MULTIDIMENSIONAL UNFOLDING: DETERMINING CONFIGURATION FROM COMPLETE RANK ORDER PREFERENCE DATA

WILLIAM L. HAYS

UNIVERSITY OF MICHIGAN

AND

JOSEPH F. BENNETT*

LINCOLN LABORATORY, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Within the model of isotonic space, a principle is presented which generalizes the unfolding technique to the multidimensional case. The availability of exhaustive configurational solutions given complete data is pointed out. Finally three criteria are suggested for the choice of a particular solution from among the set of all solutions, which are applicable in the case either of complete or incomplete data.

In an earlier paper [2], the authors discussed the estimation of the dimensionality underlying a set of rank orders. The data to which these methods are applicable are rank orders of preference given by subjects for a set of n objects. The dimensionality estimated is that required by the stimulus space in which the objects are presumably viewed by the judges. Three lower-bound criteria for dimensionality were proposed: mutual boundary, cardinality, and the existence of permutation groups. The present paper continues with the generalization of Coombs' unfolding technique [3] to the multidimensional case, which was begun in the first paper, and concerns the problem of determining configuration and arriving at a solution when the data are "complete," in the sense to be described below. In particular, a principle providing the generalization of the unfolding method will be introduced, and various criteria for determining a set of r rank order axes for the description of a configuration of n stimulus points in r dimensions will be discussed. It will be convenient to acquaint the reader with some terminology before the matter of a solution is taken up.

Some Terminology

This discussion will be couched in terms of a simple generalization to several dimensions of the model proposed by Coombs [3, 4]. It is assumed

*Deceased.

that subject i views stimulus k as a point X_{ik} in a stimulus space of some dimensionality r. Given some arbitrary origin, the components of X_{ik} are x_{ijk} , the projection or loading of stimulus k for individual i on attribute j. To simplify matters, it will be assumed that x_{ijk} is the same for each individual i, so that the projection of stimulus k on attribute j may be denoted simply by x_{jk} , a component of X_{ik} . The stimulus space will be assumed to have Euclidean properties as well. Strictly speaking, neither of these assumptions may be essential, but it has proved very difficult to develop the model without some such restrictions.

Also following Coombs, it is assumed that each individual i may be associated with an "ideal" stimulus, a real or hypothetical object which the individual would most prefer in any given stimulus space. This ideal stimulus may also be represented as a vector C_i with components c_{ij} . Once again the simplifying assumption is made that each individual is associated with one and only one such ideal point in the space.

Now the observational equation linking the judged rank order of preference to the distances within the stimulus space is given by

$$(k \cdot \succ m)_i \Leftrightarrow \sum_{j=1}^{r} (c_{ij} - x_{jm})^2 > \sum_{j=1}^{r} (c_{ij} - x_{jk})^2.$$

That is, stimulus k will be preferred to stimulus m by individual i if and only if the sum of the squared differences between a stimulus and the ideal stimulus over a set of r orthogonal reference axes is greater for stimulus m than for stimulus k. The subject prefers that stimulus which is closer to his ideal in the stimulus space.

In short, the stimuli are conceived simply as having some configuration in a Euclidean stimulus space of dimensionality r. A particular rank order of preference reflects increasing magnitudes of distance from the ideal stimulus to the respective stimulus points. Obviously, for other than the one-dimensional case, a solution to this problem of describing the stimulus configuration requires an excursion into the geometry of higher spaces. Furthermore, since the data with which we start are nothing more than a set of rank orders, a special class of such higher spaces must be considered; these are so-called *isotonic* spaces, in which every region in the space is characterized by a rank order of distances to a fixed set of points. In other words, given the set of stimulus points, each and every point in the space must show a rank order of distances to these stimulus points, and the space as a whole may be divided into isotonic regions, convex subspaces within which each and every point shows the same rank order of distances from the stimulus points. The rank order which each point in the region exhibits will be called the characteristic order for the region, and any region will be referred to by its characteristic order.

The one-dimensional case of an isotonic space is, of course, the under-

FIGURE 1

A Configuration of Four Points in Two Dimensions

lying model for the unfolding technique of Coombs [3]. The unfolding technique is based upon the principle that a sequence of $\binom{n}{2} + 1$ rank orders may be found representing the unique arrangement of the isotonic regions in a one-dimensional space; from this sequence of rank orders, the order of stimulus points on the attribute may be inferred, as well as a partial order of the distances between points.

There is no such unique sequence in the case of two or more dimensions, however, as the number of possible isotonic regions increases very rapidly both with the number of stimulus points and the number of dimensions. An example of an isotonic space for four points in two dimensions is given by Figure 1, and an example for five points also in two dimensions by Figure 2.

