Chapter 6

The Theory of Multidimensional Scaling

Though widely used until the 1950s, Thurstone's approach to psychological measurement gradually gave way in the late 1960s to a new approach: *Multidimensional scaling*. One of the features that came to disturb psychologists about Thurstone's theory was its commitment to unidimensionality. Although in relatively well controlled psychophysical experiments the unidimensionality of the stimulus set could be ensured, in other areas, like social psychology, there was little control over the dimensionality of the stimulus set. Here the stimuli were attitude statements, political candidates, and so on. It was difficult to identify the relevant dimensions, let alone control them. The methods of multidimensional scaling had the advantage of leaving the issue of dimensionality open, it was one of the parameters determined by the data rather than by the researcher.

This shift in method also involved a change in the kind of judgments used. Thurstone's method made use of pair comparison judgments and his theory was a theory of such judgments. Multidimensional scaling makes use of similarities data. The most common form of similarities data used are direct judgments of similarity or dissimilarity between stimulus pairs, although other, less direct kinds of data may also be used. For example, a theory may prescribe that the time taken to discriminate between a pair of stimuli is an index of their dissimilarity, or that the proportion of times two stimuli are confused is an index of their similarity.

The concept of similarity is a commonly used one and, at first sight, seems clear enough. We often make judgments of similarity, for example, the judgment that the English language is more similar to German than to Russian or, that the music of Mozart is more like that of Bach than like that of Delius. Such judgments are taken to be facts. That is, similarity is taken to be an objective relation between things. If it were not such an objective relation then there would be no reason to suppose that in judging similarity any two people would be judging the same kind of thing. This would make construction of a theory of similarities judgments a difficult business.

The theory of multidimensional scaling offers a simple explanation of similarities judgments. The similarity of one thing to another is said to be inversely related to the distance between them in a multidimensional attribute space. All things possess attributes, and their similarity or difference to other things is related to the difference between them on these attributes. Each attribute may be thought of as a dimension in a multidimensional space, and each object as a point whose location is defined by the object's value on each attribute. The closer two objects are in this attribute space the more similar they are and vice versa. Hence, judgments of similarity are monotonically related to distance in the multidimensional attribute space. The theory says that x will be judged more similar to y than w is to z, if and only if the distance from x to y in this space is less than that from w to z.

Simple as this idea is, it was not until the early 1960s that methods were devised for multidimensional scaling. Such methods take as their starting point the judgments of similarity or dissimilarity, and conclude by constructing a space wherein the objects judged may be located, so that judged similarity and distance within the space are inversely monotonically related. The major external constraint is to keep the dimensionality of the space as low as possible and most methods allow some slippage in the monotonicity requirement in order to attain a solution of satisfactorily low dimensionality. The methods most commonly used (due to Guttman, 1968, Kruskal, 1964 a & b, & Shepard, 1962 a & b), are complex, lengthy, and best implemented via electronic computers. No account of them will be given in this book and the interested student is referred to these articles. Simple sketches of some of the methods are given in Baird and Noma (1978) and van der Ven (1980). Our interest here is in the theory behind the method, for it is (i) only if this theory enables measurement, and (ii) only if its quantitative aspects can be experimentally tested, that the methods of multidimensional scaling will be of any value to psychology.

An illustration of the application of multidimensional scaling is the study by Wish, Deutsch, and Biener (1970) (also reported in Carroll & Wish, 1974). In that study the subjects were instructed to rate the similarities of each pair of 12 nations on a 9 point rating scale. This gave a 12 × 12 matrix of similarities ratings for each subject, which the multidimensional scaling method was then applied to. The particular method used in this study was one devised by Carroll and Chang (1970) called INDSCAL. It produces different configurations of stimuli in multidimensional space for different subjects. The presumption is that while all subjects judge similarity on the basis of the same attributes, they may differ in the importance they attach to different attributes. Figure 6.1 shows the two-dimensional solutions obtained for two of the subjects in this study (from Carroll & Wish, 1974).

Dimension one was interpreted by the authors as reflecting how procommunist

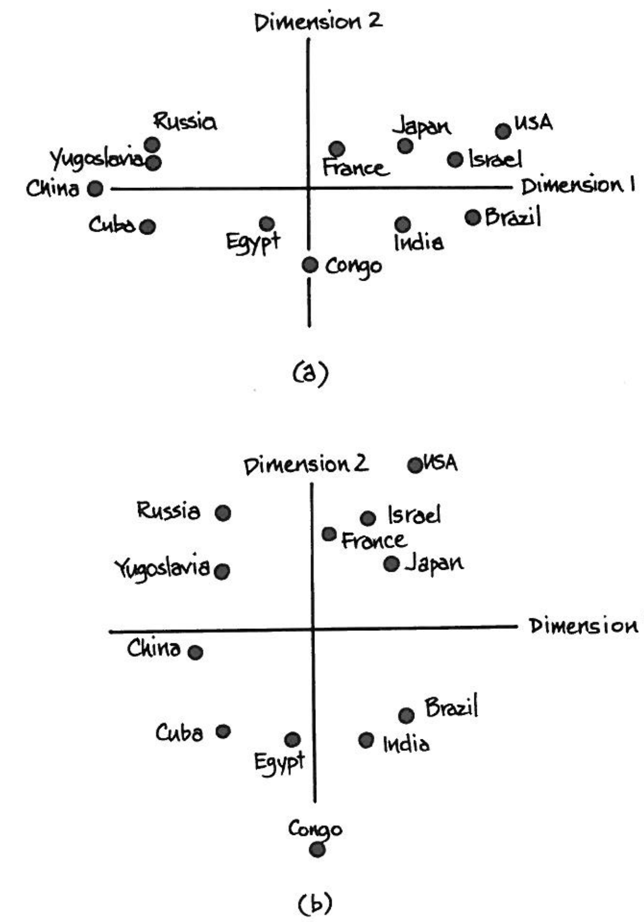

FIG. 6.1. Two dimensional solutions for two subjects (a and b) of their judgments about the relative similarities of 12 nations, using the INDSCAL program. (Based upon Fig. 10 in Carroll and Wish (1974). Adapted by permission of Academic Press and the authors.)

