

MULTIDIMENSIONAL UNFOLDING: DETERMINING THE
DIMENSIONALITY OF RANKED PREFERENCE DATA

JOSEPH F. BENNETT*

LINCOLN LABORATORIES, MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

AND

WILLIAM L. HAYS

UNIVERSITY OF MICHIGAN

A model is proposed which treats rankings given by a group of judges as representing regions in an isotonic space of dimensionality r . Three possible criteria for estimating lower bound dimensionality are discussed: mutual boundary, cardinality, and the occurrence of transposition groups. Problems associated with each criterion are mentioned.

The task of a psychological scaling technique is to search for some form of lawfulness, i.e., redundancy, in experimentally collected data. This redundancy, when it is present, permits a description of the items (also perhaps of the subjects) which is simpler than an exhaustive account of the response of every subject to every item, and yet tells the experimenter everything he wants to know about his data. If the scaling technique is the Guttman scale [5], for example, the experimenter hopes to find that the score of one of his subjects will tell him not simply how many items the subject passed but which items he passed, within some reasonable margin for error. Sometimes this scaling process is an end in itself; at other times the items or the people are scaled in order to "calibrate" them for application in some other context. In either case, the scaling technique chosen for a particular application must be appropriate to the task given the subjects, i.e., whether they were asked to agree with the items, pass them, rank them, or compare them in pairs, etc.

This paper is concerned with a scaling technique designed for the analysis of ranked preference data. The subject is asked to rank a group of items, for example, the names of hobbies. The model states that each hobby can be characterized by its position on each of several underlying attributes (e.g., scientific-artistic, solitary-gregarious, skilled-unskilled, etc.). The model states further that every subject can be characterized by his own maximum preferences on each of these attributes, and that he will rank

*Deceased.

the hobbies according to their increasing distances from the ideal hobby defined by his own maximum preference on each attribute; e.g., the scientific, gregarious, skilled subject will probably give model-airplane racing a high rank, photography a middling rank, and finger painting a low rank.

The critical problem in constructing such a model is the selection of a weighting scheme by which the subject is hypothetically supposed to combine the attributes in judging the distance of a given hobby from his own ideal hobby. The scheme proposed here is the simplest available: let the attributes be the axes of a multidimensional space, and interpret "distance" literally as the distance from the point representing the subject's ideal hobby, located by its projection on the axes, to another point representing one of the hobbies listed in the questionnaire.

Before elaborating these points, some questions which this introduction is likely to raise will be considered. First, is this model intended for application to the sort of preference data ordinarily analyzed by the method of rank-order [7]. No, it is not; there the task is to uncover an underlying order of popularity of the stimuli. The analysis presupposes that all subjects agree on this order and that all would give it, were it not for random errors. In other words, only the stimuli are being scaled, the people are not. The present model is usable only where subjects differ fundamentally in their preferences for the items, though they all view these items substantively the same within some common system of attributes. To choose another illustration, this model would be appropriate to the analysis of voters' preferences for political candidates if there were reason to suppose that each voter could be characterized by his position on, let us say, the conservatism-liberalism and isolationism-internationalism continua, and that he valued the available candidates in proportion as they approximated his own position on these continua. Within the "stimulus space," the voters might agree perfectly as to how the candidates *differ from each other*: it is required that the voters themselves differ only in their *preferences* among the candidates. In this system, as in Guttman scaling, the subjects as well as the items may be scaled.

Second, is it really necessary that these continua be agreed upon by all the subjects, that is, that everyone have the same conception of the liberality of Senator Jones or the gregariousness of photography? Yes, fundamentally it is, with some qualification since this method is nonmetric. It will deal throughout only with rankings, and will never attempt to determine *numerical* values for the distances between points; only relative distances in the form of rank orders will be considered. Within the relative freedom of a nonmetric model there is allowance for some minor differences between subjects about the structure of the space. But since the point of the method is to discover this common underlying structure on which the model supposes the subject's

responses to be based, if the structure is not there it will not be found and the method will fail. This is true of every other similar system, from Guttman scaling to factor analysis. Of course when one speaks of agreement between subjects about the structure of the space one does not mean explicit verbal agreement, any more than a subject taking an *IQ* test would be expected to be aware of the factorial compositions of the items. As in factor analysis, it is necessary to assume that the factorial composition of an item is the same for all subjects.

Origins in the Unfolding Method

It will be evident to the reader acquainted with scaling literature that the model proposed above is simply a multidimensional generalization of Coombs' method of unfolding [2, 3]. Coombs supposes that both subjects and items can be represented by points on a line segment (the single attribute under investigation). Each subject ranks all of the items in order of their increasing distances from his own position. Since he is concerned only with ordering the magnitudes of these distances and does not care about their directions, the subject might be said to pick up the continuum at his own position as one might pick up a piece of string, letting the ends swing together and fuse. The analytic task is the unfolding of these rankings (whence the name), in order to recover the original ordering of subjects and items on the continuum. The method used to accomplish this task, described in detail in the papers cited, requires no definition of "zero," "addition," or the "unit interval," and in fact employs no properties of the real line except the ordering of its points and the comparability of (some) intervals in magnitude. Internal checks provide that even these assumptions shall justify themselves in practice; that is, the unfolding method is a scaling criterion as well as a scaling method, and any given set of rankings may or may not unfold.

