JOURNAL OF CHEMOMETRICS
J. Chemometrics 2005; **19**: 180–190
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cem.921

Probabilistic scaling analyses of sensory profile, instrumental and hedonic data

David B. MacKay*
Kelley School of Business, Indiana University, Bloomington, IN 47405, USA

Received 15 June 2004; Revised 18 February 2005; Accepted 1 March 2005

Sensory and hedonic variability are fundamental psychological characteristics that must be explicitly modeled if one is to develop meaningful statistical models of sensory phenomena. Sensory objects are perceived with differential uncertainty. Subjects also differ; some are very certain about what they desire and others are not. When variability is not modeled, the variability becomes confounded with dissimilarity or disutility and the estimates lose their meaning. Probabilistic scaling models are proposed that explicitly measure sensory and hedonic variability. The resulting estimates are shown to be free of the systematic bias that characterizes traditional deterministic models. Applications of the proposed models are presented and extensions of the models that can help bridge the gap between sensory science, marketing and product development are illustrated. Copyright © 2005 John Wiley & Sons, Ltd.

KEYWORDS: multidimensional scaling; PROSCAL; Thurstonian models; unfolding

1. INTRODUCTION

Understanding the relationships of consumers' preferences and choices to the sensory characteristics of products is a primary goal of sensory science. Probabilistic multidimensional scaling (PMDS) models provide a way of obtaining this understanding. By capturing key fundamental psychological characteristics of sensory and hedonic data, PMDS models have been shown to provide excellent recovery of known parameters. Their maximum likelihood (ML) foundation permits the derivation and testing of simple models that have a high degree of extensibility.

A family of PMDS models is described in this paper that spatially represents the perception of real sensory objects and imaginary ideal objects by multivariate normal distributions. Section 2 presents the conceptual foundation of these models and describes their use with common types of sensory data. Section 3 gives a mathematical outline of their development. Section 4 describes a wine study application in which sensory profile and hedonic rating data are combined in a common space and extended to make inferences about product loyalty and to estimate 'perceptual shares' for new products.

PMDS models can also be used to compare the latent structure of products defined by sensory profiles with the latent structure of products defined by instrumental methods. A second application, in Section 5, compares the latent structures of peas, as estimated from sensory profile and instrumental methods, using traditional and PMDS procedures.

Finally, in Section 6, we discuss how the testing process employed in the applications can be generalized to other situations. Comments on the costs and benefits of industrial applications are also offered.

*Correspondence to: D. B. MacKay, Kelley School of Business, Indiana University, Bloomington, IN 47405, USA.
E-mail: mackay@indiana.edu

2. MULTIDIMENSIONAL SCALING

2.1. Deterministic MDS

Traditional MDS models are deterministic: the latent structure of sensory objects is represented by a configuration of points in a multidimensional space. The latent structure reflects a simplified organization of the manifest (directly observed) variables [1]. In the sensory sciences, MDS methods are used to model both proximity data and hedonic judgments. A number of good discussions on how MDS has been used in sensory analysis are available [2–4].

Proximity data may consist of dissimilarities or similarities. A high dissimilarity value indicates that a pair of sensory objects is dissimilar, a low value indicates that a pair is similar. Dissimilarity data may be judgments obtained directly from consumers or expert assessors. Dissimilarities may also be derived from profile judgments about the attribute values of sensory objects. The attribute valuations may be judgments obtained from consumers or expert assessors. They may also be measurements obtained from instruments. Similarities, however, are usually judgments obtained directly from consumers and traditionally vary from zero to one, with a zero indicating no similarity and a one indicating a high similarity or identity.

Numerous types of hedonic judgments, obtained from consumers, may also be evaluated with MDS models. Most common, perhaps, are liking-rating judgments. Binary preferential choices, preference ratios in which values indicate the degree to which one object is preferred to another, and other types of hedonic judgments may also be used.

The output of a traditional MDS analysis is a set of points in a low-dimensional space. When, in an analysis of proximity data, the points represent real sensory objects, the distance between a pair of points indicates how different the objects are estimated to be from one another. When, in an analysis of hedonic judgments, one point represents a real sensory object and the other point represents an imaginary ideal object, the distance between the real and ideal point indicates the degree of *dis*utility possessed by the ideal point subject(s) for the real sensory object. A common way of obtaining real and ideal point estimates is by minimizing the deviance of the distances among pairs of points and their corresponding judgments. The number of dimensions in the space indicates the complexity of the latent structure. Identification of the dimensions may be made by expert judgment or by fitting profile data to the projections of the points on the axes. A good introductory text for MDS is by Davison [5]. More comprehensive texts are those by Borg and Groenen [6] and Cox and Cox [7].

2.2. Probabilistic MDS

Deterministic models only make use of probabilistic concepts, if at all, in the process of fitting the model to the data to account for discrepancies between the data and the estimates of the model.

Probabilistic models may introduce probabilistic components in a variety of ways. With dissimilarity models, for example, one way of admitting probabilistic components is to assume that the dissimilarity judgments are normally or lognormally distributed about the true distance. A fundamentally different assumption is that it is the sensory objects that are normally perturbed and that it is through the variance attending the sensory objects that variability enters the judgment process. Comparisons of these and other means of introducing probabilistic components may be found in References [8–10].

The focus here is on probabilistic models in which variability enters the judgment process through the perception of real and, when hedonic judgments are being evaluated, ideal sensory objects. The reason for this focus is that differential percept variability seems to be a fundamental psychological characteristic of sensory phenomena. Percept variability frequently differs from object to object and from dimension to dimension. Objects that have extreme values on one dimension are often perceived more uniformly with a lower variance than objects that are less extreme. For some product sets a dominant dimension such as flavor strength may have lower variance than a less dominant dimension. The assumption of differential percept variability can, of course, be tested. Recent food industry studies finding evidence for differential percept variability include References [11–13].

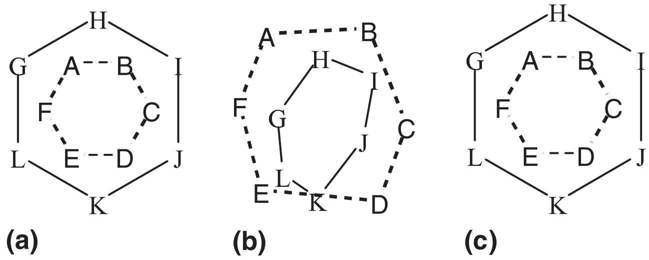

Figure 1. Results of Zinnes and MacKay's [14] hexagon simulation: (a) parametric configuration—the standard deviation of the objects in the inner hexagon is 1.6, the standard deviation of the objects in the outer hexagon is 1.0 and the width of the outer hexagon is 2.0; (b) best solution estimated by KYST; (c) best solution estimated by PROSCAL.

If differential percept variation had little or no effect on deterministic models' estimates of real and ideal sensory objects' locations, then the need to consider probabilistic models would not be compelling. The effect of differential percept variation is, however, not trivial and at times can be rather dramatic.

