New Developments in Psychological Choice Modeling
G. De Soete, H. Feger and K. C. Klauer (eds.)

PROBABILISTIC MULTIDIMENSIONAL ANALYSIS OF PREFERENCE RATIO JUDGMENTS

Joseph L. Zinnes
National Analysts, Philadelphia, PA, U.S.A.

David B. MacKay
Indiana University, U.S.A.

A probabilistic multidimensional model is described for analyzing preference ratio judgments. This model combines the unfolding model of Coombs with the probabilistic model of Hefner, in which stimuli and individuals are represented by multivariate normal distributions. A simple procedure is described for approximating the maximum likelihood estimates of the location and variance parameters of the model. Two simulations show how well this procedure works, especially when there is considerable variability in the data.

1. Introduction

For some time now we have been confronted by a seemingly insolvable problem: how can be study, in a serious and convincing way, individual choice of interesting, multi-attribute stimuli when those stimuli are clearly identifiable. The problem with identifiable stimuli is that individual choices of those stimuli cannot be replicated a large number of times. It is easy enough, at least if one has sufficient reinforcers available, to ask subjects to indicate over and over again whether they can detect a signal buried in background noise, or whether they can identify which of two tones was presented, etc. But it is quite a different matter to ask subjects

This reseach was supported by National Science Grant SES-8120871.
This paper is a revised version of an article published in *Communication & Cognition*, 1987, *20*, 17-43.

over and over again which of two specific cars they prefer or which of two specific houses they would buy. Subjects can readily identify these stimuli and therefore can readily recall their previous responses. This is precisely the same problem that occurs in the testing field.

Large numbers of replicated choices are important to the study of choice behavior because of the nature of the class of choice models that we believe are relevant. These models are inherently probabilistic and have large numbers of parameters to estimate. To estimate these parameters accurately and also to carry out sensitive statistical tests, namely those which discriminate between alternative choice models, generally require a considerable amount of replicated choices.

At present we see only one way out of this predicament and that is to study individual choice behavior by collecting numerical judgments of preferences, rather than by obtaining choice data. Unlike choice data, these numerical judgments are obtained by having subjects indicate both which stimulus they prefer *and* by how much.

The value of numerical judgment has been pointed out by numerous writers (Anderson, 1982; Eisler, 1982). They contain, under appropriate conditions, more information than simple choice responses, and therefore they make it possible, at least in principle, to obtain accurate parameter estimates using few if any replications of the individual judgments. This, at least, is our hope for the present.

That hope does, of course, depend on a leap of faith: that numerical judgments and choice responses will be compatible, that the same underlying model will apply and therefore that the estimates of the parameters obtained by numerical judgments are precisely the same as those that would have been obtained had it been possible to replicate choice responses a large number of times. This is rather a large assumption. It is one we expect to investigate more fully in the future.

Thus, in this paper, we pursue only the question of how to estimate, using numerical judgments, the parameters of a specific choice model, when it is reasonable to assume that that choice model is appropriate.

The specific numerical judgment discussed in this paper is a ratio judgment, or what we call a preference ratio judgment. We assume that stimuli are presented pairwise to the subjects and that the subjects are asked to indicate how much they prefer one stimulus over another. The instructions that we have used in our own experimental work attempt to make it clear to the subject that what is wanted is a ratio judgment. The response of two, for example indicates that one stimulus is preferred twice as much as the other. To make sure that these instructions are understood, the subjects are given warm up trials involving very simple stimuli, such are lines of different lengths. The subjects then practice making ratio judgments concerning the relative lengths of pairs of lines.

Although we confine ourselves in this paper to preference ratio judgments, it should not be concluded that this is the only type of judgment that could have been used to extract numerical information from subjects. In fact, the ubiquitous rating judgment has been used in the preference domain to do just this (Bechtel, 1976; Saaty, 1980; Scheffé, 1952; Sjöberg, 1967). Our use of the ratio judgment stems from our belief that it is the most appropriate judgment to use if our underlying preference model is indeed correct. In this model, the utility or desirability that a person has for a stimulus is represented by a Euclidean distance. Thus, choosing between two stimuli is conceptually equivalent to comparing two distances in a Euclidean space. Since distances in the model are determined only up to a multiplicative transformation, it would not make sense to compare the differences between two distances, as would be suggested by a rating judgment. This is the case, because differences are not invariant over multiplicative transformations. Their magnitude is thus totally arbitrary within the model. It would make sense, however, to determine the ratio between a pair of distances, because that value is invariant over a multiplicative transformation. Within the model, therefore, the preference ratio judgment is a meaningful judgment.

There is another issue concerning our use of the ratio judgment. Even though we assume that subjects are carefully instructed to perform a ratio judgment, it does not follow that they will actually carry out those instructions. It has been shown (Birnbaum, 1982) that under some conditions subjects apparently respond to stimulus differences even when they are instructed to respond to their ratios. Birnbaum has, however, provided some evidence to indicate that when stimuli can be represented as distances, subjects appear to respond to the stimulus ratios when instructed to do so. It is, therefore, not unreasonable for us to assume that the preference ratio instruction does indeed generate a preference ratio response, that it does generate a response based on the ratio of two distances. However, the final determination of the precise conditions under which this assumption (and the others to be described in the following section) are valid, will have to wait for more detailed experimental tests.

The preference model we use in this paper is a probabilistic, multidimensional version of Coombs' unfolding model (Coombs, 1964). The probabilistic aspects are based on a model first put forth by Hefner (1958). The essential idea of Hefner's model is to represent each stimulus in terms of a multivariate normal distribution and subjects' decision processes in terms of random samples from these distributions.

The Hefner model, or closely related models, have been used in connection with a number of different types of data. It has been used to explain same-different judgments (Zinnes & Kurtz, 1968; Zinnes & Wolf, 1977), choice response (Böckenholt & Gaul, 1984; Croon, in press; De Soete, Carroll, & DeSarbo, 1986; Suppes & Zinnes, 1963; Zinnes & Griggs, 1974), similarity judgments (MacKay & Zinnes, 1981; Zinnes & MacKay, 1983) and recognition responses (Ashby & Townsend, 1986).

The attraction of Hefner's model is its conceptual simplicity. It is a very natural and powerful extension of the single dimensional choice models of Thurstone (1927). It is powerful, because the properties of the multivariate normal are well known and therefore one can answer in detail basic questions concerning the goodness-of-fit of the model and the invariance of its parameters over different experimental conditions.

Our experimental work with preference ratio judgments has just begun (MacKay, Ellis, & Zinnes, 1986; MacKay & Zinnes, 1986). In these experiments, subjects made preference ratio judgments concerning residences that differed with respect to environment, location and economic characteristics. The Coombs-Hefner preference model discussed in this paper was applied to the data of these experiments and appeared to do well in explaining those data.

In the following sections we focus on the problem of obtaining the maximum likelihood estimates of the parameters of the Coombs-Hefner preference model, when the data consist of preference ratio judgments. In Section 2, a simple approximation of the likelihood functions is developed. In Section 3, a simple expression for the initial estimates of the parameters is worked out. In the final two sections, two simulations are described, the purpose of which is to provide some idea of the accuracy and feasibility of the maximum likelihood estimates, especially when there is considerable variability in the data.