The three criteria of dimensionality adverted to above may be illustrated from these examples. Note, for instance, that no more than four isotonic regions anywhere mutually bound in either Figures 1 or 2: this reflects two dimensions, according to the first criterion. Second, notice that only 18 of the 24 possible permutations of four objects occur as charac-
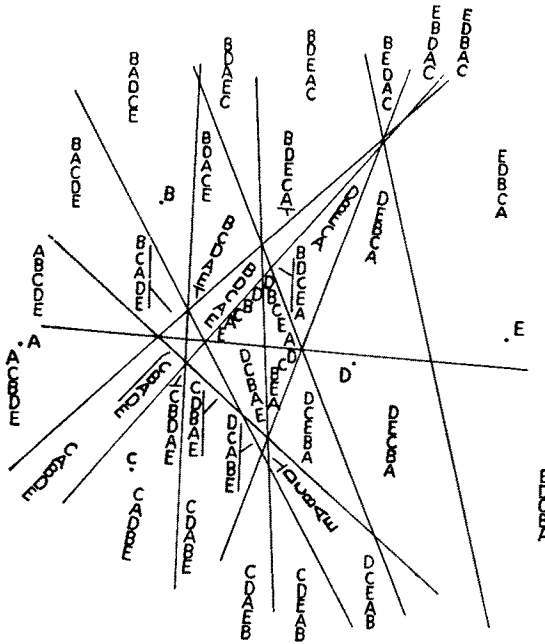
FIGURE 2

Isotonic Regions Generated by Five Points in Two Dimensions

teristic orders for regions in Figure 1, and that only 44 different orders from among the 120 possible permutations of five things occur in Figure 2: this illustrates the so-called *cardinality* criterion. Finally, and most important, note that while there are complete sets of 6 permutations in rank order for subgroups of three stimuli embedded among the rank orders, there exist no such complete sets for subsets of either four or five stimuli: this illustrates the *groups* criterion, which will be of importance in determining configuration as well as in estimating dimensionality for such data.

The isotonic regions are bounded by loci of equidistance from two stimuli—these represent regions of equal preference for the two stimuli. The loci of equidistance in two-dimensional cases as illustrated are merely the perpendicular bisectors of the lines joining pairs of stimulus points. These appear in the figures as the lines dividing the space into regions. Such loci of equidistance from pairs of stimuli will be referred to as 2-loci, and will be denoted by $H(A, B)$, where A and B refer to a particular pair of stimulus points. For any configuration there will be $\binom{n}{2}$ such 2-loci. It is essential to remember that in one dimension, a 2-locus will be a point, in two dimensions a line, in three dimensions a plane, in four dimensions a

three-space, and so on. In general, in r dimensions, the 2-locus $H(A, B)$ will be a hyperplane of dimensionality $r - 1$.

Another feature of note in any isotonic space is the fact that regions may be divided into two disjoint classes: *open* and *closed*. Closed regions are, of course, everywhere bounded by loci of equidistance, while open regions are not. The distinguishing feature of rank orders derived from open and closed regions is that *each and every open region must have a mirror image mate*, another region which has a characteristic order which is the exact reverse of the first rank order; any pair of mirror image rank orders in the data leads to the inference of the existence of a pair of open regions. On the other hand, *no closed region may have such a mirror image mate*. In any dimensionality lower than $n - 1$ for n stimuli, there must exist some closed regions in the isotonic space, and hence all $n!$ rank order permutations may not occur in the data for less than dimensionality $n - 1$. This is actually the basis for the cardinality criterion of dimensionality.

It will be noted from Figures 1 and 2, that there are points of intersection of sets of three 2-loci. These intersections are equidistant from sets of three stimuli, and as such are appropriately called 3-loci and designated $H(A, B, C)$, where (A, B, C) is any set of three stimuli. In two dimensions a 3-locus is a point, in three dimensions a line, in four dimensions a plane, and so on. In general, in r dimensions, a 3-locus will be a hyperplane of $r - 2$ dimensions.

It will be convenient to consider loci of even higher order, so that a general notation and dimensional principle will be useful. A g-locus will be the space of all points equidistant from a set of g points, $H(A, B, C, \cdots, g)$, and will always be a subspace of dimensionality $r + 1 - g$. Furthermore, it will also be useful to remember that a point which is equidistant from some g stimuli in $g - 1$ dimensions is the center of a hypersphere having the stimulus points in question on its surface. Thus, three points requiring two dimensions must lie on a circle, four points requiring three dimensions must lie on a sphere, five points requiring four dimensions must lie on a four-dimensional hypersphere, and so on. Furthermore, the converse is true— if there exists no hypersphere in $g - 1$ dimensions such that a particular set of g points may lie on its surface, then the set of points may be embedded in a space of $g - 2$ dimensions or less. As will appear in the discussion to follow, this is simply another way of phrasing the groups criterion of [2].

One final feature of the higher spaces should also be mentioned here: given a subspace T of dimensionality t and a subspace S of dimensionality s, $t \geq s$, such that S is not a subspace of T, then the intersection of S and T is a subspace of dimensionality $s - 1$. Thus, the intersection of a plane and a line (not entirely in the plane) is a point, the intersection of two non-coincident planes is a line, the intersection of a six-space and a three-space is a plane, and so on.

A General Unfolding Principle

As mentioned above, the one-dimensional unfolding technique relies on the fact that if the rank orders emanate from a one-dimensional stimulus space, it is always possible to construct a unique sequence such that each distinct rank order differs from either neighbor in the sequence by a reversal in order of only one pair of objects. The end or mirror image rankings then provide the order of the objects on the attribute.

