagonal of $\mathbf{\Sigma}_{ijk}$, $1 \leqslant i \leqslant m$, $1 \leqslant j,k \leqslant n$, $j \neq k$, and independent sampling, where the corresponding diagonals are zero. The results are given in Fig. 2.

The shapes of the bilateral–unilateral and dependent–independent sampling distributions differ dramatically. The difference in the bilateral and unilateral plots under dependent sampling enables the probabilistic unfolding analysis to estimate unique solutions for a single-ideal. Conversely, the identical nature of unilateral and bilateral distributions under independent sampling prohibits the estimation of unique solutions for a single-ideal. While the approximations are not perfect—the median absolute differences between the Jensen and Solomon approximation and the simulated distributions for the 2500 cells in each condition were 0.0056, 0.0007, and 0.0019 for the bilateral-dependent, unilateral-dependent, and independent

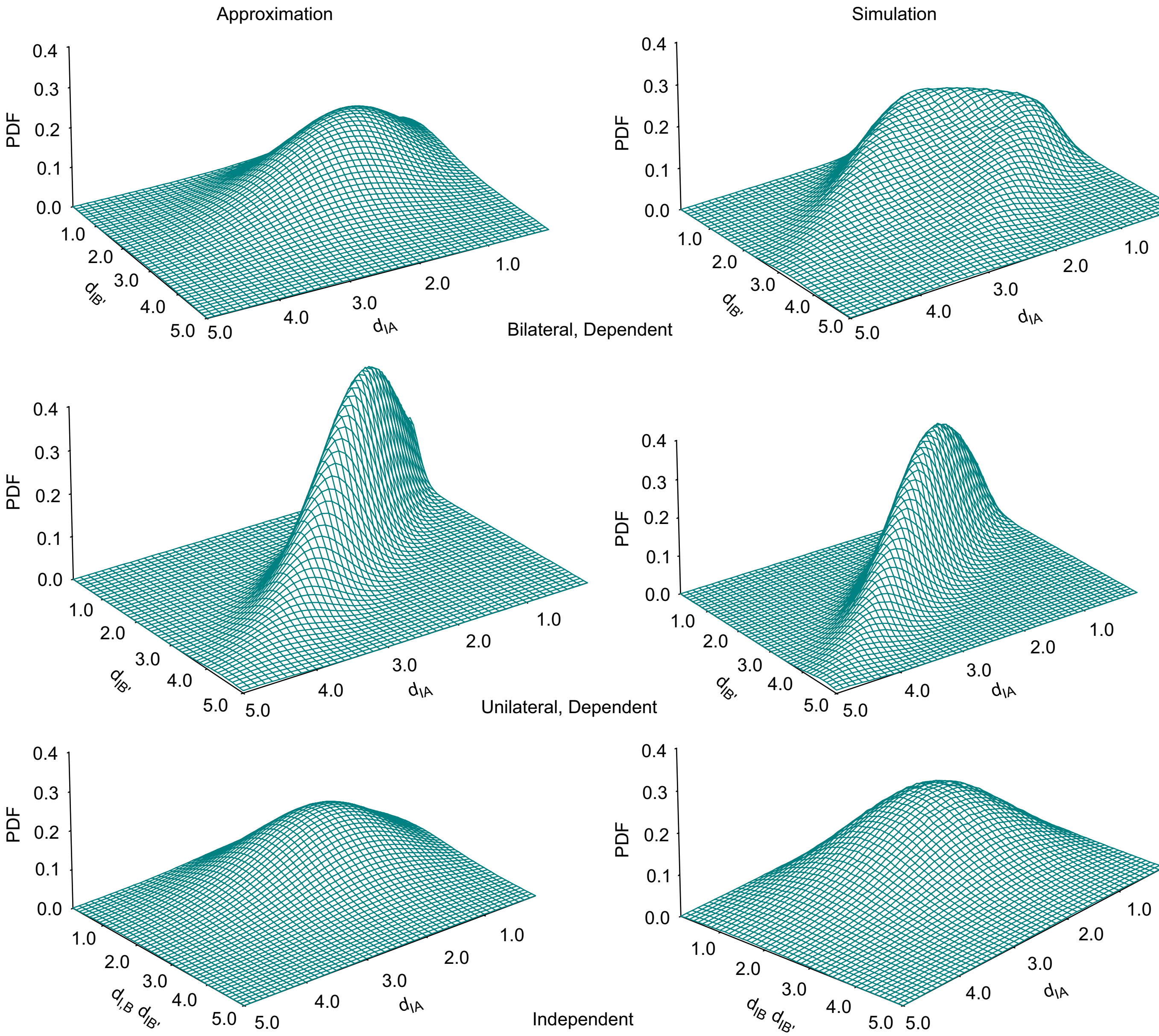

Fig. 2. Approximated and simulated PDFs.

conditions—the approximations do capture the differences in the shapes of the PDFs. More detailed comparisons of the estimated and simulated PDFs of Fig. 2 indicated that the approximation tended to overestimate PDF values near the origin when the variances were high.

3. Selected properties

Unique estimation with a single-ideal object depends upon dependent sampling and the covariances of pairs of distances from an ideal object to real objects. From (3), it is seen that under dependent sampling the covariance of squared distance pairs depends upon the coordinate variances of the ideal object (the off diagonal blocks of $\mathbf{\Sigma}$ are determined only by the coordinate variances of the ideal object) and products of differences in means which include, on expansion, products of the mean coordinates (centroids) of real object pairs. To enhance generalization, interest will be focused on correlations of distances instead of covariances of distances. We shall first look at how the correlations of squared distances are related to the variances of the objects and then see how they are related to the squared distances of the real objects' centroids to each other. This is followed by examples of parametric recovery when the data conform to the probabilistic model and when the data are transformed so they do not conform to the probabilistic model. Finally, there is a brief analysis of the accuracy with which the proposed model can correctly detect data generated by dependent and independent sampling processes.

