Multidimensional Scaling Models and Methods for
Two-Way and Three-Way Proximity and Preference Data

J. Douglas Carroll
AT&T Bell Laboratories

In its broadest definition, multidimensional scaling (MDS) includes a wide variety of geometric models for representation of psychological or other behavioral science data. Such models can include discrete geometric models such as tree structures (typically associated with hierarchical clustering) overlapping or nonoverlapping cluster structures, or other network models. More typically, however, MDS is associated with continuous spatial models for representation of data. In the broad definition of MDS, such spatial models can include such general geometric structures as what are called the vector model or the unfolding model for representation of individual differences preference (or other dominance) data or even the factor analysis model. For discussion of some of the models and associated methods included in this broad definition of MDS, see Carroll (1980), Carroll and Arabie (1980), and Shepard (1980).

For now, however, we shall focus on the most common meaning of MDS, and, indeed, focus even more narrowly on what would now be called, in the taxonomy proposed by Carroll and Arabie (1980), MDS models and methods for one-mode, two-way proximities data.

In this narrower meaning in which we now focus, MDS represents data on proximities (e.g., similarities, dissimilarities or other measures of association or "closeness") of pairs of stimuli or other objects in terms of a distance model. These proximity data can arise from direct human judgments of similarity or dissimilarity, from such data as confusability of pairs of stimuli, or from various types of derived measures of similarity or dissimilarity (e.g., a "profile dissimilarity" measure computed between stimuli over rating scales, or various measures of similarity derived from word association tasks). In some cases matrices of correlations can be used as measures of similarity (of variables, people, stimuli or other entities) so that MDS can, when used this way, be viewed as an alternative to factor analysis for deriving multidimensional structure from correlational data.

Once defined, the proximities can be thought of as comprising a two-way square table (or matrix) whose general entry is $d_{jk}$, the proximity of stimuli (or other objects) $j$ and $k$. This table will usually (but not always) be symmetric, since the proximity of $j$ to $k$, for many measures of proximity, will be the same as that of $k$ to $j$ (i.e., $d_{jk} = d_{kj}$). We call this a symmetric, one-mode, two-way proximities matrix. ("One-mode" means that both the rows and columns of the matrix correspond to the same set of objects or, put another way, the two subscripts $j$ and $k$ range over the same mode, or set of entities. "Two-way" means that, since $8$ has two subscripts, a two-way table is implied.)

The model assumed for these proximities can be expressed as follows:

$$F(b_{jk}) = d_{jk}$$

where $F$ is some function (e.g., linear, some specified nonlinear function, or in the case of nonmetric MDS, a merely monotonic function), $b_{jk}$ can be interpreted as meaning "approximately equals" (or equals, except for error terms which will not be further specified in this article) while $d_{jk}$ is the distance between points representing $j$ and $k$ in an $R$ dimensional space. The distance, $d_{jk}$, is usually, but not necessarily, assumed to be Euclidean distance, defined as follows:

$$d_{jk} = \sqrt{\sum_{r=1}^{R} (x_{jr} - x_{kr})^2}$$  (1)

where $x_{jr}$ is the coordinate of the point representing the $j$th stimulus or other object on the $r$th dimension, while $r$ ranges from 1 to $R$, the dimensionality of the space. The objective of MDS is, given the $8$'s to solve for the coordinates of the $R$ dimensional space. [Options for specific non-Euclidean distance functions exist in many of the modern programs for multidimensional scaling, but these non-Euclidean metrics in MDS are troublesome at the least, and certainly demand that very great care be taken. In general, use of such non-Euclidean options in MDS is not recommended for the unsophisticated user.]

An important distinction that is frequently made is between metric and non-metric approaches to MDS. Essentially, the distinction comes down to the form assumed for the function $F$ in Eq. (1) above. In the metric approaches, $F$ is generally assumed to be linear, although in some metric approaches $F$ is some well defined nonlinear function (e.g., exponential, polynomial of some specified degree, or a power function). Another way of stating this is that, in the metric approaches the proximities are assumed to be measured on at least an interval scale. In nonmetric MDS, the function $F$ is typically assumed only to be monotonic — generally increasing (or non-decreasing) if the $8$'s are dissimilarities and decreasing (or non-increasing) if they are similarities. Thus, nonmetric scaling assumes proximities measured only on an ordinal scale.