Coombs' procedure could perfectly well be interpreted as finding a sequence of 2-loci rather than a sequence of isotonic regions, since the reversal in order of a pair of objects for a pair of regions simply fixes such a 2-locus point. Considering the unfolding technique in this way suggests the principle which allows an extension to the multidimensional case. Since this principle is in fact the general statement of Coombs' basic idea, it seems important and nontrivial enough to state and prove.

PRINCIPLE. Given some fixed line L, and three points A, B, and C in general position in an isotonic space of dimensionality r, the line L intersects the 2-loci $H(A, B)$, $H(B, C)$, and $H(A, C)$ in that order (or the reverse) if and only if the perpendicular projections of the three points upon L are in the order CAB or the reverse.

PROOF. The necessary condition will be proved first. If the line L coincides with the first of a set of r orthogonal reference axes $(X_1, X_2, \cdots, X_r)$, such that the origin lies at the intersection of L and $H(A, B)$, then every point k on L is characterized by an r-tuple $(x_{1k}, 0, 0, \cdots, 0)$. Let the intersection of L and $H(A, B)$ be $(x_{11}, 0, \cdots, 0)$, with $x_{11} = 0$, that of L and $H(B, C)$ be $(x_{12}, 0, \cdots, 0)$, and L and $H(A, C)$ be $(x_{13}, 0, \cdots, 0)$. Finally let the point A be characterized by $(x_{1A}, x_{2A} \cdots, x_{rA})$, and similarly for B and C. Now the 2-locus $H(A, B)$ is defined by

$$(1) \quad (x_1 - x_{1A})^2 + \cdots + (x_r - x_{rA})^2 = (x_1 - x_{1B})^2 + \cdots + (x_r - x_{rB})^2,$$

where $(x_1, x_2, \cdots, x_r)$ is any point lying in the 2-locus. Similar definitions may be made for the other two 2-loci as well. Solving (1) for x_{11} and putting $(x_1, x_2, \cdots, x_r) = (x_{11}, 0, 0, \cdots, 0)$ gives

$$(2) \quad x_{1A}^2 + \cdots + x_{rA}^2 = x_{1B}^2 + \cdots + x_{rB}^2 .$$

The value of x_{12} is given by

$$(3) \quad x_{12} = \frac{x_{1B}^2 - x_{1C}^2 + \cdots + x_{rB}^2 - x_{rC}^2}{2(x_{1B} - x_{1C})}$$

and the value of x_{13} by

$$(4) \quad x_{13} = \frac{x_{1A}^2 - x_{1C}^2 + \cdots + x_{rA}^2 - x_{rC}^2}{2(x_{1A} - x_{1C})} .$$

Because the intersections of the 2-loci with L are in the order $H(A, B)$,

$H(B, C)$, $H(A, C)$ or the reverse, the absolute difference between x_{11} and x_{12} must be less than that between x_{11} and x_{13}, or $d_{13}^2 > d_{12}^2$, where

$$(5) \qquad d_{13}^2 = \frac{(x_{1A}^2 - x_{1C}^2 + \cdots + x_{rA}^2 - x_{rC}^2)^2}{4(x_{1A} - x_{1C})^2},$$

$$(6) \qquad d_{12}^2 = \frac{(x_{1B}^2 - x_{1C}^2 + \cdots + x_{rB}^2 - x_{rC}^2)^2}{4(x_{1B} - x_{1C})^2}.$$

However, subtracting $x_{1C}^2 + x_{2C}^2 + \cdots + x_{rC}^2$ from each side of (2) and squaring shows that the numerators of (5) and (6) must be equal. Hence it follows that since d_{13}^2 exceeds d_{12}^2, the denominator of (6) must exceed the denominator of (5), so that

$$(x_{1B} - x_{1C})^2 > (x_{1A} - x_{1C})^2,$$

the perpendicular projections of A and C upon L must be nearer than those of B and C. An identical argument using $x_{13} = 0$ shows that the distance between A and B projections must also be less than that between projections of B and C. Since distance is invariant under translation or rotation of axes, the necessary condition is proved. The sufficient condition is proved simply by reversing the steps of the necessary condition argument.