3.1. Object variances and squared distance correlations

Starting with the example of Fig. 1, Panel B, (3) was used to calculate the correlation of the squared distances between objects (I,A) and (I,B); (I,A) and (I,B'), and (I,B) and (I,B') with dependent sampling. The first pair is unilateral and the last two pairs are bilateral. To determine how the magnitudes of the objects' variances and covariances affect the correlation of the squared distances, real and ideal object variances were first multiplied by a constant which varied from 0.05 to 5.0. Results are shown in Fig. 3.

The negative covariance which was observed earlier for bilateral objects with the variance structure of Fig. 1, Panel A, is seen to hold only when the variances are not too large. As the variances increase, all of the correlations of squared distances among bilateral pairs increase and the differences between unilateral and bilateral pairs lessen.

As the relative variances on the ideal object diminish, the distinction between dependent and independent sampling will diminish as well.

Fig. 1, Panel C, illustrates a situation that is the opposite of Panel B, the high variances are now on the real objects and the low variance is on the ideal object. Repeating the previous analysis with this new set of object variances gives us the result shown in Fig. 4. While the pattern of the squared distance correlations is similar to that observed in Fig. 3, the magnitude of the correlations is much smaller, as expected.

3.2. Distances among real objects and squared distance correlations

In the preceding section, the centroids of the objects remained fixed. In this section, we shall take Fig. 1, Panel B, as a starting point and look at how the correlation of d_{IA}^2, d_{IB}^2 and the correlation of d_{IA}^2, $d_{IB'}^2$ vary as object A is rotated in a circle about the origin of the space. The results are shown in the polar graph of Fig. 5.

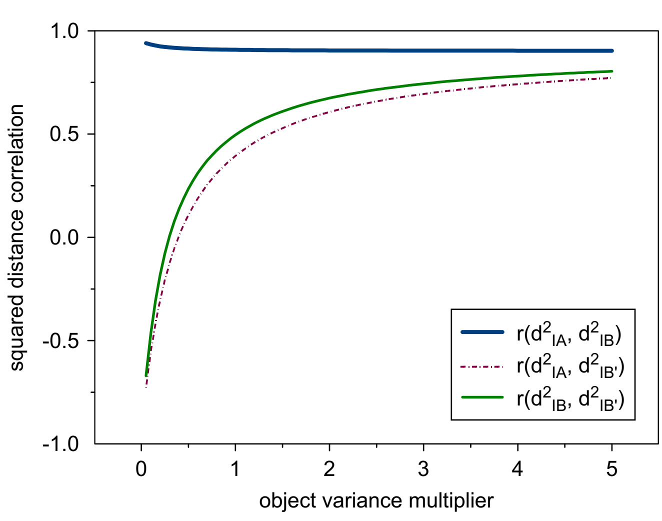

Fig. 3. Correlation of squared distances under dependent sampling with high ideal object and low real object variances.

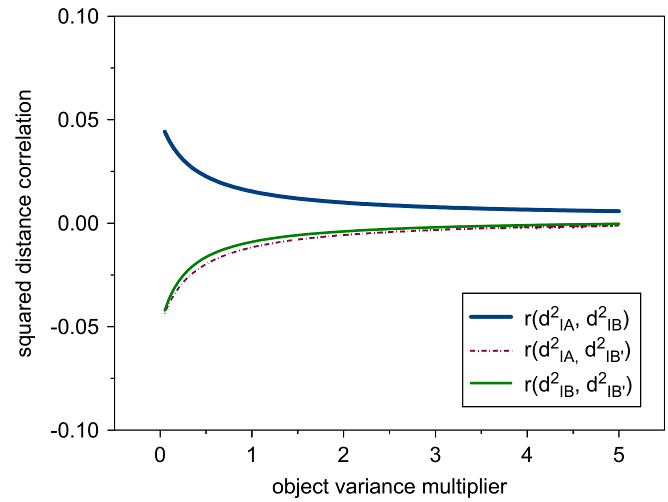

Fig. 4. Correlation of squared distances under dependent sampling with low ideal object and high real object variances.

Since object B and object B' are symmetrically placed, the plots of the correlations are symmetric as well. Due to the large variance on the ideal object, all of the correlations are positive. Looking first at the unilateral pair d_{IA}, d_{IB}, the starting point from Fig. 1, Panel B, is at the left side of the graph, with polar coordinates of (0.91, 180°). The correlation, 0.91, corresponds to the value of $r(d_{IA}^2, d_{IB}^2)$ for an object variance multiplier of 1 in Fig. 3. Keeping the multiplier fixed, as A rotates toward 0°, adjacent to B', the correlation steadily decreases until it hits a minimal value of 0.39. If the object variances were smaller, $r(d_{IA}^2, d_{IB}^2)$ would be negative at this point since A and B are now bilateral. An analogous interpretation holds for $r(d_{IA}^2, d_{IB'}^2)$. This dependence of disutility covariances upon distances among real objects provides more information for internal unfolding than is present in traditional deterministic analyses. The paucity of information in traditional unfolding analyses is a primary reason for the degenerate solutions that are frequently observed and for the few examples of internal preference analysis in the literature (Davidson, 1992).

Figs. 3–5 have shown that under dependent sampling, the covariance and correlation of disutilities have very definite patterns that, through the information conveyed in the likelihood function, should enable us to estimate the location and variance parameters of the objects when there is only a single-ideal.