or prowestern the nations were; dimension two as how economically developed or underdeveloped they were. The two subjects differed according to how much importance they attributed to these different dimensions in judging similarity. Subject (a) saw the USA as being more similar to Brazil than to Russia because he gave more weight to political than to economic differences, whereas (b) reversed the weights and, saw the USA as more like Russia than like Brazil.

The appeal of multidimensional scaling resides in its reduction of a matrix of data to a simple geometric picture that allows one (after interpretation) to see at a glance, both the dimensions used and the similarities relations involved. Such matters of apparent convenience, however, cannot hide the substantive questions that such pictures pose. Are broad political and economic differences between nations quantitative? Is similarity really like distance in a multidimensional space? Because the methods of multidimensional scaling always ensure a solution, answers to these questions must depend on research into the theory underlying the method. The fact that an interpretable solution is obtained, given a set of similarities data cannot, itself, validate the method.

THE THEORY OF ORDINAL MULTIDIMENSIONAL SCALING

The theory may be summed-up in the following five propositions.

(1) There are k variables relevant to the judgments of similarity (where k is some natural number).

(2) All of these k variables are quantitative.

(3) These k quantitative variables form a space.

(4) The distance between any pair of points, x and y, in this k-dimensional space ($d(x,y)$) is given by the following expression:

$$d(x, y) = \left[\sum_{i=1}^{k} |x_i - y_i|^r \right]^{1/r}$$

(where x_i and y_i are the values of x and y on the ith dimension and r is a real number greater than or equal to one).

(5) Judged similarity or dissimilarity between any two stimuli, x and y, is monotonically related to the distance between them in the k dimensional space ($d(x,y)$) as follows:

(a) the judged similarity of x to y increases as $d(x,y)$ decreases;

(b) the judged dissimilarity of x to y increases as $d(x,y)$ increases.

The significance of each of these propositions will be discussed in turn. In proposition 1 the k variables relevant to the judgments of similarity will, of course, be variables upon which the objects judged possess values. That proposition, is then, hardly a startling hypothesis. Every object possesses many, if not an infinite number of attributes. However, not all of these are relevant to judgments of similarity. At least two factors restrict the relevant attributes to a relatively small class. First, because of the limitations of our sensory systems, we are only sensitive to a small number of attributes possessed by objects. For example, by looking at an object we cannot usually directly perceive its temperature or density and so when judgments are based on vision only, these attributes would not be relevant. Second, whatever personal interests dominate our judgments of similarity, these may further restrict the class of relevant variables: We will attend only to some of the features we might have sensed, ignoring the rest. For example, in judging the similarities of different cheeses we might attend only to attributes of flavor, odor and texture, ignoring color, weight, and temperature. So a complete theory of similarities judgments would specify which attributes are relevant in which contexts. This the theory of multidimensional scaling does not yet attempt to do. Rather the method is seen as a means whereby these relevant attributes may be discovered.

Proposition 2 limits the range of application of the theory to situations where the k variables involved are all quantitative. Because some of the variables upon which similarity judgments are based are not quantitative (e.g., hue and texture), the theory is not completely general. This, of course, has not prevented psychologists from applying methods of multidimensional scaling in contexts where the variables involved are not quantitative, or have not been shown to be quantitative. In fact it is within this latter context that multidimensional scaling is thought to be able to make a contribution. It is felt by many psychologists that lying behind judgments of similarity between many stimuli (such as political candidates, attitude statements, word meanings, etc.) are unmeasured quantitative variables of a psychological kind. So it is supposed that multidimensional scaling provides a means not only of identifying, but also, of measuring them. Certainly, if there are such quantitative variables and the theory of multidimensional scaling is true then this supposition is true. These, however, are big "ifs" and they are precisely the ones in which we are interested. As will be shown in the next section, the theory of conjoint measurement shows how we may come to grips with them.

Proposition 3 is one that is rarely brought into the open. It is, of course, implicit in Proposition 4, for the concept of distance relative to more than one variable, only has meaning if the variables jointly constitute a space. The mere fact that objects differ on k quantitative variables does not itself imply that each object has a location in a k-dimensional space. That conclusion requires the existence of quantitative relations *between* the k variables. These quantitative relations must be of a particular kind if the concept of distance is to be applicable. Distance between points in a space is always a function of component distances within each of the dimensions. If V is any quantitative variable and x and y objects possessing values of V, then the distance between x and y on variable V is $|x_v - y_v|$, the absolute value of the difference between x and y on V. Now, the multidimensional or spatial concept of distance is some function of these component distances and so, since a function is a kind of relation, there must be some relation between the component distances. In the case of the theory of multidimensional scaling the distance is an additive power function of the component distances and this requires that the component distances stand in some kind of additive relation to one another. That is, it requires for any two variables involved, U and V, that any difference, $s_u - t_u$, on U be r times any difference on V, $p_v - q_v$ (where r is some real number).

