MULTIVARIATE BEHAVIORAL RESEARCH, *42*(1), 103–132
Copyright © 2007, Lawrence Erlbaum Associates, Inc.

Multidimensional Unfolding by Nonmetric Multidimensional Scaling of Spearman Distances in the Extended Permutation Polytope

Katrijn Van Deun
Catholic University of Leuven

Willem J. Heiser
Leiden University

Luc Delbeke
Catholic University of Leuven

A multidimensional unfolding technique that is not prone to degenerate solutions and is based on multidimensional scaling of a complete data matrix is proposed: distance information about the unfolding data and about the distances both among judges and among objects is included in the complete matrix. The latter information is derived from the permutation polytope supplemented with the objects, called the preference sphere. In this sphere, distances are measured that are closely related to Spearman's rank correlation and that are comparable among each other so that an unconditional approach is reasonable. In two simulation studies, it is shown that the proposed technique leads to acceptable recovery of given preference structures. A major practical advantage of this unfolding technique is its relatively easy implementation in existing software for multidimensional scaling.

Ordering is a natural activity. For example, competitions are held to determine who is the first, second, ... last. When making decisions, people order the dif-

*This article was completed while Willem J. Heiser was Research Fellow at the Netherlands Institute for the Advanced Study in the Humanities and Social Sciences (NIAS) in Wassenaar, The Netherlands.

Correspondence concerning this article should be addressed to K. Van Deun, Department of Psychology, Catholic University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium. E-mail: katrijn.vandeun@psy.kuleuven.be

ferent alternatives on important criteria. Also, often people are asked to order a list of objects according to their preference. The objective of this article is to obtain a low-dimensional graphical representation of the preference data where both the judges and the objects are represented in a way that the more an object is preferred, the closer it is located to the judge. Such techniques are known as multidimensional unfolding and have been primarily applied as a tool for the analysis of preference data, though they are applicable to all data that consist of full rankings. In the next section, we take up a discussion on the technique of multidimensional unfolding. From this discussion it is clear that the distance information provided by the preference data is insufficient, with the result that degenerate solutions are obtained. We then introduce a high-dimensional theoretical structure, the permutation polytope, that provides supplementary information which can be used to derive a complete matrix such that nondegenerate solutions are obtained. In two simulation studies, we evaluate the performance of our proposed technique and we illustrate it with some empirical data.

MULTIDIMENSIONAL SCALING AND UNFOLDING

In its simplest form, multidimensional scaling (MDS) is a geometric mapping technique for data that express the distances among the objects of a set. For example, subjecting the distances among the European capitals to a multidimensional scaling analysis will result in a (possibly rotated) map of Europe. More general measures of distance, called *proximities* or *(dis)similarities* can be used because MDS allows for transformation of the data. A distinction is made between metric (linear transformations with or without an intercept) and nonmetric multidimensional scaling (monotonic transformations).

By allowing for weights and a different transformation for each row, preference data can be subjected to a multidimensional scaling analysis as well. This case involves restructuring the data in a square symmetric $(n + m) \times (n + m)$ matrix with empty blocks on the diagonal and the preference data inserted in the off-diagonal part (see e.g., Borg & Groenen, 1997, p. 233). Such a multidimensional scaling of preference data is known as *multidimensional unfolding*. Although ordinary MDS is known to generally yield fine results, the special case of multidimensional unfolding is cursed by degenerate solutions. These are solutions that fit well but are not interpretable (see Van Deun, Groenen, Heiser, Busing, & Delbeke, 2005). For example, frequently a configuration is found where all judges fall together in the center point of a circle formed by the objects.

When comparing the general case of multidimensional scaling with unfolding, the main difference seems to be the large number of missing values. However, this characteristic is not the reason why degenerate solutions occur in unfold-

ing: multidimensional scaling with a large number of missing values yields robust results on the condition that the missing data are not clustered (Spence & Domoney, 1974). The real reason is that the missing data are highly structured, in such a way that adding a positive constant to all nonmissing elements tends to decrease badness-of-fit and allows for the degeneracy to occur, whereas in ordinary MDS adding a positive constant tends to increase badness-of-fit and moves the solution into higher dimensionality. A way out of the problem of degenerate solutions seems to complete the MDS matrix. For example, the proximity among judges can be expressed by the Spearman correlation between their rankings. Note that the rank scores are not comparable with the rank correlations and this can be accounted for by a block-conditional MDS analysis meaning that each block is subjected to a different transformation (see Lingoes, 1977, and Steverink, Heiser, & van der Kloot, 2002, for such block-conditional analyses). Steverink et al. (2002) developed an MDS program for the analysis of completed unfolding data that allows even for different transformation levels of the partitions forming the MDS matrix. However, the preference and subject block are subjected to the same transformation function in their examples, or the MDS matrix contains in fact only two partitions. Because the main objective of the present article is to obtain nondegenerate unfolding solutions by performing MDS on a complete matrix, we tested in a pilot study a true block-conditional approach with three partitions (a block for the preference data, one for the judges and one for the objects). This resulted in degenerate solutions with the consequence that, for our approach to work, we need to find a matrix of all comparable proximities such that an *unconditional* approach is possible. In the next section, we introduce a structure, called the permutation polytope, from which comparable distances can be derived.

THE PERMUTATION POLYTOPE

A rank vector $\boldsymbol{\pi}_i$ is a permutation of the positive integers $1, \ldots, m$ where the numbers indicate the position the object occupies in the ordering. For example, given three objects A, B, and C the ranking $\boldsymbol{\pi}_i = [2\ 3\ 1]'$ corresponds to the ordering $\{C, A, B\}$. A well-known case of ranking data is preference data where n judges $i = 1, \ldots, n$ order m objects $k = 1, \ldots, m$ according to their preference. These data can be represented as n points in m-dimensional real space where the rank scores give the coordinates: In the left panel of Figure 1 the different rankings of three objects are plotted with respect to the axes A^-, B^-, and C^-. Because the rankings all sum to $m(m+1)/2$, they lie in $(m-1)$-dimensional space. The convex hull of all possible rankings $\boldsymbol{\pi}_i$ of m objects is called the *permutation polytope* (see Marden, 1995). The right panel of Figure 1, which is a detail of the left panel, depicts the permutation polytope for the three

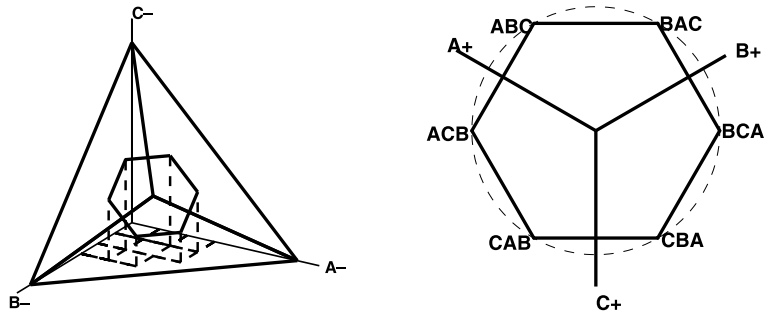


FIGURE 1 Spatial representations of ranking data for three objects. In the left panel all possible rankings of three objects are plotted in three-dimensional space; a detail is given in the right panel that depicts the two-dimensional space spanned by these rankings and labels them by their corresponding ordering.

objects A, B, and C with the rankings labelled by their corresponding ordering. Note that a high rank score on the axes of the m-dimensional space (see the left panel of Figure 1) corresponds to low preference for that object (and thus the object occupies one of the last positions in the rank order, see the right panel of Figure 1). Adjacent orderings on the permutation polytope are found by interchanging one adjacent pair of objects while orderings lying opposite to each other have a reversed order. Note that all rankings are equidistant from the center point $\mathbf{c} = [(m + 1)/2]\mathbf{1}$ with a distance of $[m(m^2 - 1)/12]^{1/2}$ (see Marden, 1995). In the tied rank approach, the center point coincides with the ranking that expresses complete indifference. Figure 2, in which the origin is translated to the center point of the polytope, illustrates the case of four objects.

In the $(m - 1)$-dimensional space containing the permutation polytope, poles of attraction can be identified (Marden, 1995). These are represented in the right panel of Figure 1 by the three halve lines and in Figure 2 by the four half lines. Each of these labelled lines represents points that are equidistant and closest to those rankings that rank the corresponding object first (they are also equidistant with respect to the rankings that rank them second, ..., and that rank them last). If we would have to place the objects in this $(m - 1)$-dimensional space, a logical choice would be the corresponding attraction pole. This leaves us with an infinite number of possibilities but we propose to represent the object by the point on the attraction pole at the same distance from the center as the rankings. The reasons for this choice will be given shortly. As a result, both the rankings and the objects lie on a $(m - 1)$-dimensional sphere. We call this a *preference sphere*.

