Food Quality and Preference 12 (2001) 427–436

Food Quality and Preference

www.elsevier.com/locate/foodqual

Probabilistic unfolding models for sensory data

David B. MacKay *

Kelley School of Business, Indiana University, 1309 E Tenth Street, Bloomington, IN 47405-1701, USA

Received 11 July 2000; received in revised form 26 December 2000; accepted 5 February 2001

Abstract

Unfolding models are conceptually appealing for the analysis of consumers' hedonic evaluations of food products. The appeal of the unfolding model is three fold; its conceptual simplicity, spatial character, and assumption of satiety — more is not always better. Unfortunately, the success of unfolding models does not always match their appeal. Reformulating unfolding models in a probabilistic framework is shown to improve their success, extend their application and further enhance their conceptual attraction. Data provided by the organizers of The Fifth Sensometrics Meeting are used to illustrate the proposed reformulation.

Keywords: Multidimensional unfolding; Hedonic ratings; Attribute ratings; Product maps; Multidimensional scaling

1. Introduction

In their simplest form, unfolding models (Coombs, 1964) start with consumers' hedonic liking ratings for real products and estimate coordinates in a multidimensional space to represent both these real products and ideal products. Coordinates are estimated so that the resulting distances between an ideal product and the real products inversely approximate the hedonic evaluations — large distances indicate minimal liking. Attribute information is not required. Identification of the dimensions of the space comes either from the analyst's knowledge of the objects being studied or from external information. Depending upon the analyst's objectives, ideal products may be estimated for individual consumers or for segments of consumers.

Differences in how traditional and the proposed probabilistic unfolding model conceptualize consumers' psychological processes are illustrated in Fig. 1. The traditional model, in the panel on the right, conceptualizes the real and ideal objects as points. In this hypothetical situation, the consumer subject prefers drink D2 to D1 since D2 is closer to the subject's ideal product *S*. If the dimensions are identified as measuring aroma and sweetness, then we conclude that the subject prefers moderate values of aroma and sweetness to strong or weak values. The proposed model, in the panel on the left, conceptualizes the real and ideal objects as multivariate normal distributions. Each symbol represents a value sampled from these distributions. Even though the centroids of the distributions in the two panels are the same, drink D1 will be preferred more often than D2 with the proposed model since there is more overlap in the distributions of D1 and *S* than there is in the distributions of D2 and *S*. Stated differently, the expected distance between D1 and *S* is less than the expected distance between D2 and *S*. The ellipses in the panel on the left represent the unit standard deviational contours of the three objects.

In the following sections, we shall discuss selected properties of the proposed model, briefly describe how the model works, contrast the application of the proposed and traditional models with consumer data on beverage attributes and likings, and discuss the limitations and benefits of using probabilistic unfolding models with sensory data.

2. Selected properties

Both economics and psychology have contributed models for helping us better understand consumer behavior. Unlike many economic models, the unfolding model has the compelling property of satiety. It allows us to model situations in which preferred products are

* Tel.: +1-812-855-1009; fax: +1-812-855-6440.
E-mail address: mackay@indiana.edu

0950-3293/01/$ - see front matter
PII: S0950-3293(01)00033-7

characterized by moderate as well as extreme values of their attributes.

Unfolding models assume that products exist in a common space, one that is usually multidimensional in nature. Differences in product preferences among consumers are assumed to be due to differences in their ideals — some consumers prefer sweet products and others prefer less sweet products. Consumers are assumed to perceive the products in a common fashion. With traditional deterministic models, the common space assumption is very severe. It requires, for example, that on each salient dimension, *all* consumers consistently view the coordinate or attribute values of the products in exactly the same order.

It does not take a lot of experience collecting perceived product attribute data to realize that these consistency assumptions do not hold, even for a single subject. Differences in product perception across subjects can be very high.

Choice consistency is also a property of traditional models. To comport with the model, if a consumer prefers product *A* to product *B* and product *B* to product *C*, then product *A* must always be preferred to product *C*. Indeed, if a consumer prefers product *A* to product *B* once, product *B* will never be chosen. In Fig. 1, the right hand panel tells us that *S* always prefers D2 to D1. For the probabilistic model illustrated in the left hand panel, D1 is usually preferred to D2 but D2 is frequently chosen.

Violations of model assumptions are not new, they happen all the time. What makes these violations worth noting is (1) that they are the rule, not the exception, and (2) that failure to account for these violations will result in systematically biased results. Fig. 2 illustrates the nature of this bias. In the left panel of Fig. 2 are the mean and variance parameters (represented by the standard deviational ellipses) for eight real objects, *A*–*H*, and four ideal objects, *I*–*L*. The first four real objects and the four ideal objects have relatively small variances while the second four real objects have relatively large variances. (While these hypothetical simulation variances may seem large, in our experience, they are not that uncommon for actual empirical studies.)