The unfolding technique for one dimension is actually a special case of this more general principle for finding projections upon lines by constructing sequences of regions (or dually, 2-loci). In isotonic space of any dimensionality, the existence of $\binom{n}{2} + 1$ regions which fit the unfolding qualifications is sufficient for the inference of the order of projections which the points have relative to some line. As an illustration of this principle, consider the following sequence of seven regions drawn from the example of Figure 1: $DACB$, $DCAB$, $CDAB$, $CDBA$, $CBDA$, $CBAD$, $BCAD$. In this sequence, $H(C, D)$ lies between $DCAB$ and $CDAB$, $H(A, C)$ falls between $DACB$ and $DCAB$, and $H(A, D)$ lies between $CBDA$ and $CBAD$; the order of these three 2-loci is thus $H(A, C)$ $H(C, D)$ $H(A, D)$ (or the reverse), so that on a line extending through these seven regions, the order of projections of the three stimulus points A, C, and D must be CAD (or the reverse). Likewise, the order of A, B, and C as projected upon such a line would be ACB or the reverse, since the order of their 2-loci is $H(A, C)$ $H(A, B)$ $H(B, C)$. An inspection of the 2-loci for all such triples of stimuli establishes that the order of projections on such a line would be $DACB$ or the reverse (obviously, since there is no fixed origin in the isotonic space, the orders of projections on any line may be read in either direction). Any line capable of being located in the space must pass through such a sequence of $\binom{n}{2} + 1$ regions, and any complete unfolding sequence of $\binom{n}{2} + 1$ regions occurring in the space must represent at least one possible line in the space.

An important feature of Coombs' unidimensional unfolding solution is the recovery of metric relations among the stimulus points, yielding an ordered metric scale of the stimuli. This metric information is inferred from the sequence of 2-loci just as is the simple order of the stimuli themselves. Unfortunately, it can be shown that the ability to obtain metric information in this way is *restricted to the unidimensional case*. While the theorem above is a complete generalization of the method for obtaining the order (or the order of projections) of the stimulus points, in other than the one-dimensional case, Coombs' method for inferring metric information does not work for the *projections* of the points upon axes in the space. An important corollary follows directly from the principle just given.

> In any isotonic space of n stimuli in r dimensions, two regions may have characteristic orders which are mirror images if and only if there exists the possibility of a line in the space such that the order of projections of the stimuli on the line is the same as the characteristic order of either of the regions (or the reverse, of course).

An Exhaustive Solution for Complete Data

The practical implication of this corollary principle is that any pair of mirror images existing in the data afford a potential solution, in that there must exist the possibility of an axis showing such an order of projections. Even more important is the fact that any potential solution must be represented by such a mirror image pair of regions in the data, when the data are complete. In this light, the question of a solution for complete data becomes rather trival. First, in the present context, let *complete data* be understood to mean sets of rank orders such that each and every isotonic region in the stimulus space has its characteristic order represented at least once in the data. Thus when complete data are at hand, all possible configurational solutions may be recovered from the data simply by finding mirror image pairs of rankings. Each mirror image pair located is one potential axis for describing the configuration; for this reason the solution from complete data may be called *exhaustive*.

In the simple example of Figure 1, the mirror image pairs of regions are *DACB–BCAD, DABC–CBAD, ADBC–CBDA, ABDC–CDBA, ABCD–DCBA, BACD–DCAB*. Thus there are six possible simple orders which may represent axes or solutions to the configuration: *DACB, DABC, ADBC, ABDC, ABCD,* and *BACD* (or their reverses). These constitute the exhaustive solution for this configuration.

The number of such potential axes varies, of course, both with the number of stimuli and the dimensionality. Actually, it is possible to calculate the maximum number of distinct such axes (i.e., distinct mirror image pairs

of regions). These maximum numbers have already been tabled in another context [1] in the form of the maximum number of open isotonic regions which may exist for given numbers of stimuli and dimensionalities; in order to convert this table into maximum number of rank order axes, one simply divides the entries by 2. Thus, for example, there are 36 different rank orders of projections possible for 5 stimuli in three dimensions, 105 different possible rank order axes for 15 stimuli in two dimensions, and just over a billion possible rank orders of projections for 30 stimuli in five dimensions! Obviously, such exhaustive solutions leave something to be desired in the way of parsimony of description. Moreover, seldom would we be interested in all solutions anyway, even if there were fairly restricted numbers of such possibilities.

Under the influence of factor-analysis methods, we have grown accustomed to the description of configurations of points requiring r dimensions in terms of a set of axes numbering fewer than r. While the usual factor analysis deals only with common factors, no such restriction exists within this model. The dimensionality estimated for an isotonic space includes both common and specific factors and, in principle, it should be possible to analyze the data for all r dimensions. Still another difference exists between metric and nonmetric approaches to this problem: metric methods such as factor analysis provide dimensions from which one may reproduce the original data, while, at this writing, there seems to be no prospect that one might reproduce an original set of rank orders in terms of some r rank order dimensions. In a sense there is more information in the data than in the rank order dimensions obtained. Failing any criterion for the reproducibility of the data in nonmetric terms, the only recourse seems to be to choose some solution from among the set of all solutions according to criteria of a best fit to the data. In order to do this, one must settle upon some criteria of goodness for choosing among all possible solutions.

The remainder of this paper will be devoted to a description of three features of the isotonic space which may serve as criteria in the choice of a rank order solution from all possible such solutions. These criteria are applicable not only in the theoretical case of complete data but also in the case of incomplete data as well, and thus they will be described in detail here.

The Idea of a Central Intersection

One requirement for a solution might be that each successive axis pass through the *center* or greatest concentration of points in the configuration. That is, the first axis should describe the *length* of the configuration through its greatest concentration; the second axis should describe the length of the configuration of projections of points on a space of one less dimension, and so on. Thus, axes may be sought which are roughly analogous to principal axes in the usual factor-analysis model.