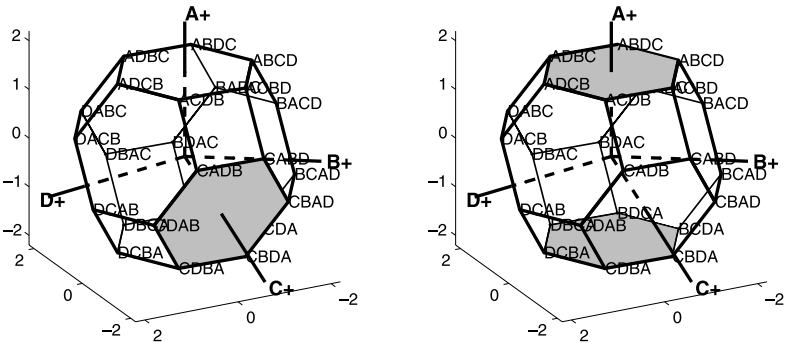
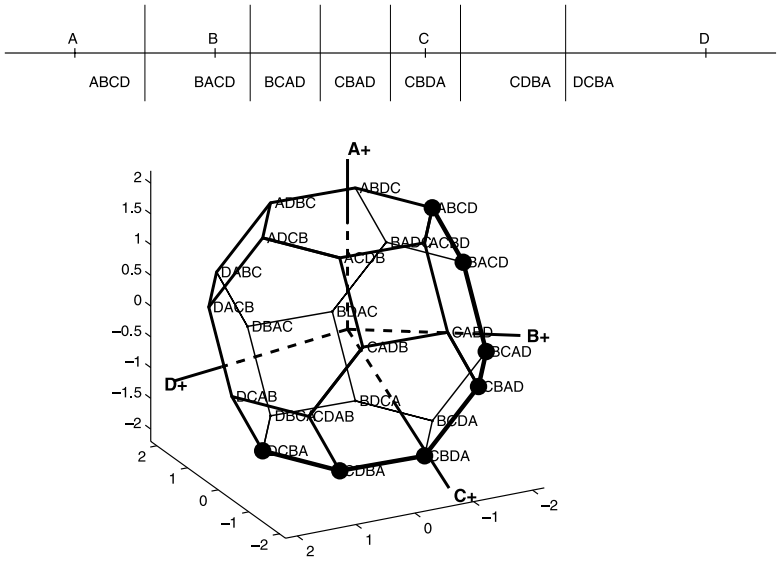


FIGURE 2 Permutation polytope and attraction poles for four objects. In the left panel the side corresponding to preference for C is shaded. In the right panel the shaded sides correspond to preference and dispreference for A.

The Rationale of Our Approach

The core idea of this article, that is, to obtain a low-dimensional unfolding space by scaling the distances derived from the permutation polytope, needs some more thought and justification than might appear at first sight. The permutation polytope is a given theoretical reference structure for ranking data that is completely determined by the number of preference items: no data (here, rankings) are needed to define it. The data add information on which rankings occur and with what frequency. In fact, the polytope supplemented with these frequencies is merely an equivalent geometric representation of the data. Although the permutation polytope is a high-dimensional structure, the data probably stem from a low-dimensional one. An illustration is given in Figure 3a where a one-dimensional preference space has a high-dimensional polytope structure (also see Heiser, 2004). It is this low-dimensional structure, that is determined by order relations between the distances both in the low-dimensional preference space and on the permutation polytope, that we want to recover. An example of an order relation between distances is that the distance from the ranking BCAD to CBAD is smaller than from the the former ranking to CBDA. Here multidimensional scaling comes into play because it is an outstanding technique to find low-dimensional representations when only the distances between the points are given. Note that projection techniques are unsuitable to find the low-dimensional representation, as they cannot remove the nonlinearity and therefore tend to overestimate dimensionality. To be able to perform this nonmetric MDS, we should know the (order of) distances between all points. In the following section we motivate our choice for the Spearman distance.



(a) Theoretical representations



(b) Non-metric MDS solution

FIGURE 3 Representations of preference rankings for four objects A, B, C and D: (a) one-dimensional theoretical preference space (top), permutation polytope (middle), and (b) one-dimensional recovered non-metric MDS representation using our approach (bottom). Figure a: Theoretical representations. Figure b: Non-metric MDS solution.

DERIVING DISTANCES FROM THE PERMUTATION POLYTOPE

So far, we have shown how preference data can be equivalently represented by a geometric structure called the permutation polytope. Also, it was made clear that the objects should be represented such that they are equidistant with respect to the rankings that pick them first, second, . . . , last. Furthermore, the discussion on multidimensional unfolding showed that to avoid degeneracies, dissimilarities are needed that are comparable and, as is explained, that have an equal range of values in the different blocks. Finally, these dissimilarities should reflect the original preference data. There are ample ways to measure the dissimilarity be-

tween rankings, or, the distance between points on the permutation polytope. However, only few of them will fulfill the requirements just mentioned: For example, it can be proven that in case of Kemeny distances either the dissimilarity measures have a different range of values for the different blocks or the objects are not equidistant with respect to the rankings that pick them first, second, ..., last. In this section we show that Spearman distances, calculated between points on the preference sphere, fulfill the preset requirements.

Distances Among Rankings

The degree in which rankings (judges) correspond with each other is often of interest and can be assessed by several measures of ordinal association. Among these are Spearman's rank correlation, Kendall's tau, Cayley's distance and Kemeny's distance (see Kemeny & Snell, 1962). For an overview, see Kruskal (1958) and Diaconis (1988). Spearman's rank correlation is probably the best known and most frequently used measure. It is defined by

$$\rho = 1 - \frac{6}{m(m^2-1)} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)^2, \tag{1}$$

with $d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)$ the Euclidean distance between the rankings $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_j$ and where the squared Euclidean distance $d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)^2$ is known as Spearman's distance. Here, we will measure the degree of noncorrespondence or dissimilarity between two rankings with this squared Euclidean distance. As can be seen in (1), this measure of dissimilarity is a monotonic transformation of Spearman's ρ. During a nonmetric multidimensional scaling, the dissimilarities are subjected to monotonic transformations and therefore using Spearman's distance as a measure of dissimilarity will yield the same results when using the Spearman rank correlation as a similarity measure.

Note that the squared norm $\|\boldsymbol{\pi}_i\|^2$ is constant, both with $\mathbf{c}$ and $\mathbf{0}$ as the origin: $\|\boldsymbol{\pi}_i\|^2$ is respectively equal to $m(m^2-1)/12$ and $m(2m+1)(m+1)/6$. With $\mathbf{c}$ as the origin, it follows from the law of cosines that the squared distance between two rankings equals $m(m^2-1)(1-\cos\theta)/6$, with θ the angle between the two rank vectors. In other words, the distance between two rankings is merely a function of the angle between the rankings.

To measure the distance between all pairs of n judges, we calculate the squared Euclidean distance between the preference rankings of the two judges composing the pair. As shown, this distance is both closely related to Spearman's rank correlation and to the cosine of the angle between the rank vectors. The same multidimensional scaling solution will be obtained with each of the three (dis)similarity measures for an identical starting configuration.

Distances Involving the Objects

We need to find a representation for the objects that fits in the approach taken so far, that is, in reference to the permutation polytope and such that the Spearman distance can be used as a measure of *comparable* dissimilarities. As already noted, in the space spanned by the permutation polytope, lines can be added of points that are closest and equidistant to all those rankings that rank the object first (see Figures 1 and 2). It is natural that the objects should be positioned on these lines. A corresponding ranking would be one with highest preference for that particular object and indifference for the remaining objects. This is a partial ranking of the *pick r out of m* type with $r = 1$. A representation of the object k by a tied ranking $\boldsymbol{\pi}_k$ is then,

$$\begin{cases} \pi_{kk'} = 1 & \text{if } k = k', \\ \pi_{kk'} = 1 + m/2 & \text{if } k \neq k'. \end{cases} \tag{2}$$

For example, with four objects A, B, C, and D, the tied ranking for B is $[3\ 1\ 3\ 3]'$. Such a representation of the objects, which was already suggested by Heiser (1999) and used by Steverink et al. (2002), is not only intuitively appealing but also supported by the results of Critchlow (1980) who extended some of the measures for ordinal association to the case of partially ranked data. He demonstrated that the tied ranks approach, in combination with Spearman's ρ as a measure of agreement, is a reasonable metric. Note that the objects are equidistant with respect to both $\mathbf{c}$ and $\mathbf{0}$ with $d(\boldsymbol{\pi}_k, \mathbf{c})^2 = m(m-1)/4$ and $d(\boldsymbol{\pi}_k, \mathbf{0})^2 = 1 + ((m-1)(m+2)^2)/4$. The distance between two objects or between an object and a judge can then again be calculated by Spearman's distance. In the Appendix, we prove that this distance between the vectors $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_k$ preserves the order of the preference scores. An important aspect of the proof is that it shows that the object vectors should have equal norms.