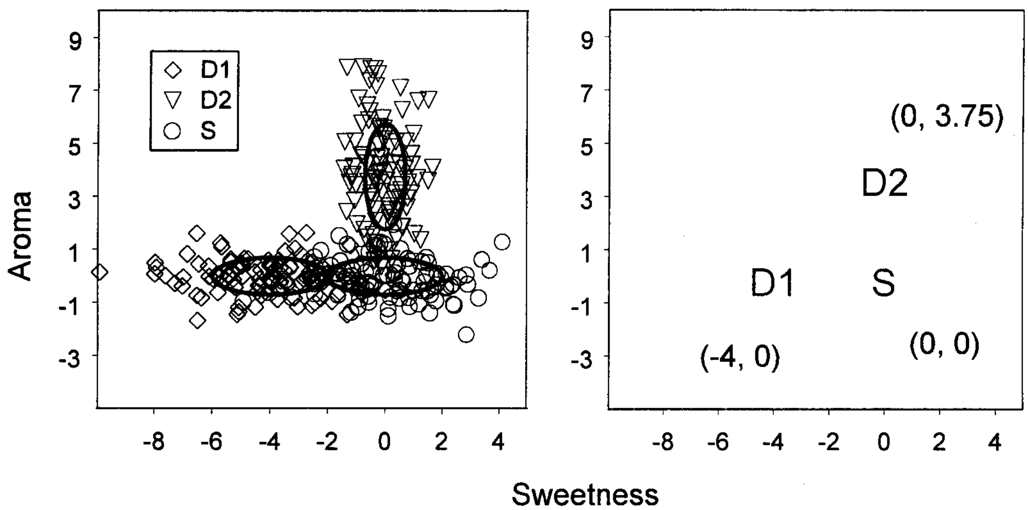

Fig. 1. Probabilistic and deterministic models of a subject *S*'s preference for drinks D1 and D2.

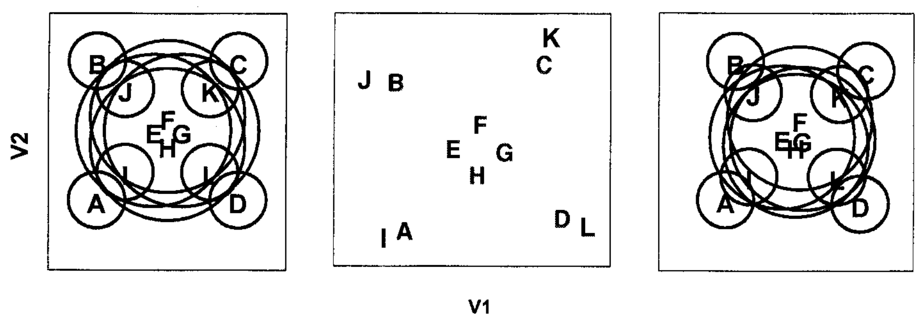

Fig. 2. Parameters, deterministic estimates and probabilistic estimates for eight real (*A*–*H*) and four ideal (*I*–*L*) objects. Estimates are from simulated hedonic judgments.

The middle panel of Fig. 2 illustrates the deterministic, nonmetric solution provided by KYST (Kruskal, Young, & Seery, 1977), a popular traditional unfolding program, when provided with the mean liking ratings for the eight real objects of 5600 simulated subjects, 1400 for each ideal object. It is observed that the traditional solution has inverted the locations of the first four real objects and the four ideal objects. This bias will not go away if the sample size increases; it will hold even for infinitely large samples. If dimensions V1 and V2 were identified, the traditional solution would tell product designers that all consumers preferred products with extreme values on both dimensions. This would clearly be bad advice.

The systematic bias shown in the middle panel has a very intuitive explanation. If two objects have fixed centroids, the expected distance and thus the average "disliking" between them will increase as one or more of the variance magnitudes increases. When judgments characterized by high variance are put into a deterministic model that does not admit variances, the effect of the variances is accommodated by estimating the points as being farther apart than the centroids actually are. Note that in this symmetric example, the ideal objects all move toward the periphery of the space. In so doing, the relatively short distances of the four ideal to the first four real products are maintained and the longer distance to the second four real products, the high variance products, is accommodated. In asymmetric situations, a common deterministic accommodation to differential variances is to produce degenerate solutions in which all the centroids of one set, the real or ideal objects, are estimated as having similar values.

The right panel of Fig. 2 shows the parameter estimates provided by PROSCAL (MacKay & Zinnes, 2000), a program for probabilistic unfolding and probabilistic mapping. The estimates are not perfect but they are very good. The correlation of inter-centroid distances in the first and last panel is 0.996. The actual variance values are 0.04 and 0.30 while the estimated variances are 0.04 and 0.28.

So far, we have seen that the inclusion of variance parameters provides a process that is more consistent with what we know of consumer behavior and that does not have the estimation bias of deterministic models that exclude variance parameters. Another advantage is that the new models permit the estimation of the probability with which a real object will be chosen. Deterministic models have no basis for estimating choice probabilities and must rely on ad hoc heuristics. To estimate choice probabilities using a probabilistic model, the analyst must first specify whether independent or dependent sampling is assumed.

To illustrate independent and dependent sampling, consider the example of Fig. 3 where a single subject *S* is making a choice among three real products, *A*, *B* and *C*. Under independent sampling, when *S* estimates the distance from the ideal to a real product, values are sampled from the distribution about *S* and the distribution about the real product and a distance between the two is calculated. When a second product is evaluated, a separate independent sample from *S* and a sample from the new real product is undertaken again. Integration over the joint density function or simulation can be used to estimate the first choice probabilities, which, in this case, are 0.72 for product *A* and 0.14 for both products *B* and *C*. Under dependent sampling, a single sample is used to draw a value from the distribution about *S*. This value is then compared to the values sampled from the distributions about *A*, *B* and *C*. Here, the first choice probabilities are 0.72, 0.08 and 0.20. Under dependent sampling, a real product's first choice probability is thus contingent upon the locations of all the products, not just upon the relation of the real product to the ideal product. If the analyst believes that independent sampling occurs in the laboratory and dependent sampling occurs in the field, then it is possible to provide estimates under independence assumptions and make first choice predictions under dependence assumptions.