In order to find such a solution within the isotonic space, the idea of a *least intersection* of a configuration may be introduced. In any r-dimensional space, the locus of equidistance from any $r + 1$ points in general position is a point. This point is the center of a hypersphere of dimensionality r. Thus, the locus of equidistance from three points in two dimensions is the center of a circle (2-sphere); the locus of equidistance from four points in three dimensions is the center of a sphere (3-sphere); the locus of equidistance from five points in four dimensions is the center of a hypersphere of four dimensions (4-sphere), and so on.

Any hypersphere of whatever dimensionality must bear one of three possible relationships to any point in the space: the point in question must either be interior to the hypersphere (fall within its surface), exterior to the hypersphere (fall beyond the space enclosed by its surface), or conjoint with the hypersphere (fall upon its surface, and thus be at a distance from its center equal to that of any other point on the surface). Furthermore, if a stimulus point X is exterior to an r-sphere, then the order of distances which is characteristic of its center must show the point X more distant from the center than any point on the surface. On the other hand, if the point X is interior to the r-sphere, then the order associated with the center must show X less distant from the center than any point on the surface. Finally, if the point X is conjoint with the hypersphere, the order associated with the center must show the X equally distant with any point on the surface.

Obviously, among n stimulus points in general position in r dimensions, any given stimulus point must be conjoint with $\binom{n - 1}{r}$ distinct r-spheres, since there will exist a center of an r-sphere for each set of $r + 1$ points. Also, any pair of points must be conjoint with $\binom{n - 2}{r - 1}$ r-spheres.

The various r-spheres defined by the points in the space will differ in the extent to which they include the entire configuration within or on their surface. Some r-spheres will have none of the remaining $n - r - 1$ points interior to its surface. At least one r-sphere will have all of the points either within or on its surface. Such an r-sphere containing all points either within or on its surface will be called an *enveloping sphere*. Given two or more enveloping spheres, the subspace formed by the intersection of the spaces they bound will contain the configuration; this subspace will be called the *central intersection* of the spheres. In Figure 3, for example, the circle defined by the points A, C, and E is an enveloping sphere, since all five points are either on or within its surface, and the same is true of the circle defined by A, B, and E. The central intersection of these two circles, as shown by the shaded area, contains the configuration.

In particular, if there exist r distinct enveloping r-spheres each generated by the same two points X and Y and some set of $r - 1$ other points,
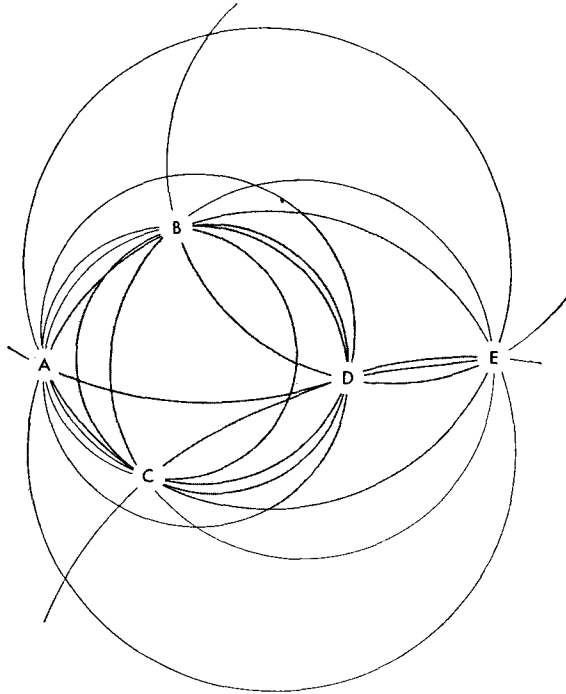
FIGURE 3

Circles Generated by Five Points in Two Dimensions Showing the Least Intersection
for the Configuration

then the central intersection of all of the $\binom{n-2}{r-1}$ spheres will be called a
least intersection. The major axis of this least intersection subspace will be
the line joining X and Y, and the minor axes of the subspace must lie in
the hyperplane of equidistance $H(X, Y)$ defined by these two points. In
Figure 3, the shaded area bounded by the two circles formed by A, B, and
E and by A, C, and E is a least intersection, and the major axis of this area
is the line $\overline{AE}$ with a minor axis defined by their perpendicular bisector.