An early example of how differential variation can affect deterministic model estimation was provided in a hexagon simulation by Zinnes and MacKay [14]. A two-dimensional configuration of 12 points which formed the centroids of 12 normal isotropic distributions was created; six formed an inner hexagon and six formed an outer hexagon. The parametric configuration is given in Figure 1(a). The standard deviations of the inner hexagon points were 1.6 on each dimension; outer hexagon standard deviations were 1.0. The width of the outer hexagon was approximately 2.0. Thirty replications of all dissimilarity judgments were simulated by calculating Euclidean distances from independently sampled co-ordinates. The simulated dissimilarities were then evaluated and the estimated configurations were compared with the parametric configuration. A widely used nonmetric MDS program, KYST [15], estimated the deterministic solution, and the probabilistic MDS program PROSCAL estimated the probabilistic solution. (KYST and PROSCAL are available without charge from References [16,17].)

The deterministic solution, Figure 1(b), recovered two hexagons, but the hexagons were inverted—the hexagon that should have been on the inside was on the outside. In contrast, the probabilistic solution of Figure 1(c) was almost perfect. Parametric recovery is thus seen to be much better for the probabilistic model. The reason for this difference is that the probabilistic model can disentangle object variance from inter-object distances. Deterministic models confound variance and distance. The large variances of the inner hexagon objects create large dissimilarity values which the deterministic models mistakenly interpret as distances. To accommodate these large distances, the points of the inner hexagon are spread apart and end up being estimated as the outer hexagon. Deterministic point estimates are thus determined as much or more by variance than by central tendency.

Variance magnitude and variance structure determine how badly deterministic models perform at recovering

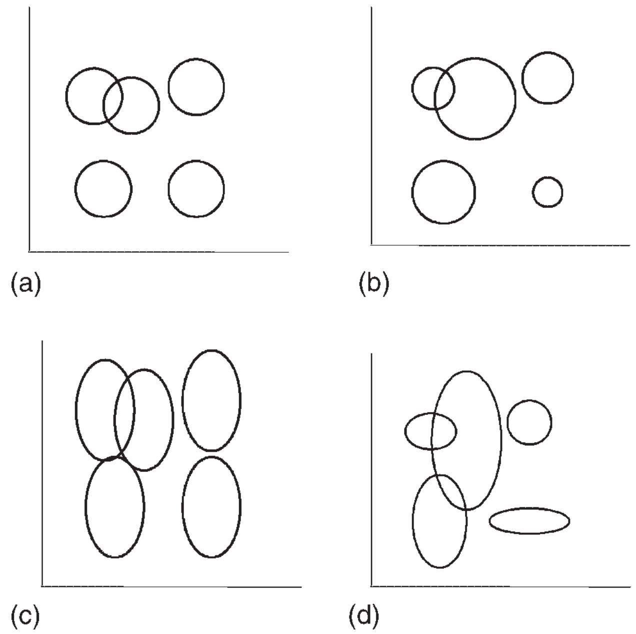

Figure 2. Four possible variance–covariance structures for five objects varying along two dimensions: (a) Case V isotropic; (b) Case III isotropic; (c) Case V anisotropic; (d) Case III anisotropic.

location parameters. For data with low variances, such as dissimilarities derived from instrumental product profiles, the difference between deterministic and probabilistic solutions may be slight. However, the error theory that accompanies probabilistic models may still offer some hypothesis-testing advantages. Data from human subjects, even from expert assessors, are usually characterized by substantial variance magnitudes. In these situations the degree of deterministic model recovery will depend upon the structure of the variances.

Four prototypical variance structures are illustrated in Figure 2. The ellipses define equal-likelihood contours. All the points on an equal-likelihood contour are one standard deviation from the centroid of the distribution. Following Thurstone [18], Case V models estimate the same variance structure for all objects, and Case III models permit different variances for different objects. Isotropic models estimate the same variance across dimensions for a particular object, and anisotropic models permit different variances. Anisotropic models may also have non-zero valued covariances.

Case V isotropic variance structures will, when evaluated with a deterministic model, have configurations that are most like the configurations estimated by a probabilistic model. This is because the expected distances between all pairs of objects are monotonically related to the corresponding distances between the distributions' centroids. (Expected values are defined in the next section.) The same is not true for anisotropic and Case III variance structures.

Probabilistic MDS models are thus seen to do a better job than deterministic models at representing and recovering the underlying psychological processes of sensory evaluation. Most of the probabilistic MDS models that have been proposed are ML models which are well suited for model testing. Many of the deterministic models are based upon decomposition theorems in linear algebra, have no error theory and must rely on heuristic rules of thumb, such as the eigenvalues-greater-than-one rule or scree tests, to evaluate hypotheses.

Paradoxically, while the mathematics of probabilistic models may seem more complicated than that of deterministic models [8], their estimates are frequently simpler. Probabilistic models are often able to represent objects in a lower-dimensional space than deterministic models. The reason for this is again the confusion of distance and variation by deterministic models. In their mistaken attempt to represent variances as distances, deterministic models employ extra dimensions.

A final advantage of probabilistic models is that their formulation leads to extensions relevant to product development and marketing. One extension is the ability to estimate a product's perceptual share or percentage of 'first choices'. This estimate comes naturally from the probabilistic formulation of the model. With deterministic models, *ad hoc* rules [19] must be used to estimate choice probabilities.

3. MATHEMATICAL DEVELOPMENT

PROSCAL is a PMDS program that represents objects by multivariate normal distributions and provides ML estimates of each distribution's mean and variance parameters. The location of an object i is represented by the estimated mean of its co-ordinates $\hat{\mu}_{ik}$ on the $k = 1, \ldots, r$ dimensions. The co-ordinates of all objects are referred to as the configuration. Objects may be real or ideal.

The probabilistic model upon which PROSCAL is based was first proposed by Hefner [20] for use with same–different judgments. For objects i, j in an r-dimensional space the distances d_{ij} were assumed to have Euclidean properties, specifically

$$d_{ij}^2 = \sum_{k=1}^{r} (x_{ik} - x_{jk})^2 \tag{1}$$

and the co-ordinates x_{ik} were assumed to be normally and independently distributed with mean μ_{ik} and variance σ^2. These assumptions have been relaxed. One is now allowed to use dissimilarity and similarity data as well as a variety of hedonic judgments—individually or in combination with one another. Variances may be uniquely specified for each object i on each dimension k in a Euclidean or city-block space. Response functions can be estimated to extend the ways proximity and hedonic data are related to the distance estimates of the ML model. For dissimilarity data δ_{ij}, one-, two- or three-coefficient linear–exponential response functions of the form $\delta_{ij} = a + bd_{ij}^c$ are used to express the relationship to the underlying latent distances d_{ij}. Default values of the three coefficients, (0, 1, 1), define an identity relationship. Complete, incomplete or replicated judgments may be used for each subject. A historical sketch of the development of probabilistic proximity models is provided in Reference [21].

The type of data, the variance structure, the dimensionality of the space, the metric (Euclidean or city-block) and the response function are the basic determinants of the ML function used in the estimation process. The simplest situation is a two-dimensional, isotropic Case V model for dissimilarities in a Euclidean space with an identity relationship as the response function.