The "classical" metric approach to MDS, as described in Torgerson's (1958) book, can be described
in terms of a sequence of steps as follows (here assuming the data symmetric, with diagonals undefined and that $F$ is linear):

1. If the data are similarities, the values are first reversed (say by multiplying by minus one) to make them dissimilarities. If the original data are dissimilarities, they are left unchanged.

2. We shall henceforth assume the $d$s are dissimilarities (sometimes called "comparative distances"). These are converted to estimated distances, by estimating an additive constant by one of a number of schemes. Thus,

$$d_{jk} = b_{jk} + \hat{c}$$

where $\hat{c}$ is the estimated constant, and $d_{jk}$ the estimated distance between objects $j$ and $k$ (now measured, presumably, on a ratio scale). The simplest, and probably most often used, procedure for additive constant estimation simply defines $\hat{c}$ to be the smallest constant that guarantees that the $d$'s satisfy the triangle inequality. The triangle inequality, which states that $d_{ik} \leq d_{ij} + d_{jk}$ for all $i$, $j$, and $k$ (i.e., that the direct distance from $i$ to $k$ must be less than the indirect distance going through any intermediate point, $j$) is one of the metric axioms, which must hold for any distance measure.

3. The estimated distances are then converted into estimated scalar products by the formula

$$\hat{\delta}_{jk} = -\frac{1}{2}(d_{jk}^2 - \frac{1}{n} \sum_{i=1}^{n} d_{ik}^2 - \frac{1}{n} \sum_{j=1}^{n} d_{jk}^2 + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^2)$$

where

$$d_{jk}^2 = \frac{1}{n} \sum_{i=1}^{n} d_{ik}^2$$

$$d_{ik}^2 = \frac{1}{n} \sum_{j=1}^{n} d_{jk}^2$$

and

$$d_{ij}^2 = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{n} d_{jk}^2$$

One way to summarize this formula is to say that one first takes the matrix of estimated distances (with zeros in the diagonal) squares all the entries, multiplies the resulting matrix of squared distance estimates by $-1/2$, and then double centers that matrix to get the matrix of estimated scalar products. Double centering a matrix is equivalent to subtracting out row and column main effects in a two-way analysis of variance, leaving what are sometimes called "interaction numbers." The estimated scalar products ($\hat{\delta}_{jk}$) are simply interaction numbers calculated from the matrix whose general entry is $-1/2d_{jk}^2$. While the analogy is not exact, these estimated "scalar products" can be interpreted in many ways as analogous to variances ($\hat{\delta}_{jj}$) and covariances ($\hat{\delta}_{jk}$ for $j \neq k$) among a set of variables. Therefore, the remainder of the "classical" approach to metric MDS uses a step essentially equivalent to factor analysis (or, more precisely, principal components analysis) of a covariance matrix. The $R$ dimensional representation then is given by the first $R$ factors (or principal components) thus derived from this matrix of estimated scalar products.