3.3. Recovery of parameters with a single-ideal

To illustrate the proposed model, 12 real objects in two dimensions with centroids randomly sampled in the range (−0.5, 0.5) and a single-ideal object with centroid (0, 0) were selected. The variances of all objects on all dimensions were set at 0.1. Fig. 6 shows the parameter space. Only a

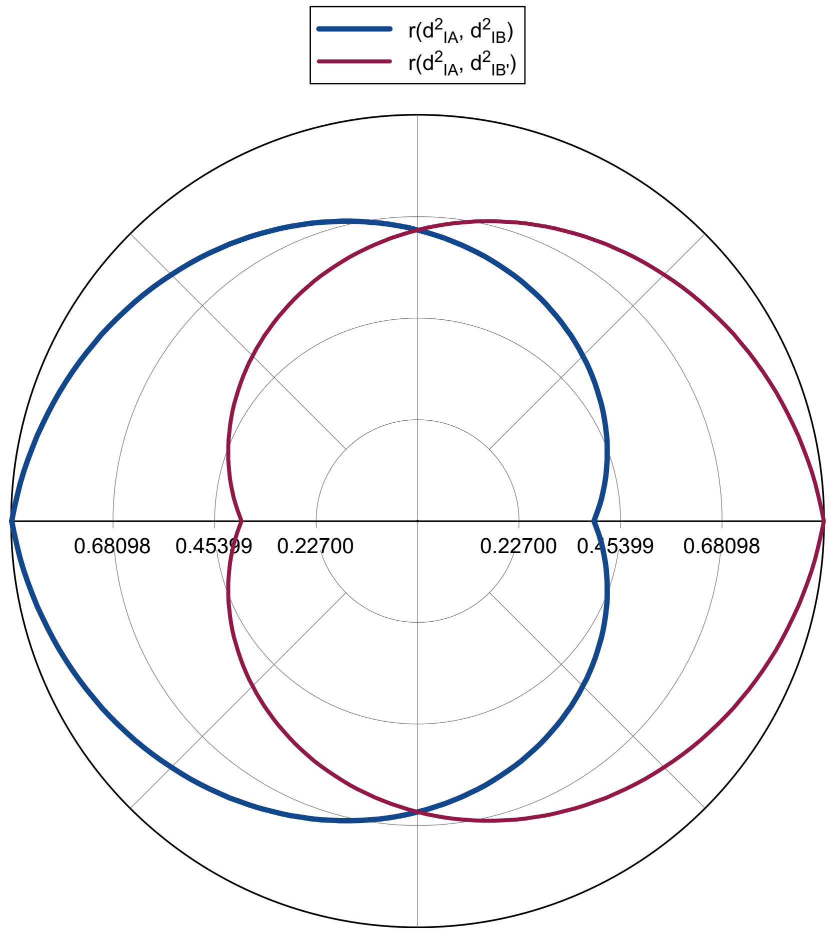

Fig. 5. Correlation of squared distances under dependent sampling as object A of Panel B, Fig. 1, rotates about the origin of the space.

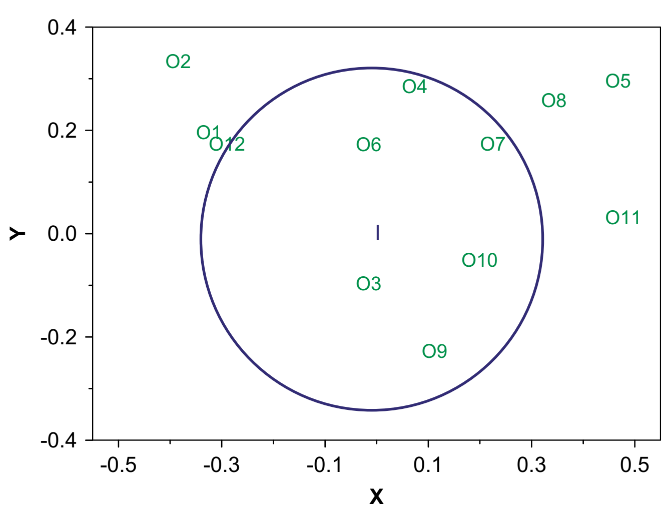

Fig. 6. Parameters for 12 real and one ideal object.

single standard deviational ellipse has been portrayed for increased legibility. This parameter space is thought to be somewhat challenging since the magnitude of the variance is greater and the number of objects is smaller than what is usually encountered in the empirical studies, primarily within the food industry, with which the author is familiar.

Five hundred sets of disutilities were generated from the parameter space and evaluated with a probabilistic MDS program, PROSCAL. The estimation process of PROSCAL is described in Section 4. The correlation of the 78 inter-centroid distances in the parameter space with the corresponding estimated values was 0.972. The estimated variance was a little high, 0.110.

The real valued simulated judgments, while admitting a large degree of variability, possess a degree of precision that would not be expected in empirical studies. To get a more realistic idea of what might be expected in an empirical study, the judgments were degraded to a nine point integer scale. (A nine point scale was chosen because it is the standard in the food industry, a common application area for unfolding models.) The estimates, shown in Fig. 7, were very close to those observed with the real valued judgments. The correlation was 0.967 and the estimated variance, after transforming the scale to that of the parameter space, was 0.118.

Another way of evaluating the estimates is through their ability to predict the first choice probabilities of the 12 real objects. A comparison of the actual first choice probabilities, calculated from the 500 sets of real valued disutilities, and estimated first choice probabilities, calculated by PROSCAL from the parameter estimates derived from the integer valued data, is given in Fig. 8. The actual first choice probabilities possess a little more variation than the estimates but the general pattern is preserved.

Traditional nonmetric MDS programs will, of course, not be able to capture the objects' locations but they should be able to capture the relative disutility of the real objects. The estimates provided by KYST (Kruskal & Wish, 1978) are shown in Fig. 9. The solution correctly identifies the low- and high-disutility objects but is unable to determine their location in the space.

It should be noted that the inability of traditional nonmetric MDS program to provide good estimates for the single-ideal situation is due to their deterministic formulation, not to the way they penalize the objective function used to obtain their estimates.