However, since one is dealing strictly with an isotonic space, in which
the only information available is in the form of rank orders for regions, the
problem remains of finding that pair of points defining the axis of a least
intersection. This may be done as follows. The center of an r-sphere is, of
course, an $(r + 1)$-locus, the point of equidistance from some set of $r + 1$
points. The $(r + 1)$-locus point does not fall into any isotonic region having
a simple characteristic order of distance; rather, such an $(r + 1)$-locus must
fall on the boundary separating a number of such regions, and consequently
have a partial ordering of distances, since it is by definition equally distant

from at least $r + 1$ points. For example, in the configuration of five points in two dimensions, the 3-locus $H(A, B, C)$ for the three points A, B, and C is the point of intersection of the three 2-loci $H(A, B)$, $H(A, C)$, and $H(B, C)$, thus falling on the boundary lines among six regions. Both D and E are exterior to the circle, so that the order of distances from the center of the circle to the five points is the partial order $(ABC)DE$. (Single parentheses enclosing a set of points in a partial order will always denote equality among the points in the set, while double parentheses will denote the set of all permutations in order of the enclosed set: thus $(ABC)DE$ is read as A, B, C equally in the first place followed by D and E, while $((ABC))DE$ would read as any of the set of six orders consisting of some permutation of A, B, and C followed by D and then E.) Note also that the six regions immediately surrounding the center of this circle are all alike in order, except that each shows a different one of the six permutations $((ABC))DE$, the positions of D and E remaining fixed. Such a set of *permuting* regions would also be used in the groups criterion for dimensionality mentioned earlier. The general principle which such sets of regions exhibit is: if there exist a set of $(r + 1)!$ regions showing exactly the same characteristic orders except for a permutation of some set of $r + 1$ stimulus points, then there exists an r-sphere to which those $r + 1$ stimuli are conjoint. The partial order of distances characteristic of the center of the r-sphere is the same as that of any of the set of permuting regions, except that all of the set of conjoint stimuli are equally distant from the center point, thus giving it a partial order of distances. In other words, the groups of permutations occurring in the data tell not only about the dimensionality, they also tell of the existence of r-spheres in the space. Figures 2 and 3 again provide an illustration. Each one of the ten circles is accompanied by a set of six permuting regions, and each set of permuting regions surrounds the center of a circle. No set of 4! permuting regions may be found, however, since 2 is the dimensionality.

Furthermore, these sets of permuting regions also give information about the positions of all of the points relative to each of the circles. It has already been mentioned that the set of $r + 1$ stimuli which permute among the $(r + 1)!$ rank orders are those which are on the r-sphere. If any stimulus falls in order below the permuting stimuli for the region orders, then that stimulus is necessarily exterior to the r-sphere. On the other hand, if any stimulus point precedes the stimuli which permute in order, then that stimulus point is interior to the circle. For instance, the regions $CBADE$, $CBDAE$, $CDBAE$, $CDABE$, $CADBE$, $CABDE$, which are members of the set $C((ABD))E$, differ only by a permutation of A, B, and D; thus they must surround the center of a circle in the example, with A, B, and D on the perimeter. Since among these regions C always precedes A, B, and D, the circle must have C as an interior point. However, E always follows the three permuting stimuli in all of the regions of the set of six, so that one must

conclude that the circle will have E exterior to it. This may be seen from Figures 2 and 3.

Since the positions of the points relative to the r-spheres may be read from the orders characterizing such permutation groups in the data, a way emerges for fixing the least intersection for a configuration of points. Recall that the least intersection is the subspace formed by the intersection of r distinct r-spheres such that each point is either conjoint with or interior to each r sphere. Then a least intersection may be determined by first finding all those permuting sets of $(r + 1)!$ orders which show the property that all stimulus points are either in the permuting set or precede the permuting set in order. That pair of stimulus points common to the permuting set for r such groups of stimuli is the major axis of the least intersection. In the example, these r-spheres are represented by the sets of six regions falling into the partial order $CD((ABE))$ and the set of six regions falling into the partial order $BD((ACE))$. The points common to the permuting set for both groups are A and E; consequently A and E describe the long axis of the least intersection for this configuration.

Moreover, the two points which define the major axis of the least intersection of the space have a property which permits them to be identified simply, without the necessity of inspecting all of the sets of $(r + 1)!$ permutation groups of regions. The circumstance that these two points are conjoint with the r-spheres forming the least intersection makes it true that among the closed regions of the space (i.e., those having no mirror image), this particular pair of points will appear in the last two places in order for the largest number of regions. On the other hand, the desired pair will appear in neither the first two places nor the last two places in any open region. Thus, the endpoints of the major axis of the least intersection may be found very easily for complete data by merely counting the number of times pairs of stimulus points appear in the last places for closed regions, minus the number of times the pair appear in an extreme position (at either end of the order) for open regions.

For example, in Table 1 based on Figure 2, it can be seen that the pair A and E occurs in last place in nine of the closed regions, and in an extreme position in none of the open regions. Thus, the major axis of the least intersection must terminate in A and E.

To recapitulate, one criterion which is proposed for choosing among rank order solutions is that the axes chosen reflect the tendency of the points to cluster along the long axis of the least intersection, so that the axes chosen may reflect the general shape of the configuration in so far as possible. This will be taken as the first requirement for a choice from among all of the available solutions for complete data.