Nonmetric MDS was first proposed by Coombs (1950, 1964). The essential assumption was that the proximities in most behavioral science applications define distances only up to an ordinal scale. Thus it is appropriate to assume $F$ to be a merely monotonic function, which we shall henceforth call $M$. The methods proposed by Coombs for implementing nonmetric MDS were heuristic, however, and not readily applicable to practical data analytic situations. The first effective algorithm for nonmetric MDS was proposed by Shepard (1962a,b) and called, by him, "analysis of proximities." Kruskal (1964a,b) then proposed a mathematically more rigorous and numerically more efficient approach, which has become the prototype of all subsequent approaches to nonmetric MDS; e.g., those of Guttman (1968), Mcgee (1966), Roskam and Lingoes (1970), or the "two-way" case of the approach of Takkunen, Young and De Leeuw (1977). Although it is a metric, rather than a nonmetric approach, and is based on a rather different criterion of fit than the others, Ramsey's (1977) maximum likelihood approach also utilizes a numerical algorithm that shares much in common with the approach first taken by Kruskal. While the details of the approach are too complex to describe adequately here, suffice it to say that Kruskal proposed using an explicit numerical method (the method of gradients, or steepest descent) to solve simultaneously for the matrix of coordinates $x_{jr}$ of the $n$ points in $R$ dimensions ($j = 1, 2, .., n; r = 1, 2, .., R$), and for the monotonic function $M$ (nonincreasing or nondecreasing, depending on whether $b$ was a dissimilarity or a similarity measure) so that $M$ and $x_{jr}$ ($j = 1, 2, .., n; r = 1, 2, .., R$) together optimize (minimize) a measure Kruskal called STRESS. STRESS (in its original, and still most popular, definition) is defined as:

$$\text{STRESS} = \sqrt{\frac{\sum_{j=1}^{n} \sum_{k=1}^{n} (d_{jk}^2 - \hat{d}_{jk}^2)^2}{\sum_{j=1}^{n} \sum_{k=1}^{n} \hat{d}_{jk}^2}}$$

where

$$\hat{d}_{jk} = M(b_{jk}) .$$

(Later a second definition of STRESS, sometimes called STRESSFORM2, was introduced as an option, in which a different normalizing factor was used to define the denominator.) This general approach has
also been adopted by Kruskal and others to provide alternatives to the classical approach to metric MDS described earlier. Essentially, this just replaces the function \( M \) by a function \( F \) that is linear or of some other specified form. Finally, to illustrate the use of MDS, we borrow a new classical illustration from Shepard (1962), involving analysis of some data on confusions of Morse code symbols due to Rothkopf (1957). The confusability matrix for the thirty-six Morse code symbols is shown in Fig. 1. Note that this

![Confusability Matrix](image)

**FIGURE 1.**

Matrix of confusabilities of 36 Morse code signals (data from Rothkopf, 1957).


is a nonsymmetric proximities matrix, since, for example, the proportion of time that a "same" judgment was made when the signal for \( A \) precedes that for \( I \) (46%) is not equal to that when the signal for \( I \) precedes that for \( A \) (64%). That is, there are systematic order effects in these "same-different" judgments. Also, the diagonal terms are not absent, as is typically the case with directly judged similarities or dissimilarities. Since confusabilities are similarity measures (i.e., presumed to relate by a nonincreasing monotonic function to distances) we would expect that the diagonals would generally be larger than the off-diagonal terms — which they are (in the high 80's or 90's, but not 100% as they would be if the subjects were performing with perfect accuracy). This example illustrates that modern approaches to MDS can be applied even to quite general types of data (not only to simple symmetric matrices of similarities or dissimilarities). In this case, the nonsymmetry and presence of diagonals were both handled quite straightforwardly — by simply defining the STRESS measure by summing over all \( j \) and \( k \) (not just for \( j < k \), as would be done if the matrix were a symmetric matrix without diagonals). Some more recent approaches attempt to account systematically for such asymmetries (see Carroll and Arabie, 1980, for a discussion of some of these).

Shepard applied nonmetric scaling to this matrix, and found that two dimensions provided a generally good representation. (The STRESS value was reasonably low, although evidence from further analyses suggests that in fact there are additional, albeit, "weak", dimensions.) The resulting two dimensional configuration, with a curvilinear coordinate system (provided by Shepard) superimposed for purposes of aiding interpretation of this solution, is shown in Fig. 2. The power of MDS techniques can be easily discerned by comparing the information derived from inspection of the original 36 x 36 matrix of confusabilities (which to most of us just provides a headache combined with what William James referred
to as a "blooming, buzzing confusion" — appropriate, perhaps, for confusions data) to that conveyed by the two dimensional representation Shepard found by use of MDS. While the two dimensions shown here may