For those not used to using probabilistic unfolding models, the solutions of probabilistic unfolding models may present some surprises. It is, for example, possible for a real product whose centroid is close to the centroid of an ideal product to have a very low probability of being chosen. One cause for this may be that the real product has a very large variance. As a result, for any one decision, the chances are high that the real object is actually a good distance away from the ideal object.

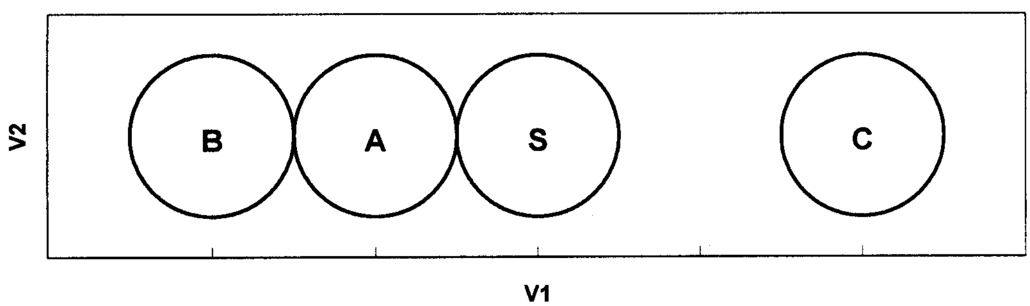

Fig. 3. Dependent and independent sampling among three objects (*A*–*C*) by one subject (*S*).

Stated differently, it is often true that expected distances are non-monotonic with distances among centroids.

The relation of variances to choice probabilities is complex. Very often, the reduction of a real product variance will increase the probability of the product being chosen. While smaller variances often lead to higher choice probabilities, a zero level of variance is not always optimal. A larger variance may extend a product's appeal to multiple segments and increase its overall first choice probability or "perceptual share." Optimizing a product's perceptual share depends both upon obtaining the optimal centroid and the optimal level of variation in perception.

The perceptual shares estimated by probabilistic unfolding should, of course, be distinguished from products' market shares. Even if the products being evaluated in an unfolding study are products on supermarket shelves, marketing and distribution variables will inevitably cause disparities between perceptual shares and market shares. Nevertheless, it is still meaningful to ask what attribute values will optimize perceptual share. Once the products and market segments have been established, it is possible to build "what-if" models from probabilistic unfolding model output to assess the perceptual shares of untested products.

Product variances are affected by many factors. When using unidentified products in the laboratory, variances are due to fluctuations in the subjects and the stimuli and to differences among the subjects. When using identified products in the marketplace, a host of marketing variables also becomes relevant. A common observation of advertising is that it will reduce the perceived variance in a product's perception. Sometimes, the introduction of a new product may draw attention to existing products and lower their variances as well. When this happens, regularity can be violated — the addition of a new element in the choice set may actually increase the choice probability of existing products.

Before leaving the topic of perceptual shares, it should be noted that the relationship of products' perceptual shares is almost never monotonic to their mean liking ratings. As will be seen later on, it is quite possible for a product with a modest liking rating to have a very high perceptual share, and vice versa. The lack of monotonicity is due to the forces exerted by different market segments and competitive products in the market place.

The last property to be discussed is the presence of an error theory. Traditional deterministic models have no error theory. To determine if a model should, for example, be estimated in one, two or more dimensions, heuristic rules of thumb must be employed. The variances of probabilistic models not only allow choice probabilities to be estimated, they also allow hypothesis tests to be made. Examples of hypothesis tests include the dimensionality of the space, equality of variances and equality of centroids. If a company is thinking of changing the ingredients of a product, hypothesis tests can be used to answer the question of whether consumers can recognize the difference.

3. Model estimation

Maximum likelihood (ML) methods are used to estimate the probabilistic unfolding models. ML estimation proceeds by maximizing the log of the likelihood function, which is the sum of the logs of the probability density functions (PDFs) over all observations. To do this, we need to know the nature of the observation or judgment and its PDF.

The simplest judgment for unfolding analysis is a liking rating. If we are in a two dimensional space, with variances σ^2 that are the same for all products on all dimensions, and we assume that the liking rating is modeled as a Euclidean distance d_{ij} between an ideal object i and a real object j with centroids on dimension k of μ_{ik} and μ_{jk}, respectively, then the calculation of the PDF for the liking rating proceeds as follows:

Let

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

where

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

Then,

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

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

$$d_{ij}^2/2\sigma^2 \sim \chi'^2_{\nu,\lambda_{ij}};\ \nu = 2, \lambda_{ij} = \sum_{k=1}^{2}(\mu_{ij}^2/2\sigma^2)$$

$$d_{ij} \sim (d_{ij}/\sigma^2)\chi'^2_{\nu,\lambda_{ij}}.$$

The PDF of d_{ij} is thus a function of the non-central chi-square distribution that has two parameters, a non-centrality parameter λ_{ij} and a degrees of freedom parameter ν that is equal to the dimensionality of the space. Closed form expressions for the non-central chi-square distribution do not exist and approximations must be used (Zinnes & MacKay, 1983). Measurement models, described below, may be used to relate d_{ij} to the subject's judgment δ_{ij}.