Albeit the axis of the least intersection will be determinate for most configurations, it is possible to construct configurations in which there will

TABLE 1

Open and Closed Regions for the Five-Point Example

Open Regions		Closed Regions	
ABCDE	EDCBA	BDACE	BCADE
ACBDE	EDBCA	BCDAE	BDCAE
BACDE	EDCAB	BDCEA	BDEAC
BADCE	ECDAB	BDECA	BDAEC
BADEC	CEDAB	BEDCA	CBDAE
BAEDC	CDEAB	DCBAE	DCABE
BEADC	CDAEB	CDEBA	CDBEA
BEDAC	CADEB	DCBEA	DBCAE
EBDAC	CADBE	DBCEA	DBECA
EDBAC	CABDE	DEBCA	DECBA
		DCEBA	EBDCA
		ECDBA	CEDBA

be fewer than r distinct r-spheres, each of which will have the property of including all of the points within or on its surface. In this situation, there will be ambiguity as to which of the pairs of points best characterize the major axis of the configuration. For example, in the configuration of Figure 2, if one ignores stimulus A, and concentrates on B, C, D, and E only as four points in two dimensions, he can see that the circle generated by B, C, and E fits the qualification for one enveloping circle, but that there is no circle among the remaining three described by B, C, and D; B, E, and D; and C, E, and D which fits this qualification. For this configuration, then, there is no special choice among the pairs BC, BE, and CE as determining the end points of a first axis. There will always, however, exist at least one enveloping r-sphere in any configuration in any dimensionality r.

A Quasi-Simple Structure Criterion

It does not seem quite enough, however, to insist that the rank order axes chosen should reflect the length, breadth, and height of the configuration. It seems desirable to seek solutions which have a certain degree of inherent parsimony of description. In other words, another aspect to a good solution should be its simplicity in some sense. This is true especially since there seems to be no good analogy to rotation within the isotonic model.

The search for such solutions in factor analysis is indissolubly linked with the name of Thurstone and the concept of simple structure [5]. While the rules for achieving simple structure seem to be very much bound up with the mechanics of factor analysis, and especially of rotation, there does seem to be one aspect which may have a rough analogy in the isotonic model. In describing the characteristics of simple structure, Thurstone ([5], p. 335)

emphasized the desirability of maximizing the number of zero loadings which any given factor should show, while, at the same time, minimizing the number of factors on which a test should show high loadings. In the isotonic model, there is, of course, no unique origin in the space, and since the possible solutions are only ordinal in character, the concept of zero loading has no special meaning for this model. However, it does seem that a pertinent part of this requirement for simple structure is not that the loadings for a number of tests are zero per se, but rather that the differences among a maximal number of points are zero when projected upon an axis. In other words, since reference axes are ways of describing the differences which exist among points anyway, maximum clarity is achieved when the various axes describe different sorts of differences among the points of the configuration, so that differences which project large upon one axis shall not project large upon others. This is emphatically not the only interpretation of the concept of simple structure by any means; it is, however, an aspect for which there is at least a distant analogy within the confines of an isotonic space. Thus, a quasi-simple structure requirement may also be imposed in the choice of a solution: each axis should be chosen in such a way that the number of zero distances among projections on each of the axes is maximal.

For n points in r dimensions, there are, in a sense, $\binom{n}{r}$ ready-made axes consisting of all of the r-loci in the space. Recall that the locus of equidistance from r points in r dimensions is a line, and this line must have projections of all of the points upon it. Hence, each of the r-loci is a potential axis. Furthermore, these r-loci do have one valuable property in the light of the quasi-simple structure notion just introduced. This is that the r points defining the locus must project onto exactly the same point upon it; that is, since they are all equally distant from any point on the r-locus, their projections onto the locus must coincide. For example, note in the five-point example that on the line $H(A, B)$ for instance, how the projections of A and of B must coincide, and so on for each pair of points defining a 2-locus line. Consequently, this quasi-simple structure criterion may be approached by taking the r-loci themselves as the axes. This reduces the choice of the set of r axes to some extent, but there are still $\binom{n}{r}$ such loci from which to choose.

How may one determine the order of projections upon an r-locus from complete data? The answer is, by finding permutation groups in sets of $r!$ open regions. According to the corollary of the theorem, open regions must describe possible orders of projections, and permuting sets of $r!$ open regions must thus describe the order of projections upon the line described by an r-locus. In the example, the line $H(A, B)$ has an order of projections

$(AB)CDE$, which is reflected by the set of open regions $ABCDE$ and $BACDE$, a permuting set $((AB))CDE$ in A and B, and by their mirror image mates $EDCBA$ and $EDCAB$ which also constitute a permuting set $EDC((AB))$ in A and B. Hence the order of projections on the 2-locus $H(A, B)$ is $(AB)CDE$. Similar inspections of open regions permuting in A and E show $B(AE)DC$ for $H(A, E)$, $BA(CD)E$ for $H(C, D)$, and so on.

Each of the r-loci axes will then have the property of zero distance among projections for at least r stimuli. The choice among these axes can be made by finding that axis which tends to parallel the major axis of the least intersection.

Orthogonality of Axes

A third desideratum in choosing a solution from among the exhaustive possibilities is that the axes chosen be more or less orthogonal to each other. Obviously, the r-loci will not in general be orthogonal, so that in a choice among these for the set of axes, only approximate orthogonality can eventuate at best. However, there is an advantage in choosing those r-loci which will have even this approximate degree of orthogonality, in that we may at least be sure that our description of the space is as nonredundant as possible.