not account for all the structure in this matrix, they certainly do account for a very major portion of it. By simple inspection (aided, of course, by Shepard's curvilinear coordinates) we see that the two dimensions can be called "number of components" (going systematically from the one component to the five component signals as we move vertically up the two dimensional diagram), and "dots vs. dashes" (or dot-dash ratio, as it is sometimes called), contrasting signals (on the left) comprising all or mostly dots with those (on the right) made up of all or mostly dashes, with a relatively "smooth" transition between. These two dimensions account for the major features of the original confusability matrix. To confirm this one can note, by comparing the two dimensional configuration with the confusions data, that those signals most often confused are, indeed, generally those that are closest in this configuration, although that correspondence is far from perfect. We can see the degree of this correspondence more clearly in Fig. 3, which plots the original data, percent "same" judgments (the ordinate), against the interpoint distances (the abscissa) in the Euclidean two dimensional space. Since this was a nonmetric MDS analysis, we do not expect that plot to be linear, and, indeed it is not! It appears to be more nearly a negative exponential, or so-called "exponential decay" function. Since there is good theoretical reason to expect such a negative exponential function for confusions data, this fact is of interest in itself. (Shepard's original motivation for his work on nonmetric scaling, in fact, had almost as much to do with determining the shape of these functions relating distances to data as it did with determining the dimensions underlying the stimuli.) The best fitting monotonic function relating the data to the distances was also fit by the procedure. This is shown in the figure by the jagged step-like function neatly interspersed among the data points on this scattergram (this entire diagram is often referred to as a "Shepard diagram").

FIGURE 2.


FIGURE 3.


One aspect that is very important to keep in mind in the case of two-way MDS, at least when the Euclidean metric is assumed (as it almost always is in practice) is that, while the data are sufficient to determine the configuration of points, they do not, in general, determine the appropriate orientation of coordinate axes. This is because Euclidean distances are invariant under rigid rotation of the coordinate system.
Shepard had to rotate the essentially arbitrary coordinate system in terms of which the solution originally emerged from the computer before drawing the slightly curved coordinates in terms of which he interpreted the solution. (Just as an ordinary roadmap would be just as valid if drawn in terms of a northwest-southeast and southwest-northeast coordinate system as in terms of a north-south and east-west system — the particular coordinate system we use is convenient for interpretation of the map, but the distances remain unchanged when we effect such a change in coordinates.) While this process of rotating coordinates to effect optimal interpretability is relatively straightforward in two dimensions (where, after all, one can choose precisely what is going on) in three, four or nine or ten dimensions, this becomes very hard indeed (often prohibitively so). In fact, it is very likely the difficulty of finding the optimal rotation of coordinates in high dimensional representations that has provided a strong bias in favor of small dimensional solutions (usually two or three dimensional) in two-way MDS applications. We shall now see how effective use of individual differences among subjects (or subgroups of subjects, or other "data sources") can be used to overcome this often insuperable problem by uniquely orienting the coordinate system.

INDIVIDUAL DIFFERENCES MULTIDIMENSIONAL SCALING. While multidimensional scaling (MDS) was originally intended to apply to the case of a single one-mode, two-way proximities matrix, it very quickly became of interest to deal with the problem of individual differences in proximities judgments, or other proximities data. The data in this case generally comprise a number of symmetric proximities matrices, one for each subject (or other data source). This is generally called two-mode (stimuli and subjects) but three-way (stimuli × stimuli × subjects) data (see Carroll and Arsole, 1980).