As it was the case for the judges, here too the order of the distances from a judge to the different objects depends only on the cosine of the angle of the rank vectors. This follows from the equal length of the object vectors. We prove this result with the law of cosines,

$$\begin{aligned} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2 &< d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_{k'})^2 \\ \|\boldsymbol{\pi}_i\|^2 + \|\boldsymbol{\pi}_k\|^2 - 2\|\boldsymbol{\pi}_i\|\|\boldsymbol{\pi}_k\|\cos\theta_{ik} &< \|\boldsymbol{\pi}_i\|^2 + \|\boldsymbol{\pi}_{k'}\|^2 \\ &\quad - 2\|\boldsymbol{\pi}_i\|\|\boldsymbol{\pi}_{k'}\|\cos\theta_{ik'} \\ \cos\theta_{ik} &> \cos\theta_{ik'}. \end{aligned} \tag{3}$$

Note that changing the length of the object vectors does not change the preference orders. We can use this property to place the objects on the attraction pole at a distance $\|\boldsymbol{\pi}_i\|$ from the origin. In this case, using the squared Euclidean distance as a measure of dissimilarity results in a measure that depends only on the angle between the vectors. Again, comparable dissimilarities are obtained because it seems reasonable to state that the smaller the angle between two vectors, the more the rankings are similar, and this without considering if these refer either to two judges, two objects, or a judge and an object.

So far, we have shown that the position of the object points with respect to the permutation polytope, should be on the attraction poles and at the same distance from the origin for all objects. Furthermore, two positions were found for which the squared Euclidean distances between the points lead to comparable dissimilarities, both within and between blocks. A first one was the one defined by the tied rank vectors of (2). Note that these rankings are a convex combination of rankings that rank the object first, so that these points are positioned on the intersection of the attraction pole with the permutation polytope. Another possible position was obtained by positioning the object points on the attraction pole and at the same distance from the origin as the points representing the judge. All points lie then on a sphere. A case can be made to scale the object vectors such that $\|\boldsymbol{\pi}_k\| = \|\boldsymbol{\pi}_i\|$ in order to avoid degenerate solutions. Already in 1962, Shepard reported a degenerate solution when (unconditionally) scaling a complete matrix with larger between than within set proximities. Calculating Spearman's distance with $\mathbf{c}$ as the origin of the Euclidean space so that $\|\boldsymbol{\pi}_i\|^2 = (m(m^2-1))/12$ and $\|\boldsymbol{\pi}_k\|^2 = (m(m-1))/4$ leads to the following results:

$$d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)^2 = \frac{m+1}{6} m(m-1)(1-\cos\theta_{ij}) \tag{4}$$

while

$$d(\boldsymbol{\pi}_k, \boldsymbol{\pi}_{k'})^2 = \frac{3}{6} m(m-1)(1-\cos\theta_{kk'}) \tag{5}$$

and

$$d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2 = \frac{3}{6} m(m-1)\left[\frac{m+4}{6} - \sqrt{\frac{m+1}{3}}\cos\theta_{ik}\right]. \tag{6}$$

When comparing equations (4) and (5) it is clear that for increasing m the distances among judges become relatively longer than those among objects. A comparison of (5) and (6) shows that for increasing m the distances between a judge and an object becomes longer than the distance among objects. A nonmetric

unconditional multidimensional scaling solution can be expected to degenerate, with the objects clustered together. We actually obtained this type of solutions for empirical data. For this reason, the length of the object vectors, with $\mathbf{c}$ as the origin, is set equal to $(m^{1/2}(m^2-1)^{1/2})/(12^{1/2})$. All distance formulas become then equal to (4). Despite these precautions, the distances can still be smaller among judges than from a judge to the different objects: this situation can occur when the judges have similar preferences. An additional adjustment is needed then, which shifts the distribution of the preference distances to those smaller distances. For a further discussion of this problem, we refer to the section on the Delbeke data.

In summary, the objects should be positioned on their attraction pole and with norm $\|\boldsymbol{\pi}_i\|$. As a result, all points lie on a sphere with center $\mathbf{c}$. Only then the squared Euclidean distance is a measure of dissimilarity with the required properties of preserving the preference orders, of being overall comparable, and of avoiding larger between than within set distances. An unconditional multidimensional scaling analysis can then be performed on these dissimilarities to obtain a non-degenerate unfolding configuration. Note that all dissimilarities between objects are equal on the preference sphere: to avoid equidistance among the objects in the solution configuration, the primary approach to ties should be used, which allows ties to become untied.

Our unfolding technique is given by the following steps:

1. Add m additional rows $\boldsymbol{\pi}_k$ to the preference data that represent the objects by tied ranks as in (2).
2. Center the data with respect to $\mathbf{c}$.
3. Scale the object vectors by multiplying the additional rows by $((m+1)^{1/2})/3^{1/2}$.
4. Derive the $(n+m)\times(n+m)$ matrix of dissimilarities by calculation of the squared Euclidean distance between the centered (and scaled) rank vectors.
5. Perform a nonmetric multidimensional scaling analysis on the quantities obtained in the previous step.

Thus our approach to unfolding consists of some easy-to-take data preprocessing steps, followed by an ordinary multidimensional scaling. We applied our unfolding technique to the theoretical example of Figure 3a. The resulting configuration, depicted in Figure 3b, perfectly recovers the one-dimensional preference space in a non degenerate manner. By using Spearman distances, we were able to describe the permutation polytope in a way that nonmetric MDS can recover the underlying low-dimensional structure. In the following sections we show that this approach gives satisfying results, both in simulations and with empirical data.

SIMULATION STUDIES

The objective of the simulation studies presented here was to evaluate how well both the underlying low-dimensional space and the rank orders were recovered by our unfolding procedure. We also wanted to know which factors were influential in this respect. Two studies were conducted, because the low-dimensional hypothetical structure underlying preference data can be defined in two ways that lead to two different data-generating mechanisms. The first hypothetical low-dimensional structure for preference data, is one where only the objects hold a fixed position: in such a space, only a restricted number of preference orders can occur. In this case, the rankings were perturbed. The structure underlying the second simulation study, is one that positions both judges and preference items in a low-dimensional preference space, and here we perturbed the distances.

Simulation Study 1

The first simulation study is based on the hypothesis that a preference space is made up by the positions of the objects only. With m objects, $m!$ preference orderings can be given but only a limited number of them can be represented in a low-dimensional Euclidean space (Coombs, 1964). This number depends on the scatter of the objects throughout the Euclidean space. We illustrate the situation for the four objects A, B, C, and D in Figure 4. Between each pair of objects a line can be drawn of points that are at the same distance of both objects. It follows then that the subjects located at the same side of the line as the object, prefer this object over the object at the other side. By drawing these lines of equidistance for each pair of objects, regions are formed that hold points with the same distance ordering of the four objects. In Figure 4, six lines of equidistance can be drawn. These lines separate 18 regions, 17 of which are shown.[1] Each of these regions correspond to points that all have the same preference ordering and are called isopreference regions. With four objects, 24 different orderings can be obtained but in two dimensions maximally 18 of them can be represented (see Bennett & Hays, 1960). This maximum number of regions for a given dimensionality and a given number of objects is called the cardinality (Bennett & Hays, 1960).

In the first simulation study, two-dimensional configurations with four, five, six, or seven objects and the maximum number of isopreference regions were used. These configurations were obtained by generating random coordinates for the objects and by counting the number of isopreference regions using the linear programming function of MATLAB (2002) as follows: First, define each of the $m!$ rankings by a set of inequalities that express the position of the points in

[1]The 18th is DABC and falls outside the range of the plot.

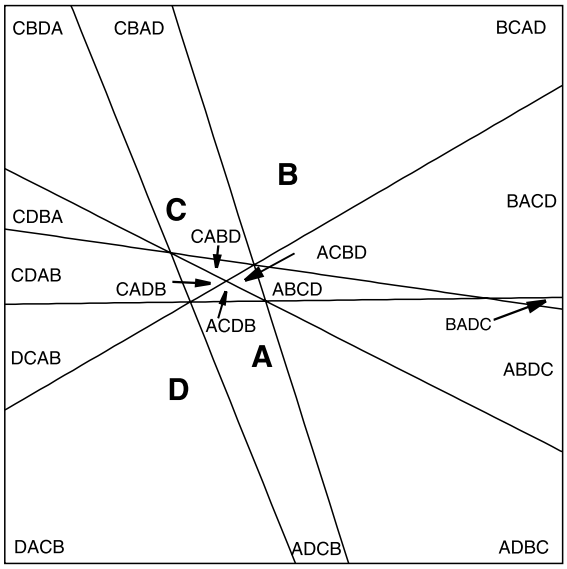


FIGURE 4 Isopreference regions for four objects in two dimensions.

the corresponding isopreference region with respect to each line of equidistance, and second, count the number of sets of inequalities for which MATLAB finds a point that satisfies the set. In this way, four configurations that reached the cardinality were constructed for each number of objects.[2] Error was added to the admissible rankings of the configurations using Mallows θ model (see Critchlow, 1980; Mallows, 1957). According to this model the probability of a ranking $\boldsymbol{\pi}$ given a central ranking $\boldsymbol{\pi}_0$ is exponentially distributed as follows,

$$p(\boldsymbol{\pi}) = Ce^{-\lambda d(\boldsymbol{\pi},\boldsymbol{\pi}_0)}, \tag{7}$$

with $d(\boldsymbol{\pi}, \boldsymbol{\pi}_0)$ the Euclidean distance between the rankings, $\lambda \geq 0$ a dispersion parameter and $C = (\Sigma e^{-\lambda d(\boldsymbol{\pi},\boldsymbol{\pi}_0)})^{-1}$ where the sum is taken over all $m!$ rank orders. C is a constant of proportionality that ensures that the probabilities add up to 1. When λ equals zero, all rank orders have equal probability (the null model) while for increasing λ the distribution becomes more and more peaked around $\boldsymbol{\pi}_0$. Random ranks are obtained under (7) by picking a rank vector with a cumulative probability (almost) equal to a random number generated under

[2]Note that under random generation of the object coordinates, it is most unlikely to find a configuration that does not reach the cardinality.

the uniform distribution. Note that several ranks are equidistant to the central rank and this should be taken into account, for example, by a random choice among the equidistant ranks. With this sampling procedure it is possible to pick a ranking that is not admissible for the given object configuration.