This is easily seen by considering a hypothetical space constituted by the variables height and weight. Each person, possessing both a height and a weight, will have a point in this space. But how are distances within the space related to distances within height and weight. If we let a difference of one inch equal a difference of one pound, then the distance between points a and b equals that between points c and d (see Fig. 6.2a). However, if the space is reconstituted so that a difference of one centimeter equals a difference of one kilogram then, as Fig. 6.2b shows, the distance from c to d exceeds that from a to b. As is obvious from this figure, in order to have a height/weight space (and hence height/weight distances between people) differences in height must equal some factor times difference in weight.

The physical space wherein we live, is of course, our paradigm of a space, and within that space the three dimensions, height, width, and depth, are quantitatively comparable because they each involve the same quantity (length) and differ only in direction. In recent times, it has been discovered that these three dimensions are yoked with time to form a four dimensional space. That is, the spatial dimensions (height, width, and depth) and time are interrelated so that there is such a thing as the space–time distance between any two events. Similarly, if attribute spaces for similarity judgments exist then the attributes (or variables)

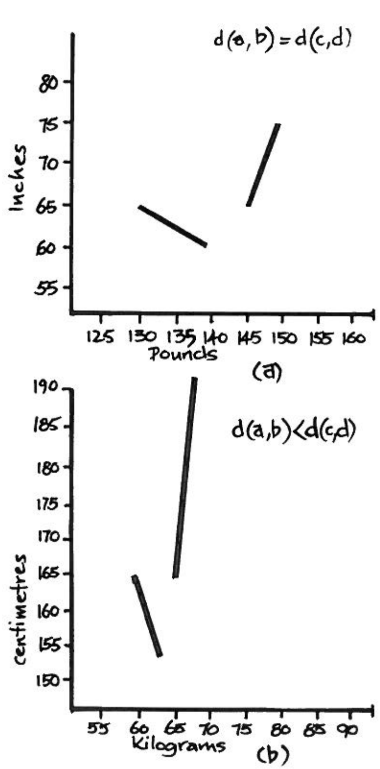

FIG. 6.2. (a) A height-weight space in which 1 inch equals 1 pound. Then $d(a,b) = d(c,d)$. (b) A height-weight space in which 1 centimeter equals 1 kilogram. Then $d(a,b) < d(c,d)$. Since the heights and weights of a, b, c and d remain unchanged, this illustrates the fact that in such spaces distance is relative to the quantitative relationship between height and weight.

involved must be interrelated so that there is a similarity-distance between any two objects. Of course, it may be that such attributes are only interrelated via the subject who makes similarities judgments: The subject's mind or nervous system interrelates the attributes. But then there would be no objective similarity which the subject was judging and it would be a strange coincidence if all subjects were then judging in the same way. Perhaps such thinking lay behind Carroll and Chang's INDSCAL method (mentioned earlier). It allows for subjects to weigh attributes differently. But if similarity is a thoroughly subjective concept, more leeway than this might be needed to account for all individual differences in its judgment.

Proposition 4 construes the concept of distance in multidimensional space in accordance with what Beals, Krantz, and Tversky (1968) called *the power metric*. This relates the distance between any points, x and y, to the values that x and y have upon the dimensions involved. The power metric relation has three components.

(i) *Intradimensional Subtractivity:* The distance between x and y is related to *the difference between x and y within each dimension.*

(ii) *Interdimensional Additivity:* Each dimension makes a contribution to distance (viz. $|x_i - y_i|^r$) and distance is a function of the *sum* of these dimensional contributions.

(iii) *Power:* The distance between x and y is related to *the intradimensional differences raised to some power* (equal to or greater than one) and *the interdimensional sum raised to the inverse power.*

Of course, the familiar Euclidean distance concept is the special case of the power metric resulting when the power exponent is two. That is,

$$d(x, y) = \left[\sum_{i=1}^{k} |x_i - y_i|^2 \right]^{1/2} .$$

So the power metric is really a generalization of this familiar distance concept, in that 2 is replaced by $r(\geqslant 1)$.

Does this very general distance concept exhaust all the possibilities? This depends on what properties the distance between any two points is expected to have. Generally, it is expected to have three properties, that are taken to be the defining properties of a distance metric (or measure). These are:

(a) *Positivity:* for any points x and y, $d(x,x) = 0$ and if $x = y$ $d(x,y) > 0$;

(b) *Symmetry:* for any two points x and y, $d(x,y) = d(y,x)$; and

(c) *The Triangle Inequality:* for any three points x, y and z, $d(x,y) + d(y,z) \geqslant d(x,z)$.

Now Beals, Krantz, and Tversky (1968) note that the power distance metric is not the only possible distance relation between two points in a multidimensional space. Another possibility is what they call the *exponential metric:*

$$d(x, y) = \log_p \left[1 + \sum_{i=1}^{k} (p^{|x_i - y_i|} - 1) \right]$$

(where p is a real number greater than one). This distance measure also satisfies (a), (b), and (c) above and so one would like to know if anything other than unfamiliarity caused its exclusion from the theory of multidimensional scaling. These authors go on to state that one major difference between the power and exponential metrics is that the power metric satisfies a condition they call *segmental additivity* while the exponential does not. The idea behind this condition is that the power metric distance between any two points is composed of additive parts or segments. Taking any two points x and y and the path between them whose distance is $d(x,y)$ then segmental additivity means that for any point z on that path $d(x,z) + d(z,y)=d(x,y)$. Given the exponential metric, no two points satisfy segmental additivity unless they differ on only one dimension. So as long as segmental additivity is thought to be important the power metric will be preferred to the exponential. Interestingly, Tversky and Krantz (1970) have proved that given segmental additivity and both (i) and (ii) above, the power distance relation is the only distance metric available.