PREFSCAL (Busing, Groenen, & Heiser, 2005) is an example of a sophisticated deterministic scaling program with a rich variety of options for determining initial solutions and customizing loss functions. A variety of PREFSCAL models were used to analyze the simulated judgments. Solutions obtained from different program options varied significantly, but all were bad. Correlations of the 78 inter-centroid distances in the parameter space

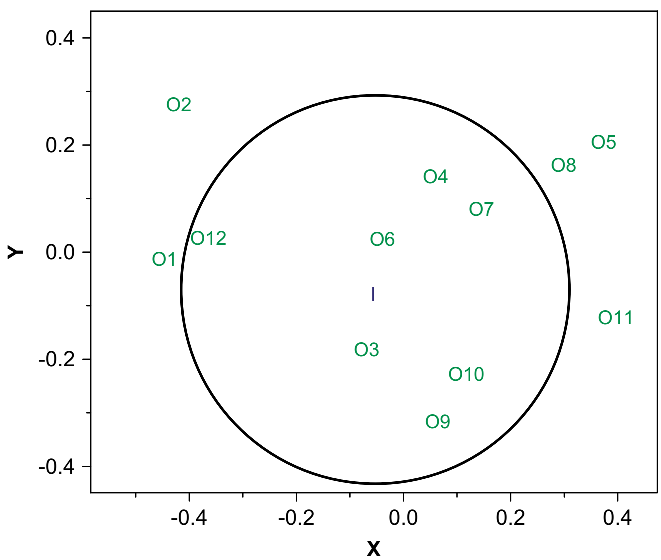

Fig. 7. Estimated parameters from integer scale data for 12 real and one ideal object.

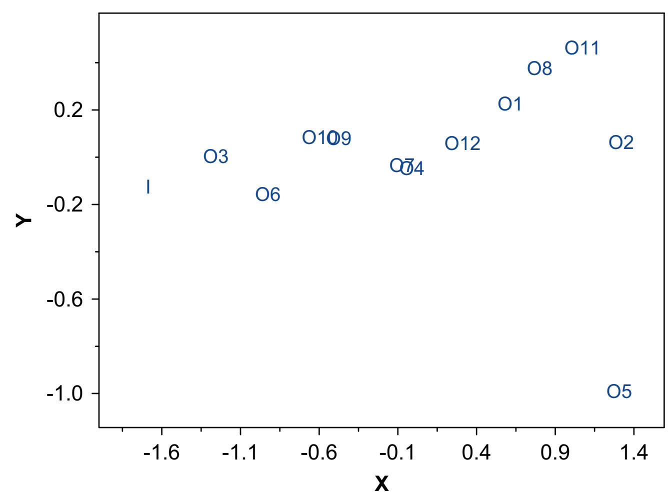

Fig. 9. Nonmetric estimates of locations for 12 real and one ideal object.

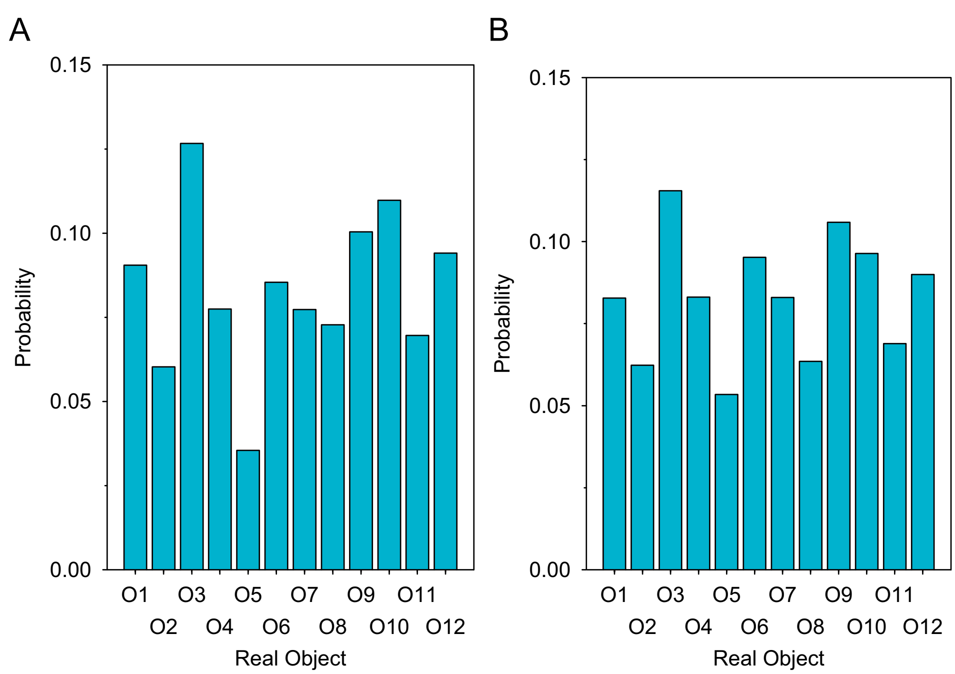

Fig. 8. Actual (Panel A) and estimated (Panel B) first choice probabilities for 12 real objects.

with the corresponding estimated values ranged from −0.085 to 0.087.

3.4. Classification of independent and dependent samples

In many if not most empirical unfolding applications, rows of the proximity matrix represent responses from different subjects. Dependent sampling would, in this situation, intuitively seem to be a reasonable process. There may, though, be situations where independent sampling is the rule. Hypothesis tests are needed to distinguish dependent and independent sampling processes.

From (1) and (2), it is obvious that dependent and independent sampling are based upon the same estimated parameters. The only difference is whether the off diagonal variance blocks should be the same as the ideal object (dependent sampling) or zero (independent sampling). When the numbers of parameters differ, information criterion or likelihood ratio tests may be used to test models. When the numbers of parameters are the same, a comparison of likelihoods can be used to determine which model has more support.

A simple Monte Carlo experiment was run using the parameters of Fig. 6. Ten sets of dependent and independent samples of 25, 50, and 100 integer valued disutility vectors were generated and evaluated with a dependent and independent sampling model. The percentages of correct sampling classifications (based upon the likelihoods) were calculated. The process was then repeated for a low-variance sample, where the variances of the objects were 0.05. Results are shown in Fig. 10.

All of the dependent sampling data sets were correctly classified—both for the low- and high-variance condition. The independent sample data did not fare as well. For the small sample size condition, the percentage of correct classifications were 60% and 70% for the high- and low-variance conditions. Sample sizes of 50 and 100 disutility vectors were all classified correctly.