It is apparent that the only change introduced by the present authors is the replacement of Coombs' line segment by a space of dimension r , which may equal one and may be greater than one. The following section will explain a number of ways of determining r from the data. It will be assumed that distances in this r -dimensional space can be compared in magnitude, so that it is meaningful to say "it is farther from A to B than from C to D ," where A , B , C , and D are points in the space.

Some Definitions

Let the following definitions refer to the familiar Euclidean space, although (as the reader will see) all the Euclidean properties will not be needed.

The set C of subjects $c_1, \dots, c_i, \dots, c_m$ and the set Q of items or objects $q_1, \dots, q_i, \dots, q_n$ are regarded as sets of points in a space of dimension r . When a system of reference axes is inserted in the space, each subject

or object point can be characterized by an ordered r -tuple of real numbers, e.g., $q_i = (q_{1i}, q_{2i}, \dots, q_{ri})$, which are its projections on the axes. The task of each subject is to rank the whole set of objects according to their increasing distances from his own position.

Consider what position in space a subject's ideal must occupy in order to present a particular ranking of the objects. It is clear that relative to any two objects, A and B , a subject may make one of three reports: A is preferred to B , symbolized $A \succ B$ (which means that A is closer to the subject than is B); the converse, $B \succ A$; and third, that he is undecided between them, symbolized $A = B$, which means that the subject's ideal is equidistant from A and B . In one-space (that is, an infinite straight line), there is only one point exactly equidistant from two distinct points, which is called their midpoint. All the points on the line to one side of that midpoint will be closer to one of the objects, and all on the opposite side will be closer to the other object. Only a subject with an ideal lying exactly on the midpoint would report $A = B$, that is, indecision between the two. (See Figure 1.) In the unidimensional case explored by Coombs, the whole one-space is segmented by these midpoints into regions, within each of which every subject will report the same ranking of the objects, although within any one region different subjects' ideals may lie at different absolute distances from the objects.

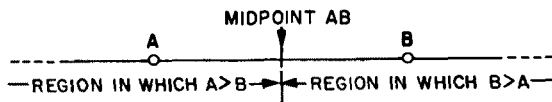


FIGURE 1
Midpoints and Regions Generated by Two Stimulus Points

In two-space (a plane), the locus of equidistance from two objects will be a line, the perpendicular bisector of the line segment connecting those objects. All subjects with ideals in the plane on one side of this line, the side containing the object A , will report $A \succ B$; all those on the opposite side will report $B \succ A$; only those exactly on the line will report $A = B$. In general the locus of equidistance from any two distinct objects will be a surface of dimension one less than the dimension of the total space in which the objects occur, i.e., a hyperplane. Hence the locus of equidistance from two distinct objects A and B is the boundary hyperplane $H(A, B)$. Since it will also be convenient to speak of loci of equidistance from sets of more than two objects, this notation will be generalized: $H(A, B, \dots, N)$ will symbolize the locus of equidistance from a set of n objects (A, B, C, \dots, N) in r dimensions, which will be a subspace (hyperplane) of some dimensionality less than r .

Between three objects (A, B, C) in two-space not all on the same line,

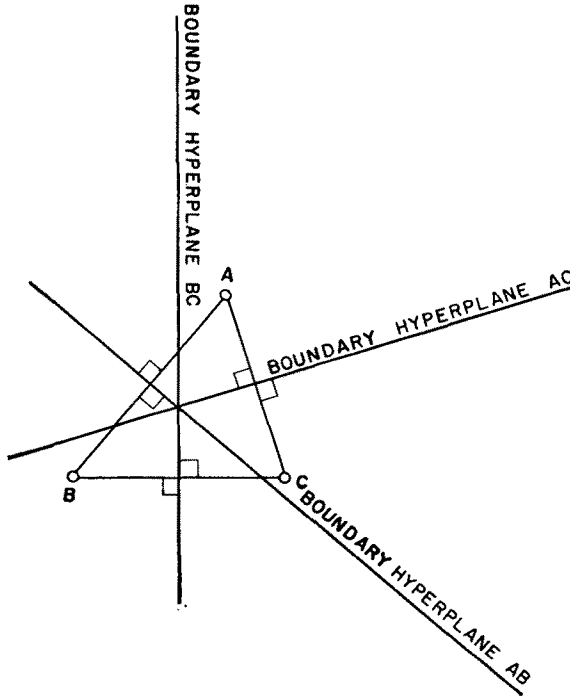


FIGURE 2

Boundary Hyperplanes Generated by Three Points in Two Dimensions

there will be three such boundary hyperplanes, $H(A, B)$, $H(A, C)$, $H(B, C)$, which will meet in a point equidistant from all three objects, $H(A, B, C)$ (see Figure 2). In the same fashion, any four non-coplanar objects in three-space (A, B, C, D) determine a point equidistant from all of them, $H(A, B, C, D)$ through which all six of their boundary hyperplanes of the form $H(A, B)$ must pass. In a space of r dimensions, there will always be one and only one point equidistant from $r + 1$ objects which are scattered—that is, which do not all lie on a subspace of the r -space.