Proposition 5 links distance to similarity and dissimilarity judgments and so completes the psychological theory. Because the core of the theory is Proposition 4, it is interesting to observe what implications 4 has for judgments of similarity and dissimilarity given 5. Beals, Krantz, and Tversky (1968), Tversky and Krantz (1970), and Tversky and Gati (1982) have traced many of these implications. Some of the simpler points they make are mentioned here to give something of the flavor of multidimensional scaling as a psychological theory.

Consider the three components of the power metric, (i), (ii), and (iii) above. The implication of intradimensional subtractivity for judgments of similarity and dissimilarity is that it is the *differences* between things that are all important rather than what things have in common. This is brought out clearly by considering a simple unidimensional example. In Fig. 6.3 the difference between lines a and b equals that between lines c and d.

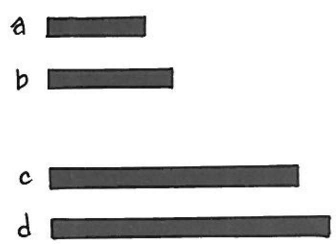

FIG. 6.3. According to the principle of intradimensional subtractivity similarity depends only on differences and not on common features. Hence, line *a* is as similar to line *b* as *c* is to *d*, for the difference between *a* and *b* equals that between *c* and *d*, even though *c* and *d* have more in common than do *a* and *b*.

This means that a is as similar to b as c is to d. This is despite the fact that c and d have much more in common (the length c) than do a and b. If judgments of similarity did not conform to this expectation then intradimensional subtractivity would be refuted as a basis for similarities judgments. Tversky (1977) has proposed another theory of similarities judgments which takes common features as well as differences into account.

Interdimensional additivity means that the contribution of any one dimension (i.e., the difference between x and y on that dimension) to judged similarity or dissimilarity is independent of the contribution of each other dimension. These contributions merely add together. They do not interact or modify each other. Now interaction between the dimensions in producing judgments of similarity is a logical possibility. In Fig. 6.4, interdimensional additivity implies (granting that height and width are the relevant dimensions) that a and b are as similar to one another as are c and d, and that a and c are as similar to one another as are b and d. Yet it is conceivable that these two dimensions could interact: for example, as width is increased (from 2 to 5 cm) a fixed height difference (4cm−3cm) contributes more to dissimilarity. Then a and b would be judged as being more similar than are c and d. Krantz and Tversky (1975) present some

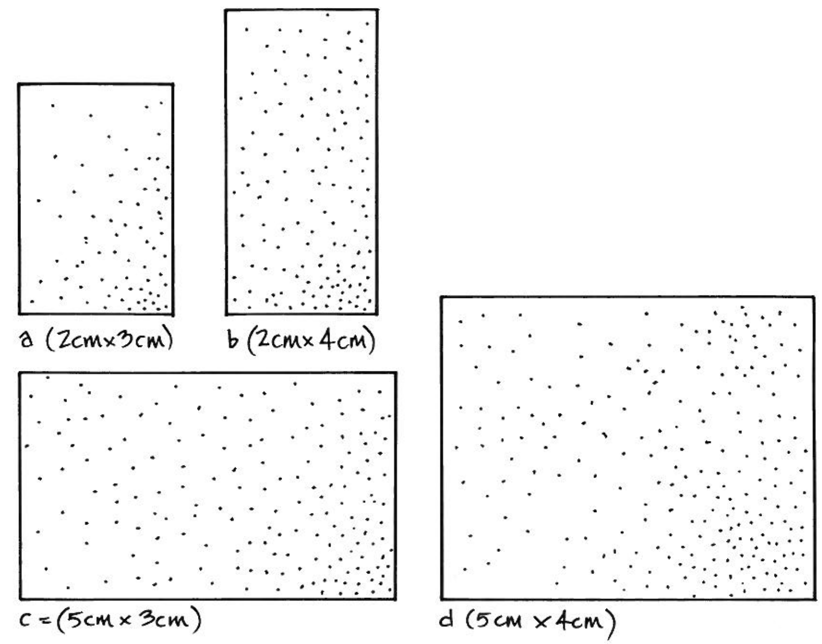

FIG. 6.4. According to the principle of interdimensional additivity the contributions from different dimensions to similarity do not interact with one another. Hence, if in the case of the similarity of rectangles width and height are the relevant dimensions, the similarity of *a* to *b* equals that of *c* to *d* and the similarity of *a* to *c* equals that of *b* to *d*.

experimental results that suggest that in judgments of the similarity of rectangles, the two dimensions of height and width do interact rather than simply add together.

Component (iii) of the power metric has interesting implications for judgments of similarity. The effect of the exponent, r, is to weight each of the intradimensional differences by an amount equal to that difference raised to the power of $r - 1$. That is, $|x_i - y_i|^r = |x_i - y_i|^{r-1} \times |x_i - y_i|$. Now, the larger $|x_i - y_i|$ is, the larger $|x_i - y_i|^{r-1}$ will be (for $r \geqslant 1$), so larger differences get weighted by proportionally larger weights. That is, the effect of r is to give more weight to big intradimensional differences in producing

$$\sum_{i=1}^{k} |x_i - y_i|^r .$$

When $r = 1$ all weights are equal (i.e., 1) but as r approaches infinity the weight given to larger differences increases. Some idea of this effect may be obtained by considering the set of all points, y, such that $d(x,y) = 1$ in two dimensional space, for some different values of r (Fig. 6.5) (Coombs, Dawes, and Tversky, 1970). Or, another way is to consider two points, x and y, in two dimensional space, and compute the distance between them for various values of r (Fig. 6.6). Obviously, as r increases, so $d(x,y)$ approaches 4 which equals $|x_1 - y_1|$, the larger of the two differences, that is, the contribution of $|x_1 - y_1|^r$ to $d(x,y)$, relative to that of $|x_2 - y_2|^r$, becomes progressively greater. This also shows how quickly, in two dimensions, the value of r reaches a point where increments have an insignificant impact.