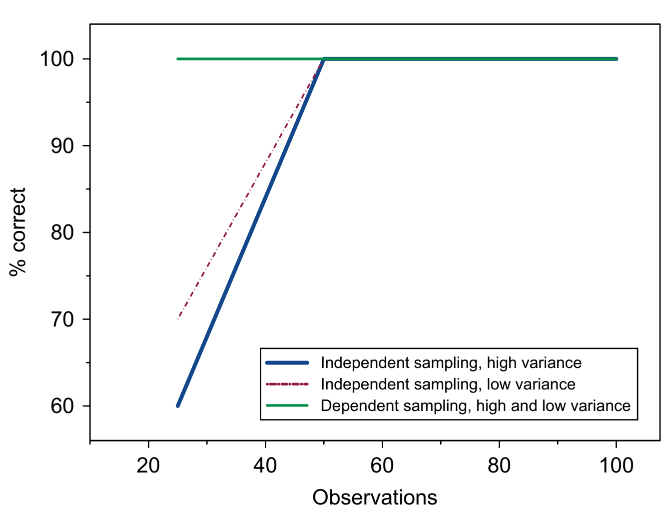

Fig. 10. Sample classifications for simulated integer valued disutilities.

4. Estimation algorithm

PROSCAL, a probabilistic MDS program that is available without charge from http://www.proscal.com, was used to estimate the single-ideal unfolding models. User guides, available at the website, describe the estimation process and program options in considerable detail. Equality constraints may be placed by the user on centroids and variances. In the above examples, all variances were constrained to have the same value but it is possible, though not usually advisable, to estimate unique variances for all objects on each dimension. Information criterion or likelihood ratio tests may be used to determine the appropriate degree of model complexity.

Initial estimates are formed from disutility judgments by first estimating an additive constant which will transform the minimal disutility judgment to a small positive value. (If liking ratings are being used, they must be first transformed by the user to disutility judgments by reverse scaling or some other means.) Arithmetic means of disutilities associated with each real object for a given ideal are calculated and I-scales (Coombs, 1950) are formed for each ideal object. The I-scales are double centered and a singular value decomposition is performed to arrive at the first estimate of an initial configuration. This first initial estimate is modified by setting the coordinates of the ideal objects equal to the coordinates of the real objects for which they have the lowest disutility. In situations like the single-ideal, where there is an insufficient number of ideal objects to estimate an r-dimensional solution, the ideal object is placed at the centroid of the space and the real objects are systematically placed in the space so that their distances to the ideal are proportional to their disutility. Initial variance estimates are isotropic, even when an anisotropic solution is to be estimated. The first initial variance magnitudes are proportional to the mean squared differences of the disutility judgments for objects and the estimated distance of the real object from its ideal object. The first initial variance estimates are modified so that the minimum ideal object variance is equal to the minimum real object variance.

Given the initial location and variance estimates, a two stage ML estimation process begins in which variances are estimated holding location estimates constant and locations are then estimated holding variance estimates constant. The iterative two stage process continues until no appreciable gain is made in the likelihood function. Optimization at each stage is carried out using Chandler's (1969) direct search algorithm, STEPIT. (A wide variety of optimization algorithms have been tried. The criterion function, log likelihoods calculated from (4) and summed over the p m-variate observation vectors, appears to be too non-linear for most gradient based methods.)

PROSCAL admits a variety of judgment types, including similarities, dissimilarities, preference ratios, binary preferential choices, and disutility judgments. There are no limits on the number of judgments accepted by PROSCAL. Currently, no more than six spatial dimensions and 60 objects are allowed. When different types of data are combined, linear exponential transformations are used to enhance the conformity of the measurement scales. Additional transformations are used with similarities (MacKay & Lilly, 2004). Transformations may also be employed when disutility judgments are used by themselves. Transformation constants are estimated in the second ML stage, along with the coordinates that define the centroids of the objects.

Hypotheses are tested using information criterion statistics, such as CAIC (Bozdogan, 1987) or BIC (Schwarz, 1978). CAIC penalizes likelihoods by the number of parameters estimated in the model and takes the form

$$\mathrm{CAIC} = -2 \ln L + ck,$$

where L is the likelihood, k is equal to the number of independently estimated parameters, and c is the cost of adding an additional parameter to the model.

A formula for computing the number of free parameters k in a model is

$$k = u + q + t + (s - 1) - r - r(r - 1)/2 - 1,$$

where u is the number of coordinates being estimated, q the number of variances being estimated, t the number of measurement model parameters being estimated, s the number of segments estimated when a mixture model is used, one otherwise, r the dimensionality of the space.

The last three terms of the equation are subtracted for the centering, rotation, and scale invariance of the solution. When anisotropic or city-block models are used, the rotational invariance term should be omitted. When multiple-ideal object solutions are estimated, mixture models may be used to assign subjects to ideal objects. (See McLachlan & Peel, 2000; Wedel & DeSarbo, 1995 for a discussion of mixture models.) For the cost,

$$c = \ln(S) + 1,$$

where S is the number of judgments. BIC is very similar, the only difference being that $c = \ln(S)$.

Low values of CAIC or BIC indicate the superior model.

5. Two empirical examples

Two examples, the first employing liking ratings and the second based on liking rankings, were used to illustrate the proposed probabilistic model.

5.1. A liking ratings example

Liking ratings for the 12 cereals identified in Table 1 were collected from 46 student subjects on a 0–100 graphic rating scale. Eleven of the 12 cereals were readily available in local stores. The 12th, Quisp, was not. Subjects were given an opportunity to look at the cereal boxes before providing liking rating judgments. The cereals were not tasted. Eleven of the 12 boxes provided images of the cereal, the 12th, again Quisp, did not. The collection of liking ratings was preceded by a warm-up task to familiarize subjects with the rating scale. The judgments were reverse scaled to turn them into disutilities before being submitted to PROSCAL.