These observations imply a general conclusion: in a space of r dimensions, if a locus of equidistance exists for some set of n points in general position (i.e., well scattered throughout the space), the locus will be a subspace of dimensionality $r - n + 1$. While this is actually a rather trivial result, it may be informative to sketch out a proof. It will be necessary to draw on a well-known feature of the geometry of higher spaces. The intersection of two r -dimensional subspaces S and T , neither of which is a subspace of the other, is a subspace of both S and T having $r - 1$ dimensions. The general outline for a proof using a form of induction argument on n , the number of points, would go as follows.

For a set of two points, A and B , embedded in r dimensions, consider a line L joining A and B . Now let there be some set of r orthogonal reference axes (X_1, X_2, \dots, X_r) such that X_1 is collinear with L . Let the origin be at the midpoint between A and B on L . The subspace everywhere orthogonal to X_1 must be a hyperplane of dimensionality $r - 1$. Now consider any point p which lies in this $r - 1$ hyperplane, such that p projects onto the origin on X_1 . It follows that the squared distance from p to A equals that from p to B as projected on X_1 . Furthermore, A and B each project onto the origin on the $r - 1$ hyperplane, so that the squared distances between p and A and p and B are equal as projected on the hyperplane. Thus, any point p in the hyperplane lies, by definition, in the locus of equidistance $H(A, B)$. Furthermore, *only* points in the hyperplane of $r - 1$ dimensions fit the definition of $H(A, B)$; otherwise the point p could not project equally distant from A and B on axis X_1 . Since the origin and particular set of reference axes utilized are completely irrelevant for distance among points in a Euclidean space, the locus $H(A, B)$ is a hyperplane of dimensionality $r - 1$; that is, $r - n + 1$ dimensions, $n = 2$.

Now assume the proposition true for $n - 1$ stimulus points in general position in r dimensions. A set of n stimulus points (A, B, \dots, Q, R) in general position in r dimensions may be divided into two overlapping sets of $n - 1$ stimulus points, (A, B, \dots, Q) and (B, C, \dots, Q, R). By assumption, the locus $H(A, B, \dots, Q)$ must be a space of dimensionality $r - (n - 1) + 1$, or $r - n + 2$, and $H(B, C, \dots, R)$ a space of the same dimensionality. Suppose that $H(A, B, \dots, Q)$ and $H(B, C, \dots, R)$ intersect. Since each of these loci is a subspace of dimensionality $r - n + 2$, their intersection must be a subspace of $r - n + 1$ dimensions. By the transitivity of the relation of equality, this subspace must also be the intersection of all the remaining loci of equidistance of $n - 1$ points drawn from the original set of n points, so that the $r - n + 1$ space is by definition the locus $H(A, B, \dots, Q, R)$. If the loci $H(A, B, \dots, Q)$ and $H(B, C, \dots, R)$ do not intersect, then by the same transitive property of equality, the locus of equidistance from the set of r points may not exist. This proves the proposition.

Regions

Any one boundary hyperplane $H(A, B)$ divides the whole space into two half-spaces, within each of which all points satisfy the same distance relation relative to the two objects A and B generating the hyperplane. Call one of these zones, that including the object A , the isotonic (same-ordered) region AB , (meaning $A \cdot \succ B$) and the other, the isotonic region BA . If the boundary $H(B, C)$ passes through the region AB , every point within the region AB which is also on the B side of the hyperplane $H(B, C)$ will lie in the order $A \cdot \succ B \cdot \succ C$ from the three objects, so that this isotonic

region may be designated ABC . Indeed, any isotonic region is the set of all points on the indicated side of each of several boundary hyperplanes, or (equivalently) the intersection of a particular set of larger isotonic regions, or (equivalently) the set of all points satisfying a given set of distance relations relative to the objects. For this last purpose, a convenient notation is that of partially ordered sets or posets, in which the relation $A \succ B$ is indicated by writing AB . The division of a two-space by three objects is illustrated in Figure 3. It is evident that every isotonic region, being bounded by hyperplanes, is everywhere convex, that is, any two points within the region can be connected by a straight line which does not pass outside the region.

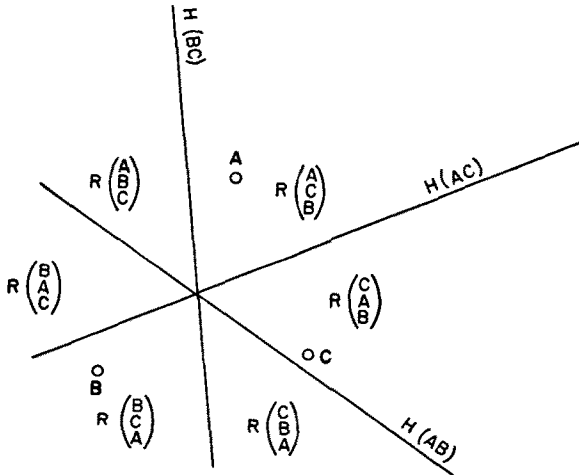


FIGURE 3
Regions Generated by Three Points in Two Dimensions

The minimum cells of the space, the regions out of which all other regions are constructed, will be those specified by a complete or simple ordering of all the objects, of the sort that the subjects in the space will make. Call these elemental regions. Thus region $A(BC)$ in Figure 3 is made up of the union of two elemental regions, ABC and ACB . The notation $A(BC)$ will be adopted here to symbolize the union of the regions ABC and ACB —similarly the notation $A(BC)D$ will mean the union of the regions $ABCD$ and $ACBD$, $(ABC)D$ will indicate the union of the regions $ABCD$, $BACD$, $BCAD$, $CBAD$, $CABD$, $ACBD$, and so on.