2. The Preference Model

In the unfolding model of Coombs (1964), subjects and stimuli are both represented as points in an r-dimensional Euclidean space. The preference of the subjects are assumed to be determined by the distances between the subject points, called "ideal points", and the stimulus points. The smaller the distance d_{ij} between ideal point i and stimulus point j, the more desirable is stimulus S_j to subject P_i.

To this deterministic model of Coombs we add the probabilistic assumptions of Hefner (1958). In particular, we let the r-dimensional random vectors $\mathbf{X}_i = (X_{i1}, \ldots, X_{ir})$, $i = 1,..., m$, be associated with the m ideal points and assume that they have an r-variate normal distribution with mean vector $\mathbf{u}_i = (u_{i1}, \ldots, u_{ir})$ and covariance matrix $\sigma_i^2\mathbf{I}_r$. Similarly, for the stimulus points, the r-dimensional random vectors $\mathbf{X}_j = (X_{j1}, \ldots, X_{jr})$, $j = m+1, \ldots, m+n$ are associated with the n stimulus points and it is assumed that they also have an r-variate normal distribution with mean vector $\mathbf{u}_j = (u_{j1}, \ldots, u_{jr})$ and covariance matrix $\sigma_j^2\mathbf{I}_r$.

The notation is intended to indicate that the variances of the components of each stimulus point do not differ from dimension to dimension, but that on any given dimension, the variances of the components of different stimuli may differ. Thus within a single dimension, these assumptions are precisely those of a Thurstone (1927) case 3 pair comparison model. The same is true for the variances of the ideal points. It would be desirable to formulate more general assumptions concerning the covariance matrices, but doing this might tend to increase the number of parameters that would have to be estimated.

Under these assumptions, the interpoint distance d_{ij} is a random variable. On each trial, its value is determined by subject i sampling from the ith ideal and jth stimulus distribution and "calculating" the Euclidean distance between the two sample points. Thus, in terms of the r-dimensional random vectors $\mathbf{X}_i$ and $\mathbf{X}_j$, the distance d_{ij} is given by

$$d_{ij}^2 = (\mathbf{X}_i - \mathbf{X}_j)'(\mathbf{X}_i - \mathbf{X}_j). \tag{1}$$

In contrast, the *true* distance D_{ij} is not a random variable, but is defined in terms of the mean vectors $\mathbf{u}_i$ and $\mathbf{u}_j$ by

$$D_{ij}^2 = (\mathbf{u}_i - \mathbf{u}_j)'(\mathbf{u}_i - \mathbf{u}_j). \tag{2}$$

It may also be noted that the true distance D_{ij} does not correspond to the expected value of the distance d_{ij} and, in fact need not be monotonically related to it (Zinnes & MacKay, 1983).

It will be useful to define the *joint* variance σ_{ij}^2 by the equation

$$\sigma_{ij}^2 = \sigma_i^2 + \sigma_j^2 \tag{3}$$

which can be conceptualized as the variance of the difference between the components X_{ik} and X_{jk} on each of the r dimensions. This term, which appears in many of the equations that follow, should not be confused with the variance of the distance d_{ij}. That variance, unfortunately, has a considerable more complex expression.

To deal with preference ratio judgments, the experimental task of interest here, we use a direct adaptation of the decision rule of the Coombs unfolding model. It is assumed that subject i reports the preference ratio R_{ijk} when the ratio of the distances d_{ij} and d_{ik} equals R_{ijk}, that is, when

$$R_{ijk} = \frac{d_{ij}}{d_{ik}}. \tag{4}$$

Since the interpoint distances d_{ij}, $i = 1, \ldots, m$, and $j = m+1, \ldots, m+n$, are random variables, their values and that of the ratio R_{ijk}, can be expected to change with replications. The decision rule given in (4) only asserts that the subject accurately reports the ratio as it is perceived on each trial.

It will be helpful to make one more assumption. This assumption concerns the independence of the distances d_{ij} and d_{ik} when the subject judges the stimulus pair S_j and S_k. We shall assume that the subject randomly selects two *independent* samples from his ideal point distribution, one of which is used to determine the distance d_{ij} and the other the distance d_{ik}. Under these conditions, the two distances d_{ij} and d_{ik} will be independent random variables.

Whether this assumption is plausible or reasonable would depend on the specific details of the experimental procedure. If the two stimuli to be judged are presented sequentially or in widely separated spatial positions, the subject would have a tendency to evaluate each of the stimuli independently. This might also happen when the stimuli are complex, requiring the subject to spend a significant amount of time considering each stimulus separately. In any event, we assume in what follows that the two-sample, independence case applies and therefore that the ratio judgment R_{ijk} is based on the ratio of two independent random variables. Whether our results can be generalized to the one-sample, dependent case remains to be seen.

3. The Likelihood Function

We consider first the problem of evaluating the probability density function of the ratio judgment R_{ijk}. This density function is needed because it forms the basis of the likelihood function that is to be maximized.

Under the assumptions stated thus far, it follows that the "standardized" squared distance d_{ij}^2/σ_{ij}^2 has the noncentral chi-square distribution $\chi'^2(\nu,\lambda_{ij})$, where the degrees of freedom equals the dimensionality of the space r and the noncentrality parameter λ_{ij} equals

$$\lambda_{ij} = D_{ij}^2/\sigma_{ij}^2 \tag{5}$$

(Hefner, 1958; Zinnes & MacKay, 1983). Because of this and the independence assumption stated previously, we can immediately conclude that the ratio of the standardized squared distances

$$\frac{d_{ij}^2 / \sigma_{ij}^2}{d_{ik}^2 / \sigma_{ik}^2}$$

has the doubly noncentral F distribution $F''(\nu_j,\nu_k,\lambda_{ij},\lambda_{ik})$ (Bulgren, 1971; Suppes & Zinnes, 1963; Zinnes & Griggs, 1974). The two noncentrality parameters of this distribution λ_{ij} and λ_{ik} are defined in (5), as they are for the noncentral chi-square distribution, while the degrees of freedom ν_j and ν_k are both equal to the dimensionality of the space r.

These results indicate that there is a close relationship between the probability density function of the ratio judgment R_{ijk} and the probability density function of the doubly noncentral F distribution. Specifically, letting $g(R_{ijk})$ be the desired density function of R_{ijk}, then

$$g(R_{ijk}) = 2R_{ijk}\left[\frac{\sigma_{ik}^2}{\sigma_{ij}^2}\right] h''\left[\frac{d_{ij}^2/\sigma_{ij}^2}{d_{ik}^2/\sigma_{ik}^2}\right], \tag{6}$$

where $h''(\cdot)$ is the density function of the doubly noncentral F distribution $F''(\nu_j,\nu_k,\lambda_{ij},\lambda_{ik})$.