Table 1
Cereals used in empirical study

Abbreviation	Cereal
AJ	Apple Jacks
CC	Captain Crunch
CH	Cheerios
CF	Corn Flakes
FF	Frosted Flakes
GN	Grape Nuts
LC	Lucky Charms
Qu	Quisp
RB	Raisin Bran
SW	Shredded Wheat
SK	Special K
To	Total

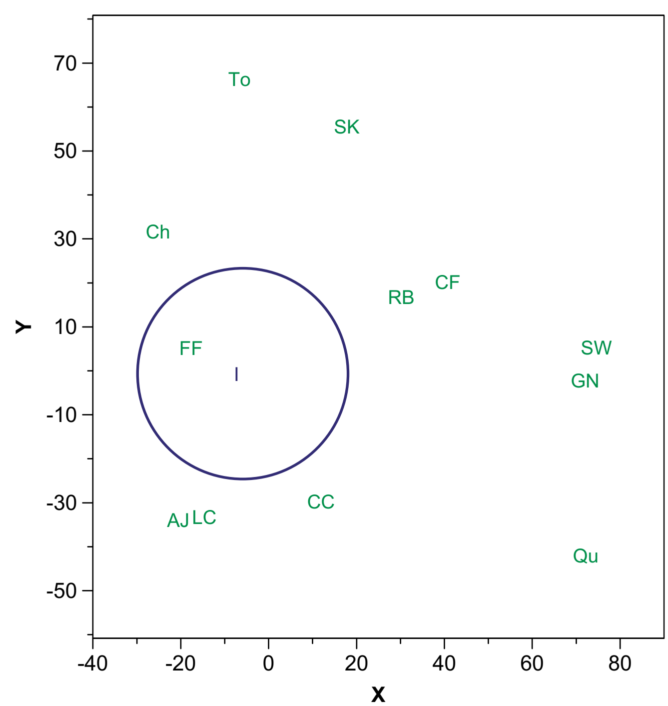

Fig. 11. Single-ideal solution for 12 cereals.

The single-ideal Case V isotropic two-dimensional solution estimated by PROSCAL is shown in Fig. 11. The configuration comports nicely with the physical properties of the cereals. The Y-axis distinguishes three classes of cereals—health oriented cereals are at the top, high-sugar cereals are at the bottom, and the ones in the middle are traditional lightly sweetened cereals. The X-axis distinguishes cereals by texture. The nine cereals on the left have traditional flake or nugget textures. Two of the cereals on the right, Shredded Wheat and Grape Nuts, have non-traditional textures. The third cereal on the right, Quisp, was a mystery for most subjects since they had not been previously exposed to the cereal and no image of the cereal was on the package. However, the package design did indicate that the target market was children and the subjects correctly deduced that it would fall in the very sweet category. The CAIC value for the analysis was 5387, 28 parameters were estimated—26 coordinates, one variance, and one measurement model parameter. The correlation of actual and estimated first choice percentages was 0.92. A comparable analysis under the assumption of independent sampling had a CAIC value of 5412, indicating support for dependent sampling. A three-dimensional dependent sampling analysis, with 41 estimated parameters, had a CAIC value of 5395, indicating support for the two-dimensional solution.

To check the stability of the solution, a k-means cluster analysis of the proximity matrix was performed and the subjects were divided into three groups of 30, 14, and two subjects. The two-dimensional triple-ideal dependent sampling solution is shown in Fig. 12. The CAIC value of 4793, with 32 estimated parameters, indicated that it was a better solution.

The triple-ideal solution indicates, not surprisingly, two primary groups of subjects, those who prefer high-sugar cereals and those who prefer healthier cereals. A much smaller group expressed a stronger preference for cereals with a non-traditional texture. Of more interest for this study, though, is the strong similarity of the configurations of the real objects in the single-ideal and triple-ideal solutions, again suggesting the ability of a single-ideal solution to capture the latent structure of the proximity matrix.

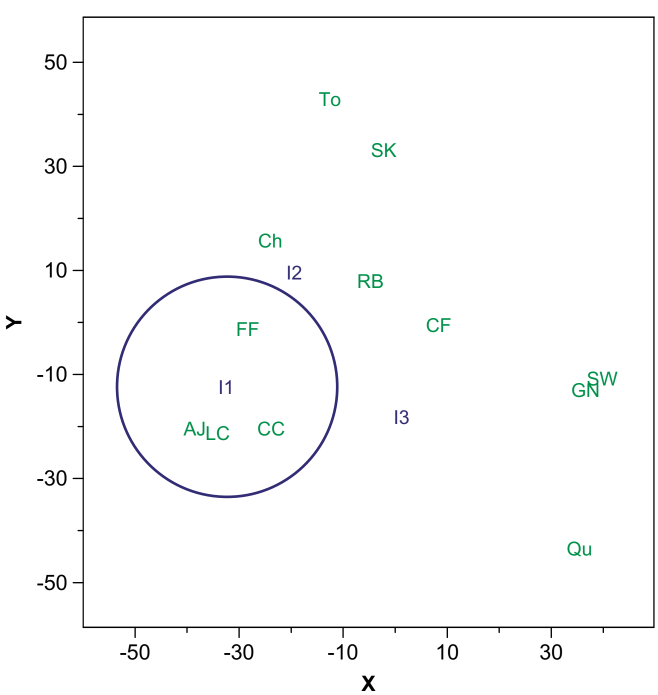

Fig. 12. Triple-ideal solution for 12 cereals.

5.2. A liking rankings example

The PREFSCAL model mentioned in Section 3.1 illustrates its SPSS implementation with ranking data on 15 breakfast items by 42 subjects. Ranking data provide a more rigorous test for the PROSCAL model since the assumptions of the PROSCAL model are less likely to be valid for ranking data than they are for ratings data. The data are from a study by Green and Rao (1972). Abbreviations used for the 15 breakfast items are given in Table 2.

The PREFSCAL analysis estimated 15 points for the real objects and 42 ideal points, one for each subject. Fig. 13, Panel A, illustrates the PREFSCAL estimates of the 15 breakfast cereals.