For the simulation study, the number of objects and λ were varied, and the dimensionality was set at 2 (see Van Blokland-Vogelesang, 1989, for the one-dimensional case): m took values 4, 5, 6, and 7 whereas λ took values zero, one half, 1, and 2. The cardinality equalled 18, 46, 101, and 197, respectively. Each of the possible rankings in the configuration was used as a central ranking in the Mallows distribution from which five rankings were sampled. Note that the sampling procedure allows for rankings that are not admissible. These form the input data for our unfolding analysis (e.g., a data set of size $18 \times 5 = 90$ rows and 4 columns for $m = 4$). For every data set, rankings were generated under each of the four dispersion levels. Having obtained the rankings, the dissimilarity matrix was derived from the corresponding preference sphere and subjected to a nonmetric MDS analysis. We were mainly interested in the recovery of the admissible rankings but we also considered the recovery of the positions of the objects by our unfolding technique. For the MDS analysis we used the SMACOF algorithm[3] (see De Leeuw & Heiser, 1980) with a multistart procedure based on 20 random starts for each data set and with 200 iterations per random start. In this way we tried to account somewhat for the fact that MDS is known to yield local optima (see Groenen, 1993) and that early stopping can result in a configuration that is on its way to become degenerate.

To measure the recovery of the rankings, we used index $C1$ (Carroll, 1972) which measures the proportion of preference orders that is correctly reproduced by the distances. This proportion is computed on the $m(m-1)/2$ pairwise comparisons per subject and $C1$ is the average over subjects. Note that $C1$ will have values between approximately 0.50 and 1 as values below 0.50 indicate an inverse relation between the distances and the proximities. As a measure of recovery of the positions, we will use the Procrustes statistic (see, e.g., Cox & Cox, 1994). This statistic is the sum of the squared Euclidean distances between the points in the solution configuration and in the true configuration after bringing the two sets of points into optimal correspondence by a Procrustean similarity transformation. To obtain a normalized version, the Procrustes statistic is divided by the sum of squared norms of the points in the true configuration,

$$L(\mathbf{X}, \mathbf{X}_0) = \frac{\text{tr } (\mathbf{X} - \mathbf{X}_0)'(\mathbf{X} - \mathbf{X}_0)}{\text{tr } \mathbf{X}_0'\mathbf{X}_0}, \tag{8}$$

[3]The SMACOF algorithm is implemented in the PROXSCAL program available from SPSS, version 10 and higher.

with $\mathbf{X}$ the matrix of transformed coordinates of the solution and $\mathbf{X_0}$ of the initial configuration. $L(\mathbf{X}, \mathbf{X_0})$ takes values between zero and one. For ease of interpretation, we use a derived measure, $R^2 = 1 - L(\mathbf{X}, \mathbf{X_0})$, which is similar to the proportion of variance accounted for: a value of zero represents a complete misfit between the configurations while a value of one represents a perfect fit.

The results of the simulation study are summarized by the boxplots in Figures 5 and 6. Of particular interest are the conditions with $\lambda = 0$ because these consist of data that were generated without any relation to the true data. As can be expected, $C1$ is approximately equal to .50. R^2 however, is for most data sets much larger than zero, but this is due to the fact that only few points (4 up to 7) are used in the Procrustes procedure: with so few points, it is always possible to find a Procrustean similarity transformation that brings the points into correspondence. That the recovery of the object point coordinates is rather good, often even for the random case, can be inferred from the configurations with the highest ($R^2 = 0.98$, $C1 = 0.91$) and lowest ($R^2 = 0.07$, $C1 = 0.50$) R^2 value (see Figure 7). What the boxplots show, is that λ has a strong influence on the recovery of the true data, especially with respect to the recovery of the rankings. When $\lambda \geq 1$, both $C1$ and R^2 are larger than .70 and with more than 5 objects even larger than .80 which indicates a good recovery of the underlying structure.

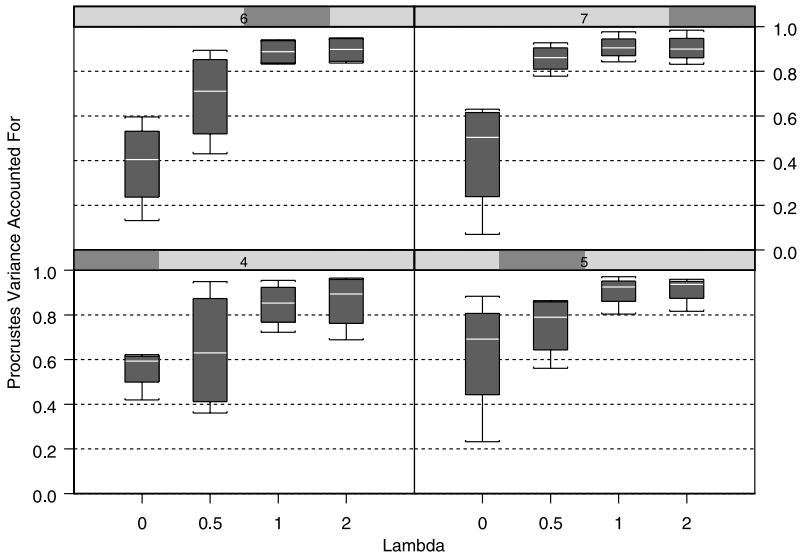


FIGURE 5 Box plots for the derived Procrustes statistic in function of the dispersion. The different panels refer to a different number of objects.

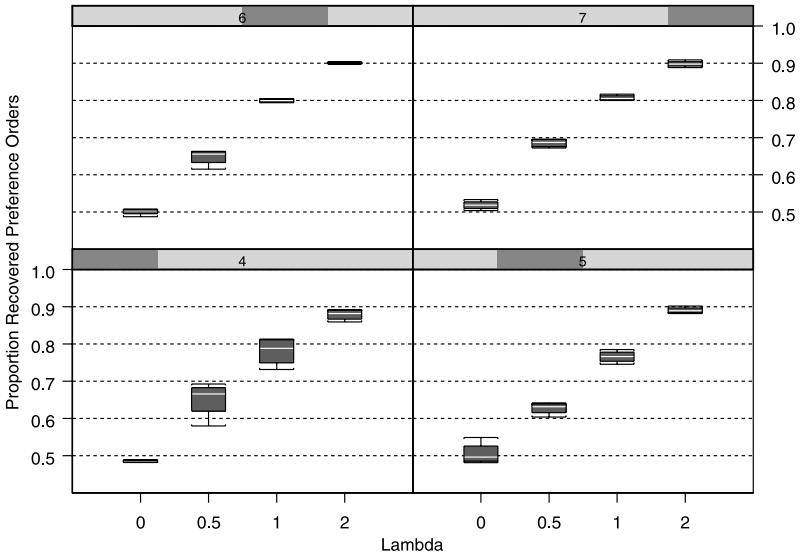


FIGURE 6 Box plots for the proportion of recovered preference orders in function of the dispersion and the number of objects.

The influence of the number of objects, the dispersion level, and of the data set on the proportion of recovered preference orders and on the Procrustes statistic, was evaluated with an analysis of variance (ANOVA) in which we treated the data set as a random factor and the number of objects and the dispersion level as fixed factors. Details are reported for index $C1$ whereas only the effect size, ω^2, is reported for the Procrustes statistic (see Table 1). Two effects reached

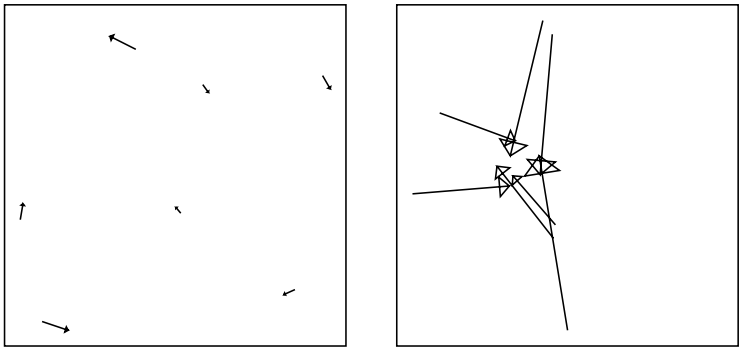


FIGURE 7 Solutions with the highest (left panel) and lowest (right panel) value for the Procrustes statistic. Each arrow is drawn from the true object point to the solution point.