Assuming

$$x_{ik} \sim N(\mu_{ik}, \sigma^2) \tag{2}$$

then let

$$d_{ijk} = x_{ik} - x_{jk} \sim N(\mu_{ijk}, 2\sigma^2)$$

where

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

Then

$$d_{ij}^2 = \sum_{k=1}^{2} d_{ijk}^2$$

and, from Reference [22],

$$d_{ij}^2/2\sigma^2 \sim \chi'^2_{\nu,\lambda_{ij}} \tag{3}$$

where $\chi'^2_{\nu,\lambda_{ij}}$ is the non-central chi-square distribution with $\nu = 2$ degrees of freedom and a non-centrality parameter

$$\lambda_{ij} = \sum_{k=1}^{2} \left(\mu_{ijk}^2/2\sigma^2\right)$$

Then, using the transformation technique [23] to find the probability density function (PDF) of d_{ij}, we take

$$\left|\frac{\partial}{\partial d}\left(d^2/2\sigma^2\right)\right|$$

and get

$$d_{ij} \sim \left(d_{ij}/\sigma^2\right)\chi'^2_{\nu,\lambda_{ij}} \tag{4}$$

The non-central chi-square PDF does not have a closed form, but good, rapidly converging approximations are available [14]. Letting $f(d\,|\,\boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p=2)$ be the generalized form of the Euclidean ($p=2$) PDF of d_{ij} in (4), where $\boldsymbol{\mu}$ is the vector of co-ordinate differences, $\boldsymbol{\Sigma}$ is the variance–covariance matrix and r is the dimensionality of the space, the ML objective may be conveniently expressed as maximizing

$$\ln L = \sum_{d} \ln f(d\,|\,\boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p=2) \tag{5}$$

In this simple case, $\boldsymbol{\Sigma}$ is a scalar matrix.

The expected value $E(d_{ij})$ can be approximated [24] by

$$E(d_{ij}) \approx \sigma_{ij}\left(\frac{2a-(1+b)}{2}\right)^{1/2}$$

where $a = \nu + \lambda_{ij}$ and $b = \lambda_{ij}/(\nu + \lambda_{ij})$.

Proceeding to the anisotropic and Case III conditions, we can temporarily assume that the covariances are zero and generalize (3) to

$$d_{ijk}^2 \Big/ \sum_{k=1}^{r} \sigma_{ijk}^2 \sim \chi'^2_{\nu,\lambda_{ij}}$$

where

$$\sigma_{ijk}^2 = \sigma_{ik}^2 + \sigma_{jk}^2$$

To accommodate the situation where the covariances are not zero, we note (dropping the i, j subscripts) that the distribution of d^2 is a specific example of a quadratic form

$$Q(\mathbf{W}) = \sum_{k=1}^{r} \sigma_k^2\left(w_k - \sqrt{\lambda_k}\right)^2$$

in normal variables distribution, where $\mathbf{W}$ is an r-element vector and $w_k \sim N(0, 1)$. Johnson and Kotz [25] (chap. 29) provide the transformations for expressing the non-zero-covariance case as a function of independent variables. The PDF f of d follows trivially for the Euclidean metric as

$$f(d\,|\,\boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p=2) = f(Q(\mathbf{W}))2d$$

where $f(Q(\mathbf{W}))$ is the PDF of the quadratic form.

To find $E(d)$, we again go to Reference [25] for the characteristic function of $Q(\mathbf{W})$ which then allows $E(d^2)$ to be calculated. Jensen and Solomon [26] derive moments of transformations of the type $(Q(\mathbf{W})/E(d^2))^h$ which, letting $h = ½$, allow $E(d)$ to be derived.

For the city-block ($p=1$) metric, we start with the observation that the difference on the RHS of

$$d_{ij} = \sum_{k=1}^{r} |x_{ik} - x_{jk}|$$

unlike in the Euclidean case, must be positive. Assuming, as with the Euclidean model, that the co-ordinates are normally distributed, the absolute differences of the city-block metric then follow a folded normal distribution. PDFs of one-dimensional folded normal distributions are provided by Leone *et al.* [27], from which we get

$$f(d) = \frac{1}{\sqrt{2\pi}\sigma}\left[\exp\left(-\frac{(d-\mu)^2}{2\sigma^2}\right) + \exp\left(-\frac{(d+\mu)^2}{2\sigma^2}\right)\right]$$

PDFs of distances derived from multidimensional folded normal distributions are provided by MacKay [28].

To obtain the expected values, an easy approach is to calculate the moment-generating function

$$m(t) = \int_0^\infty \exp(td)f(d)\mathrm{d}d$$

for the one-dimensional case, then take the log for the cumulant generation function,

$$c(t) = \ln\left\{\frac{1}{2}\exp\left(\frac{1}{2}t(-2\mu+\sigma^2 t)\right)\left[1+\exp(2\mu t) - \mathrm{erf}\left(\frac{\mu-\sigma^2 t}{\sqrt{2}\sigma}\right) + \exp(2\mu t)\mathrm{erf}\left(\frac{\mu+\sigma^2 t}{\sqrt{2}\sigma}\right)\right]\right\}$$

and evaluate by setting $t = 0$. Since the cumulants of a sum of independent random variables are equal to the sum of the cumulants, the mean (first cumulant) is readily found by summing over r dimensions to be

$$E(d\,|\,\boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p=1) = \sum_{k=1}^{r}\left[\frac{\sqrt{2/\pi}\sigma_k}{\exp(\mu_k^2/2\sigma_k^2)} + \mu_k \mathrm{erf}\left(\frac{\mu_k}{\sqrt{2}\sigma}\right)\right]$$

Up to this point we have just considered dissimilarities. Similarities, which are usually scaled to range between zero and one [7], are commonly related to dissimilarities by one of two functions: an exponential function in which $s(d) = \exp(-d)$ or a Gaussian function in which

$s(d) = \exp(-d^2)$. Theoretical and empirical reasons for these two functions are given in References [29,30]. The exponential function is generally associated with city-block distances and the Gaussian function with Euclidean distances, though Shepard [31] has proposed that the exponential function should be used with all metrics. Using the transformation technique, as before, to express the PDF of similarities as a function of the PDF of dissimilarities, we get for the exponential relationship

$$f(s \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p) = |1/s| f(-\ln(s) \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p)$$

and for the Gaussian relationship

$$f(s \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p) = \left|1/\left(2s\sqrt{-\ln(s)}\right)\right| \times f\left(\sqrt{-\ln(s)} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, r, p\right)$$

The impact on the likelihood of the linear–exponential response functions that are used to relate the judgments to the underlying latent distances is incorporated in a similar fashion through the use of the transformation technique. For simplicity we have been defining the PDFs of similarities and dissimilarities as functions of $\boldsymbol{\mu}, \Sigma$, r and p, but a vector $\mathbf{v}$ of ancillary parameters should also be added. Elements of $\mathbf{v}$, in addition to the three linear–exponential response function coefficients, could include response bias parameters and terms for modeling probabilistic components that enter the judgments directly instead of indirectly through the sensory objects.