Generalizing the above results to higher dimensions is trivial; generalizing to other variance structures, different types of judgments, and alternative metrics is more complicated. As an example of different variance structures, consider the four situations presented in Fig. 4. Following Thurstone's (1927) nomenclature, we distinguish between Case 5 models where, as above, the variances are the same for all objects and Case 3 models where the variances may differ from object to object. We also distinguish between isotropic models where the variances for any one object are the same on all dimensions and anisotropic models where the variances may differ from dimension to dimension. PDFs for anisotropic liking ratings, which follow the quadratic forms in normal variables distributions, are described in MacKay (1989).

The selection of an appropriate variance structure is up to the analyst. In the testing of unidentified food products, for example, it is often appropriate to use a Case 5a model. On the other hand, if the products are identified products with which the consumers are familiar, then Case 3a models may be a better choice. Likelihood ratio tests may be used to determine if the greater fit available by using a Case 3a model outweighs the loss in degrees of freedom. The modeling of variances is very flexible. It is a simple matter to add terms that account for things such as the order bias of objects and the distance magnitudes of judgments.

Instead of using liking ratings, other types of preference judgments may also be used. A type of judgment that is more powerful than liking ratings is the preference ratio. For preference ratios, subjects evaluate pairs of real objects and, for each pair, indicate the preferred item in the pair and state how many times it is preferred over the less preferred object. Graphic rating scales are frequently used to obtain these judgments. To obtain the PDF of preference ratio judgments involves finding the distributions of ratios of distances (MacKay, 2001; MacKay & Zinnes, 1995; Zinnes & MacKay, 1987). The extra power in preference ratios comes from the fact that they are conjoint judgments — each judgment involves an evaluation of two real objects as well as one ideal object. Solutions obtained with preference ratio judgments tend to be more interpretable than solutions obtained from liking ratings, which are examples of disjoint judgments — each judgment involves the sampling of values from two objects that are in different sets.

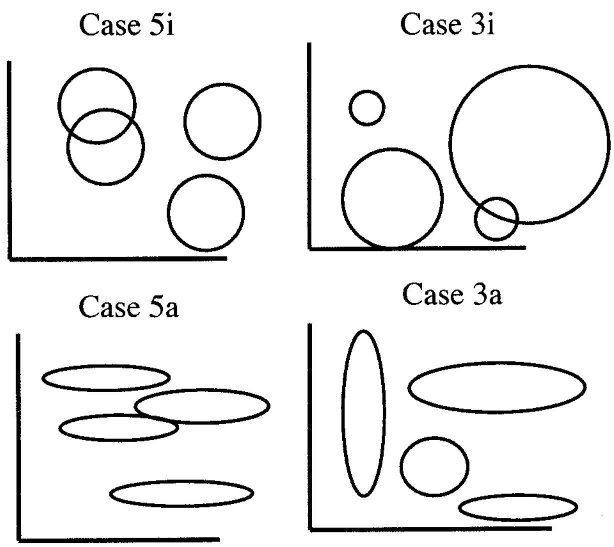

Fig. 4. Standard deviational ellipses for four types of variance structures.

Finally, another consideration is the type of metric one chooses to use. Euclidean metrics are used most often but density functions can also be formulated for the city-block metric (MacKay, 2001). It is commonly proposed that city-block metrics may be more appropriate for modeling the judgments of experts while Euclidean metrics are more appropriate for modeling the judgments of novices. Testing this proposition with nonmetric models is almost impossible, even using rule of thumb heuristics, since the presence of just a small amount of object variance quickly turns the selection of a metric into a situation that is little different than a coin toss. Probabilistic models, on the other hand, are very successful at testing metric properties.

Estimation of a probabilistic unfolding model is not restricted to one type of data. Thus, it is possible to combine both liking ratings and preference ratios in the estimation process. It is also possible to combine other types of data. A good candidate for inclusion is the dissimilarity judgment — a judgment where large values indicate that two real objects are very dissimilar and small values indicate that two real objects are very similar. Dissimilarities and similarities play a pivotal role in much theorizing in psychology (Goldstone, 1994). Dissimilarity judgments are conjoint judgments and can add a lot of explanatory power to an unfolding analysis. Dissimilarity judgments may be made directly by consumers or they may be derived from consumers' attribute evaluations. Dissimilarity judgments and liking ratings do not share the same scale. To simulta-

Table 1
Probability density functions for different types of judgments, metrics and variance structures[a]

	Liking ratings		Preference ratios	
	Euclidean	City-block	Euclidean	City-block
Case 5i	χ'^2	FN	F''	RFN
Case 5a	Q	FN	Q	RFN
Case 3i	χ'^2	FN	F''	RFN
Case 3a	Q	FN	Q	RFN

[a] χ'^2 Non-central chi-square; F'', doubly non-central F distribution; Q, quadratic forms in normal variables distribution; FN, folded normal distribution; RFN, ratio of folded normal variables distribution.

neously evaluate dissimilarity judgments and liking ratings, a measurement model must be defined so that the two types of judgments may be evaluated in the same space; linear and exponential transformations are most commonly used. Likelihood ratio tests or expert judgment may be used to select the measurement model.

Table 1 summarizes how the attributes described in this section are related to the choice of a PDF in the construction of a probabilistic unfolding model. The papers referenced in this section describe how these PDFs are defined and calculated.