Any region will wholly contain any other region whose defining poset wholly satisfies the poset of the first, with the addition of some further refinement; thus the region AB must contain the region $A(BC)$ if the latter exists. On the other hand, two regions cannot intersect, that is, have some

subregion in common, if their posets contain both a relation and its complement. Thus the region $A(BC)$ cannot intersect the region BA because no point in the space could obey simultaneously the contradictory orders of distance AB and BA , that is, lie simultaneously on both sides of the hyperplane AB .

When $k + 2$ objects (or more) occur in a k -space, then certain of the $(k + 2)!$ possible elemental regions will disappear. For example, in Figure 4, representing three objects in one-space, there is no region ACB and no CAB , i.e., the region $(AC)B$ present in Figure 3 has disappeared. It is not an accident that those two particular elemental regions are missing. Their absence was dictated by the dimension of the space and the configuration of the objects in a way which the following sections will define. This dependence will become the analytic tool of the method, making it possible to reconstruct dimension and configuration from experimentally obtained rankings.

Determining Dimension

The first problem to be considered is that of finding r , the dimension of the space, given the rankings produced by the subjects. It is not claimed that the three methods suggested here exhaust all the possibilities of dimensional analysis, or even that they are best; they simply recount as much as the authors know about the problem at the time of writing.

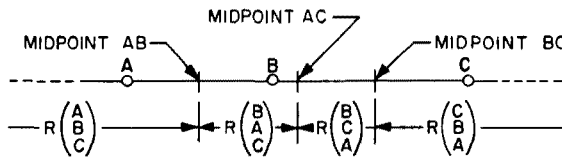


FIGURE 4

Rank Orders Associated with Regions Generated by Three Stimulus Points in One Dimension

Dimension by Mutual Boundary

In the light of the definition of an isotonic region as the set of all points satisfying a given set of distance-relations relative to the objects, the origin of the term boundary hyperplane is clear. Every isotonic region is bounded by segments of boundary hyperplanes, that is, by $(r - 1)$ -dimensional cells. In Figure 4, for example, the region BAC is bounded by the hyperplanes $H(A, B)$ and $H(A, C)$. In Figure 3, region $A(BC)$ is bounded by segments of the hyperplanes $H(A, B)$ and $H(A, C)$. Note that it is not bounded by the hyperplane $H(B, C)$. This is because the poset defining the region $A(BC)$ specifies no distance-relation between B and C , so that a moving point inside $A(BC)$ might reverse its distance-relations with B and C —that is, it might pass through the boundary hyperplane $H(B, C)$ —without leaving

the region $A(BC)$. Of necessity, every hyperplane bounding a given region is represented by a pair of objects which are adjacent in the defining poset of that region. (Two elements of a poset will be said to be adjacent if one precedes the other in rank and if there is no other element in the poset immediately succeeding the first and preceding the second.) However, the converse that every hyperplane defined by a pair of objects adjacent in the poset actually bounds the region is not always true. For example, in Figure 4, region ABC is bounded only by the hyperplane $H(A, B)$. In all such cases, any remaining hyperplane defined by an adjacent pair, such as BC in Figure 4, constitutes the boundary of a larger region containing that region.

Two isotonic regions bound if they do not intersect, that is, there is no point common to both of them, and the same $r - 1$ cell is part of the set of $r - 1$ cells bounding each. Thus in Figure 4 region ABC bounds the three regions BAC , $B(AC)$, and BA . Since all isotonic regions are everywhere convex, two isotonic regions can bound only on a single hyperplane or on entirely coincident hyperplanes. Hence two isotonic regions must bound, granted they both exist, if the defining poset of one can be transformed into the poset of the other by the reversal of a single adjacent pair of objects; this adjacent pair represents the hyperplane which separates them. The possibility of coincident hyperplanes does not permit concluding that isotonic regions bound only if such a single-pair transposition of their defining posets is possible. If hyperplane coincidence were to occur, passing from one region into the other would mean simultaneous passage through two (or several) hyperplanes, and consequently the simultaneous reversal of one (or several) adjacent pairs. However, the existence of coincident hyperplanes could be

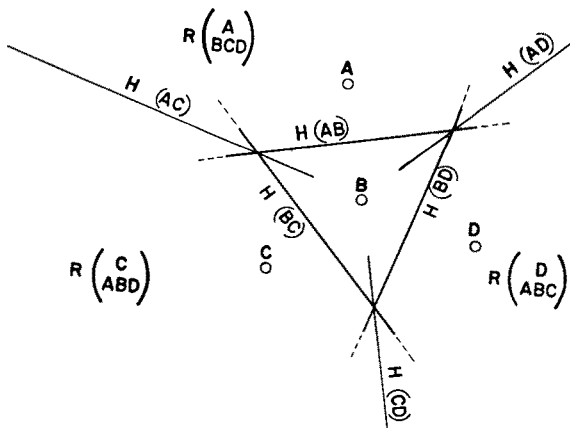


FIGURE 5
Regions Generated by a Configuration of Four Points in Two Dimensions

readily detected in the total set of rank orders returned by all the judges by a complete mutual implication of the defining posets of their isotonic regions.