The power metric requires that $r \geqslant 1$ and Tversky and Gati (1982) have shown an interesting implication of this. Given four points in two dimensional space, as shown in Fig. 6.7, the fact that $r \geqslant 1$ means that $d(c,e)+d(e,a) \geqslant d(c,a)$ (if $r = 1$ then $d(c,a) = d(c,e) + d(e,a)$), so it cannot be true that either

$$d(c,b) > d(c,e) \text{ and}$$

$$d(b,a) > d(e,a)$$

$$\text{or } d(c,b) > d(e,a) \text{ and}$$

$$d(b,a) > d(c,e).$$

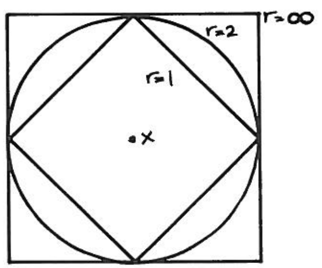

FIG. 6.5. The locus of points y such that $d(x,y) = 1$ for $r = 1$, $r = 2$ and $r = \infty$ using the power metric. (Based upon Fig. 3.16 in Coombs, Dawes, & Tversky, 1970. Adapted by permission of authors.)

r	$\|x_1-y_1\|^r$	$\|x_2-y_2\|^r$	$d(x,y)$
1	4	2	6
2	16	4	4.47
3	64	8	4.16
4	256	16	4.06
10	10.5^6	1024	4.00

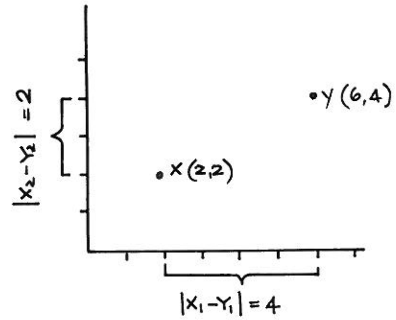

FIG. 6.6. As r increases the contribution of $|x_1 - y_1|^r$ to $d(x, y)$ increases relative to the contribution of $|x_2 - y_2|$.

That is, as Tversky and Gati (1982) put it, the center path cannot exceed the corner path. If $r < 1$ then it will and in so doing the triangle inequality is violated. This means that one of the conditions that a distance metric must satisfy is false in this context and so similarity judgments no longer behave as if reflecting such a metric. So if a, b, c, and e represent stimuli in a two dimensional attribute space then the dissimilarity of c to b cannot exceed that of c to e, while the dissimilarity of b to a exceeds that of e to a, and neither can the dissimilarity of b to a exceed that of c to e, while that of c to b exceeds that of e to a. Tversky and Gati (1982) review a number of experiments in which this prediction has been falsified.

If similarity and dissimilarity are monotonically related to $d(x,y)$ further consequences flow from the fact that $d(x,y)$ is a metric. First, in accordance with (a), the similarity of all things to themselves must be equal and greater than the similarity of any two different things. Second, in accordance with (b) similarities

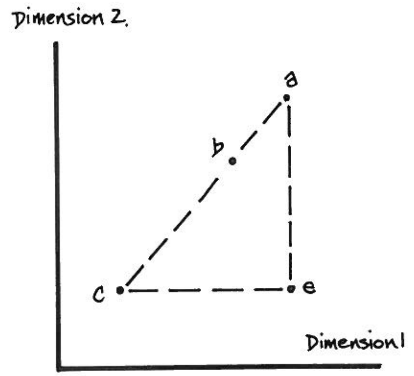

FIG. 6.7. In this two dimensional attribute space a and e take the same value on dimension 1 and c and e take the same value on dimension 2. Hence, if $r \geqslant 1$ then $d(c, e) + d(e,a) \geqslant d(c, b) + d(b, a)$.

must be symmetric: The similarity of object x to y must equal the similarity of y to x. Tversky and Gati (1978) present evidence suggesting the occurrence of systematic asymmetries in certain kinds of similarities judgments.

Anyone proposing to use the methods of multidimensional scaling should be aware of these predictions and the experimental evidence relating to them. Here, however, it is not my aim to review this evidence, but to show how the theory of conjoint measurement may be applied to the theory of multidimensional scaling. This provides theoretical justification for the methods of multidimensional scaling as measurement procedures and also, shows how the specifically quantitative parts of the theory may be tested.

THE APPLICATION OF CONJOINT MEASUREMENT

There are two applications of the theory of conjoint measurement to multidimensional scaling. The first is to each dimension separately and it tests two things: (i) that the dimension is quantitative and (ii) the proposition of intradimensional subtractivity. The second application is to dimensions pair by pair and, subject to confirmation of the first test, it tests interdimensional additivity, so that the dimensions together form a space.

Note that these tests require the prior identification of the dimensions or relevant stimulus attributes. However, they require no more than that values of these attributes be classifiable. It is not necessary that they be ordered.