Closed form solutions to the probabilistic unfolding model are not available. Iterative numerical estimation procedures are used instead. We have tried a large number of non-linear bound constrained optimization methods with the probabilistic unfolding model. Theoretical evaluations (Easley, 1996) suggest that gradient-based optimization methods have problems with the likelihood surface of the unfolding model. In general, we find that direct search methods, though slower, give the best results.

An alternating ML method is used to obtain the estimates. On the first iteration, variance parameters are updated, on the second iteration centroid parameters are updated, and on the third iteration measurement model parameters are updated. The alternating procedure is repeated until satisfactory convergence is reached.

4. Application

The organizers of The Fifth Sensometrics Meeting provided information pertaining to the evaluation and liking of 28 fruit-flavored beverages. Data were provided for two sets of consumers, one using nine and the other five point scales, as well as data from a trained sensory panel. Each of the two sets of consumers was, in turn, distinguished on the basis of which of two geographies they represented.

Differences exist in how analysts choose to treat sensory panel and consumer panel data. The view of this author is that unfolding models are models of psychological processes and are better estimated only with consumer data. Once the unfolding model is estimated, it is then perfectly appropriate to use sensory panel data or instrument data to determine how the psychological evaluation of the products is related to the objects themselves. This evaluation is not undertaken in this paper.

Quality problems existed with both sets of consumer data. The nine point scale data were thought to have fewer problems and are used here. A little over five percent of the observations were discarded due to quality considerations. A higher discard rate could be justified.

An early decision that must be made in an unfolding analysis is whether ideal objects should be defined for each consumer, for one composite consumer, or for segments of consumers. An advantage of the probabilistic unfolding model used here is that it can be used when only one ideal object is to be estimated (MacKay, Easley, & Zinnes, 1995). Traditional models require multiple ideal objects. Since we wanted to compare the results to those obtained with traditional methods, a multiple ideal solution was sought.

Definition of the ideal objects depends upon the objectives of the analyst. If the objective were to compare geographies one and two, then a two ideal model would be appropriate. If the objective were to model the entire market, then a decision must be made on the number of ideal object market segments and the composition of the consumers for each segment.

To determine the number and composition of the ideal segments, the hedonic ratings were standardized for each subject across the evaluated beverages and submitted to sequential K-Means cluster analyses. The cluster analyses determined which consumers were members of each ideal cluster. Unfolding analyses with different numbers of ideal objects were conducted and, as described below, a four ideal object solution was chosen as being sufficient to capture the complexity in the data set. Ideal objects one through four contained approximately 36, 20, 27 and 17% of the sample, respectively. (A variety of statistical and data mining procedures for determining the appropriate number and composition of clusters exists. See Arabie, Hubert, & De Soete, 1996, for an overview of much of the recent literature.)

Having defined the composition of the ideal object market segments, the data were first evaluated with KYST, the traditional deterministic unfolding model referenced earlier. The two dimensional KYST solution is shown in Fig. 5. (Rule of thumb guidelines were ambiguous as to whether a two- or three-dimensional solution should be used. The two dimensional solution is presented for ease of interpretation. The two dimensions in Fig. 5 were also the dominant dimensions in the three dimensional solution.) Letters *a*–*z* and the symbols & and * represent the real objects (beverages). The numbers 1–4 represent the centroids of the four ideal objects. (For those who may choose to evaluate these data for themselves, the beverages were ordered by their sample number, not their sequence number.)

The solution of Fig. 5 is typical of the problematic solutions frequently encountered using traditional unfolding analysis. The ideal objects are on the left side of the plot and the beverage objects are on the right side of the plot. The values of the beverages on the dimensions tell us nothing about their attributes. Beverage values on dimension V1 are essentially measures of liking — *x* is the most liked beverage and *u* is the least liked. The correlation of the mean overall liking measure provided in the data set and the values on V1 is −0.99.

Two assumptions made in the probabilistic scaling of the data were that the real objects would best be modeled by a Case 5 anisotropic variance structure and that the ideal objects would best be modeled by a Case 3 anisotropic variance structure. Similar real objects, especially ones that are not identified, often have similar variance structures. On the other hand, consumer segments often value attributes differently and this is reflected in the dimensional differences in their variances. A Euclidean metric was also assumed.

The initial probabilistic unfolding solution obtained from PROSCAL, is given in Fig. 6. Since all of the real objects had the same standard deviations, only one ellipse, for product *o*, is shown. While the probabilistic solution does not display the degenerate nature of the solution shown in Fig. 5, it is not as interpretable as one might like. Interpretation difficulties are caused in part by the significant correlation ($P<0.05$) of the centroids on dimensions V1 and V2.

To enhance the interpretability, a simultaneous scaling was undertaken of the liking ratings and dissimilarity values derived from the intensity scales of 10 attributes — color, cloudiness, aroma strength, grape flavor strength, sweetness, tartness/sourness, natural taste, thickness, smoothness and aftertaste strength — provided in the dataset. (The data may, of course, be disguised in undisclosed ways.) The derived dissimilarity judgments were standardized by subject. The simultaneous scaling caused a slight decline in the fit of the expected liking ratings — the correlation of the expected utilities and the mean liking ratings declined from 0.94 to 0.85. However, the dimensions were much more interpretable. The solution is given in Fig. 7.