The value of the concept of boundary lies in its potential application as an initial dimensional criterion to indicate whether the space may be of dimension one, two, or higher. In a space of dimension one, it is not possible for more than two regions to bound one another, and no more than four isotonic regions may all bound one another in two-space. (Of course this famous "four-color" conjecture has not actually been proved [4], but it may probably be used without any great anxiety of its imminent disproof.) The two-space limit is illustrated in Figure 5, where four regions all bound one another along cells of hyperplanes. The reader may convince himself that it is impossible to add a fifth region which will bound all of the existing four.

A hypothetical example may help to clarify the criterion of mutual boundary. Imagine that a group of judges, in evaluating four objects, return the following rank orders.

<i>ABCD</i>	<i>ACBD</i>	<i>ADBC</i>	<i>BCAD</i>	<i>CDAB</i>
<i>ABDC</i>	<i>ACDB</i>	<i>ADCB</i>	<i>CBAD</i>	<i>CDBA</i>
<i>BACD</i>	<i>CABD</i>	<i>DABC</i>	<i>CBDA</i>	<i>DCAB</i>
	<i>CADB</i>	<i>DACB</i>		<i>DCBA</i>

The presence of *ABCD* and *BACD* in these data is evidence, for example, that $A(BCD)$ bounds $B(ACD)$. Note that in four cases, sets of three regions all mutually bound. These are $C(DA)B$, $D(CA)B$, $A(DC)B$; $C(AB)D$, $C(AD)B$, $C(BD)A$; $A(CD)B$, $A(BD)C$, $A(BC)D$; $A(BC)D$, $B(AC)D$, $C(AB)D$. This excludes at once the possibility that the solution is unidimensional. On the other hand, a two-dimensional solution is not excluded, since there is no instance where five or more regions mutually bound. In several instances sets of four regions do mutually bound—for example, $(AB)(CD)$, $(CB)(AD)$, DB , and $(CA)BD$ all bound one another.

It has not so far proved possible to extend the mutual boundary approach into spaces of dimension higher than two. Fortunately the existence of more powerful methods renders this unnecessary. The method may continue to serve as a quick check of the possibility of a solution in one or two dimensions, and as a first outline of the configuration of objects in the space.

Dimension by Cardinality

The boundary method just outlined operates by the determination of a lower bound to the dimensionality of the space; that is, by concluding that a given set of rankings must have been given by judges in a space of dimension "at least k , and perhaps higher, but not less than k ." In the same fashion, the following method will serve to impose a lower bound on the possible dimension of the space by a comparison of the total number of different

rankings returned by the judges, with $C(n, k)$, the maximum possible number of elemental isotonic regions that can be generated by n objects in k dimensions. Thus Coombs has shown that four objects can generate no more than seven rank orders in one-space; if eight or more distinct rankings were experimentally obtained, they could not all be accommodated in a uni-dimensional solution.

In one-space, there will be one more elemental region than boundary hyperplanes (i.e., midpoints). Since there will be $\binom{n}{2}$ boundary hyperplanes for n objects, the maximum cardinality for n objects in one-space is

$$(1) \quad C(n, 1) = \binom{n}{2} + 1.$$

Special circumstances, such as the coincidence of boundary hyperplanes, can serve only to lower the number of regions in the space. Consequently, in determining $C(n, k)$, assume as before that the objects are scattered so that no such special configuration can occur.

In determining $C(n, k)$ for spaces of dimension $k - 1$, Professor K. E. Leisenring of the University of Michigan (personal communication) has pointed out that when the objects are scattered, every boundary hyperplane will be subdivided into isotonic regions by boundary hyperplanes in the same fashion as a complete space of one fewer dimension and containing one fewer object. Leisenring's observation may be employed to determine $C(n, k)$ by considering that the addition of a new object will add as many new boundary hyperplanes as there were objects already in the space—one between each of the previous objects and the new member. Each of these new hyperplanes will, in turn, be responsible for creating as many new regions as it intersects, since each of the intersected regions will be cut into two. But these regions may be readily counted by the number of $(k - 1)$ cells in each new hyperplane, which, by Leisenring's observation, will be equal to $C(n - 1, k - 1)$. Consequently,

$$(2) \quad C(n, k) = C(n - 1, k) + (n - 1)C(n - 1, k - 1).$$

No satisfactory nonrecursive expression for (2) has been found. However, Dr. R. M. Thrall in a personal communication has pointed out the identity of the values obtained from it with sums of absolute values of Stirling numbers of the first kind [6], for which also no general expression exists. The relation may be written

$$(3) \quad C(n, k) = \sum_{m=n-k}^n |S_n^m|,$$

where S_n^m is a Stirling number. Some values of $C(n, k)$ are given in Table 1. The reader will note that when $k = n - 1$, $C(n, k) = n!$. That is, a space of

dimension $n - 1$ can always account for all possible orderings of n objects, just as in components analysis a factor space of n dimensions can always account for all the variance in n tests. Of course such a solution is trivial; a solution is not counted successful unless it accomplishes some economy of