TABLE 1
Analysis of Variance for the Proportion of Recovered Preference Orders

Source	*df*	*MS*	*F*	ω^2_{C1}	$\omega^2_{R^2}$
Between data sets					
O	3	0.0031	6.27**		
R(O)	12	0.0005	1.25		0.03
Within data sets					
L	3	0.4562	1152.82**	0.97	0.51
O*L	9	0.0006	1.49		
Residual	36	0.0004			

Note. O represents the number of objects, L the dispersion level and R the data set where R(O) indicates the nested structure. ω^2 is reported for values of at least .01.

* $p < .05$. ** $p < .01$.

the significance level of .01, these are the dispersion level and the number of objects: the proportion of recovered preference orders increases in an almost linear way for increasing λ and increasing object size (see Figure 6). The large influence of λ is confirmed by ω^2: the dispersion level accounted for 97% of the variance. Note that $C1$ correlates almost perfectly with the Euclidean distance, between the true and reproduced rankings, normalized by the maximum possible distance between the two rankings ($r = -.9975$). Regarding the derived Procrustes statistic R^2, only about 50% of the variance is explained which can be explained by the fact that the measure only takes the object points into account and is therefore not very informative. Both measures of recovery have a high correlation ($r = .74$) so that the configurations that recover the rankings well are also the configurations that recover the positions of the objects well.

Simulation Study 2

The second simulation study was designed from the viewpoint that preference is determined by only a few attributes (Coombs, 1964): both the judges and the objects are positioned on these, representing respectively the subjective ideal and the commonly perceived object's value. Our main interest was in the recovery of these positions by our proposed unfolding procedure and in the factors which influence the recovery. Second, we were also interested in the recovery of the rankings. As influencing factors, we considered the amount of error in the data, the dimensionality of the true configuration, and the size of the data which depends on the number of objects and the number of judges. A better reproduction can be expected under lower levels of noise and more constraints which corresponds here with a larger number of judges and objects and a lower dimensionality of the configuration.

The data were generated in the following way: the true configuration was generated using a uniform distribution on [−1,1]. Then noise was added under a lognormal distribution by multiplying the distances with the exponential of a random number generated under the normal distribution with mean zero and variance equal to the error percentage (see Ramsay, 1977). These perturbed distances were then used to derive the preference rankings which served as the input for our unfolding technique. By combining an error level (5 or 25%), a number of objects (10 or 20), the ratio of subjects to objects (1, 2 or, 5) and, the dimensionality (2, 3 or, 4 dimensions), 36 conditions were formed and per condition 20 data sets were generated. Here too, we used the SMACOF algorithm for the MDS analysis. In this case, a solution was found based on multistart procedure with 50 random starts for each data set and with 1,000 iterations per random start.

A first impression of the results can be obtained from the boxplots in Figures 8 and 9. The first boxplot summarizes the values found for the derived Procrustes statistic in function of the dimensionality of the configuration, the error level, and the number of objects whereas the second does the same for the proportion of recovered preference orders. Both measures indicate a good fit of the solution configuration to the true configuration: the minimal values are .40 for R^2 and .70 for $C1$ whereas most values are above .80 for both measures. Unless the error level and dimensionality are both high, the solutions obtained by our proposed algorithm seem to recover the true configuration well both with respect to the position of the points and the preference orders. As it was the case for the previous simulation study, the measures have a high correlation ($r = .80$).

In this simulation study, we were mainly interested in the recovery of the configuration coordinates and only secondary in the recovery of the rankings. Therefore, Table 2 presents details of an ANOVA on the derived Procrustes statistic R^2 and only effect sizes for index $C1$. All factors were treated as fixed. At a significance level of .01, all main effects and a number of interaction effects were significant. In total, about 71% of the variance is accounted for. The largest effect is found for the dimensionality which accounts for 29% of the variance. The error level and number of objects account both for about 15% whereas the number of subjects and the interaction between the error level and the dimensionality of the configuration both account for about 5%. As expected, the configuration was better reproduced in case of lower levels of error and more constraints. Note that the effect of the number of subjects is quite small in comparison to the effect of the number of objects which is as large as the effect of the error level. Turning to the ω^2 values for $C1$, we see that in total about 87% of the variance is accounted for with a contribution of 52% by the error level alone. Other important effects are the dimensionality (15%) and the interaction between the dimensionality and the error level (9%).

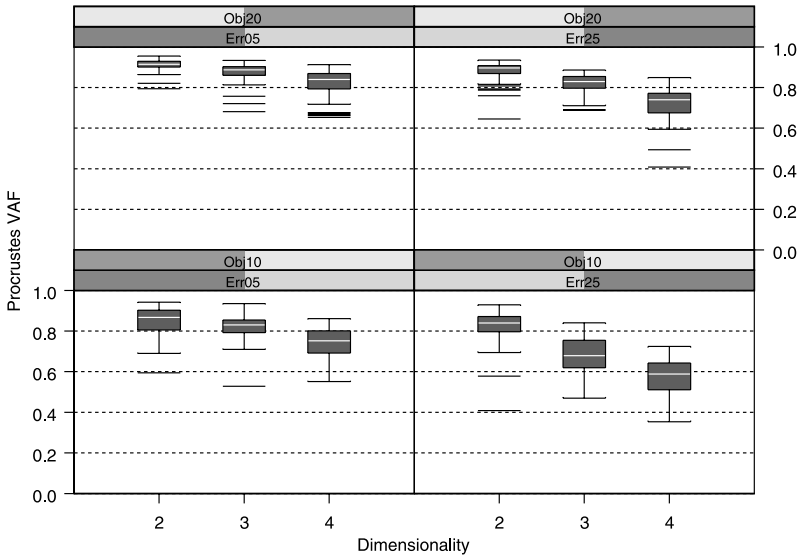


FIGURE 8 Box plots for the derived Procrustes statistic in function of the dimensionality. The different panels refer to different error levels and a different number of objects.

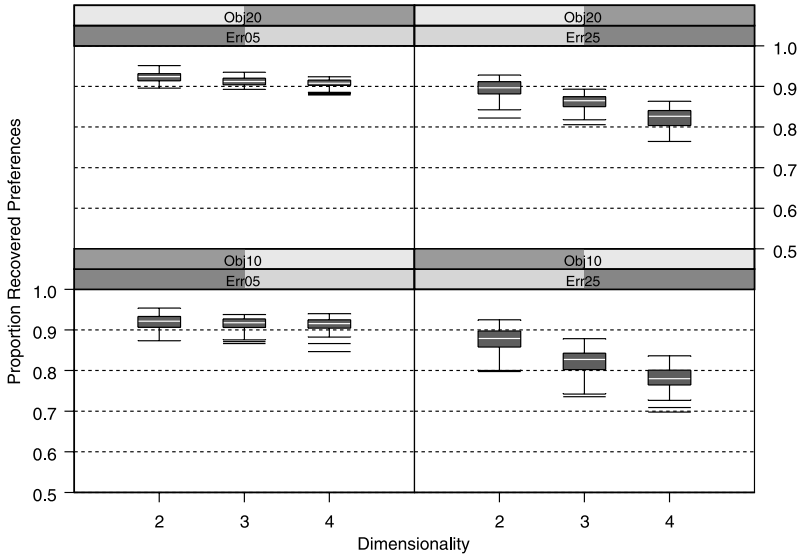


FIGURE 9 Box plots for the proportion of recovered preference orders in function of the dimensionality. The different panels refer to different error levels and a different number of objects.

TABLE 2
Analysis of Variance for R^2 and Effect Sizes for R^2 and C1

Source	*df*	*MS*	*F*	$\omega^2_{R^2}$	ω^2_{C1}
O	1	1.446	391.06**	0.16	0.03
S	2	0.227	61.48**	0.05	0.05
E	1	1.368	370.22**	0.15	0.52
D	2	1.340	362.38**	0.29	0.15
O*S	2	0.000	0.10		
O*E	1	0.099	26.82**	0.01	0.03
E*S	2	0.006	1.56		
O*D	2	0.043	11.65**		
S*D	4	0.006	1.74		
E*D	2	0.182	49.11**	0.04	0.09
E*O*D	2	0.032	8.61**		
E*O*S	2	0.006	1.48		
O*D*S	4	0.027	7.43**	0.01	
E*S*D	4	0.005	1.26		
O*E*D*S	4	0.006	1.48		
Residual	684	0.004			

Note. E represents the error level, O the number of objects, D the dimensionality and S the size of the subjects-to-objects ratio. ω^2 is reported for values of at least .01.

* $p < .05$. ** $p < .01$.