Equation (6) shows that if will be sufficient to focus our attention on developing a procedure for evaluating the function $h''(\cdot)$ of the doubly noncentral F distribution, in order to obtain a simple procedure for evaluating the density function $g(R_{ijk})$. We consider next, therefore, the F'' distribution.

The exact expression of the density function of F'' distribution has been worked out (Bulgren, 1971; Kendall & Stuart, 1961, p. 252), but it is not expressible in closed form. It contains a doubly infinite series of terms, which for some values of the parameters — namely those in the tails of the distribution — converge extremely slowly. For practical applications, it is essential to find a simple, approximate expression of this density function.

Two simple possible approximations immediately suggest themselves. One approach uses the central chi-square to approximate the noncentral chi-square distribution (Patnaik, 1949), this approach makes it possible to convert the F'' distribution to the central F distribution. The other approach uses a normal distribution to approximate the noncentral chi-square distribution. Although this latter approach has been used successfully to approximate the cumulative distribution function of the F'' distribution (Zinnes & Griggs, 1974), it did not seem to work as well to approximate the density function of this distribution. Consequently, our discussion here is confined to the former approach, based on using the central chi-square to approximate the noncentral chi-square distribution.

From the Patnaik approximation, it follows that if s_j has the noncentral chi-square distribution $\chi'^2(\nu_j,\lambda_j)$ then s_j/ρ_j will have approximately the central chi-square distribution $\chi^2(\nu_j^*)$ where the degrees of freedom ν_j^* equal

$$\nu_j^* = \frac{(\nu_j + \lambda_j)^2}{\nu_j + 2\lambda_j} \tag{7}$$

and the multiplicative factor ρ_j is given by

$$\rho_j = \frac{\nu_j + 2\lambda_j}{\nu_j + \lambda_j}. \tag{8}$$

Thus, to obtain the central F approximation of the F'' distribution, we start with the distribution function of the F'' distribution

$$H''(f \mid \nu_1,\nu_2,\lambda_1,\lambda_2) = p\left[\frac{s_1/\nu_1}{s_2/\nu_2} < f\right]. \tag{9}$$

Multiplying both sides of the inequality in (9) by $\rho_2\nu_2^*/\rho_1\nu_1^*$, and making use of (7), (8) and the definition

$$a_j = \frac{\nu_j + \lambda_j}{\nu_j} \tag{10}$$

reduces (9) to

$$H''(f \mid \nu_1,\nu_2,\lambda_1,\lambda_2) = p\left[\frac{s_1/\rho_1\nu_1^*}{s_2/\rho_2\nu_2^*} < \frac{fa_2}{a_1}\right]. \tag{11}$$

Now we can make direct use of the Patnaik approximation. According to this approximation, the left-hand side of the inequality in (11) has approximately a central F distribution. Therefore, (11) can be written as follows

$$H''(f \mid \nu_1,\nu_2,\lambda_1,\lambda_2) = H\left[\frac{fa_2}{a_1} \mid \nu_1^*,\nu_2{*}\right] \tag{12}$$

where $H(\cdot)$ is the distribution function of the central F distribution $F(\nu_1^*,\nu_2^*)$. The degrees of freedom ν_1^* and ν_2^*, which in general will not be equal to each other and will have noninteger values, are given by (7). The final result is obtained by differentiating (12) with respect to f, which gives the approximation

$$h''(f \mid \nu_1,\nu_2,\lambda_1,\lambda_2) = \frac{a_2}{a_1}h\left[\frac{fa_2}{a_1} \mid \nu_1^*,\nu_2^*\right]. \tag{13}$$

This equation expresses h'', the density function of the F'' distribution, in terms h, the density function of the central F distribution. It may be noted that the function h, which is the key element of (13), has a simple closed form expression and therefore (13) does indeed provide a straightforward procedure for evaluating h'' for any of the values of its four arguments.

To summarize, (13) which gives the approximation that is fundamental to evaluating the likelihood function to be maximized, replaces the density

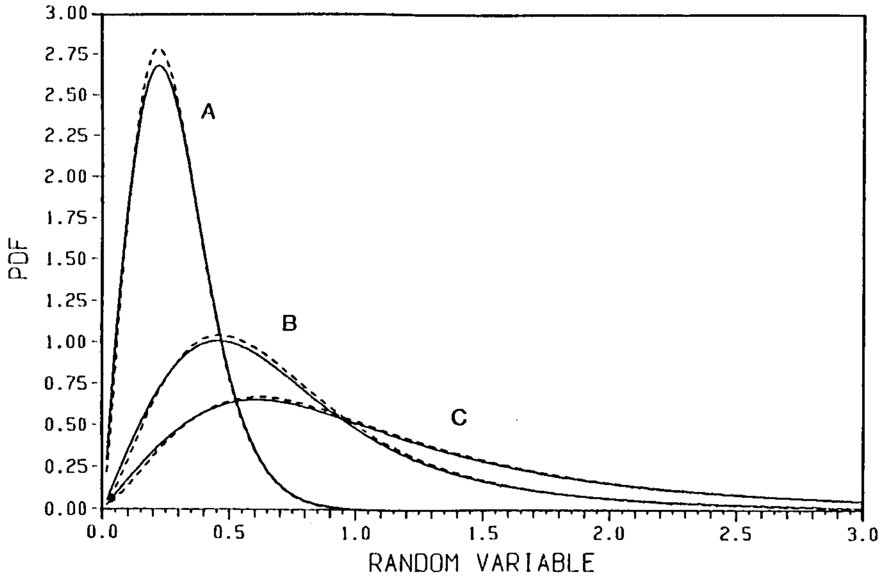

Figure 1. The central F approximation of the doubly noncentral F distribution. The degrees of freedom ν_1 and ν_2 are both equal to 2. The solid lines are the exact values, the dashed lines are the central F approximation. For curve A: $\lambda_1 = 1$, $\lambda_2 = 30$, for curve B: $\lambda_1 = 1$, $\lambda_2 = 4$; for curve C: $\lambda_1 = 1$, $\lambda_2 = 1$.

function of the F'' distribution, having equal degrees of freedom and integer values, with the density function of the central F distribution, having unequal degrees of freedom and noninteger values.

Some idea of the accuracy of the approximation given in (13) is shown in Table 1 and Figure 1. Table 1 gives both the approximate and the exact values of the function $h(f \mid \nu_1, \nu_2, \lambda_1, \lambda_2)$ where $f = 1$, $\nu_1 = \nu_2 = 2, 4, 8$ for a number of different values of λ_1 and λ_2. The absolute and relative errors, given in columns 5 and 6 of the table, suggest that the approximation is quite good. The absolute errors do not exceed .02 and the relative errors do not exceed 6 percent. Furthermore, the larger relative errors seem to occur only for values in the tails of the distribution, where according to the last column of the table, convergence of the infinite series in the exact expression tends to be slowest.