TABLE 1
Maximum Number $C(n,d)$ of Rank-Orders
Generated by n Objects in d Dimensions

n	d				
	1	2	3	4	5
1	1	1	1	1	1
2	2	2	2	2	2
3	4	6	6	6	6
4	7	18	24	24	24
5	11	46	96	120	120
6	16	101	326	600	720
7	22	197	932	2.556×10^3	4.320×10^3
8	29	351	2.311×10^3	9.080×10^3	2.221×10^4
9	37	583	5.119×10^3	2.757×10^4	9.485×10^4
10	46	916	1.037×10^4	7.364×10^4	3.430×10^5
11	56	1.376×10^3	1.953×10^4	1.773×10^5	1.079×10^6
12	67	1.992×10^3	3.466×10^4	3.921×10^5	3.030×10^6
13	79	2.796×10^3	5.857×10^4	8.080×10^5	7.735×10^6
14	92	3.823×10^3	9.491×10^4	1.569×10^6	1.824×10^7
15	106	5.111×10^3	1.484×10^5	2.898×10^6	4.021×10^7
16	121	6.701×10^3	2.251×10^5	5.125×10^6	8.368×10^7
17	137	8.637×10^3	3.323×10^5	8.726×10^6	1.657×10^8
18	154	1.097×10^4	4.791×10^5	1.438×10^7	3.140×10^8
19	172	1.374×10^4	6.765×10^5	2.300×10^7	5.728×10^8
20	191	1.701×10^4	9.376×10^5	3.585×10^7	1.010×10^9

description by accommodating all or nearly all the data in comparatively few dimensions.

There is an obvious similarity between the cardinality criterion just discussed and the dimensionality criterion proposed by Bennett [1] for the problem of analyzing rank-orders of *subjects* given by *tests*, within the general model of factor analysis. Bennett's model is restricted exclusively to the

case in which the rank orders given by tests are monotone, as seems proper in the factor analysis context. This restricts the possible isotonic regions strictly to open regions, by the condition that for two subjects A and B (not stimulus objects as here) $H(A, B)$ shall pass through the origin of the space.

On the other hand, there is no such monotonicity assumption made in the model of preferential choice used here, as, unlike test performance, individual preference may reasonably be thought of as non-monotone. Thus, the possibility of both open and closed regions is allowed. In consequence, other things being equal, the maximum cardinality numbers given in Table 1 are always greater than or equal to those given in Bennett ([1], p. 388). However, Table 1 from [1] could be used if there were some way of knowing from the data that only open regions were represented, as his table does give the maximum number of such open regions which may occur in this model. Since each mirror-image region in a space must have a mirror image, another region which has the exact reverse of its rank order, and since only open regions show such mirror images, one might supplement the cardinality criterion in the following way. Count the number of mirror-image pairs of rankings existing in the data, and compare them with one-half the number of regions listed in Bennett's table for the appropriate n and a given r . If this number is exceeded, then the dimensionality is greater than r .

In the hypothetical example given previously, eighteen rankings were returned by the subjects. Reference to Table 1 shows this to be just the maximum allowable number for four objects in two dimensions; hence there is no reason as yet to believe that the dimension of the space is higher than two.

The method of cardinality has the peculiar feature that it does not consider what rankings were returned. This gives it a great advantage in simplicity and ease of application, but of course it also makes it a very insensitive test, likely to give an optimistically low estimate of dimension when applied to a chaotic set of rankings. Furthermore, Table 1 indicates that in successively higher dimensions the number of different possible rankings goes up very rapidly with the number of items, and in most practical instances is likely to exceed the whole size of a sample of judges which an experimenter might use. This means that the method of cardinality can be applied directly (in the form of a direct comparison of the number of distinct rankings obtained with the entries in Table 1) only to data (such as repeated psychological judgments) in which the number of experimentally independent rankings is large, preferably much exceeding the factorial of the number of objects ranked. In most other circumstances, one cannot expect to have enough experimentally independent rankings to exhaust all the permissible rankings tabulated in the appropriate cell in Table 1.

Dimensions by Groups

A space of r dimensions contains every possible ranking of any $r + 1$ objects, or expressed another way, contains the complete transposition group of simply ordered sets of those objects, provided only that the objects are scattered. This observation is independent of the presence in the space of objects other than the $r + 1$; that is, the regions in question may be either elemental regions, or the $(r + 1)!$ larger regions determined by the simple orderings of the $r + 1$ objects. On the other hand, a space of dimension r cannot contain the complete transposition (permutation) group over $r + 2$ objects or more.