Both simulation studies show that our proposed unfolding procedure recovers both the low-dimensional configuration and the preference orders well when the error level is not too high and when there are enough constraints (more specifically, for low dimensionality and a large number of objects). The recovery of the coordinates of the configuration depends largely on the amount of constraints whereas the recovery of the preference orders is highly dependent on the error level. Note that these results are generalizable, because they were obtained under two different data generating mechanisms, each corresponding to a different hypothetical low-dimensional structure for preference data.

APPLICATIONS

In this section we illustrate our unfolding technique for two classical examples in the unfolding literature: the breakfast data of Green and Rao (1972) and data on preference for different family structures (Delbeke, 1968). A comparison is made with PREFSCAL, currently the only program that guarantees non-degenerate solutions for ordinal unfolding (see Busing, Groenen, & Heiser, 2005). First we give a short description of PREFSCAL and we illustrate the differences in performance with our approach through the illustrations. Occasional limitations

to our approach and how these can be circumvented are illustrated with the data on preference for family structures.

PREFSCAL

PREFSCAL (Busing et al., 2005) is an algorithm for multidimensional unfolding, both metric and nonmetric, that avoids degenerate solutions by using a penalized approach to the Stress function. The key in this algorithm is that it minimizes a loss function that includes a penalty component. The penalty is an inverse function of the coefficient of variation of the disparities, such that it has a high value when the variation is small. Strong penalties lead to transformations with equal increments, which corresponds to metric unfolding and has the consequence that Stress barely improves. Note that the function that is currently minimized in the PREFSCAL program is an adapted form of the penalized Stress proposed in Busing et al. (2005).

Some differences between PREFSCAL and our approach are illustrated. It becomes clear that our method outperforms PREFSCAL when the latter is used with default values for the penalty parameters. Setting stronger penalties will solve the degeneracies but it will also drive the solution to a metric one and result in higher Stress values. Our method, on the other hand, is straightforward in use and not hampered by the subjective issue of picking the "right" values. We also illustrate the difference in interpretation of the configurations: our approach allows not only for relative but also for absolute statements.

Breakfast Data

Green and Rao (1972) asked 21 MBA students and their wives to order 15 breakfast items according to their preference. As a first step in our approach to unfolding, the partial rankings representing the objects (see the definition in (2)) were added as 15 additional rows to the preference data, all data were centered, and the rows representing the objects were scaled such that they have the same length as the preference rankings (as described previously). From the resulting $(42+15)\times 15$ augmented data matrix, a 57×57 matrix of dissimilarities was derived by calculating the squared Euclidean distances between all pairs of points. It is this matrix that served as the input for a nonmetric multidimensional scaling analysis. To illustrate the flexibility of our unfolding approach, we used several programs for MDS: the SMACOF algorithm (see De Leeuw & Heiser, 1980) that minimizes normalized Stress, KYST with the option of minimizing Stress-2 and the MDS procedure implemented in the SAS package with the option of minimizing S-Stress. Note that KYST can be downloaded for free and that the SAS analysis is similar to an ALSCAL analysis (see the chapter on the MDS procedure in SAS Institute Inc., 1999). These results are compared with

TABLE 3
Some Goodness of Fit Measures for the Four Different Algorithmic Solutions of the Breakfast Data

Measure	*SMACOF*	*PREFSCAL*	*KYST*	*ALSCAL*
Av. recovered preference	0.7855	0.7785	0.7871	0.7762
Av. Pearson correlation	0.7422	0.7418	0.7500	0.7117
Av. Spearman correlation	0.7329	0.7283	0.7423	0.7164
Raw normalized stress	0.0298	0.0355	0.0306	0.0367
Kruskal's Stress1	0.1759	0.1910	0.1779	0.2037
Kruskal's Stress2	0.4895	0.4500	0.4767	0.5121
S-Stress1	0.2647	0.2937	0.2691	0.2724

a PREFSCAL analysis where the penalty parameters were set equal to their default values. Note that PREFSCAL minimizes raw normalized Stress when no penalty is set. To account for the problem of local minima in MDS, we used a multistart procedure with 250 random starts in SMACOF and PREFSCAL. As this type of initialization is not straightforward in KYST and SAS, we opted for the available rational start in these cases. Strong convergence criteria were used to ensure that the solutions are stable. Busing et al. (2005) performed unfolding analyses of the same data with several algorithms, including KYST and ALSCAL: even when using rational starts, they found degenerate solutions.

The final configurations[4] are depicted in Figure 10. In all cases, an interpretable structure appears: the horizontal dimension distinguishes between the hard (at the left) and the soft items, while the vertical dimension distinguishes the buttered items (bottom) from those without butter or margarine (see also Heiser & Busing, 2004 for a detailed analysis of these data). Note that the PREFSCAL solution and the ALSCAL solution show a tendency to clustering. Some measures of fit of the solutions to the original preference data are given in Table 3. The first three measures indicate how well the distances fit the original preference orders, whereas Stress measures the lack-of-fit of the data to the monotonically regressed distances. We calculated each measure per judge and an overall measure is obtained either by averaging or, for the last three Stress values, by taking the root mean square Stress. The SMACOF and KYST solutions have the highest fit on all values except Stress-2. However, the differences between the different solutions are rather small. Except for Stress-2, the PREFSCAL and ALSCAL solutions have somewhat lower fit and higher Stress. An important observation is that—at least with this example that is well-known in the literature for its susceptibility for degenerate solutions—the different solutions based

[4]The configurations were drawn with equal vertical and horizontal scales and with a range equal to the maximal range of either the vertical or the horizontal dimension.

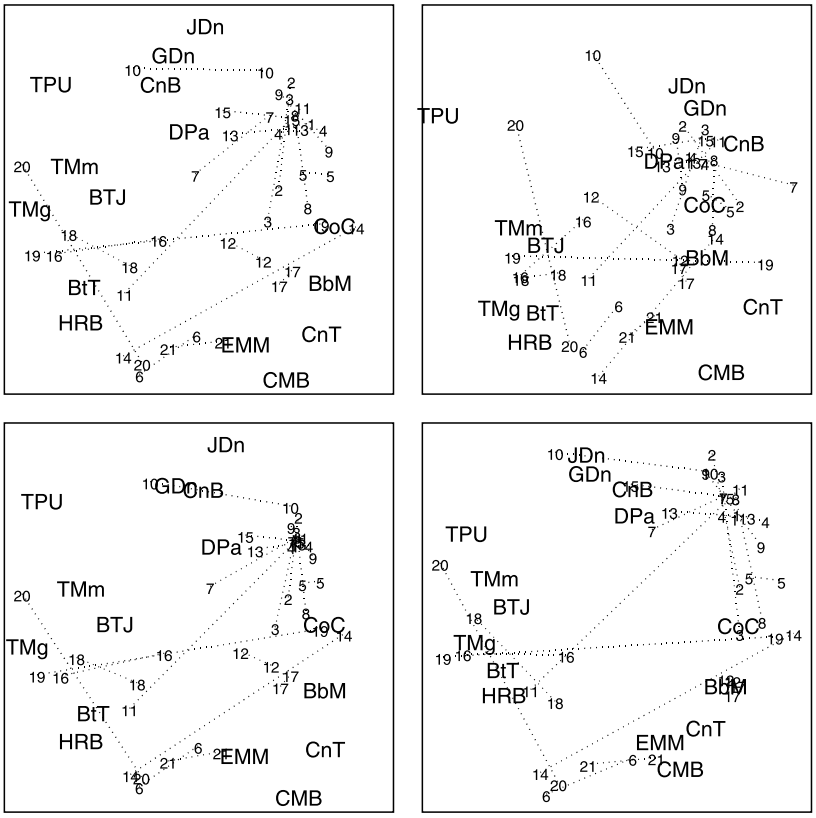


FIGURE 10 Unfolding solutions for the breakfast data: SMACOF on complete data (top left), PREFSCAL solution (top right), KYST with Stress-2 on complete data (bottom left), and ALSCAL on complete data (bottom right). The couples are labelled by the numbers; the breakfast items are labelled by the three letter codes: hard rolls and butter (HRB), buttered toast (BtT), corn muffin and butter (CMB), English muffin and margarine (EMM), cinnamon toast (CnT), blueberry muffin and margarine (BbM), coffee cake (CoC), toast and margarine (TMg), buttered toast and jelly (BTJ), toast and marmalade (TMm), Danish pastry (DPa), toast pop up (TPU), cinnamon bun (CnB), glazed donut (GDn), and jelly donut (JDn).

on the completed MDS matrix, are very similar, which makes the approach truly flexible and at the same time robust over different algorithmic solutions. Furthermore, the configurations are nondegenerate, are well interpretable, and fit the preference data well. In this case, our approach outperformed PREFSCAL. With stronger penalties, configurations are obtained that are very similar to the configurations obtained with our unfolding technique but with higher Stress and with transformation functions that are nearly linear.