Figure 1 offers additional support for the approximation given in (13). Unlike Table 1, this figure attempts to show how the accuracy of the approximation affects the evaluations of the function $g(R_{ijk})$ for the entire range of values of the variable $R_{ijk} = d_{ij}/d_{ik}$. Three different distributions are plotted in this figure for the three different values of λ_{ij}: 1, 4 and 30. To highlight any major weakness of the approximation, the degrees of freedom ν_1 and ν_2 were both set equal to 2 for all three cases. Distributions with larger degrees of freedom tend to be more symmetric and therefore tend to be easier to approximate.

Table 1. Exact and approximate values of the probability density function $h(l \mid \nu_1, \nu_2, \lambda_1, \lambda_2)$

Degree of freedom		$h(l)$		Error[a]		
λ_1	λ_2	Exact	Approx.	Absolute	Percent	K^b
			$\nu_1 = \nu_2 = 2$			
.8	1.0	.5184	.5311	-.0128	-2.46	44
.1	1.0	.4920	.5009	-.0089	-1.80	28
1.0	2.0	.5294	.5497	-.0203	-3.94	58
1.0	3.0	.5158	.5384	-.0226	-4.39	76
1.0	4.0	.4890	.5109	-.0218	-4.47	86
1.0	8.0	.3352	.3427	-.0074	-2.22	130
1.0	12.0	.1982	.1952	-.0030	1.52	173
			$\nu_1 = \nu_2 = 4$			
.8	1.0	.7596	.7640	-.0044	-.57	52
.1	1.0	.7403	.7436	-.0033	-.44	33
1.0	2.0	.7609	.7697	-.0087	-1.15	70
1.0	3.0	.7393	.7515	-.0122	-1.65	86
1.0	4.0	.7037	.7180	-.0143	-2.04	97
1.0	8.0	.5032	.5140	-.0108	-2.14	149
1.0	12.0	.3138	.3149	-.0011	-.34	192
1.0	20.0	.0978	.0917	.0061	6.23	288
			$\nu_1 = \nu_2 = 8$			
.8	1.0	1.0980	1.0992	-.0012	-.11	64
.1	1.0	1.0843	1.0852	-.0009	-.09	42
1.0	2.0	1.0937	1.0967	-.0030	-.27	89
1.0	3.0	1.0684	1.0736	-.0051	-.48	107
1.0	4.0	1.0283	1.0356	-.0072	-.71	121
1.0	9.0	.7246	.7344	-.0098	-1.35	191
1.0	15.0	.3815	.3819	-.0004	-.11	273
1.0	25.0	.0973	.0913	.0059	6.09	415

[a]The error equals the exact value minus the approximate value.
[b]Indicates the number of terms summed to obtain the exact values of the density function given in the table.

From the discrepancy between the exact and approximate values in this figure (the difference between the solid and dashed lines), it is evident that the approximation has its largest absolute error at the middle of the distribution, especially when the distribution is highly skewed. In general, however, the dashed lines (the approximate values) follow the solid lines (the exact values) quite closely, even in the tails of the distribution. For the level of accuracy typical of judgmental data, it would appear that (13) provides a reasonable level of accuracy. This is particularly encouraging, and to some extent surprising, since the approximation used in (13) is quite simple.

4. Starting Values

Even though a simple approximation was developed in the previous section for the density function of the observation R_{ijk}, the likelihood function containing products of these density functions will still tend to be quite complicated. It is, therefore, not likely that a simple, closed-form solution exists that maximizes this function and consequently, it will be necessary to use iterative methods to obtain the maximum likelihood (ML) estimates of the unknown parameters.

There are a number of standard iterative procedures that can be used (e.g., Chandler, 1969; IMSL, 1979). It does not seem to make a great deal of difference which one is selected, provided the iterative process starts with reasonably good parameter estimates. We consider next, therefore, procedures for obtaining good starting value (SV) estimates for both the coordinates and the uncertainty values of the stimulus and ideal points. Our concern here is to develop quick and simple procedures that can be expected to produce moderately accurate parameter values.

SV estimates of the coordinates. For the purpose of obtaining these initial estimates, we assume that the joint uncertainty value $\sigma_{ij}^2 = \sigma_i^2 + \sigma_j^2$ is small relative to the distance D_{ij}^2. Well established metric, nonprobabilistic procedures can then be utilized. Although the accuracy of these SV estimates will depend very much on the validity of this assumption, it should be noted that these estimates are only the initial values of the iterative process. The final values, the ML estimates, do not require this assumption.

Metric analyses of the unfolding model start generally with a set of I scales. The I scale for single subject P_i consists of the set of distances between ideal point i and each one of the stimulus points. To obtain these I scales in the present case, some preliminary analysis of the ratio judgment R_{ijk} is necessary.

Define

$$R^*_{ijk} = \log R_{ijk}$$

and

$$D^*_{ij} = \log D_{ij}.$$

Then the problem of estimating the distances D_{ij} $(j = m+1, \ldots, m+n)$ for the subject P_i consists of solving the linear systems of equations

$$R^*_{ijk} = D^*_{ij} - D^*_{ik} \quad (j = m+1, \ldots, m+n) \tag{14}$$

for D^*_{ij}. The least squares solution of (14) is

$$\hat{D}^*_{ij} = R^*_{ij\cdot} + D^*_{i\cdot} \tag{15}$$

where we have let

$$R^*_{ij\cdot} = (1/n) \sum_k R^*_{ijk} \tag{16}$$

$$D^*_{i\cdot} = (1/n) \sum_k D^*_{ij} \tag{17}$$

and have taken $\hat{D}^*_{ij}$ to be the least squares estimate of D^*_{ij}.

To be precise, it should be made clear that the solution given in (15) is only an approximation, and that this is true even when the ratio judgment R_{ijk} is averaged over an infinite number of replications. This follows from the fact that the expected value of the log of the ratio judgment R_{ijk} does not equal the true ratio $\log(D_{ij}/D_{ik})$ and will only approach the true value in the limit when the ratios D_{ij}/σ_{ij} and D_{ik}/σ_{ik} both increase without bounds. This latter point is discussed more fully in the next section.

Several different metric approaches can be used to analyze these I scales (Bechtel, 1976; Carroll, 1972; Schönemann, 1970). We have found it very effective to use Schönemann's procedure to solve simultaneously for the coordinates of the stimulus and ideal points, but then to discard the solution for the ideal points. With the coordinates of the stimuli now treated as known, it is possible to set up, for each subject, a linear regression problem to solve for the coordinates of the ideal points and, incidentally, for the subject-specific constant $D_{i\cdot}^{*}$, which appears in (15). Both the unknown coordinates and the antilog of the unknown constant $D_{i\cdot}^{*}$ turn out to be equal to the regression weights of the linear regression equation (see, for example, (26) in Zinnes & MacKay, 1983).

It should be mentioned that this procedure for solving for the stimulus coordinates has had to depend on group data. In essence we have had to assume that the stimulus coordinates do not differ from subject to subject. It would have been desirable to have been able to solve for the stimulus coordinates separately for each subject, but, under the assumptions of the present model, this does not seem to be possible.

SV estimates of uncertainty values. To determine the SV estimates of the uncertainty values σ_i ($i = 1, \ldots, m$) and σ_j ($j = m+1, \ldots, m+n$), it will also facilitate matters to proceed as we did in the previous section.