Just as there is a feature of the space which allows one to approach the quasi-simple-structure criterion easily, so are there guides to orthogonality as well. If any set of r points in general position defines an r-locus line, the subspace of $r - 1$ dimensions in which the r points are embedded is everywhere orthogonal to the line of the r-locus. Also, if two points define a line in the space, then the 2-locus defines a hyperplane of $r - 1$ dimensions such that any line in the hyperplane is orthogonal to the line between the two points. For example, in Figures 1 and 2, note how each 2-locus H is a perpendicular bisector of, and hence orthogonal to, a line.

Now suppose that the first axis is chosen to be that r-locus which is approximately parallel to the major axis of the least intersection. The remaining $r - 1$ axes should be approximately parallel to the minor axes of the least intersection if they are to be orthogonal to the first; in other words, the remaining axes should be guided by the 2-locus which is orthogonal to the axis of the least intersection. In the example, A and E were, of course, found to define the major axis; thus, the 2-locus $H(A, E)$ is orthogonal to this line. By taking the second axis parallel to this 2-locus, one insures that it will be approximately orthogonal to the first axis: in this instance, the second axis must show A and E projecting on the same point. Actually, in the example, the only possible 2-locus which would then qualify as an axis would be $H(A, E)$ with order of projections $B(AE)DC$. With dimensionality higher than two, however, there would be a choice among a number of r-loci showing A and E adjacent in order, where A and E are the endpoints of the major axis.

The procedure for locating axes beyond the first follows the same general plan. Only r-loci showing the endpoints of the first axis adjacent are considered. Then, always omitting one of the stimulus points which served as endpoints for the first axis, that pair of stimulus points is found which serve as axis points for a least intersection enclosing the largest number of points which do not fall between the points in question on the first axis. For instance, in the example, if the four points A, B, C, D are considered, then A and D form the major axis for those four points. However, A and D are not taken as endpoints for the second dimension, since both B and C fall between A and D on the first axis. On the other hand, if B, C, D, E are examined, it is found that only the circle generated by B, C, and E has all four stimuli either conjoint or interior to the circle. In such a case, any of the pairs BC, BE, or CE could serve as the major axis for this subset of points. However, D falls both between B and E and C and E on the first axis, so that only B and C apparently qualify as endpoints on the second axis. This is a trivial finding for a two-dimensional example, of course, since there is only one rank order which qualifies in the first place. Nevertheless, the procedure would be the same for higher dimensionalities, always locating higher axes in terms of central intersections for reduced numbers of points such that there is minimal duplication of previous axes' rank orders.

The end results of an analysis based on these criteria would be a set of rank orders, representing projections upon axes which parallel major axes of least intersections, which show the property of quasi-simple structure, and which are approximately orthogonal to each other. Obviously it is not possible to plot such axes and perform any sort of rotational operations upon such a solution. All possible solutions are immediately at hand in such data, and if the solution obtained is unsatisfactory for some reason, there are certainly others which may be chosen by abandoning one or all of these criteria. However, the criteria proposed here do seem to have some recommendation on common-sense grounds as well as by analogy to current practice in factor analysis. On the other hand, only properties of the isotonic space itself are relied on in these criteria for selecting among the possible solutions, and these criteria are presented here as isotonic principles *sui generis*, and not as approximations to results which might be found by metric methods. The analogies drawn to principal axes and simple structure are meant only to be expository and suggestive rather than exact.

The Problem of Incomplete Data

Any method for multidimensional unfolding has very limited practical utility as long as it is limited to the case of complete data. The number of isotonic regions which may exist even for a small number of stimuli in small dimensionality grows truly astronomical very quickly with increase in either. As was pointed out in [2], complete data cannot possibly be obtained except

in the most trivial cases. Our purpose in limiting this discussion to the case of complete data was, however, simply to make an exceedingly complicated topic a jot more comprehensible. In applications to real data some complexities do arise, mainly due to the fact that one does not necessarily have the exhaustive solution already implicit in the data in the incomplete case, and thus other steps must be introduced to supply the information given in the complete case by the mirror image pairs. However, the general ideas both for dimensionality and for a configurational solution may be applied to the case of either complete or very incomplete data, so that the principles enunciated here will be the basis for future discussion of the incomplete data situation.

REFERENCES

[1] Bennett, J. F. Determination of the number of independent parameters of a score matrix from the examination of rank orders. *Psychometrika*, 1956, **21**, 383–393.
[2] Bennett, J. F. and Hays, W. L. Multidimensional unfolding: determining the dimensionality of ranked preference data. *Psychometrika*, 1960, **25**, 27–43.
[3] Coombs, C. H. Psychological scaling without a unit of measurement. *Psychol. Rev.*, 1950, **57**, 145–158.
[4] Coombs, C. H. A theory of psychological scaling. *Engng. Res. Bull.* No. 34, Ann Arbor, Mich.: Univ. Michigan Press, 1952.
[5] Thurstone, L. L. *Multiple-factor analysis*. Chicago: Univ. Chicago Press, 1947.