An advantage in interpretation of our approach, is that it allows for absolute interpretations of the inter-rater agreement as measured by the Spearman distances. This possibility is due to the fact that the dissimilarities are all comparable and subjected to an unconditional analysis: the average distance among the preference items corresponds approximately to orthogonal vectors and this knowledge can be used as a benchmark. Note that classic unfolding programs only fit the between-set distances such that the distances among points of one set can only be interpreted in a relative way. We make this issue concrete with the breakfast data by taking a closer look to the distances among the subject points in Figure 10: the couples are labeled and their members are joined by a dotted line. Most husbands and wives seem to have more similar preference ratings than random pairs. This finding was confirmed by a measure of diversity of Feigin and Alvo (1986): the proportion of diversity in rankings within the couples amounts to 36% whereas it amounts to 74% between the couples. With the permutation test we found a p value (based on 5000 samples) of .0002 under the null hypothesis of complete agreement between the couples. Green and Rao came to the same conclusion with a sign test on the distances (see Green & Rao, 1972, p. 63). So far, this is a relative interpretation because we only compared distances among judges. That there is a strong agreement between the members of a couple, which is an absolute interpretation, follows from the fact that the distances among couple points are on average considerably shorter than among breakfast points. In fact, the average rank correlation among couples equals 0.40.

Delbeke Data

When discussing the choice of the length of the object vectors, the importance of having the same range of distances in the different blocks of the dissimilarity matrix was stressed. In order to avoid degenerate solutions this length was set equal to the norm of the rank vectors. However, in case there is considerable agreement among the different rankings, the distances among judges will be smaller than the distances in the preference and object block and this characteristic will lead to solutions where the two sets of points are separated from each other. While testing our approach to different empirical data sets, this situation occurred for the Delbeke (1968) data on preference for family compositions (see the left panel of Figure 11). This configuration is not degenerate but lacks *intermixedness* (see Kim, Rangaswamy, & DeSarbo, 1999). Furthermore, the structure of the family compositions is somewhat distorted: in Figure 11 lines are drawn that represent the total number of children (e.g., the line 02-11-20 indicating two children) and that represent the sex-bias (e.g., the line 01-12-23 indicating one girl more than boys). One would expect a structure like the one in the right panel. Nevertheless, as can be seen in the column labelled UNFOLD in Table 4, the solution fits the data well and is informative: the family items

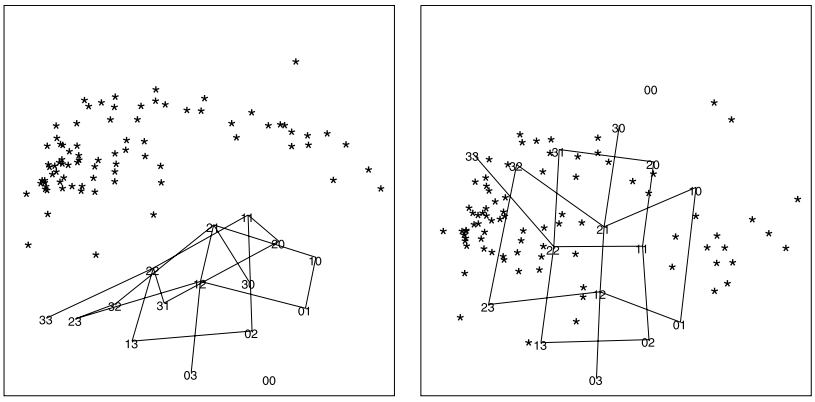


FIGURE 11 Unfolding configurations for the Delbeke data. The stars represent the judges, the labelled point the family structures with the first number indicating the number of boys and the second the number of girls. In the right panel a constant was subtracted from the dissimilarities in the preference block. (a) Theoretical representations. (b) Non-metric MDS solution.

are ordered from left to right by the total number of children and the bottom items consist of families with at least two girls more than boys (except for the item 00). It is also clear that most subjects prefer large families and that the families with more girls are unpopular. Knowing that the cosine of the angle between a pair of objects equals $-1(m-1)^{-1}$, the average distance between the family items corresponds to an angle of 94 degrees. Compare this distance with the distance among most of the judges: it is clear then, that most of the judges agree considerably. Note that subjecting these data to a PREFSCAL analysis with default penalty parameters results in a truly degenerate solution: Except for the items 00 and 33, all other items fall together in one point (not shown). Although Stress is very low, the distances do not fit the data very well (see the column labeled PREFSC1 of Table 4).

TABLE 4
Some Goodness of Fit Measures for the Unfolding, the Intermixed Unfolding, and the Prefscal Solutions of the Delbeke Data. The First Prefscal Solution Is Obtained with Default Values for the Penalty Parameters, the Second with a Stronger Penalty

Measure	*UNFOLD*	*INTERM*	*PREFSC1*	*PREFSC2*
Av. recovered preference	0.8605	0.8603	0.7498	0.8860
Av. Pearson correlation	0.8688	0.8651	0.5133	0.9086
Av. Spearman correlation	0.8629	0.8664	0.6280	0.9088
Raw normalized stress	0.0075	0.0164	0.0011	0.0129

In case an intermixed configuration is desired, this can be obtained by an additional step in our unfolding technique but at the cost of loosing some information. Once the first four steps (the Distances Involving the Objects section) are performed, we force the mean of the preference dissimilarities to be equal to the mean value of the subject block by a simple trick: a constant c is subtracted from the dissimilarities in the preference block, where we chose

$$c = \frac{\sum_{i,k} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2}{nm} - \frac{\sum_{i<j} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_j)^2}{2^{-1}n(n-1)}. \tag{9}$$

In this way the mean dissimilarity in the preference block equals the mean dissimilarity in the judges block but the dissimilarities in different blocks are no longer comparable. As before, the resulting dissimilarity matrix is subjected to an unconditional MDS analysis. Applying our extended unfolding technique to the data on family preference, results in the configuration in the right panel of Figure 11. Now, the configuration is intermixed and the items are nicely structured though less informative: where all distances in the left panel of Figure 11 are comparable, this is in the right panel only the case for distances in the same block. So, the information that there is a strong agreement between most of the judges is lost in the intermixed configuration. The preference data are still well fitted (see the column labeled Interm in Table 4). A very similar configuration is obtained with PREFSCAL when using strong penalties. As can be seen in the column PREFSC2 of Table 4, this solution fits the data very well but has higher Stress than the degenerate solution and the nonintermixed solution obtained with our approach.

The idea of obtaining intermixed configurations by subtracting a constant from the $d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2$ is connected with two other approaches to unfolding that try to avoid degenerate solutions. PREFSCAL, which we already discussed, avoids degenerate solutions by making solutions with a large additive constant unattractive. Kim et al. (1999) proposed to force intermixed configurations by pulling preferred objects close to the subject while dispreferred objects are pushed away from the subject. This is done for each subject. We countered the occurrence of such a constant by subtracting it. Although the idea of subtracting a constant as defined in (9) seems to work well, more research is needed to evaluate this approach and to compare it with similar ones based on, for example, another constant.

DISCUSSION

Two issues discussed in this article should stand out. First, we introduced an alternative geometric representation of ranking data (the so-called permutation

polytope) that clearly shows that ranking data contain more information than what is used in MDU: some rankings are closer or, more similar, than others and poles of attraction can be defined that are strongly associated to the objects. Second, we showed that classical MDU does not work as it results in degenerate solutions: The idea of doing a row-conditional MDS analysis of a matrix with a lot of missing values gives too much freedom and more constraints are needed in order to obtain interpretable configurations. Here, we chose to impose more constraints by completing the MDS matrix on the basis of information that can be derived directly from the preference data after representing these data in the preference sphere. As we showed, alternative ways of analyzing completed matrices by using a block conditional approach result in degenerate solutions. This leads us to the conclusion that the derived MDS matrix should consist of all comparable dissimilarities. Moreover, for the same reason of avoiding degeneracies, the range of values should be the same for the different blocks.

The permutation polytope is a high-dimensional curved structure and our objective is to obtain a low-dimensional representation of the polytope and the attraction poles. A nonlinear mapping is needed then, and this can be accomplished by a nonmetric MDS of the distance relations on the polytope. For reasons discussed in this article, these dissimilarities were calculated using Spearman distances on the preference sphere. As these describe the polytope and attraction poles well, and in a way that constrains the MDS configuration sufficiently, well-fitting and nondegenerate configurations are obtained.

In two simulation studies, we evaluated the recovery of given preference structures by our proposed unfolding technique under different conditions. Both simulation studies indicated (a) that the coordinates are recovered well when the error level is not too high and when there are enough constraints and (b) that the rankings are recovered well when the error level is not too high. Overall, the proposed procedure performed well.

The performance of our technique was also illustrated for real data and tested against PREFSCAL. First we analyzed the breakfast data (Green & Rao, 1972) and compared the results with those obtained by PREFSCAL with default penalty parameters: the proposed procedure resulted in a better structured configuration, gave slightly better measures of fit to the data, and had lower Stress. We also illustrated the interpretational advantage of our approach. Application of the proposed unfolding procedure to the data on preference for family structures (Delbeke, 1968), showed that in case of relative homogeneity of preference rankings, solutions are obtained where the two sets of points are separated in the configuration or, in case of PREFSCAL with default penalties, are truly degenerate. We illustrated how an intermixed configuration can be enforced by our approach: the resulting configuration fits the preference rankings well but loses the unconditional character of the interpretation. However, further research is needed: for example, we do not know how homogeneous the rankings should be

to obtain nonintermixed or even degenerate solutions and we need to evaluate if the subtraction of a constant is sufficient to overcome this problem. From these two illustrations, it seems that our approach avoids degenerate solutions in a straightforward manner, whereas the default penalties of PREFSCAL are not degeneracy proof. Furthermore, Stress was higher for the nondegenerate PREFSCAL solutions than for the solutions found with our approach, unless we forced intermixedness.