We assume that the uncertainty values are small compared to the interpoint distances and we also make use of the log transformation. This transformation is especially useful here, because the expected value and variance of the log of a random variable having the central F distribution have, at least approximately, very simple expressions. In particular, if log f has the central F distribution $F(\nu_1^*, \nu_2^*)$, then

$$E(1/2 \log f) \simeq 1/2 \left[\frac{1}{\nu_2^*} - \frac{1}{\nu_1^*} \right] \tag{18}$$

$$\mathrm{Var}(1/2 \log f) \simeq 1/2 \left[\frac{1}{\nu_1^*} + \frac{1}{\nu_2^*} \right] \tag{19}$$

when ν_1^*, and ν_2^* are large (Kendall & Stuart, 1963, p. 379). These results are directly relevant here, because as we have seen in Section 3, the central F distribution is closely related to the distribution of the ratio judgment R_{ijk}. We begin, therefore, with an attempt to use (18) and (19) to obtain an approximate expression for the mean and variance of R_{ijk}^*.

From the definition of R_{ijk} and letting

$$f_{ijk}'' = \frac{d_{ij}^2}{d_{ik}^2} \frac{\sigma_{ik}^2}{\sigma_{ij}^2},$$

we can write

$$R_{ijk}^2 = \frac{\sigma_{ij}^2 \, f_{ijk}''}{\sigma_{ik}^2}. \tag{20}$$

The notation f_{ijk}'' is appropriate here, because, as noted earlier, under the assumptions of the model, f_{ijk}'' has the doubly noncentral F distribution $F''(\nu_1, \nu_2, \lambda_{ij}, \lambda_{ik})$.

Equation (13) now provides the motivation to define

$$f_{ijk} = \frac{a_{ik} \, f_{ijk}''}{a_{ij}} \tag{21}$$

where, as in (10),

$$a_{ij} = \frac{\nu_i + \lambda_{ij}}{\nu_i} \tag{22}$$

because from (13) we know that f_{ijk} will have approximately the central F distribution $F(\nu_{ij}^*, \nu_{ik}^*)$ with the degrees-of-freedom parameters determined, as in (7), by

$$\nu_{ij}^* = \frac{(\nu_i + \lambda_{ij})^2}{\nu_i + 2\lambda_{ij}}. \tag{23}$$

Substituting (21) in (20) gives

$$R_{ijk}^2 = \frac{\sigma_{ij}^2 \, a_{ij} \, f_{ijk}}{\sigma_{ik}^2 \, a_{ik}} \tag{24}$$

which expresses the square of the ratio judgment R_{ijk}^2 directly in terms of a variable having approximately a central F distribution.

We are now in a position to apply the approximation of (18) and (19), along with (24) to obtain the mean and variance of log R_{ijk}. It is appropriate to apply (18) and (19) in the present case, because under our present assumption concerning the size of σ_{ij} relative to D_{ij}, the noncentrality parameter λ_{ij} will be large. And, from (23), it is clear that then the degrees of freedom ν_{ij}^* will also be large, and, in fact, in the limit

$$v_{ij}^{*} = \frac{\lambda_{ij}}{2}. \tag{25}$$

This fullfils the conditions required by (18) and (19).

Thus, applying the operator 1/2 log to both sides of (24), we obtain the expected value

$$E(1/2 \log R_{ijk}^{2}) = 1/2 \log \left[\frac{\sigma_{ij}^{2} a_{ij}}{\sigma_{ik}^{2} a_{ik}} \right] + E(1/2 \log f_{ijk}),$$

which, from (18), (25) and the assumption that λ_{ij} is large, reduces approximately to

$$E(R_{ijk}^{*}) = \log \left[\frac{D_{ij}}{D_{ik}} \right]. \tag{26}$$

We can proceed similarly to obtain the variance of R_{ijk}^{*}, but this time making use of (19) instead of (18). Equation (24) then becomes

$$\begin{aligned} \mathrm{Var}(R_{ijk}^{*}) &= \mathrm{Var}(1/2 \log f_{ijk}), \\ &= 1/2 \left[\frac{1}{v_{ij}^{*}} + \frac{1}{v_{ik}^{*}} \right], \\ &= \frac{1}{\lambda_{ij}} + \frac{1}{\lambda_{ik}}, \\ &= \frac{\sigma_{ij}^{2}}{D_{ij}^{2}} + \frac{\sigma_{ik}^{2}}{D_{ik}^{2}}. \end{aligned} \tag{27}$$

Equation (27) is the basic result we need. It suggests that an estimate of the joint variance σ_{ij}^{2} can be obtained by solving a simple, linear system of equations, provided that we have estimates of the left hand side of (27), namely, the variance of R_{ijk}^{*} and have estimates of the true distance D_{ij}. The latter estimates present no problem, since these interpoint distances can be calculated directly from the SV estimates of the coordinates determined in the previous section. For an estimate of $\mathrm{Var}(R_{ijk}^{*})$ we can make use of the fact that under our present assumptions, the approximation given in (26) is still valid and therefore in the limit,

$$\mathrm{Var}(R^*_{ijk}) = E\left[R^*_{ijk} - \log\frac{D_{ij}}{D_{ik}}\right]^2.$$

Consequently, we can use for an estimate of $\mathrm{Var}(R^*_{ijk})$

$$s^{*2}_{ijk} = \frac{1}{n_{ijk}}\sum\left[R^*_{ijk} - \log\frac{D_{ij}}{D_{ik}}\right]^2 \tag{28}$$

where the summation is taken over the n_{ijk} replications of R_{ijk} and the distances D_{ij} and D_{ik} are, as before, calculated from the SV estimates of the coordinates.

Keeping these estimates in mind, we return to the system of equations given in (27). Since these equations are linear in σ^2_{ij}/D^2_{ij} the least squares solution is

$$s^2_{ij} = D^2_{ij}(s^{*2}_{ij\cdot} - s^{*2}_{i\cdot\cdot}), \tag{29}$$

where we are using s^2_{ij} as the estimate of σ^2_{ij} and

$$s^{*2}_{ij\cdot} = \frac{1}{n-2}\sum_k s^{*2}_{ijk} \tag{30}$$

$$s^{*2}_{i\cdot\cdot} = \frac{1}{(n-1)(n-2)}\sum_{k>j} s^{*2}_{ijk}. \tag{31}$$

In order to arrive at estimates of the uncertainty values σ_i and σ_j for the subjects and the stimuli, the estimates of the joint variances s^2_{ij} given in (29) can be carried out one step further. To do this, however, requires distinguishing between two cases.

Case 1. Assume that the subject and stimulus uncertainty values are unique to each subject. Then the best that can be done with the estimate of the joint variances s^2_{ij} is to set the subject uncertainty estimate s_i equal to some small arbitrary value and to solve for the stimulus uncertainty s_j using

$$s^2_j = s^2_{ij} - s^2_i. \tag{32}$$

There is, however, the possibility that one of the uncertainty estimates might turn out to be negative. This can be avoided by letting the uncertainty parameter for subject P_i be defined by

$$s_i = 1/2 \min_j s_{ij}. \tag{33}$$

This solution has the convenient property of equating the uncertainty estimate for subject P_i with the smallest stimulus uncertainty estimate for this subject.