The first approach to individual differences multidimensional scaling was the "points of view" approach proposed by Tucker and Messick (1963). "Points of view" attempted to represent actual subjects by a set of prototypical or "idealized" individuals. It does not, however, provide a generally parsimonious representation of the data, nor a well defined prediction of the observed data for individual subjects from which a goodness of fit measure could be defined. This, and other considerations, led to the approach embodied in what is now generally called the INDSCAL (for INDividual Differences multidimensional SCALing) model (Bloom, 1968; Horan, 1969; Carroll and Chang, 1970). INDSCAL accounts for individual differences in proximity data in terms of a process assuming a common set of dimensions underlying the stimuli (or other objects) while assuming that different individuals have different patterns of saliences or importance weights for these common dimensions (assuming for now that the "three-way" proximities data correspond to similarity judgments by different human subjects). The INDSCAL model can be expressed mathematically as follows:

\[ d_{jk} \approx d_{jk}^{opt} = \sqrt{\sum_{r=1}^{R} w_{jr}(x_{jr} - x_{kr})^2} \] (2)

Here \( d_{jk} \) is the similarity or dissimilarity judgment of subject \( i \) for stimuli \( j \) and \( k \). \( F_i(d_{jk}) \) is a function transforming \( \delta_{jk} \) into an (approximate) distance \( d_{jk}^{opt} \) between the points representing stimuli \( j \) and \( k \) for that same subject \( i \), where that distance is expressed as a weighted Euclidean distance given by the formula on the right. Comparing this equation for \( d_{jk}^{opt} \) with the standard formula for Euclidean distance reveals that this formula differs only by the introduction of the subject weight \( w_{jr} \). The stimulus coordinates \( (x_{jr}, x_{kr}) \) being the coordinate of stimulus \( j \) on the \( r \)th stimulus dimension) are the same for all subjects, only the weights for the \( R \) dimensions differ from subject to subject.

More graphically, the INDSCAL model can be explained in terms of the hypothetical example shown in Fig. 4. In this case the hypothetical group stimulus space, with stimuli (labelled by letters A through I) arbitrarily arrayed in a square lattice is shown in the upper left of the figure. In the INDSCAL model there is also a second "space," often called the subject space, whose coordinates correspond to the weights, or perceptual saliences, of the subjects on the dimensions of the group stimulus space. In Fig. 4 the subject space for nine hypothetical subjects (labelled 1 through 9) is shown in the upper right.

We may think of the effect of these differential weights as producing, for each subject, a "private perceptual space" by rescaling (stretching and contracting) the dimensions of the group stimulus space. In the illustration, for example, subject 3 has equal weights for the two dimensions, that subject's private perceptual space would therefore look exactly the same as the group stimulus space (except for an overall scale factor which could stretch or contract both dimensions uniformly, leaving their relative saliences unchanged).

The private spaces for subjects 2 and 4 are shown in the lower left hand and right hand corners of Fig. 3, respectively. Subject 2, who weights dimension 1 more highly than dimension 2, has a perceptual space compressed along the dimension 2 axis or (equivalently) stretched in the dimension 1 direction. The reverse applies to subject 4, who has a higher weight on the second than on the first dimension. These differential transformations, of course, lead to quite different patterns of distances, and therefore of predicted similarity or dissimilarity judgments, for the different subjects.

An important property of INDSCAL is one often called "dimensional uniqueness." This means that,
Hypothetical example illustrating the INDSCAL model.

Weights from the subject space are applied to the dimensions of the "group" stimulus space to produce private perceptual spaces for subjects 2 and 4. (While technically it is the square roots of the weights that are applied, the first power weights have been used in this illustration to accentuate the effects of differential weighting.)

Unlike ordinary (two-way) MDS, the dimensions are uniquely determined, and cannot be rotated or otherwise transformed without changing the solution in an essential way. (They can be permuted in order, though, and the order of points on a dimension can be reflected, if desired.) Psychologically this means that INDSCAL dimensions are assumed to correspond to "fundamental" physiological, perceptual or cognitive processes whose strength, or "salience," may differ from individual to individual (whether because of genetic or environmental differences, or simply because of differing interpretations of instructions or the like). Mathematically, a rotation or other transformation of the coordinate system will change the family of permissible transformations of the group stimulus space, and thus the family of possible individual metrics or distance functions (the $d_{ij}$'s). (This can be seen quite clearly in the hypothetical example by noting that, say, a 45° rotation would change the family of private perceptual spaces from one of different rectangles to one of different rhombuses, or diamond shaped structures). Statistically, a rotation or other transformation of axes will generally deteriorate the fit of the data to the INDSCAL model. The property of dimensional uniqueness proves to be one of the most important aspect of INDSCAL, as it obviates the problem of rotation of axes in almost all cases. This problem is one that (particularly in the case of higher dimensional solutions) frequently complicates and often considerably reduces the utility of MDS analysis. The value of this property should become clearer in the discussion of an application to some color data to follow.