When it comes to consumers' hedonic judgments, PROSCAL represents the judgments by using an unfolding model [32]. Unfolding models, which view utility as an inverse function of the distance from a real object to an ideal object, are particularly appropriate for sensory phenomena, because they capture another key psychological characteristic, satiety. Except for variables such as 'off-taste', most sensory variables are characterized by having their most preferred values in the middle, as opposed to the extremes, of a sensory continuum.

One of the most common hedonic judgments is the liking rating. When liking ratings are scaled so that low values indicate a high degree of liking, it should be obvious that a liking rating is just another distance measure, one that is between a real and an ideal object as opposed to being between two real objects. Everything that has been written above about the modeling of dissimilarity judgments thus applies to liking ratings as well.

Other hedonic judgments, such as preference ratios and binary preferential choices, are more complex, since they involve not just distances but ratios of distances. Preference ratio judgments are obtained by asking subjects to indicate which one of a pair of sensory objects they prefer and by how much they prefer that one over the other. The judgment ρ_{ijk} of object j over object k for ideal object i is modeled as an inverse function of the distance of d_{ij}/d_{ik}. A high preference ratio is associated with a low distance ratio, since the distances are measures of disutilities. For binary choices the binomial distribution is used to estimate the probability of choosing object j over object k as the probability of d_{ij}/d_{ik} being less than one. In the Euclidean metric the key to modeling distance ratios is the realization that a ratio of quadratic forms is itself an indefinite quadratic form. This can be seen by expressing the cumulative distribution function (CDF) of the squared distance ratio ρ^2 in the following way:

$$F(\rho^2) = P\left(\frac{\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}}{\mathbf{X}^{\mathrm{T}}\mathbf{B}\mathbf{X}} \leq \rho^2\right) = P\left(\mathbf{X}^{\mathrm{T}}\left(\mathbf{A} - \mathbf{B}\rho^2\right)\mathbf{X}\right) \leq 0$$

where $\mathbf{X}$ is a $2r$-element vector of co-ordinate differences, the first r elements for the numerator and the second r elements for the denominator, and $\mathbf{A}$ and $\mathbf{B}$ are $(2r \times 2r)$ matrices of the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \end{pmatrix}$$

Derivations of the likelihoods and expected values of distance ratios in city-block and Euclidean spaces may be found in References [28,33].

Dependent and independent sampling are issues that need to be addressed when considering probabilistic formulations of hedonic judgments. The essential distinction concerns the number of times an ideal distribution is sampled when a consumer makes a judgment. With liking ratings, for example, dependent sampling occurs when a consumer is assumed to sample just once from an ideal distribution and to use that sampled value when making all other liking-rating judgments. Independent sampling occurs when the consumer is assumed to draw new values from the ideal distribution each time an evaluation is made. First-choice probabilities are contingent upon the relative locations of real sensory objects with dependent sampling, but not with independent sampling. For some types of hedonic judgments, PROSCAL allows the modeling of either dependent or independent sampling. For other types of judgments, only independent sampling methods are allowed. The mathematical development of independent and dependent sampling is covered in Reference [34].

When using hedonic data, consumers are usually assigned to segments. Many different bases may be used for segmentation. One basis is geography. Persons living in one place are assigned to one segment while persons living in other places are assigned to other segments. Socio-economic variables and purchase history are also frequently used for consumer segmentation. When the hedonic judgments themselves are to be the basis of the segmentation, a simple procedure when the number of sensory objects is small is to assign people with the same first choice to the same segment. When the number of objects is large, a k-means cluster analysis of the hedonic data is frequently done to make initial segment assignments. PROSCAL can refine the initial assignments by computing mixture ML estimates using an expectation–maximization algorithm [35] which automatically assigns subjects to s segments. The resulting $s-1$ mixture model parameters must then be included in the vector $\mathbf{v}$ of ancillary parameters. Individual ideal point models [13] may be estimated as well, but this is not advised, since in situations where the number of parameters increases with the number of observations the ML estimates need not be consistent.

The PROSCAL model uses a least squares procedure to obtain initial values and an alternating direct search procedure to find the ML estimates. (Gradient methods do not do well with highly non-linear objective functions.) The alternating estimation process sequentially estimates the variances, centroids and response function parameters while holding the other two sets of estimates fixed. The process continues until the log likelihood function ceases to improve. While this method does not guarantee an optimal solution, a recent comparison of its performance using a single initial estimate with a deterministic model using 100 different initial estimates to avoid local optima problems was exceptionally favorable [21].

4. A WINE STUDY

Data for this study came from the Chardonnay Wine Tasting Exercise at the 2001 Pangborn Sensory Science Symposium in Dijon. Three California wines priced at $4, $7 and $19 and two French wines priced at $5 and $14 were evaluated. Thirty assessors evaluated the five wines on 15 sensory attributes, and 454 consumers evaluated the five wines on liking-rating scales. For each assessor a lower half matrix of 10 Euclidean distances was calculated on the basis of his/her sensory attribute evaluations. The consumers were divided into five segments on the basis of their most preferred wine.

4.1. Model selection

The probabilistic analysis began with Case V and Case III solutions in two dimensional isotropic and anisotropic spaces. To compare the solutions, use was made of Bozdogan's [36] CAIC criterion.

The CAIC criterion is but one of several approaches now being used to account for complexity when selecting a model. (See the March 2000 issue of the *Journal of Mathematical Psychology* for a review of contemporary model selection approaches.) CAIC penalizes likelihoods by the number of parameters estimated in the model and takes the form

$$\text{CAIC} = -2\ln L + cK$$

where L is the likelihood, K is equal to the number of independently estimated parameters and c is the cost of adding a parameter to the model. The number of independently estimated parameters for the isotropic Euclidean space is

$$K = m + q + n - r - r(r-1)/2 - 1$$

for m co-ordinates, q unique variances, n response model coefficients and r dimensions. The last three terms are subtracted for the centering, rotation and scale invariance of the solution. The rotational invariance term is omitted when the city-block metric is used and when an anisotropic Euclidean space is used. (The directionality of anisotropic space solutions fixes the orientation of the solution.) For the cost, $c = \ln(S) + 1$, where S is the sample size. The cost formula gives the CAIC statistic its consistency property [36]. The solution with the lowest CAIC value is selected.

Other criteria, such as traditional likelihood ratios, may also be used instead of CAIC. Our preference for information criterion statistics such as CAIC or the Bayesian criterion BIC [37] (which is very similar) is due to their sensitivity to differences in model complexity and to their ability to compare non-nested models. An example of where the ability to compare non-nested models is convenient is when one is comparing models in Euclidean and city-block spaces.