The estimates of uncertainty parameters given in (32) and (33) have severe limitations. Because of their non-uniqueness properties, it is not meaningful to compare the subject and stimulus uncertainty values over different subjects. It would only be meaningful to compare the stimulus uncertainty values within a single subject.

Case 2. Assume that the stimulus uncertainty values do not differ for different subjects. Possible subject differences would then be solely reflected by differences of the subject uncertainty values. Thus, for this case there are exactly m subject uncertainty values and n stimulus values to estimate. The relevant equations for doing this are the equations:

$$s_i^2 = s_j^2 = s_{ij}^2, \quad i = 1, \ldots, m, \quad j = m+1, \ldots, m+n \tag{34}$$

which is just a simple linear system of equations in the unknowns s_i^2 and s_j^2. Consequently, the least-squares solution is

$$s_i^2 = s_{i\cdot}^2 - \frac{s_{\cdot\cdot}^2}{2} \tag{35}$$

$$s_j^2 = s_{\cdot j}^2 - \frac{s_{\cdot\cdot}}{2} \tag{36}$$

where

$$s_{i\cdot}^2 = (1/n) \sum_i s_{ij}^2 \tag{37}$$

$$s_{\cdot j}^2 = (1/m) \sum_i s_{ij}^2 \tag{38}$$

$$s_{\cdot\cdot}^2 = (1/mn) \sum_i \sum_j s_{ij}^2. \tag{39}$$

This solution has the property of equating the average stimulus variance with the average subject variance, that is, letting

$$(1/m) \sum s_i^2 = (1/n) \sum s_j^2. \tag{40}$$

While this solution is convenient, it also has the undesirable property of allowing some variance terms to take on negative values. To avoid this, the minimum subject variance could be equated with the minimum stimulus variance. This is analogous to the approach taken in Case 1. It can be accomplished by adding and subtracting a constant to the variance estimates obtained from (35) and (36). Specifically, if we let

$$m_i = \min_i s_i^2, \quad i = 1, \ldots, m \tag{41}$$

$$m_j = \min_j s_j^2, \quad j = m+1, \ldots, m+n \tag{42}$$

then if m_i is less than m_j, we can define the new estimate $s_i'^2$ and $s_j'^2$ in terms of the previous estimates, obtained from (35) and (36), by the equations

$$s_i'^2 = s_i^2 + (1/2)\ |\, m_i - m_j \,| \tag{43}$$

$$s_j'^2 = s_j^2 - (1/2)\ |\, m_i - m_j \,| \,. \tag{44}$$

If the converse should be the case, namely that m_j is less than m_i, the same result can be achieved by interchanging the plus and minus signs in (43) and (44).

The fact that the estimates of the uncertainty values are nonunique, as they were in Case 1, means that here too there are limitations as to which uncertainty values can be meaningfully compared. In the present case, it would be meaningful to compare the stimulus uncertainty values among themselves and similarly to compare the subject uncertainty values among themselves. It would not, however, be meaningful to compare the stimulus uncertainty values with the subject uncertainty values. This is to say, that within the framework of the probabilistic unfolding model we have assumed, it is not possible to discriminated between the variability due to the subjects and that due to the ideal points. And this is even true in the present case, where we have assumed that the subjects do not differ with respect to the uncertainty values associated with the stimuli.

5. Simulation I

In the previous sections, we have been concerned with developing a simple procedure for obtaining ML estimates of the parameters of a probabilistic, multidimensional choice model, using as data pairs ratio judgments. This procedure has consisted primarily of obtaining a simple, although approximate, expression for the likelihood function that is to be maximized and of obtaining a simple, although approximate, expression for the estimates of the parameters that can be used as the initial values of an iterative process. To determine how effectively and accurately this procedure works, two simulation studies were performed.

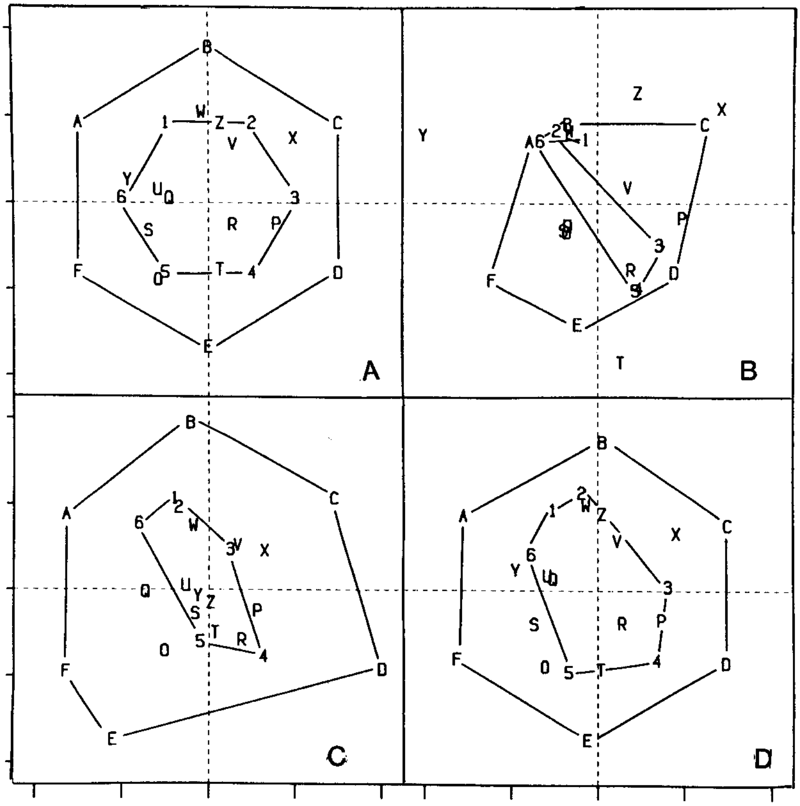

Figure 2. The original and recovered configurations of Simulation 1, Series 4. Panel A shows the original configuration. The 12 stimulus points are labeled A through F and 1 through 6; the 12 ideal points are labeled O through Z. The 6 stimulus points on the inner hexagon had an uncertainty value of 1.2. The remainder 6 stimulus points and 12 ideal points had an uncertainty value of .1. Panel B shows the configuration recovered from a nonmetric analysis; panel C the SV configuration, and panel D the ML configuration.

Simulation I contained 24 points, 12 of which were treated as stimuli and the remaining 12 as ideal points. The 12 stimulus points were located on the vertices of two hexagons, one of which was completely contained within the other. The 6 vertices of the inner hexagon lie on a unit circle, while those on the outer hexagon lie on a circle of radius 1.732. The ideal points were randomly located throughout both hexagons. This arrangement of both the stimuli and ideal points is shown in panel A of Figure 2.