Several things about Fig. 7 are worth noting. The magnitudes of the variances are quite substantial and ignoring them, by using a deterministic model, results in the degenerate solution of Fig. 5. In Fig. 7, the ideal objects are integrated in the space of the real objects. Variances for the four ideal segments are smaller than the variances for the real objects. This finding, which is quite common, contradicts assumptions made in the wandering ideal point model (De Soete, Carroll, & DeSarbo, 1986), a probabilistic unfolding model that assumes that all the variance is on the ideal objects and that the real objects are deterministic.

Identifying the dimensions, the V1 dimension has high correlations with four intensity scales — cloudiness (0.93), aroma (0.92), grape flavor strength (0.92) and thickness (0.97). These correlations are much higher than the correlations the intensity scales have with each other and thus seem to capture a basic psychological dimension of the real products that we shall call strength. The second dimension, V2, seems to be

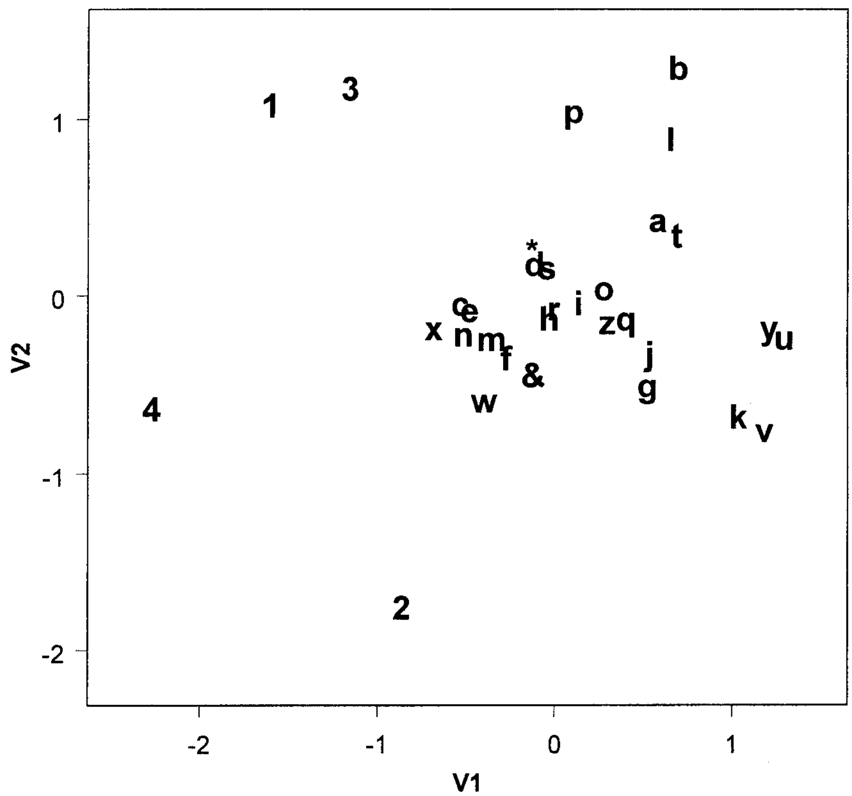

Fig. 5. Deterministic unfolding solution for 28 beverages (*a*, ... *z*, &, *) and four consumer segments (1–4).

inversely related to sweetness with beverages at the low half of the scale having significantly higher ($P<0.001$) sweetness intensity ratings than beverages in the upper half.

The four segments are seen to differ more in their preferences for sweet products than for strong products. A phenomenon observed here, which is often true in similar applications, is that consumers who prefer

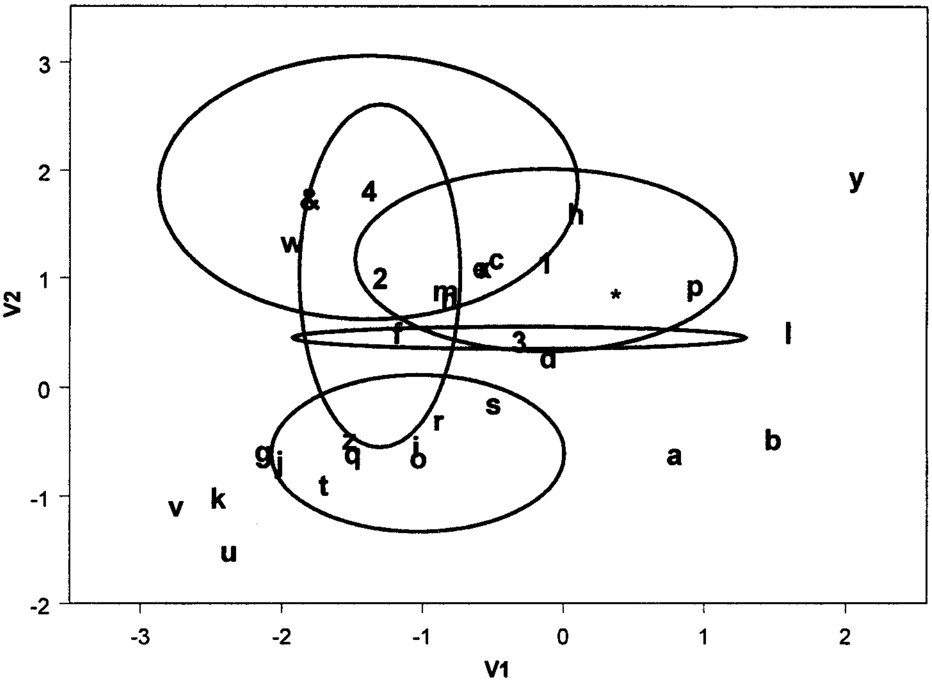

Fig. 6. Probabilistic unfolding solution for 28 beverages (*a*, ..., *z*, &, *) and four consumer segments (1,...,4).