These ideas may be summarized in the following proposition. A set of r stimulus points (A, B, \dots, Q, R) may be embedded in a space of no fewer than $r - 1$ dimensions if and only if a set of $r!$ regions exists characterized by all $r!$ permutations in order of the stimulus set. A proof of the necessary condition might be carried out by induction on r , the number of stimuli and $r - 1$, the number of dimensions. This is trivial for $r - 1 = 1$, since two points may always be put onto a line with center halfway between them, generating two regions showing the two permutations.

Now assume the proposition to be true for any $r - 1$ stimuli in $r - 2$ dimensions. Consider the set of r stimuli (A, B, \dots, Q, R), for which, by hypothesis all $r!$ permutations in order exist as regions. If this is true, then all $(r - 1)!$ permutations in order of the set of $r - 1$ points (A, B, \dots, Q) must also exist as regions, and this set of $r - 1$ points is embedded in no less than $r - 2$ dimensions. Now suppose that stimulus point R is also embedded in the same subspace of $r - 2$ dimensions so that all $r!$ permutations correspond to regions in $r - 2$ space. This would contradict the cardinality rule given above, which shows that the maximum number of regions for r stimuli in $r - 2$ dimensions is always less than $r!$. Hence, the dimensionality would be greater than $r - 2$, and the set of r points may all be embedded in a space of no fewer than $r - 1$ dimensions.

For the sufficient condition, a proof may once again be outlined using an induction argument on r and $r - 1$. The case of $r = 2$ is trivial. Now assume the sufficient condition true for $r - 1$ points in $r - 2$ dimensions. By hypothesis, the set of r points (A, B, \dots, Q, R) requires $r - 1$ dimensions. Thus, some subset of $r - 1$ points from this set, say (A, B, \dots, Q), requires $r - 2$ dimensions, so that by assumption all permutations in order for this set must exist as regions in the $r - 2$ space. The locus $H(A, B, \dots, Q)$ will exist as a point in the $r - 2$ space, by the rule of $r - n + 1$ for the dimensionality of a locus of n points. By the same rule, when the set of $r - 1$ points is embedded in an $r - 1$ space, the locus of equidistance would be a line L bounding all $(r - 1)!$ regions showing the permutations in order for the set.

Since the addition of stimulus point R to the space requires, by hypothesis, dimensionality $r - 1$, so that R may not be embedded in the $r - 2$

space, the intersection of $H(A, R)$, for example, and $H(A, B, \dots, Q)$ must exist—otherwise the required dimensionality would be $r - 2$. Furthermore, by the transitivity of the equality relation, this point of intersection must also be the point of intersection of $H(A, B, \dots, Q)$ with all of the remaining loci $H(B, R), H(C, R), \dots, H(Q, R)$, so that by definition the intersection is the locus $H(A, B, \dots, Q, R)$. Now any region showing a permutation of (A, B, \dots, Q) such as $ABC \dots Q$ must be bounded by the line L . The intersection of L with $H(A, R)$ creates two new regions such as $RABC \dots Q$ and $ARBC \dots Q$, the intersection of $H(BR)$ with L another two new regions such as $ARBC \dots Q, ABRC \dots Q$, and so on. Since the intersections of a basic region showing a particular permutation of A through Q with the $r - 1$ hyperplanes of the form $H(A, R)$ must generate r distinct new regions, each showing the same permutation of A through Q but a different position of R , and since there were originally $n - 1$ basic regions showing permutations of the set of $r - 1$ stimulus points, exactly $r(r - 1)!$ or $r!$ regions must be generated showing all permutations in order of (A, B, \dots, Q, R) .

This fact can be made the final (and perhaps the most useful) criterion

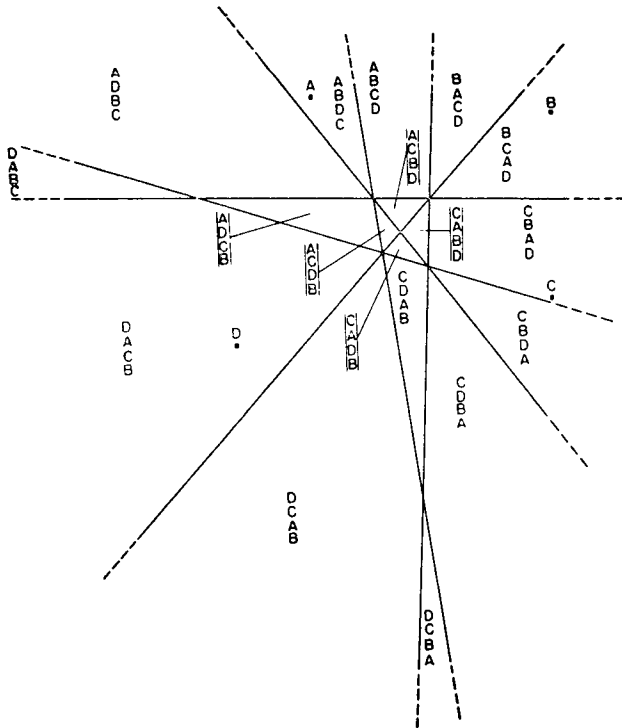


FIGURE 6
Regions Generated by a Second Configuration of Four Points

of dimension. One may be certain that the minimum dimension of the space in which a complete solution may be realized must be one less than the number of elements in the largest transposition group present in the experimental data.