Test 1

The first test of the theory of multidimensional scaling using conjoint measurement is similar in form to the test of Thurstone's theory described in the last chapter. If a set of stimuli are chosen such that they differ with respect to just one of the relevant attributes (i.e., they differ along just one of the dimensions), then the distance between any two of them, x and y, is

$$d(x, y) = \left(|x_i - y_i|^r \right)^{1/r}$$
$$= |x_i - y_i|$$

(where i is the dimension along which they differ). For any other dimension j, the value of $|x_j - y_j|$ is zero, for by hypothesis the stimuli do not differ on that dimension. Hence, $|x_i - y_i|$ is the only nonzero difference contributing to distance and thus the only one to count.

Given a set of dissimilarities (or similarities) judgments on such a set of stimuli, the order of the stimuli along the dimension (i) may be inferred. Let a and b be the most dissimilar pair of stimuli. Then the order of the dissimilarities from a (or, alternatively, from b), from least to most dissimilar, gives the order of the stimuli along the dimension (see Fig. 6.8a). This is because the order of

the dissimilarities between stimuli is the same as the order of the distances between them (Proposition 5 of the theory). While this does not enable one to infer direction along the dimension (i.e., which end is the greater and which the lesser), this is not important for this particular test of the theory.

Given, then, that the stimuli are ordered, they may be partitioned into two disjoint subsets, the first n and the last m (for completeness of testing n and m should be as large as possible, and the difference between n and m should be as small as possible). An $n \times m$ dissimilarities matrix may then be constructed in which the rows are the first n stimuli and the columns the last m. The dissimilarities judgments within this matrix should satisfy the hierarchy of cancellation conditions. The reason why is as follows.

Let the distance between any stimulus pair, x and y, be $d(x,y)$. Because of the method of construction, either *all* row stimuli are greater on dimension i than *any* column stimulus or vice versa. Without loss of generality assume the former (i.e., $x_i > y_i$). Then

$$d(x,y) = x_i - y_i$$
$$= x_i + (-y_i).$$

Hence, the distance between row and column stimuli is an additive function of their values on dimension i and so these distances must conform to the hierarchy of cancellation conditions. However, according to Proposition 5, judgments of dissimilarity are monotonic with distance, so the corresponding dissimilarities matrix must also satisfy this hierarchy of ordinal conditions.

These predictions depend not only on the assumption that dimension i is quantitative (2) (and, of course, 5 as already mentioned), but also upon the assumption of intradimensional subtractivity, one of the three components of the power metric (4).

The kind of situation envisaged here is illustrated in Figures 6.8 (i), (ii), and (iii). Fig. 6.8(i) shows that $d(a,f)$ is the largest distance and the order of the stimuli along the dimension is given by the order of their distances from a. Figure 6.8(ii) illustrates single cancellation or independence. In row a the order of the columns must be f, e, d (from greatest to least). Similarly, the same order on the columns must prevail in the other rows (b and c). Finally, Fig. 6.8 (iii) illustrates double cancellation. If $d(c,e) > d(b,d)$ and $d(b,f) > d(a,e)$ then, of course, $d(d,f)$ must exceed $d(a,c)$ and so, $d(c,f) > d(a,d)$. These two cancellation conditions are also illustrated in Fig. 6.9, which shows the order relations in the resulting dissimilarities matrix.

Test 2

The second application of conjoint measurement to the theory of multidimensional scaling requires a set of stimuli differing with respect to just two relevant dimensions. Call these dimensions i and j. Then the distance between any two of these stimuli, x and y, is

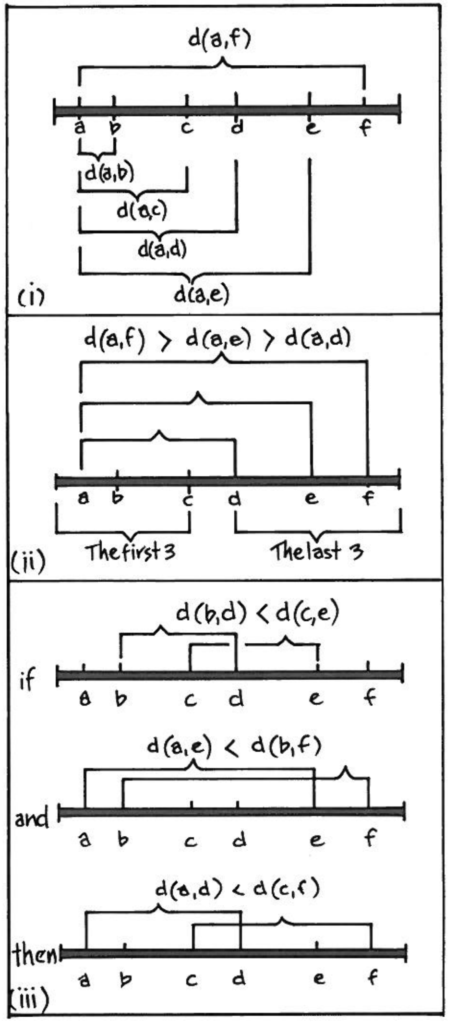

FIG. 6.8. (a) If *a, b, c, d, e* and *f* are points along a single dimension and *d(a, f)* is the greatest distance between any pair of these points then the order of the points along the dimension is given by the order of their distances from *a* (or *f*). If the order of the first 3 points from *a* is *a, b, c* and the order of the last 3 from *f* is *f, e, d* then (b) from single cancellation it follows that *d(a,f)* > *d(a,e)* > *d(a, d)*, etc.; and (c) from double cancellation it follows that whenever *d(b,d)* < *d(c,e)* and *d(a,e)* < *d(b,f)* then *d(a,d)* < *d(c,f)*.