Table I. CAIC analysis for four models of wine profile data

Model	m	q	n	r	K	S	ln(L)	CAIC
Case V isotropic	10	1	2	2	9	300	−485.445	1031.22
Case V anisotropic	10	2	2	2	11	300	−482.442	1038.63
Case III isotropic	10	5	2	2	13	300	−484.446	1056.04
Case III anisotropic	10	10	2	2	19	300	−482.440	1092.25

The CAIC analysis for the four solutions is given in Table I. A two-parameter linear–exponential function was used to relate the derived dissimilarities to the underlying latent distances. All the log likelihoods are of a similar magnitude, though, as expected, the two isotropic models are both slightly lower than the corresponding anisotropic models and the two Case V models are both slightly lower than the corresponding Case III models. Given the small differences in the log likelihoods, the best CAIC score is determined largely by the complexity of the model. In this case the simplest model, Case V isotropic, is the winner. The solution, shown in Figure 3, was then compared with one- and three-dimensional Case V isotropic models. The CAIC score for the two-dimensional model was best.

For the consumers' liking ratings a similar four-step analysis was undertaken. The model with the lowest CAIC score was the two-dimensional Case V anisotropic model. The solution, shown in Figure 4, is highly similar to a 180° rotation of the dissimilarities analysis of Figure 3. The primary difference is whether C4 or F5 should be next to C19. All the ideal distributions are close to the corresponding real distributions.

4.2. Common space analysis

The similarity of Figures 3 and 4 raises the question of what would happen if the two sets of data were analyzed simultaneously. Since the derived dissimilarities and liking ratings are independent of one another, the common space

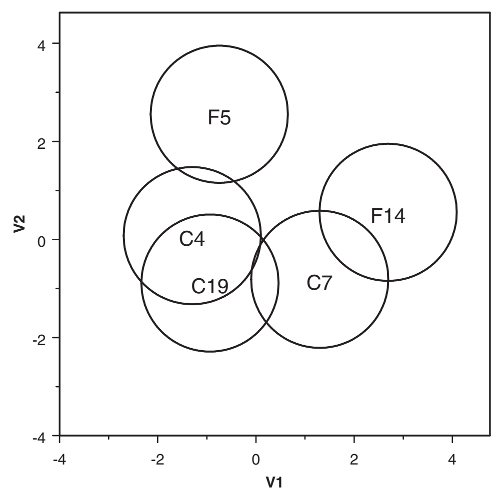

Figure 3. Case V isotropic solution for derived dissimilarities in the wine study. For the centroid labels the letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price.

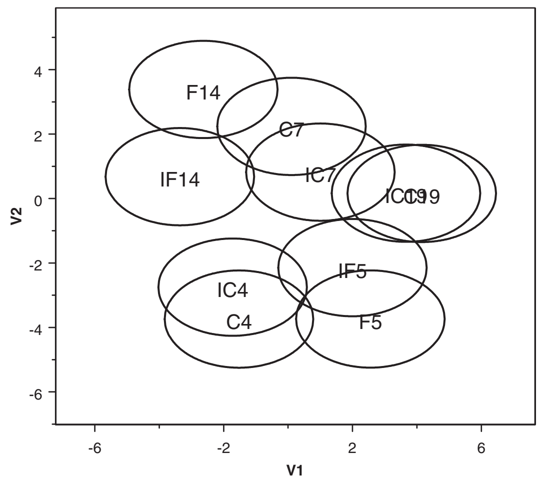

Figure 4. Case V anisotropic solution for liking ratings in the wine study. For the centroid labels the letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price. Ideal centroids begin with the letter I.

likelihood L_c is simply $L_c = L_h L_p$, where L_h is the likelihood of the hedonic judgments and L_p is the likelihood of the proximity judgments. The likelihood L_h is calculated over the hedonic data and L_p is calculated over the proximity data. A two-coefficient response function was applied to the derived dissimilarities to make their scale comparable to the scale of the hedonic judgments.

Another four-step analysis for the combined data sets was conducted and a Case III isotropic space model was selected as the one with the lowest CAIC score. The solution, shown in Figure 5, bears some resemblance to but is distinctively different from the liking-ratings solution of Figure 4. In our experience it is quite common for a common space analysis to result in a richer solution than is provided with either data set by itself. However, the question of whether the common space solution is justified must always be addressed.

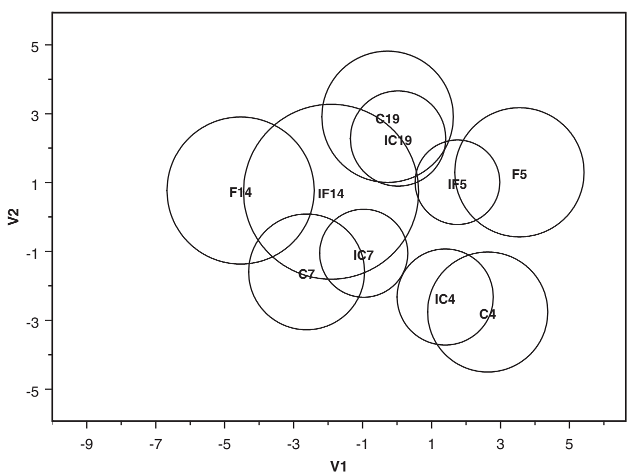

Figure 5. Case III isotropic solution for common space of derived dissimilarities and liking ratings in the wine study. For the centroid labels the letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price. Ideal centroids begin with the letter I.

Sometimes the differences in the proximity and hedonic data are so great that they should not be combined.

Fortunately, it is possible to test the meaningfulness of the common space solution by comparing its CAIC score with the CAIC score for the two independent hedonic and proximity analyses. The sum of the hedonic and proximity log likelihoods will be higher than the common space likelihood, but the common space likelihood has more degrees of freedom, since it only estimates one set of real object parameters. The CAIC score of the common space for the wine data was 10 826 and the CAIC score of the two independent analyses was 10 866, thus giving support to the common space solution. When the common space CAIC score is greater than the CAIC score for the two independent analyses, the hedonic analysis must proceed without the benefit of the proximity data.

Alternative methods are available for identifying the dimensions of the common space solution. The simplest objective measure is to correlate the projections of the centroids on the two axes with the means of the attribute profiles from the expert assessors. The horizontal axis, V1, had negative correlations of less than -0.9 ($p < 0.05$) with the attributes of body, butter and oaky-smokey. The vertical axis, V2, had a negative correlation of less than -0.9 ($p < 0.05$) with the flavor apricot. The projections on both axes were also correlated with the mean liking ratings. Both correlations were close to zero and insignificant—a good indication that in the common space solution the real objects' configuration is capturing their sensory structure and not their hedonic structure.

4.3. Perceptual share estimation

Perceptual share is defined as the estimated percentage of consumers who will choose a product as their first choice. Perceptual share may be estimated by segment or for the entire market. In this wine study, segments are defined on the basis of the wines that subjects said they liked the best. From an inspection of Figure 5 it is obvious that there is a lot of variability in the perception of both the real objects and the ideal objects for each segment. Consumers in segment IF14 appear to be very uncertain about what they prefer and consumers in segment IF5 appear to be fairly certain about what they prefer. Brand loyalty should thus differ by segment.