Returning again to the hypothetical example, the complete group over all four objects is not present, since such a group would have 24 elements and only 18 are present. Searching for groups over subsets of three objects, one finds, for example, every permutation of A , B , and C :

region ABC , represented by $DABC$, $ADBC$, $ABDC$, and $ABCD$;
 region BAC , represented by $BACD$ alone;
 region BCA , represented by $BCAD$ alone;
 region CBA , represented by $DCBA$, $CDBA$, $CBDA$, and $CBAD$;
 region CAB , represented by $DCAB$, $CDAB$, $CADE$, and $CABD$;
 region ACB , represented by $DACB$, $ADCB$, $ACDB$, and $ACBD$.

The same is true of each of the other subsets of three objects, as the reader will discover on examination. Since there is considerable evidence that the solution is two-dimensional, it is possible to attempt a geometric realization of the space. Such a construction is given in Figure 6. Using this construction one may reexamine the previous examples and relate them to the properties of the figure.

Of these three criteria for dimensionality, the groups criterion just discussed is the most practical and, to large extent, the most sensitive for use with preference data. Obviously, with a large number of stimulus items, say 10, it is seldom possible to accumulate enough data to establish the lower limit of the dimensionality at nine since this would require some 10! distinct rankings.

However, it is possible to make a good lower bound estimate from more limited data using a modification of this criterion. Given n objects (A , B , \dots , Q , R), the required dimensionality is $n - 1$ if there exist *two* complete sets of $(n - 1)!$ permutations in order for some subset of $n - 1$ objects, say (A , B , \dots , Q), such that the remaining object R *precedes* all of the remaining objects in each permutation in the first set, and *follows* all of the remaining objects in each permutation of the second set. This is seen to be true if it is recalled that the locus of equidistance from $r - 1$ points in $r - 2$ dimensions is a point, so that there may be only one set of "permuting" regions bounding such a point. If there are two distinct sets of this kind, then the locus in question must be at least a line, and the dimensionality must be at least $r - 1$. This modification requires only $2(r - 1)!$ rank orders to establish a lower bound dimensionality of $r - 1$, rather than the complete $r!$.

Finally, a very stringent criterion is obtained by a combination of the principles utilized in the cardinality and the groups criteria. It will be recalled

that the locus of equidistance between two points in r dimensions is a space of dimensionality $r - 1$. Each and every pair of regions with rank orders differing only by a reversal in order of *one* pair of objects must be separated only by the locus. Furthermore, for n points in r dimensions, the number of such pairs differing only by a reversal in the same object pair (the same hyperplane bounds each pair) must be the cardinality of $n - 1$ points in $r - 1$ dimensions. Thus, if a set of such pairs is found, and the number of such pairs exceeds the maximum cardinality of $n - 1$ points in $r - 1$ dimensions, then the dimensionality must be at least $r + 1$.

For instance, in Figure 6, there are exactly four pairs of regions differing only in the order of A and B : $ABCD-BACD$, $CBAD-CABD$, $CDAB-CDBA$, and $DCBA-DCAB$. Thus the cardinality of these pairs of regions agrees exactly with the cardinality for three objects in one dimension—there is no evidence from these orders that the dimensionality exceeds two. However, had a fifth pair such as $DBAC-DABC$ existed, then the dimensionality required would be three.

This last criterion is, of course, the most sensitive of those discussed here, particularly with small numbers of stimulus objects. Actually, it may be too sensitive for use with fallible data such as would reasonably be obtained, since the estimated dimensionality depends somewhat more upon single occurrences among the rank orders than one would expect with the groups criterion alone. Thus, in practice, the application of the groups criterion is perhaps the method of choice for a “manageably low” dimensionality estimate.

Incidentally, the method for describing configuration for such data is also based principally on the incidence of such permutation groups among the rank orders, so that once a preliminary estimate has been made for the required dimensionality, there are continual checks upon the estimate in the remainder of the procedure.

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] Coombs, C. H. Psychological scaling without a unit of measurement. *Psychol. Rev.*, 1950, 57, 145–158.
- [3] Coombs, C. H. A theory of psychological scaling. Engineering Res. Inst. Bull. No. 34, Ann Arbor: Univ. Michigan Press, 1952.
- [4] Errara, A. *Du coloriage des cartes et de quelques questions d'analyse situs*. Paris: Gauthier-Villars, 1921.
- [5] Guttman, L. A basis for scaling qualitative data. *Amer. sociol. Rev.*, 1944, 9, 139–150.
- [6] Jordan, K. *The calculus of finite differences*. New York: Chelsea, 1947.
- [7] Thurstone, L. L. Rank order as a psychophysical method. *J. exp. Psychol.*, 1931, 14, 187–201.

Manuscript received 1/3/59

Revised manuscript received 8/30/59