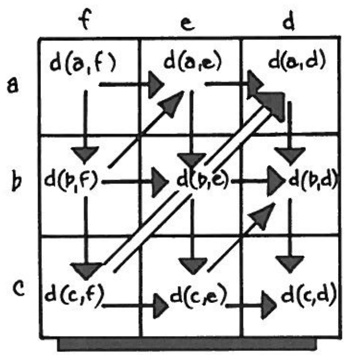

FIG. 6.9. If *a, b, c, d, e,* and *f* are points along a single dimension in that order then these are the order relations upon the distances between them, the double cancellation prediction, of course, being a conditional statement.

$$d(x, y) = \left(|x_i - y_i|^r + |x_j - y_j|^r \right)^{1/r}.$$

Each dimension contributes its own component to the distance between *x* and *y*, the contribution of dimension *i* being $|x_i - y_i|^r$ and the contribution of dimension *j* being $|x_j - y_j|^r$. The components contributed by *i* may be thought of as a variable in their own right (call this variable "*A*") and likewise the components contributed by *j* (call that variable "*X*"). Then, labeling the distance resulting from the power metric "*D*", the relationship between these three variables is

$$D = (A + X)^{1/r}$$

$$\text{or } D^r = A + X.$$

That is, D^r is an additive function of the *i* and *j* components and, hence, must conform to the hierarchy of cancellation conditions. Since for the power metric, $r \geqslant 1$, it follows that D^r is an increasing monotonic function of *D* (i.e., the order upon *D* and D^r must be the same). Furthermore, it follows from Proposition 5 that judged dissimilarities and distances must be in the same order. Hence, judged dissimilarities must also conform to the hierarchy of cancellation conditions.

This prediction depends on *i* and *j* being quantitative variables (2), on distance, and dissimilarity being monotonically related (5), and on two components of the power metric, intradimensional subtractivity and interdimensional additivity. While it was assumed that $r \geqslant 1$, this component of the powermetric is not actually required. Providing $r \neq 0$ the above mentioned prediction follows.

The kind of situation that might be used to test this prediction is illustrated in Fig. 6.10. The nine stimuli *F, G, H, I, J, K, L, M,* and *N* form a rectangular array in which *F, G,* and *H* are the same on dimension 2, as are *I, J,* and *K,* and

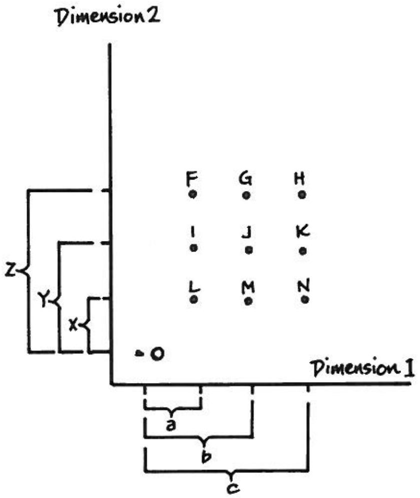

FIG. 6.10. The nine stimuli, *F, G, H, I, J, K, L, M,* and*N*, form a rectangular array in a two dimensional attribute space, with the stimuli within each triple (*F,I,L*), (*G,J,M*), and (*H,K,N*) having the same value on dimension 1 and the stimuli within each triple, (*F,G,H*), (*I,J,K*), and (*L,M,N*) having the same value on dimension 2. Hence, the distance between each of these nine and *O* is an additive function of one of *a, b, c* (to the power of *r*), and one of *x, y,* or *z* (to the power of *r*).

L, M, and *N*. Also, *F, I,* and *L* are the same on dimension 1, as are *G, J,* and *M*, and *H, K,* and *N*. Thus, the proximity of these nine stimuli to one outside of the rectangle, such as *O* (whose values on dimensions 1 and 2 are both lower than that of any of the stimuli in the rectangle), is determined by the requisite composite of one of *x, y,* and *z* with one of *a, b,* or *c*. For example,

$$d(O,L)^r = a^r + x^r$$

$$d(O,J)^r = b^r + y^r$$

$$\text{and } d(O,H)^r = c^r + z^r.$$

Thus, the dimensional components constituting the set of distances are, for dimension 1, a^r, b^r, and c^r, and for dimension 2, x^r, y^r, and z^r. Each dissimilarity judgment then is monotonically related to the sum of a dimension 1 component plus a dimension 2 component, as shown in Table 6.1.

It is the order relations between the cells of a dissimilarities matrix such as this that must satisfy the cancellation conditions (in this case single and double cancellation).