Since the distributions and segment sizes are known, it is possible to estimate—either by numerical approximation or by simulation—the percentage of consumers in each segment who will be expected on a new trial to choose their segment's brand. Weighting these percentages by the sizes of the segments gives us the weighted aggregate perceptual share. Results are given in Table II. Two points are of particular interest. The first is the exceptionally close fit of the weighted aggregate and actual shares. The second is the variation in brand loyalty which is indicated by the values along the diagonal of the matrix. Products C4 and C19 are seen to have high repeat purchase estimates, 0.57 and 0.54, product F14 (the one with the high variance on the ideal) has a low repeat rate of 0.26 and the two remaining products, C7 and F5, are in the middle with estimated repeat rates of 0.46 and 0.43.

Table II. Estimated perceptual shares by segment for the wine study

	Product				
Segment	C7[a]	C4	C19	F14	F5
IC7	0.46[b]	0.17	0.15	0.14	0.08
IC4	0.19	0.57	0.07	0.03	0.15
IC19	0.10	0.05	0.53	0.11	0.21
IF14	0.29	0.09	0.27	0.26	0.10
IF5	0.07	0.17	0.29	0.04	0.43
Weighted aggregate	0.20	0.22	0.28	0.10	0.21
Actual	0.20	0.23	0.28	0.09	0.21

[a]For the row and column labels the letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price. Segments, consisting of consumers expressing the highest liking for the same wine, begin with the letter I.
[b]46% of the consumers who were in segment IC7, consumers whose highest liking was for wine C7, are estimated as choosing product C7 on a new occasion.

While deterministic models do not allow the explicit estimation of repeat purchase rates, it is possible to make inferences about brand loyalty from the relative positioning of the point estimates. To illustrate, consider the configuration in Figure 6, which is the deterministic unfolding solution estimated by KYST from the consumers' liking ratings.

The ideal point for consumers preferring wine F14 is close to the real point for wine F14. Both the real and ideal points for F14 are far from the other points, indicating that F14 consumers are expected to have a high degree of loyalty. Two of the ideals, IC7 and IF5, are closer to other products than to their own brands, thus indicating a low degree of loyalty. The customers for C19 and C4 are in between these extremes and are thus evaluated as having a moderate degree of brand loyalty. As with the simulation, variance appears to be a major driver in determining the location of the points in the deterministic analysis.

To evaluate the very different brand loyalty estimates of PROSCAL and KYST, we go back to the original liking rating

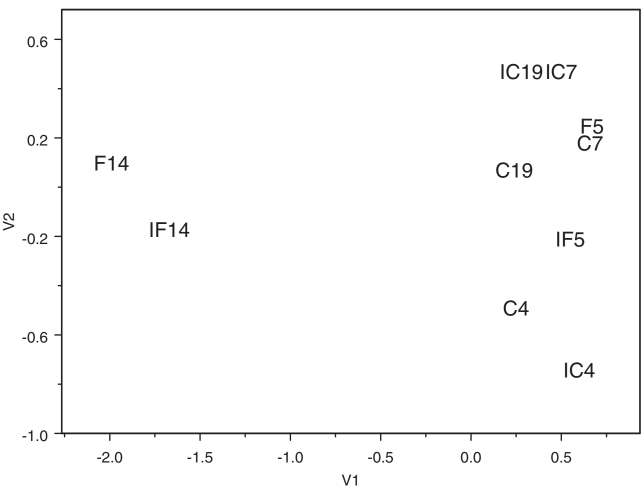

Figure 6. Deterministic unfolding solution for liking-rating data. The letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price. Ideal points begin with the letter I.

data and, for each consumer, calculate the second moment not about the mean but about the score for the most preferred brand, *i.e.*, the brand that defines the segment. A high value will indicate a high degree of brand loyalty, because it shows that the other brands are far away from the score of the most preferred brand. The equality of the means of the second-moments was tested with ANOVA ($p < 0.03$). The brand with the least loyalty (lowest second moment value) is F14. Brands C4 and C19 have the highest loyalty (highest second-moment value) and brands C7 and F5 are in the middle. This ordering is exactly the same as that derived from PROSCAL and is completely the opposite from that derived by KYST.

4.4. New product analysis

Perceptual share can be used as a criterion for evaluating specific new product strategies and for estimating an optimal new product entry. Perceptual share can also be used as an input for market share predictions which incorporate economic and competitive effects that are beyond the sensory realm.

As an example of a specific new product strategy evaluation, let us assume that the five Chardonnay wines in this study are the major brands in the market of interest and that two strategies are being considered by a manufacturer that is new to this market. The first strategy is a copycat strategy in which the manufacturer tries to copy the dominant wine, C19. The second strategy is a compromise strategy in which the manufacturer tries to create a product that is positioned in the middle of the ideal distributions for products C4, C7, C19 and F5. Let us also assume that the manufacturer is able to successfully create these new products and that the standard deviational ellipsoids for the new products are equal to C19 for the first strategy and equal to the means of the standard deviations for the four products of the second strategy. Higher standard deviation values may be assumed if there is evidence that the new products will be less successful in defining their images.

Table III gives the estimated perceptual shares for the two strategies. As expected, the copycat strategy will draw most of its share from the copied product. The share contributions from the different segments are much more even when a compromise product strategy is employed. Overall, the compromise strategy is estimated as being almost 50% better than the copycat strategy.

Generalizing from the discrete what-if modeling strategy, it is possible to estimate perceptual shares throughout

Table III. Estimated perceptual shares for two new wine product strategies

	Segment					
Strategy	IC7[a]	IC4	IC19	IF14	IF5	Total
Copycat	0.16[b]	0.06	0.35	0.21	0.21	0.21
Compromise	0.31	0.27	0.29	0.22	0.34	0.30

[a]For the column labels the letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price.
[b]16% of segment IC7's perceptual share will go to the copycat product.

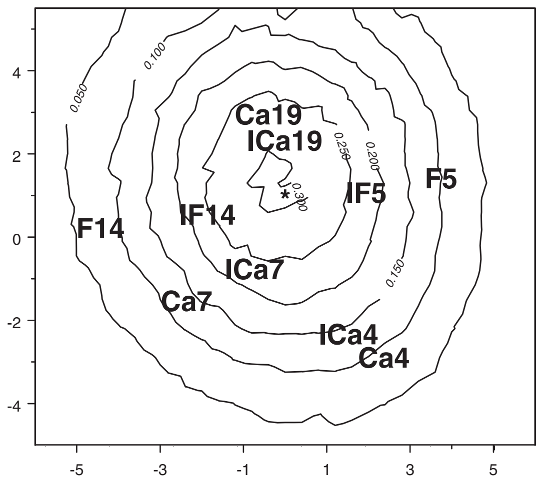

Figure 7. Predicted new product perceptual shares for an optimal single-product strategy. For the centroid labels the letters C and F indicate the place of origin, C for California and F for France, and the number indicates the price. Ideal centroids begin with the letter I. The maximum predicted perceptual share (0.320) location is indicated by the asterisk.

the space and to use contour mapping to find the part(s) of the space with the highest potential. An analysis based upon the assumption that the magnitude of the standard deviation was the same as that of the nearest real sensory object was conducted. The results are illustrated in Figure 7. From this we see that the compromise strategy was very near the optimal strategy. The estimated perceptual share for the optimal strategy was 0.32. If desired, sensitivity analyses can be conducted under less optimistic or more optimistic uncertainty assumptions. It is also possible to predict the perceptual shares that would result under alternative competitive reactions to the new product introduction.