The PREFSCAL analysis was then redone for a single-ideal object (Fig. 13, Panel B) and compared to the corresponding results obtained by PROSCAL (Fig. 13, Panel C). The locations of the real objects estimated in the single-ideal solution of PROSCAL are very close to the estimates of the real objects for the 42 ideal solution by PREFSCAL while the single-ideal solution of PREFSCAL bears no resemblance to the 42 ideal solution of PREFSCAL. Several of the real objects in the single-ideal PREFSCAL solution were estimated as having near identical locations. A 42 ideal solution (not shown) was also estimated by PROSCAL. A comparison of the CAIC and BIC statistics for the two PROSCAL solutions both indicated that the single-ideal solution was better.

6. Discussion

Most of the analyses in this paper have been two-dimensional. Given enough data for the additional number

Table 2
Breakfast items used in empirical liking rankings study

Abbreviation	Cereal
TP	Toast pop-up
BT	Buttered toast
EMM	English muffin and margarine
JD	Jelly donut
CT	Cinnamon toast
BMM	Blueberry muffin and margarine
HRB	Hard rolls and butter
TMd	Toast and marmalade
BTJ	Buttered toast and jelly
TMn	Toast and margarine
CB	Cinnamon bun
DP	Danish pastry
GD	Glazed donut
CC	Coffee cake
CMB	Corn muffin and butter

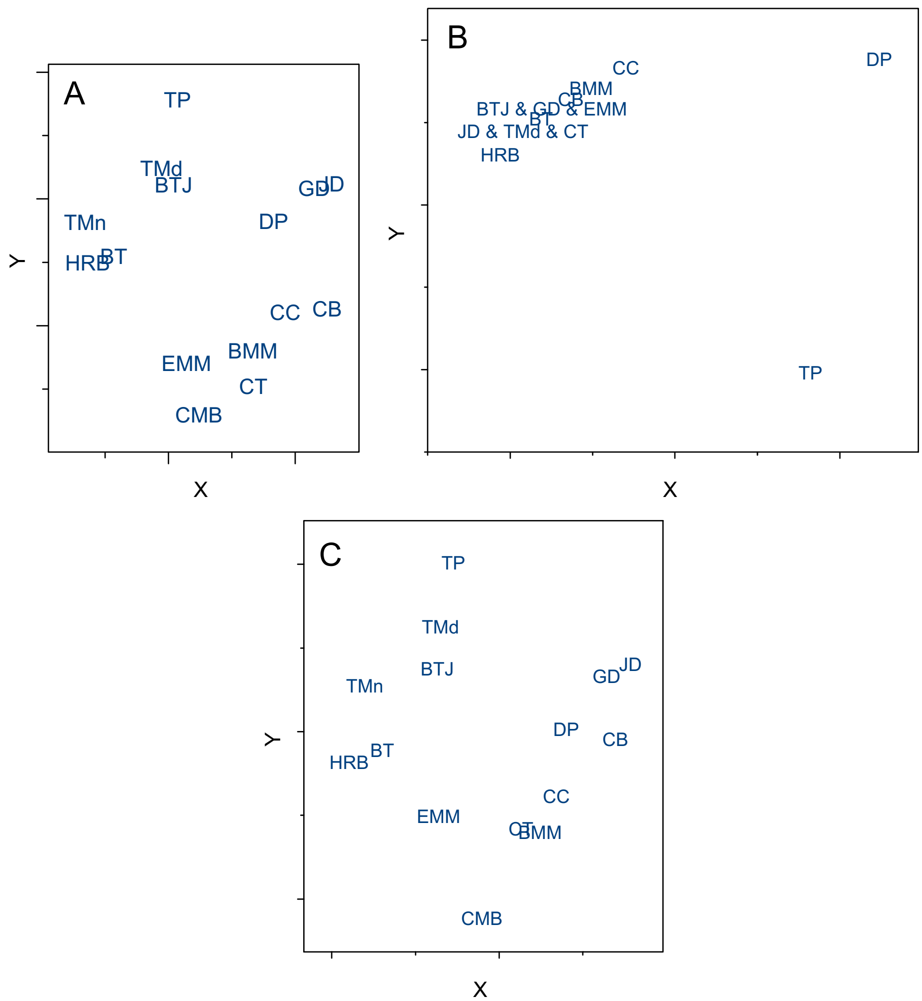

Fig. 13. Coordinates and centroids for three breakfast item solutions. In Panel A the coordinates of the real objects are from a PREFSCAL analysis with 42 ideal objects. In Panel B the coordinates of the real objects are from a PREFSCAL analysis with one ideal object and in Panel C the centroids of the real objects are from a PROSCAL analysis with one ideal object.

of estimates, one should expect that analyses in spaces of higher dimensionality should be better than those reported here because the quality of the approximation improves as the number of terms in the summations increase. Attention has also been restricted to Euclidean spaces. Probabilistic city-block models have been developed for independent judgments (MacKay, 2001), but dependent judgments are, at this point, limited to Euclidean spaces.

Ongoing studies suggest that the advantages demonstrated here of probabilistic dependent sampling models over deterministic nonmetric models will also hold when multiple-ideal objects are evaluated. Probabilistic independent sampling models are, of course, not able to estimate single-ideal unfolding models either. In multiple-ideal situations, when the variances on the ideal objects are small, the advantage of probabilistic dependent sampling over probabilistic independent sampling diminishes, even when the data are known to be generated by a dependent sampling process. When the variances on the ideal objects are large, advantages for probabilistic dependent sampling are substantial.

Internal unfolding models are conceptually appealing. One attractive feature of unfolding models is their ability to incorporate satiety, an ability not present in many of the models derived from economic theory which assume that more is always better. A second attractive feature is their estimation of the latent space that underlies the hedonic judgments. Explaining preferences from independently measured object attributes, a procedure often used with external unfolding models, assumes an ability to correctly choose and measure the attributes. The spatial nature of unfolding models also enhances their interpretability.