The proposed approach to multidimensional unfolding has practical relevance because it is an unfolding technique that gives nondegenerate solutions and that can be implemented in an extremely easy way: the calculation of the dissimilarities is straightforward and once they are available, any program for nonmetric multidimensional scaling of a square symmetric matrix is in principle suitable to proceed with the analysis in order to obtain degeneracy-free solutions for unfolding data.

ACKNOWLEDGMENTS

We thank the editor and three reviewers for their valuable comments and suggestions to improve the quality of this work.

REFERENCES

Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determinining the dimensionality of ranked preference data. *Psychometrika*, *25*, 27–43.

Borg, I., & Groenen, P. J. F. (1997). *Modern multidimensional scaling*. New York: Springer-Verlag.

Busing, F. M. T. A., Groenen, P. J. F., & Heiser, W. J. (2005). Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. *Psychometrika*, *70*, 49–76.

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

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

Cox, T. F., & Cox, M. A. A. (1994). *Multidimensional scaling*. London: Chapman and Hall.

Critchlow, D. E. (1980). *Metric methods for analyzing partially ranked data*. Heidelberg: Springer-Verlag.

Delbeke, L. (1968). *Construction of preference spaces: an investigation into the applicability of multidimensional scaling models*. Leuven, Belgium: Leuven University Press.

De Leeuw, J., & Heiser, W. J. (1980). Multidimensional scaling with restrictions on the configuration. In P. R. Krishnaiah (Ed.), *Multivariate analysis* (Vol. 5, pp. 501B522). Amsterdam, The Netherlands: North-Holland Publishing.

Diaconis, P. (1988). *Group representations in probability and statistics*. Hayward, CA: Institute of Mathematical Statistics.

Feigin, F. D., & Alvo, M. (1986). Intergroup diversity and concordance for ranking data: An approach via metrics for permutations. *The Annals of Statistics, 14,* 691–707.

Green, P. E., & Rao, V. R. (1972). *Applied multidimensional scaling: A comparison of approaches and algorithms.* New York: Holt, Rinehart & Winston.

Groenen, P. J. F. (1993). *The majorization approach to multidimensional scaling: Some problems and extensions.* Leiden: DSWO Press.

Heiser, W. J. (1999). *Ideaalpunten in de Kemeny ruimte*. Unpublished manuscript.

Heiser, W. J. (2004). Geometric representation of association between categories. *Psychometrika*, *69*, 513–546.

Heiser, W. J., & Busing, F. M. T. A. (2004). Multidimensional scaling and unfolding of symmetric and asymmetric proximity relations. In D. Kaplan (Ed.), *The SAGE handbook of quantitative methodology for the social sciences* (pp. 25–48). Thousand Oaks, CA: Sage.

Kemeny, J. G., & Snell, J. L. (1962). *Mathematical models in the social sciences*. MA: Blaisdell Waltham.

Kim, C., Rangaswamy, A., & DeSarbo, W. S. (1999). A quasi-metric approach to multidimensional unfolding for reducing the occurrence of degenerate solutions. *Multivariate Behavioral Research, 34*, 143–180.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association, 53,* 814–861.

Lingoes, J. C. (1977). A general survey of the Guttman-Lingoes nonmetric program series. In J. C. Lingoes (Ed.), *Geometric representations of relational data* (pp. 263–282). Ann Arbor, MI: Mathesis Press.

Mallows, C. L. (1957). Non-null ranking models: I. *Biometrika*, *44*, 114–130.

Marden, J. I. (1995). *Analyzing and modeling rank data.* London: Chapman and Hall.

MATLAB. (2002). *MATLAB: The language of technical computing, release 13.* The MathWorks, Inc.

Ramsay, J. O. (1977). Maximum likelihood estimation in multidimensional scaling. *Psychometrika, 42,* 241–266.

SAS Institute Inc. (1999). *SAS/STAT User's guide*. Cary, NC: SAS Institute Inc.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika*, *27*, 219–246.

Spence, I., & Domoney, D. W. (1974). Single subject incomplete designs for nonmetric multidimensional scaling. *Psychometrika*, *39*, 469–490.

Steverink, H. M., Heiser, W. J., & van der Kloot, W. A. (2002). *Avoiding degenerate solutions in multidimensional unfolding by using additional distance information.* Unpublished manuscript.

Van Blokland-Vogelesang, A. W. (1989). Unfolding and consensus ranking: a prestige ladder for technical occupations. In G. De Soete, H. Feger, & K. C. Klauer (Eds.), *New developments in psychological choice modelling* (pp. 237–258). Amsterdam: North-Holland.

Van Deun, K., Groenen, P. J. F., Heiser, W. J., Busing, F. M. T. A., & Delbeke, L. (2005). Interpreting degenerate solutions in unfolding by use of the vector model and the compensatory distance model. *Psychometrika*, *70*, 23–47.

APPENDIX

Proof of Preference Order Preservation

Here, we prove that the distance between the vectors $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_k$ preserves the order of the preference scores by showing that when judge i prefers object k to object k', the distance from $\boldsymbol{\pi}_i$ to $\boldsymbol{\pi}_k$ will be smaller than from $\boldsymbol{\pi}_i$ to $\boldsymbol{\pi}_{k'}$. Recall that $\boldsymbol{\pi}_k$ is some permutation of $[1 \;\; 1 + m/2 \ldots \;\; 1 + m/2]'$ and $\boldsymbol{\pi}_i$ of

$[1\ \ 2\ \ 3\ \ \ldots\ \ m]'$. When judge i ranks object k on the l-th position,

$$\begin{aligned} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2 = &\left[1 - \left(\frac{m}{2} + 1\right)\right]^2 + \left[2 - \left(\frac{m}{2} + 1\right)\right]^2 \\ &+ \cdots + \left[(l-1) - \left(\frac{m}{2} + 1\right)\right]^2 + [l-1]^2 \\ &+ \cdots + \left[(l+n) - \left(\frac{m}{2} + 1\right)\right]^2 \\ &+ \cdots + \left[m - \left(\frac{m}{2} + 1\right)\right]^2. \end{aligned} \tag{10}$$

Suppose judge i ranks object k' on the $(l+n)$-th position, then

$$\begin{aligned} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_{k'})^2 = &\left[1 - \left(\frac{m}{2} + 1\right)\right]^2 + \left[2 - \left(\frac{m}{2} + 1\right)\right]^2 \\ &+ \cdots + \left[(l-1) - \left(\frac{m}{2} + 1\right)\right]^2 + \left[l - \left(\frac{m}{2} + 1\right)\right]^2 \\ &+ \cdots +_{[}\, (l+n) - 1]^2 + \ldots + \left[m - \left(\frac{m}{2} + 1\right)\right]^2. \end{aligned} \tag{11}$$

Subtracting (10) from (11) gives

$$\begin{aligned} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_{k'})^2 - d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2 &= \left[l - \left(\frac{m}{2} + 1\right)\right]^2 + [(l+n) - 1]^2 \\ &\quad - [l-1]^2 - \left[(l+n) - \left(\frac{m}{2} + 1\right)\right]^2 \\ &= l^2 - 2l\left(\frac{m}{2} + 1\right) + \left(\frac{m}{2} + 1\right)^2 + (l+n)^2 \\ &\quad - 2(l+n) + 1 - l^2 + 2l - 1 - (l+n)^2 \\ &\quad + 2(l+n)\left(\frac{m}{2} + 1\right) = \left(\frac{m}{2} + 1\right)^2 \\ &= nm, \end{aligned} \tag{12}$$

with $nm > 0$ so $d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_{k'})^2 > d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2$. This proves that representing the objects by tied ranks, in combination with the Spearman metric, preserves the preference orders.

In the proof, the object vectors have all equal length. If we allow for different lengths, $\|\boldsymbol{\pi}_k\| = \gamma_k$ and $\|\boldsymbol{\pi}_{k'}\| = \gamma_{k'}$, the following result is obtained:

$$\begin{aligned} d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_{k'})^2 - d(\boldsymbol{\pi}_i, \boldsymbol{\pi}_k)^2 = (\gamma_{k'}^2 - \gamma_k^2)\left[\frac{m(m-1)}{4}\right] \\ + (\gamma_{k'} - \gamma_k)\left[m\left(l - \frac{m+1}{2}\right)\right] \\ + nm\gamma_{k'}. \end{aligned} \tag{13}$$

The preference orders are preserved if (13) is positive for all m, l, n, k, k', and i which generally holds for $\gamma_k = \gamma_{k'}$ but not when the object vectors have different lengths.