5. A PEA STUDY

Sensory and instrumental data on peas [38] were made available for contributors to this issue. Two replicates of complete sensory profiles on six sensory attributes for 60 objects were provided by 10 assessors. Instrumental data were in the form of 116 averaged values for each object from a near-infrared (NIR) spectroscopy analysis.

The following comparative analysis was undertaken to see if PMDS would provide any advantage over a traditional analysis in comparing independent latent structures estimated from the sensory and instrumental data. The comparison is a challenging one for PMDS, since only a single sample of averaged NIR data is available.

Principal component analysis (PCA) was selected as the traditional candidate to compare with PMDS. PCA was chosen over principal component regression and partial least squares regression, because our interest is not in predicting one set of variables from another set of variables but in comparing independently estimated latent structures from two different sets of data. Correlations were used as PCA input. Comparisons of PCA with PMDS and deterministic MDS are provided in References [6,7,12].

For the PMDS analysis the sensory profile data were centered for each replicate of each judge and Euclidean distances were calculated among all pairs of peas for each of the 20 sets of profile data. Euclidean distances were also calculated for all pairs of peas from the single set of NIR profile data. For PCA the centered sensory profile matrices for the 20 assessments may be summed, averaged or extended by adjoining the 20 60×6 profiles horizontally into a 60×120 matrix or vertically into a 1200×6 matrix. When different assessors make judgments or when there is inconsistency in the assessment process, summation or averaging is not appropriate [39]! Selection of a procedure to adjoin the profiles depends upon the focus of the study. In sensory analyses, where the focus is on obtaining latent space representations for the real products, a horizontal adjoining is more common [40,41]. A PCA analysis was also conducted on the single set of NIR profile data.

5.1. Dimensionality and variance structure

For the derived sensory and NIR dissimilarity matrices, PMDS analyses were conducted in one to three dimensions for Case V and Case III isotropic and anisotropic variance structures. CAIC was again used as a criterion for selecting the best sensory and the best NIR solution. For both data sets the best criterion (lowest CAIC score) was obtained for two-dimensional Case V anisotropic space models. For both data sets the first (horizontal) dimension had the lowest estimated variance, indicating greater sensitivity. Dominance of the first dimension was also indicated by the ratio of the range of values on the first to the second dimension, which was 3.32 for the sensory data and 7.99 for the NIR data.

For the sensory PCA analysis the number of components to extract varies widely with the criterion selected. Using the eigenvalue-greater-than-one criterion results in the selection of 17 components, while a scree test indicates one component using Cattell and Jaspers' [42] suggestion that, when plotting eigenroots against the root's component number, the number immediately before the straight line begins should be the extracted number of roots. The 17-component solution was used because of the relatively high amount of variance (28%) explained by components 2–17. (The first component explained 58% of the variance.) For the NIR PCA analysis, both the eigenvalue-greater-than-one criterion and the scree test indicated only one component. In terms of the general structure of the space the PMDS analyses are thus congruent in their interpretations, while the PCA analyses under the eigenvalue-greater than-one rule are not.

5.2. Configurations

Two-dimensional configurations (not shown) were obtained from the PMDS analyses, one for the sensory data and one for the NIR data. Expected distances were calculated among all pairs of points in each configuration and correlated with one another as a measure of the congruence of the configurations. The correlation, 0.665 ($p < 0.01$), was just slightly higher than the correlation of the two sets of input data, 0.661, which serves as a benchmark. (The distances used for the benchmark were calculated directly from the mean sensory profile and the NIR profile.)

To form a corresponding criterion for the PCA analyses, weighted Euclidean distances among all pairs of peas were calculated from the estimated principal component scores, with the weights determined by the percentage of variance accounted for by each component [43]. The correlation, just below the benchmark at 0.632, was close to that obtained with PMDS. (With unweighted Euclidean distances, which are often used, the correlation goes down to 0.075.) However, it should be remembered that the two PMDS analyses required the estimation of only four dimensions, while the PCA analyses required the estimation of 18.

While the PMDS analysis indicated that both sets of data had the same structure with respect to dimensionality and variance, the question of whether the configurations are the same remains to be tested. Our approach to this problem is similar to the one followed in the previous example, where we looked at whether proximity and hedonic data could be estimated in a common space. Specifically, we start by combining the two sets of data, estimating a single solution, calculating a CAIC statistic and comparing this with what is obtained if the two sets of data are analyzed separately. To make the sensory and NIR data comparable, PROSCAL standardized all the lower-half distance matrices that were used as input data to have a root mean square value of 1.0. The sum of the log likelihoods for the two independent analyses was −1024, while the log likelihood of the combined data sets was much lower, −7559. The enhanced degrees of freedom afforded by the combined data set analysis were not enough to offset the difference in log likelihoods. The CAIC score for the sum of the two independent analyses was 3430 and for the combined data sets the CAIC score was 15 810. This result is consistent with the moderate, though significant, correlation of the expected distances from the sensory and NIR analyses. We thus conclude that different latent structures underlie the two data sets.

6. DISCUSSION

In our theoretical development we have seen that the PMDS model PROSCAL is able to successfully capture sensory object variability and thus avoid the deterministic dilemma of confounding variance and distance. In our examples we have explored the hypothesis-testing capability of probabilistic models and seen how they can evaluate whether sensory and hedonic data sets should be combined and whether sensory and instrumental data sets have similar underlying latent structures.

Testing with ML models is often a process of comparing constrained models with unconstrained (or less constrained) models. In our two examples, the common space analysis and combined data set analysis, estimates are constrained in the sense that two data sets (sensory and hedonic for the first application, sensory and instrumental for the second) share common parameter estimates. Similarly, we can think of a Case V analysis as being a constrained version of a Case III analysis and we can think of an isotropic analysis as being a constrained version of an anisotropic analysis.

Constraints can be used in other productive ways as well. If a manufacturer is testing a new product which is to be a replacement for an existing product, then the test of whether the new product is seen as being identical to the existing product is conducted by doing an analysis where the co-ordinates of the replacement and existing product are constrained to be the same. The CAIC score for this analysis is then compared with the CAIC score for an analysis where the co-ordinate constraint is not imposed and the two products are allowed to have their own estimates.

In product development, experimental designs are frequently used to design the products being tested. Suppose we have a simple 4×3 design in which subjects are asked to provide proximity evaluations of the 12 products. We would expect that the final configuration should have a lattice shape but it probably will not be exact. Departure from a true lattice shape may be due to measurement error or it may be due to a fundamental difference in the psychological and physical spaces. To see if measurement error is the cause, impose the constraints that all co-ordinate estimates in each row be the same and that all co-ordinate estimates in each column be the same. The log likelihood for the unconstrained solution will be higher than the log likelihood for the constrained solution, but the difference in the number of estimated co-ordinates, seven for the constrained solution vs 24 for the unconstrained solution, will cause the CAIC score for the constrained solution to be better (lower) if there is no fundamental difference in the psychological and physical spaces.