Values of the uncertainty parameter were also assigned to each of the 24 points. The 6 stimuli on the outer hexagon and the ideal points were given the uncertainty value of .1. The 6 stimuli on the inner hexagon were assigned a series of larger values. In Series 1, each one of the six points on the inner hexagon was assigned an uncertainty value of .3. In the remaining 3 series, 2-4, these points were assigned the values .6, .9, and 1.2, respectively. Since the points on the inner hexagon lie on a unit circle, it can be seen that in Series 3 and 4, these points actually had substantial amounts of uncertainty. These large uncertainty values were selected for the points on the inner hexagon, because such values tend to make it very difficult for the estimation procedure to recover the underlying stimulus configuration (Zinnes & MacKay, 1983).

The simulated data in each of the four series consist of 12 sets of 66 pairwise ratio judgments of the 12 stimuli, one set for each of the 12 subjects. This corresponds to a complete set of data for each subject, when it is assumed that the subject do not replicate the judgments, that is, do not judge each set of stimuli more than once.

Preference ratio judgments were constructed by randomly sampling from the bivariate normal distribution, this means invariances were determined in each of the four simulations as indicated previously. In some cases, however, it turned out to be highly desirable to make a slight modification of some of the preference ratios calculated by this process. There were a few instances in which the values of the interpoint distances obtained from these random samples prove to be extremely small or extremely large, thus resulting in either a very small or very large ratio. Since these extreme values had a strong biasing influence on the estimates of the uncertainty parameter, it was considered to be desirable to place upper and lower bounds on the numerical values of the preference ratios. Consequently, a lower bound of .1 and an upper bound of 10 was arbitrarily imposed on these ratios. Calculated ratio values below the lower bound of .1 were set equal to .1 and those above 10, were set equal to 10. The effect of using the upper and lower bounds is explored in the following section.

The sets of data generated in each of the four simulations were analyzed using three different estimation procedures: the SV and ML procedures discussed in the previous section, as well as by KYST, a typical nonmetric (NM) procedure (Kruskal, Young, & Seery, 1973). The KYST analysis was performed by converting the data to *I* scales and using the standard options of that program that are relevant for analyzing *I* scales. This includes the options: "Split by rows", stress 2 (which gave better results that stress 1), "lower corner matrix", a value of STRMIN equal to .0001, and a starting configuration determined by the TORSCA option, which gave lower stress values than those obtained using the true parameter estimates. In all cases, the KYST analysis terminated normally, before reaching 200 iterations.

It will be recalled that there are two types of parameters to estimate: the coordinates and uncertainty values. Since the 12 stimuli and 12 ideal points are embedded in a 2-dimensional space, there are $48 + 3$ or 51 parameters to estimate in all. However, the actual number of independent parameters is somewhat less than this, because of the uniqueness properties of these parameters. The preference ratio judgment is, in fact, invariant over translation, rotation and stretching of the coordinate axes. Because of this, the number of independent coordinates to estimate is actually equal to 47.

Table 2 shows the degree to which the configuration of stimulus and ideal points was recovered by each of the estimation procedures. Two different measures of recovery, R and D^2, are shown in this table: R is the correlation between corresponding interpoint distances in the true and estimated configuration; D^2 is the sum of squared differences between optimally aligned coordinates of the true and estimated configuration. The origins of the coordinate axes for both configurations were placed at the centroid. Both R and D^2 were calculated using the 12 stimulus points and the 12 ideal points.

Table 2 makes it clear that the accuracy of the three estimation procedures, while quite good at low levels of uncertainty, deteriorates as the level of uncertainty increases. This is to be expected, because the higher

Table 2. Hexagon example: recovery of distances and coordinates

		Correlation $(R)^a$	
σ	SV	ML	Nonmetric
.3	.960	.998	.953
.6	.941	.994	.916
.9	.865	.986	.659
1.2	.785	.968	.579
		Squared differences $(D^2)^b$	
.3	.852	.067	1.201
.6	1.353	.153	2.041
.9	3.630	.354	7.329
1.2	6.875	.926	9.372

[a] R is the correlation between corresponding interpoint distances of the true and estimated configurations. It includes points of both stimuli and individuals.
[b] D^2 is the sum of squared differences between optimally aligned coordinates of the true and estimated coordinates. It includes the points of both stimuli and individuals.

levels of uncertainty would tend to produce ratio judgments having a higher degree of variability. It is therefore reasonable to expect the standard errors of the coordinate values to be a function of the uncertainty values.

It is also evident from this table that the rate of deterioration differs for the three estimation methods. The accuracy of the ML estimates seem to decrease only slightly with increases in the level of uncertainty. The deterioration of the SV estimates is somewhat greater and that of the NM (the nonmetric estimation procedure) greater still. These conclusions are consistent with both the R and D^2 statistics. In general, it appears that the ML estimates of the coordinates are actually quite good, even when the levels of uncertainty are substantial.

These conclusions are also evident from the plots shown in Figure 2. This figure shows the configurations recovered from the three different estimated procedures for Series 4, the one containing the highest level of uncertainty. The plots in this figure show the locations of both the 12 stimulus and the 12 ideal points. For comparison purposes, Figure 2 also shows, in panel A, the location of the stimulus and ideal points in the true configuration.

Except for stimulus point 2, located on the inner hexagon, it can be seen that the ML configuration, is as expected from Table 2, exceedingly accurate, even at this high level of uncertainty. The SV and the NM configurations are, as expected from Table 2, considerably less accurate at this high level of uncertainty. This is particularly true of the NM configuration, where the inner hexagon is quite highly distorted. The positions of stimulus points 1 and 2 are in fact reversed from their true position, while stimulus points 1 and 5 actually coincide in the NM configuration.

The ideal points, in the three recovered configurations shown in Figure 2, have properties that are very similar to those of the stimulus points. The accuracy of the ideal points in the ML configuration is substantially better than that in the SV and NM configurations. In fact, the locations of the ideal points in the NM configuration are especially poor. The ideal points T, X, Y and Z, while located within the outer hexagon in the true configuration, have actually been placed well outside this configuration in the NM configuration. In addition, the ideal points U, Q, S, and O, while having very distinct positions in the true configuration, actually coincide in the NM configuration.

Table 3. Hexagon example: recovery of uncertainty values

True		SV		ML	
σ_1	σ_2	σ_1	σ_2	σ_1	σ_2
.3	.1	.368	.095	.303	.088
.6	.1	.441	.113	.606	.089
.9	.1	.500	.152	.884	.089
1.2	.1	.507	.293	1.123	.088

Note: σ_1 and σ_2 are the uncertainty values of the coordinates of the inner and outer hexagons, respectively.