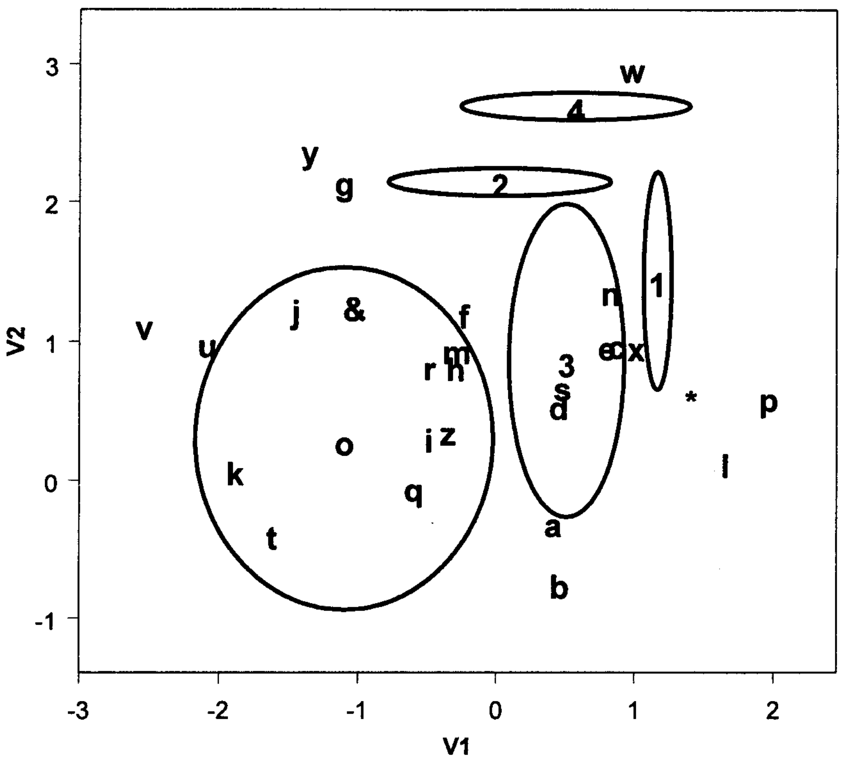

Fig. 7. Probabilistic dissimilarity scaling and unfolding solution for 28 beverages (*a*, ..., *z*, &, *) and four consumer segments (1,...,4).

products with relatively extreme dimensional values are also more sensitive to that dimension, as is indicated by the smaller standard deviation values. Thus, consumers in segments 2 and 4 have marked preferences for less sweet beverages and also have smaller variances on that dimension. Subjects in set 1 have relatively strong preferences for strong products and have a low standard deviation on that dimension. Subjects in segment 3 prefer products with more moderate values on both dimensions and have the largest standard deviations of any segment on all dimensions.

Segments 2 and 4 have centroids that are close to each other and they also have similar variance structures. This raises the question of whether these two segments should be treated as one. A three-segment model is a constrained version of the four-segment model and will have a lower log likelihood. It will also have four fewer parameters. To test the null hypothesis of no difference in the two models, we use a likelihood ratio test which makes use of the fact that two times the difference in the log likelihoods of the two models is asymptotically chi-square distributed with the degrees of freedom equal to the difference in the number of free parameters estimated by the two models. The difference in the log likelihoods was 43.0 ($P<0.001$) and the four-segment model was retained. [A likelihood ratio test was also used to see if it would be better to go to a more complex five-segment model. The five-segment model was not significantly better ($P>0.95$).]

Looking at the original liking rating data, the product with the highest mean liking rating was product *x*. From Fig. 7, the reason for the appeal of *x* becomes clear, it is well positioned with respect to the two largest segments, segments 1 and 3. Products *u* and *v*, which have the lowest overall liking of any of the 28 real products, are the most distant from all segments.

Probabilistic scaling enables us to do more than just better understand hedonic liking ratings. It allows us to make predictions about choice behavior, which may be quite different. Consider, for example, product *w*. Based on the original hedonic measures, *w* was the eleventh most liked product. If you use the PROSCAL estimates to compute the mean expected distance of *w* from each of the four ideal distributions, you also find that *w* has the eleventh lowest expected distance (disutility) value. So, in terms of liking, the model agrees completely with the input data. The PROSCAL estimates can, though, be used to estimate the probability of each real product being a first choice for subjects in each segment. If you do this and then calculate the probability of being the overall first choice by weighting the results of each segment by its sample size, then product *w* ends up having the second highest first choice probability of all the products. (Product *n*, with a strong appeal to Segment 1, has the highest first choice probability.)

Why do we get what might seem to be such counter-intuitive results? Looking at Fig. 7 the answer is clear. Product *w* dominates segment 4. Not only is its centroid close to the centroid of segment 4, there are no other products between it and the centroid of segment 4. Product *w* is also attractive to members of segment 2. Other products, such as *c, d, e* and *s* are as close or closer to the centroids of larger segments than segments 2 and 4 but they have more competition and must split the first choice vote with competitors. This discussion of course assumes that consumers are able to make choices among these 28 real products.