As this discussion of constraints indicates, the number of decisions that must be made by an analyst when using a program such as PROSCAL is greater than what is required when using deterministic MDS programs. However, a PMDS approach is able to provide additional insights that can lead to better decisions. An unexpected benefit in our application of these models is that they frequently provide a bridge that enhances communication between sensory analysts and brand managers in marketing.

REFERENCES

1. Rivisk E. Understanding latent phenomena. In *Multivariate Analysis of Data in Sensory Science*, Næs T, Risvik E (eds). Elsevier: Amsterdam, 1996; 5–35.
2. Bieber SL, Smith DV. Multivariate analysis of sensory data: a comparison of methods. *Chem. Sens.* 1986; **11**: 19–47.
3. MacFie HJJ, Thomson DMH. Multidimensional scaling. In *Sensory Analysis of Foods*, Piggott JRR (ed.). Elsevier: London, 1984; 351–375.
4. Popper R, Heymann H. Analyzing differences among products and panelists by multidimensional scaling. In *Multivariate Analysis of Data in Sensory Science*, Næs T, Risvik E (eds). Elsevier: Amsterdam, 1996; 159–184.
5. Davison M. *Multidimensional Scaling*. Krieger: Malabar, FL, 1992.
6. Borg I, Groenen P. *Modern Multidimensional Scaling: Theory and Applications*. Springer: New York, 1997.
7. Cox TF, Cox MAA. *Multidimensional Scaling*. Chapman and Hall: Boca Raton, FL, 2001.
8. Young FW, Hamer RM. *Multidimensional Scaling: History, Theory and Applications*. Erlbaum: Hillsdale, NJ, 1987.
9. Bossuyt P. *A Comparison of Probabilistic Unfolding Theories for Paired Comparisons Data*. Springer: Berlin, 1990.
10. MacKay DB. Alternative probabilistic scaling models for spatial data. *Geogr. Anal.* 1983; **15**: 173–186.

11. MacKay DB. Probabilistic unfolding models for sensory data. *Food Qual. Prefer*. 2001; **12**: 427–436.
12. MacKay D, O'Mahony M. Sensory profiling with probabilistic multidimensional scaling. *J. Sens. Stud*. 2002; **17**: 461–481.
13. Ennis DM. Foundation of sensory science. In *Viewpoints and Controversies in Sensory Science and Consumer Product Testing*, Moskowitz HR, Munoz AM, Gacula MC (eds). Food and Nutrition Press: Trumbull, CT, 2003; 391–432.
14. Zinnes JL, MacKay DB. A probabilistic multidimensional scaling approach: complete and incomplete data. *Psychometrika* 1983; **48**: 27–48.
15. Kruskal JB, Wish M. *Multidimensional Scaling*. Sage: Beverly Hills, CA, 1978.
16. The Netlib Repository Website. [Online]. Available: http://www.netlib.org/mds/ [2 June 2004].
17. The Proscal Website.[Online]. Available: http://proscal.com/ [2 June 2004].
18. Thurstone LL. A law of comparative judgment. *Psychol. Rev*. 1927; **34**: 273–286.
19. Lehmann DR. Evaluating marketing strategy in a multiple brand market. *J. Business Admin*. 1971; **3**: 15–26.
20. Hefner RA. *PhD Thesis*, Extensions of the Law of Comparative Judgement to Discriminal and Multidimensional Stimuli, University of Michigan, Ann Arbor, MI, 1958.
21. MacKay DB, Lilly B. Percept variance, subadditivity and the metric classification of similarity and dissimilarity data. *J. Classif*. 2004; **21**: 185–206.
22. Suppes P, Zinnes JL. Basic measurement theory. In *Handbook of Mathematical Psychology*, Vol. I, Luce RD, Bush RR, Galanter E (eds). Wiley: New York, 1963; 1–76.
23. Mood AM, Graybill FA, Boes DC. *Introduction to the Theory of Statistics*. McGraw-Hill: New York, 1974.
24. Patniak PB. The non-central chi-square and F-distributions and their applications. *Biometrika* 1949; **36**: 202–232.
25. Johnson NL, Kotz S. *Continuous Univariate Distributions—2*. Wiley: New York, 1970.
26. Jensen DR, Solomon H. A Gaussian approximation to the distribution of a definite quadratic form. *J. Am. Statist. Assoc*. 1972; **67**: 898–902.
27. Leone FC, Nelson LS, Nottingham RB. The folded normal distribution. *Technometrics* 1961; **3**: 543–550.
28. MacKay DB. Probabilistic multidimensional scaling using a city-block metric. *J. Math. Psychol*. 2001; **45**: 21–37.
29. Shepard RN. Stimulus and response generalization: deduction of the generalization gradient from a trace model. *Psychol. Rev*. 1958; **4**: 242–256.
30. Kornbret DE. Theoretical and empirical comparison of Luce's choice model and logistic Thurstone model of categorical judgment. *Percept. Psychophys*. 1978; **24**: 193–208.
31. Shepard RN. Toward a universal law of generalization for psychological science. *Science* 1987; **237**: 1317–1323.
32. Coombs CH. *A Theory of Data*. Mathesis Press: Ann Arbor, MI, 1976.
33. MacKay DB, Zinnes JL. Probabilistic multidimensional unfolding: an anisotropic model for preference ratio judgments. *J. Math. Psychol.* 1995; **39**: 99–111.
34. MacKay DB, Easley RF, Zinnes JL. A single ideal point model for market structure analysis. *J. Market. Res*. 1995; **32**: 433–443.
35. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. R. Statist. Soc. B* 1977; **39**: 1–38.
36. Bozdogan H. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika* 1987; **52**: 345–370.
37. Schwarz G. Estimating the dimension of a model. *Ann. Statist*. 1978; **6**: 461–464.
38. Næs T, Kowalski BR. Predicting sensory profiles from external instrumental measurements. *Food Qual. Prefer*. 1989; **4/5**: 135–147.
39. Gorsuch RL. *Factor Analysis*. Erlbaum: Hillsdale, NJ, 1983.
40. Bro R. *PhD Thesis*, University of Amsterdam/Royal Veterinary and Agricultural University, Holbæk, Multi-way Analysis in the Food Industry: Models, Algorithms, and Applications, 1998.
41. Dijksterhuis G. Procrustes analysis in sensory research. In *Multivariate Analysis of Data in Sensory Science*, Næs T, Risvik E (eds). Elsevier: Amsterdam, 1996; 185–219.
42. Cattell RB, Jaspers J. A general plasmode (No. 30-10-5-2) for factor analytic exercises and research. *Multivar. Behav. Res. Monogr*. 1967; **67–3**.
43. Rummel RJ. *Applied Factor Analysis*. Northwestern University Press: Evanston IL, 1970.