Distance of a subject from the origin in the subject space is roughly a measure of the proportion of the variance of the data for that subject accounted for by the multidimensional solution comparable to the concept of communality in factor analysis). Although subjects 3 and 7 have the same pattern of dimension weights, a higher proportion of the variance in subject 3's data could be accounted for by the hypothetical INDSCAL solution. The data for subjects closer to the origin are generally less well accounted for by the INDSCAL analysis, so that the dimensions of the group stimulus space in toto are less salient for them. The lower "communality" for subjects closer to the origin may be due to idiosyncratic dimensions not uncovered in the $R$-dimensional solution or to lower reliability (more random error) in their data.

Subject 9, who is precisely at the origin, is completely "out of" this analysis. Either he or she is responding completely randomly or is simply "marching to a different drummer" (responding reliably to a completely different set of dimensions). We could tell which is the case by trying analyses in higher dimensionality, and seeing if a dimension or dimensions emerge accounting for this subject's data. The possibility of zero weights allows as a special case the situation in which two or more groups of subjects have completely different perceptual spaces, with no necessary communality between them. This makes it possible to accommodate the "points of view" model as a special case of INDSCAL.

An Illustrative Application of INDSCAL to Some Color Dissimilarities. We now illustrate the INDSCAL model by an example which entails reanalysis of some data on color perception collected by Helm (1964). Helm and Tucker (1962) had analyzed these data via a "points of view" analysis and found that about ten points of view (that is, ten different perceptual spaces) were required to account for the individual differences among their 14 subjects.

Our analysis in terms of the INDSCAL model has quite nicely accounted for these individual differences in terms of a single two-dimensional solution. Moreover, the unique dimensions from the INDSCAL analysis shown in Fig. 5 were interpretable without rotation of axes. Dimension 1 corresponds essentially to a "blue vs. yellow" (or, more accurately perhaps, a "purple-blue vs. green-yellow" factor, and dimension 2 to a "red vs. green" (or "purple-red vs. blue-green")
The inspection of the private spaces provides insight into the way in which INDSCAL uses individual differences in perception to determine a unique orientation of the coordinate axes. If the coordinate axes were oriented in a different way (say, by a 45° rotation of the axes of the group stimulus space), the private spaces for color deficient subjects could not be compressed along a line from red to green, but would have to be compressed in some other direction.

The INDSCAL method, as described in Carroll and Chang (1970), makes metric assumptions (i.e., assumes that all the functions $F_i$ in Eq. (1) are linear), and entails a kind of three-way generalization of the "classical" method for two-way MDS. The most efficient approach for implementing this analysis is provided by a program called SINDSCAL, due to Pruzansky (1975). Recently, various alternative nonmetric and metric algorithms have been developed for fitting the INDSCAL model. Most prominent are Takane, Young and deLecuyr's (1977) ALSCAL (in its three-way implementation), and Ramsay's (1977) MULTISCAL procedure, which provides alternative methods for fitting the INDSCAL model, optimizing a maximum likelihood criterion. For a general discussion and comparison of computer programs for both two-way and three-way MDS see the book by Schiffman, Reynolds and Young (1981).

For discussion of other models and related methods see Carroll and Arabie (1980) or Carroll and Wish (1974). Of particular interest among these alternative models and methods is one usually called Three-Mode Scaling due to Tucker (1972), which is an adaptation of Tucker's (1964) three-mode factor analysis model and method to the case of three-way proximities data.