Product *w* is also a good vehicle for understanding the practical importance of the distinction between independent and dependent sampling. If dependent sampling is assumed, then product *w* will get 45% more of segment 1's first choices than it will under independent sampling. This is because under dependent sampling, when a high value of dimension V2 is sampled for segment 1, that value is used for making comparisons with all 28 beverages. Product *w* is likely to be close to segment 1 when this occurs and other beverages are likely to be farther away. Under independent sampling, the values sampled for segment 1 vary for each of the 28 beverages. The advantage product w enjoys when a high value is sampled on V2 for segment 1 for its evaluation may be diminished when other beverages are evaluated because the sampled value for segment 1 will change.

5. Discussion

The preceding paragraph illustrates an aspect of probabilistic scaling that may be troublesome, the sensitivity of the results to assumptions made by the analyst. Our experience indicates that the complexity of the PROSCAL model does result in a steeper learning curve than is present with simpler deterministic models. However, for an experienced analyst, the decisions that are most appropriate for a given class of data soon become apparent. The ability to perform hypothesis tests and compare, for example, the likelihoods of solutions involving dependent and independent sampling, can be very helpful in deciding what assumptions should be made.

Features common to many unfolding programs that are not part of the method described here include the presence of dimensional weighting parameters (Carroll, 1972), latent class modeling (Poulsen, Brockhoff, & Erichsen, 1997), and reparameterization (DeSarbo & Cho, 1989). The advantages of dimensional weights are largely captured by the use of a Case 3 variance structure. Subjects are more sensitive to dimensions on which they have low variances. Latent class modeling is something that needs to be pursued further. If unique, stable, computationally efficient latent class models can be developed for the probabilistic unfolding model, then the need to do a separate cluster analysis to define segments would be eliminated. For the moment, though, the multistage approach used here seems to be more practical. Reparameterizing the model so that the objects' coordinates are directly expressed in terms of physical or sensory panel attribute data is mathematically simple but psychologically destructive. It seems more prudent to explore the relationship of psychological and physical attributes in separate analyses.

While the probabilistic unfolding model is more complicated than deterministic unfolding models, it has the advantage of accounting for the inherent variability of the judgment process — a variability which, when not accounted for, may lead to severely biased and often meaningless solutions. Important side benefits from using the probabilistic unfolding model include the ability to test hypotheses and the ability to calculate first choice probabilities under independent and dependent sampling conditions. These properties, combined with the unfolding model's ability to account for satiety — a fundamental characteristic of food preferences, make the probabilistic unfolding model a very promising tool for sensory analysis.

References

Arabie, P., Hubert, L. J., & De Soete, G. (1966). *Clustering and classification.* River Edge, NJ: World Scientific Publishing.

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

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

DeSarbo, W. S., & Cho, J. (1989). A stochastic multidimensional scaling vector threshold model for the spatial representation of "pick any/N" data. *Psychometrika, 54*(1), 105–129.

De Soete, G., Carroll, J. D., & DeSarbo, W. (1986). The wandering ideal point model: a probabilistic multidimensional unfolding model for paired comparisons data. *Journal of Mathematical Psychology, 30*(1), 28–41.

Easley, R. F. (1996). *Complex group decision-making: a psychometric/artificial intelligence approach.* Doctoral dissertation. Indiana University, Bloomington IN.

Goldstone, R. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments & Computers, 26*(4), 381–386.

Kruskal, J. B., Young F. W., & Seery, J. B. (1977). *How to use Kyst-2a, a very flexible program to do multidimensional scaling and unfolding.* Unpublished manuscript, Bell laboratories.

MacKay, D. B. (1989). Probabilistic multidimensional scaling: an anisotropic model for distance judgments. *Journal of Mathematical Psychology, 33*(2), 187–205.

MacKay, D. B., & Zinnes, J. L. (1995). Probabilistic multidimensional unfolding: an anisotropic model for preference ratio judgments. *Journal of Mathematical Psychology, 39*(1), 99–111.

MacKay, D. B. (2001). Probabilistic multidimensional scaling using a city-block metric. *Journal of Mathematical Psychology, 45*(2), 249–264.

MacKay, D. B., Easley, R. F., & Zinnes, J. L. (1995). A single ideal point model for market structure analysis. *Journal of Marketing Research, 32*(4), 433–443.

MacKay, D. B., & Zinnes, J. L. (2000). *PROSCAL: a program for probabilistic scaling.* Unpublished manuscript, Indiana University.

Poulsen, C. S., Brockhoff, P. M. B., & Erichsen, L. (1997). Heterogeneity in consumer preference data — a combined approach. *Food Quality and Preference, 8*(5-6), 409–417.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*(4), 273–286.

Zinnes, J. L., & MacKay, D. B. (1983). Probabilistic multidimensional scaling: complete and incomplete data. *Psychometrika, 48*(1), 27–48.

Zinnes, J. L., & MacKay, D. B. (1987). Probabilistic multidimensional analysis of preference ratio judgments. *Communication and Cognition, 20*(1), 17–44.

David MacKay is a Professor of Marketing and Geography at the Kelley School of Business of Indiana University. He received his BS in Agricultural Economics from Cornell University and his PhD in Marketing from Northwestern University. His research interests are in probabilistic scaling, product design and positioning, and group decision making. Executable programs, documentation and data for the PROSCAL package referenced in this paper may be downloaded from http://proscal.com.