Table 3 gives some indication of how well the uncertainty values are recovered by two of the estimation procedures, the SV and the ML procedures. Estimates using nonmetric methods are not shown in this table because those methods are purely deterministic and therefore do not provide for an estimation of variances. The uncertainty values shown in this table are consistent with the coordinates that were obtained in the process of calculating the sum of squared differences D^2. In other words, the same multiplicative factor was applied to the joint variances σ_{ij}^2 as were applied to the coordinates in the process of optimally aligning the estimated configuration with the true configuration. This is appropriate, because the joint variances σ_{ij}^2 are only determined up to a multiplicative transformation. To arrive at the uncertainty values of the stimuli and the ideal points, the standardization discussed in the previous section was used. The minimum uncertainty values of the stimulus was set equal to the minimum uncertainty value of the ideal points. In the present case, this effectively means setting the uncertainty value of the ideal points equal to the uncertainty value of the stimulus points on the outer hexagon.

Table 3 shows that the accuracy of the uncertainty estimates depends on the magnitude of the true value. The large uncertainty values are not estimated as well as the smaller values, although the ML estimates of the larger uncertainty values are actually quite good. In contrast, the SV estimates do deteriorate substantially at the higher levels of uncertainty. This is to be expected, since they were derived by assuming that the interpoint distances are large relative to the sizes of the joint variances.

6. Simulation II

In the previous simulation, lower and upper bounds were placed on the ratio judgments, to avoid the effects of extreme values biasing the parameter estimates. To determine whether, in general, such limits should be used, an additional simulation study was performed, one in which the number of replications of the preference ratio judgments was varied. Four levels of replications were used: stimuli on the inner hexagon were set equal to .3. As in the previous simulation, the stimuli on the outer hexagon were assigned an uncertainty value of .1, as were the ideal points also.

The configuration of stimuli and ideal points used in this simulation was identical to the one used in the previous simulation.

The data from the simulation were analyzed using two different approaches. In the No Limits approach, the ratio judgments obtained from random samples of coordinates were not modified, even when under some conditions they produced extreme values. In the Fixed Limits approach, lower and upper bounds of .1 and .10 were imposed on the ratio judgments.

Table 4. Hexagon example: SV estimates of the Uncertainty parameter for different number of replications

Replications	No limits[a]		Fixed limits[b]	
	σ_1	σ_2	σ_1	σ_2
1	.252	.187	.368	.095
2	.238	.145	.338	.090
4	.300	.116	.321	.070
8	.293	.112	.334	.068

Note: The correct values of σ_1 and σ_2 are .3 and .1, respectively.
[a] No upper limit was placed on the ratio judgments to be analyzed.
[b] Lower and upper limits of the ratio judgments to be analyzed were set at .1 and 10, respectively.

The results for both approaches are shown in Table 4, for both the SV and ML estimation procedures. The values in this table show quite unmistakably that under the No Limits approach, the estimates of the uncertainty values become increasingly more accurate as the number of replications increase. This is not the case, however, for the Fixed Limits approach. Although the estimate of σ_1 improved slightly with increases in replication, the estimate of σ_2 becomes appreciably worse. Furthermore, although the No Limits approach is slightly better than the No Limits approach at low levels of replication, it is clearly the case that the No Limits approach is superior at the higher levels of replication. Thus, a concern for possible biasing effects of extreme values of the preference ratio appears to be justified only when there are few if any replications of the preference ratio judgments.

These results seem very encouraging. Even though the data of these simulation studies are highly artificial, they seem to show that the maximum likelihood estimation procedure does amazingly well in recovering the location and the uncertainty values of the stimuli and the ideal points, even in the presence of a high degree of uncertainty. Of course, these results depend upon the validity of the underlying probabilistic model with which we are working. If the model were inappropriate in some situation, one could not expect the maximum likelihood estimation procedure, or any estimation procedure, to accurately estimate the unknown parameter values.

It should also be recalled that the SV procedure was only intended to provide the starting values for the ML iterations. The fact that the SV estimates are as good as they are, given their relative simplicity, is also encouraging. Thus under normal conditions, when the data do not contain substantial amounts of variability, the ML iterations should converge fairly rapidly.

References

Anderson, N. H. (1982). *Methods of information integration theory.* New York: Academic Press.

Ashby, F. G., & Tonwsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154-179.

Bechtel, G. G. (1976). *Multidimensional preference scaling.* The Hague: Mouton.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In E. Wegener (Ed.), *Social attitudes and psychological measurement.* Hillsdale, NJ: Erlbaum

Böckenholt, I., & Gaul, W. (1984). A multidimensional analysis of consumer preference judgments related to print ads. Methodological advances in marketing research in theory and practise. EMAC/ESOMAR Symposium, Copenhagen.

Bulgren, W. G. (1971). On representations of the doubly non-central *F* distribution. *Journal of the American Statistical Association, 66,* 184-186.

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

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

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

Croon, M. (in press). A comparison of statistical unfolding models. *Psychometrika.*

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

Eisler, H. (1982). On the nature of subjective scales. *Scandinavian Journal of Psychology, 23,* 161-171.

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

IMSL (1979). *IMSL library reference manual.* New York: International Mathematical and Statistical Libraries, Inc.

Kendall, M. G., & Stuart, A. (1961). *The advanced theory of statistics* (Vol. 2). New York: Hafner.

Kendall, M. G., & Stuart, A. (1963). *The advanced theory of statistics* (Vol. 1, 2nd ed.). New York: Hafner.

Kruskal, J. B., Young, F. W., & Seery, J. B. (1973). How to use KYST, a very flexible program to do multidimensional scaling and unfolding. Bell Telephone Laboratories, Murray Hill, NJ.

MacKay, D. B., & Zinnes, J. L. (1981). Probabilistic scaling of spatial judgments. *Geographical Analysis, 13,* 21-37.

MacKay, D. B., & Zinnes, J. L. (1986). Probabilistic multidimensional scaling of spatial preferences. In R. Golledge & H. Timmerman (Eds.), *Behavioral modeling: Approaches in geography and planning.* New York: Croom-Helm.

MacKay, D. B., Ellis, M., & Zinnes, J. L. (1986). Graphic and verbal presentation of stimuli: A probabilistic MDS analysis. *Advances in Consumer Research, 13,* 529-533.

MacKay, D. B., & Zinnes, J. L. (in press). Probabilistic multidimensional scaling of residential preferences: An experimental evaluation. *Cahiers de Géographie de Besançon.*

Patnaik, P. B. (1984). The non-central chi-square and *F* distributions and their approximations. *Biometrika, 36,* 202-232.

Saaty, T. L. (1980). *The analytic hierarchy process.* New York: McGraw-Hall.

Scheffé, H. (1959). *The analysis of variance.* New York: Wiley.

Schönemann, P. H. (1970). On metric multidimensional unfolding. *Psychometrika, 35,* 349-366.

Sjöberg, L. (1967). Successive intervals scaling of paired comparisons. *Psychometrika, 32,* 297-308.

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

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

Zinnes, J. L., & Kurtz, R. (1968). Matching, discrimination, and payoffs. *Journal of Mathematical Psychology, 5,* 392-421.

Zinnes, J. L., & Griggs, R. A. (1974). Probabilistic multidimensional scaling unfolding analysis. *Psychometrika, 39,* 327-350.

Zinnes, J. L., & Wolff, R. P. (1977). Single and multidimensional same-different judgments. *Journal of Mathematical Psychology, 16,* 30-50.

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