The conceptual appeal of internal unfolding models is not reflected in their use. Reformulating unfolding models in a probabilistic framework can, as shown here, extend their scope and their success. Probabilistic models also eliminate or mitigate many of the deterministic model properties that have caused some researchers to eschew distance models. Further research is needed in developing probabilistic models and establishing their properties. Finally, there is an obvious need for rigorous experimental evaluation.

Acknowledgments

Partial support for this work was provided by a grant from the Kelley School of Business at Indiana University. Bryan Lilly, now at the University of Wisconsin Oshkosh, collected the cereal study data that are evaluated in Section 5. The author is also grateful for the comments of the reviewers.

References

Anderson, N. (1981). *Foundations of information integration theory*. New York: Academic Press.

Bechtel, G. G. (1968). Folded and unfolded scaling from preferential paired comparisons. *Journal of Mathematical Psychology*, *5*, 333–357.

Böckenholt, U. (2006). Thurstonian-based analyses: past, present and future utilities. *Psychometrika*, *71*, 615–629.

Bossuyt, P. (1990). *A comparison of probabilistic unfolding theories for paired comparisons data*. Berlin: Springer.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.

Busing, F. M. T. A., Groenen, P. J. K., & Heiser, W. J. (2005). Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. *Psychometrika*, *70*, 71–98.

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. I (pp. 105–155). New York: Seminar Press.

Carroll, J. D., & De Soete, G. (1990). Fitting a quasi-Poisson case of the GSTUN (general stochastic tree unfolding) model and some extensions. In M. Schader, & W. Gaul (Eds.), *Knowledge, data and computer-assisted decisions* (pp. 421–430). Amsterdam: North-Holland.

Chandler, J. P. (1969). STEPIT—Finds local minima of a smooth function of several parameters. *Behavioral Science*, *14*, 81–82.

Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, *57*, 145–158.

Coombs, C. H. (1958). On the use of inconsistency of preferences in psychological measurement. *Journal of Experimental Psychology*, *55*, 1–7.

Coxon, A. P. M. (1982). *The user's guide to multidimensional scaling*. Exeter, NH: Heinemann.

Davidson, M. L. (1992). *Multidimensional scaling*. Malabar, FL: Krieger.

De Soete, G., Carroll, J. D., & DeSarbo, W. S. (1986). The wandering ideal point model: A probabilistic multidimensional unfolding model for paired comparisons data. *Journal of Mathematical Psychology*, *30*, 28–41.

DeSarbo, W. S., De Soete, G., & Eliashberg, J. (1987). A new stochastic multidimensional unfolding model for the investigation of paired comparisons in consumer preference/choice data. *Journal of Economic Psychology*, *8*, 357–384.

Ennis, D. M. (1993). A single multidimensional model for discrimination, identification and preferential choice. *Acta Psychologica*, *84*, 17–27.

Green, P. E., & Rao, V. (1972). *Applied multidimensional scaling*. Hinsdale, IL: Dryden Press.

Hefner, R. A. (1958). *Extensions of the law of comparative judgment to discriminable and multidimensional stimuli*. Doctoral dissertation, University of Michigan.

Jensen, D. R., & Solomon, H. (1994). Approximations to joint distributions of definite quadratic forms. *Journal of the American Statistical Association*, *89*, 480–486.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional Scaling*. Beverly Hills, CA: Sage Publications.

MacKay, D. B. (1989). Probabilistic multidimensional scaling: An anisotropic model for distance judgments. *Journal of Mathematical Psychology*, *33*, 187–205.

MacKay, D. B. (2001). Probabilistic multidimensional scaling using a city-block metric. *Journal of Mathematical Psychology*, *45*, 249–264.

MacKay, D. B. (2005). Probabilistic scaling analyses of sensory profile, instrumental and hedonic data. *Journal of Chemometrics*, *19*, 180–190.

MacKay, D. B. (2006). Chemometrics, econometrics, psychometrics—How best to handle hedonics? *Food Quality and Preference*, *17*, 529–535.

MacKay, D. B., Easley, R. F., & Zinnes, J. L. (1995). A single-ideal point model for market structure analysis. *Journal of Marketing Research*, *32*, 433–443.

MacKay, D. B., & Lilly, B. (2004). Percept variance, subadditivity and the metric classification of similarity, and dissimilarity data. *Journal of Classification*, *21*, 185–206.

MacKay, D. B., & Zinnes, J. L. (1995). Probabilistic multidimensional unfolding: An anisotropic model for preference ratio judgments. *Journal of Mathematical Psychology*, *39*, 99–111.

Mathai, A. M., & Provost, S. B. (1992). *Quadratic forms in random variables: Theory and applications*. New York: Marcel Dekker.

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Ramsay, J. O. (1980). The joint analysis of direct ratings, pairwise preferences and dissimilarities. *Psychometrika*, *45*, 149–166.

Schönemann, P. H., & Wang, M. M. (1972). An individual difference model for the multidimensional analysis of preference data. *Psychometrika*, *37*, 257–309.

Schwarz, G. (1978). Estimating the dimensions of a model. *The Annals of Statistics*, *6*, 461–464.

Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, & E. Galanter (Eds.), *Handbook of mathematical psychology*, Vol. I (pp. 1–76). New York: Wiley.

Takane, Y., Young, F. W., & De Leeuw, J. (1977). Nonmetric individual differences in multidimensional scaling: An alternating least squares method with optimum scaling features. *Psychometrika*, *42*, 7–67.

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

Wedel, M., & DeSarbo, W. S. (1995). A mixture likelihood approach for generalized linear models. *Journal of Classification*, *12*, 21–55.

Wilson, E. B., & Hilferty, M. M. (1931). The distribution of chi-square. *Proceedings of the National Academy of Sciences*, *17*, 684–688.

Zinnes, J. L., & Griggs, R. A. (1974). Probabilistic, multidimensional unfolding analysis. *Psychometrika*, *48*, 27–48.