While space and time to not permit a full exposition of the wide range of MDS models and methods for multidimensional analysis of preference data, we end this paper with a brief mention of probably the most popular and widely used model and method of this type, called MDPREF (for, simply, Multi-Dimensional PREFerence analysis). MDPREF assumes a model sometimes called a vector model for preference (or other dominance) data. In this model each stimulus (or other object) is represented as a point in a multidimensional space, while each subject (or other data source) is represented by a vector (or directed line segment) in this same space. Orthogonal projections of stimulus points onto a subject's vector approximate (as well as possible, in a well defined statistical sense) the order of preference judgments for that subject. The

---

**FIGURE 5.**

A two-dimensional INDSCAL analysis of Helm's (1964) data on color perception produced the group stimulus space shown in (A) and the subject space shown in (B). Private perceptual spaces for two color-deficient subjects (CD1 and CD4a) and for two subjects with normal color vision (N10a and N7) were derived from the group stimulus space by applying the subjects' respective weights (shown in the subject space) to the dimensions. The coding of colors (of constant saturation and brightness) is as follows: R=red; Y-yellow; GY(1)=green yellow; GY(2)-green yellow with more green than GY(1); G=green; B=blue; PB=purple blue; P(1)=purple; P(2)=purple with more red than P(1); RP=red purple.

Factor. This accords very well with physiological and psychophysical evidence strongly suggesting the existence of "blue-yellow" and "red-green" receptors.

Included among Helm's subjects were four who were deficient, to varying degrees, in red-green color vision. In the INDSCAL analysis this deficiency is reflected in the fact that these subjects all have lower weights for the "red-green" factor (dimension 2) than do any of the normal subjects. The effect of these differential weights can be seen in Fig. 2 by comparing the private perceptual spaces for the color-deficient with those for the normals. The spaces for the color-deficient subjects are compressed in the "red-green" direction, relative to the spaces for the normals, reflecting the fact that red and green (for example) are much more similar to each other for these subjects than they are for the normals.
MDPREF method is a particular method for fitting this vector model to preferential choice data for each of a number of paired comparisons preference or dominance matrices, one for each subject, (or simply to a subjects by stimuli matrix of preference scores). MDPREF is an internal analysis of such preference data (meaning that it utilizes only the preference data to fit this joint model entailing both stimulus points and subject vectors). An external analysis of preference data, in which the coordinates of stimulus points are given a priori (say, by an independent MDS analysis of proximity data), is possible by such technique as one called PREFMAP (for PREference MAPping). MDPREF actually utilizes the singular value decomposition (s.v.d.) of a normalized preferences score matrix (either given or derived) to attain its representation of stimulus points and subject vectors. A full discussion of MDPREF, PREFMAP, and other approaches to multidimensional preference analysis (as well as of other more general models for individual difference in preference judgments such as the "unfolding model," postulating different "ideal points" rather than vectors for different subjects, and some generalization of that model) see Carroll (1972, 1980) or Coombs (1964).

Figure 6, below, illustrates the vector model assumed in MDPREF via a hypothetical example in two dimensions. Letting A, B, C, D, E represent stimuli (or other subjects) with the two vectors shown represent two subjects (or other data sources). The orders of projection of stimuli on those subjects' vectors are also indicated (near the end or terminus of each vector). These orders approximate the preference order for those two subjects.

**FIGURE 6.**
Vector model for preference illustrated. Projections of stimulus points onto a subject's vector are assumed to define preference scale for that subject.

**REFERENCES**


Horan, C. B. Multidimensional scaling: Combining observations when individuals have different perceptual structures. Psychometrika, 1969, 34, 139-165.


Tucker, L. R. Relations between multidimensional scaling and three-mode factor analysis. Psychometrika, 1972, 37, 3-27.


Supplementary References


Kruskal, J. B., & Carmone, F. Use and theory of MONANOVA: A program to analyze factorial experiments by estimating monotone transformation of the data. Unpublished manuscript, Bell Laboratories, 1968.


Takane, Y., & Carroll, J. D. How to use MAXSCAL-4.1: A program to fit the simple (unweighted) Euclidean distance model to various types of nonmetric dissimilarity data by maximum likelihood. Unpublished manuscript, Bell Laboratories, 1980.