AN EXAMPLE OF CONJOINT MEASUREMENT

As an illustration of these two tests of multidimensional scaling theory a sample of 130 psychology students at the University of Sydney were instructed to make the appropriate dissimilarities judgements about pairs of stimuli from Fig. 6.11. These stimuli are schematic pictures of flower pots and they vary systematically along two dimensions: leaf size/shape and pot size/shape. They have been so contrived that judgments on pairs within the subset, *E, F, G, H, I,* and *J* enable Test 1 to be carried out. These six stimuli differ only according to leaf size/shape and dissimilarity judgments based on these differences should conform to single and double cancellation. A similar prediction applies to judgments on the set *B, D, H, K, M,* and *P*. Finally, the differences between each of the nine stimuli *A,*

TABLE 6.1
The Distance Between Each of the Nine Stimuli in Fig. 6.10 and *O* is an Additive Function of a Row Component (From Dimension 2) and a Column Component (From Dimension 1)

		Dimension 1 components		
		a^r	b^r	c^r
Dimension 2 components	x^r	$d(O,L)^r$	$d(O,M)^r$	$d(O,N)^r$
	y^r	$d(O,I)^r$	$d(O,J)^r$	$d(O,K)^r$
	z^r	$d(O,F)^r$	$d(O,G)^r$	$d(O,H)^r$

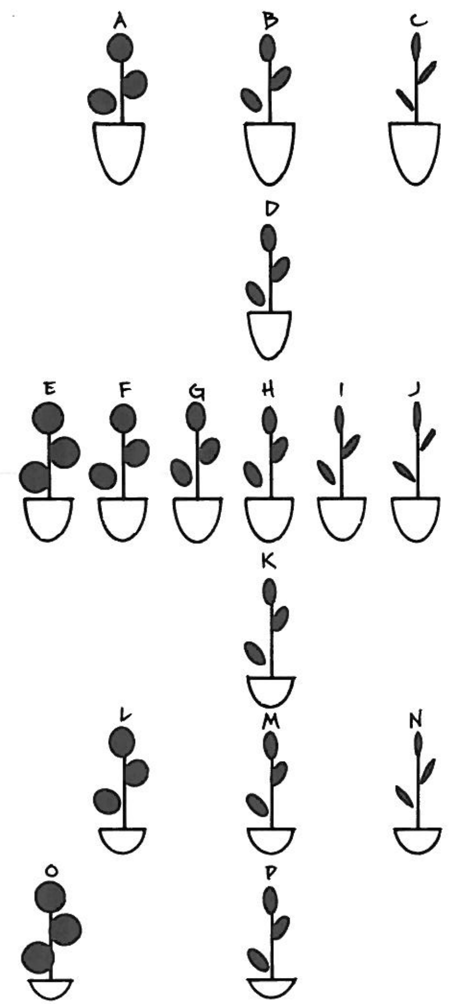

FIG. 6.11. The schematic flower pots used as stimuli in the test of the theory of multidimensional scaling.

B, C, F, H, J, L, M, N, and stimulus *O* form a matrix like that illustrated in Table 6.1 and, so, enable Test 2 to be conducted.

The relevant pairs of stimuli were presented to the subjects in a test booklet and the subjects were instructed to indicate the degree of dissimilarity between each stimulus pair on a 21-point rating scale. In order to test the conjoint measurement conditions various pairs of ratings must then be compared. For example, one may be interested in how the pair (*G,J*) is rated compared to the pair (*G,I*). Within each subject's data, the difference between the ratings on these two stimulus pairs were classified as "greater than 0" (>), "equal to 0" (=) or "less than 0" (<), and then across subjects the frequencies within these categories were obtained ($N_>$, $N_=$, $N_<$) relative to these two pairs of stimuli. The significance of the difference between such ratings across all subjects was then assessed by performing sign tests (i.e., χ^2 tests on the difference between the two frequencies $N_>$ and $N_<$). The results are shown in Fig. 6.12. Part (i) shows the results for Test 1. An arrow between two cells (e.g., between *EJ* and *EI*) indicates that significantly more subjects rated the stimulus pair at the arrow's tail (e.g., *EJ*), as more dissimilar than the stimulus pair at the arrow's head (e.g., *EI*) than did vice versa. On the assumption that individual differences are due to error, the

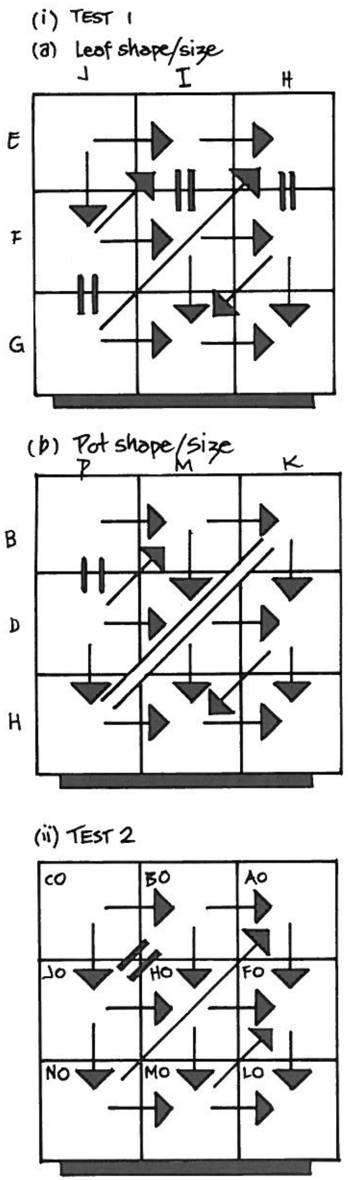

FIG. 6.12. The order relations inferred from the dissimilarity judgments of the 130 subjects used. (i)(a) fails single cancellation but satisfies double cancellation (pattern 3 of Table 5.3); (i)(b) also fails single cancellation and satisfies double cancellation (pattern 4 of Table 5.3); and (ii) satisfies both single and double cancellation (pattern 6 of Table 5.3).

arrows in Fig. 6.12 may be taken as indicative of the true dissimilarity ordering for the population.

The orders given in (i)(a) and (b) depart from that required by single cancellation. However, two alterations to the order in (i)(a) (replacing the arrow between *EJ* and *FJ* by "=" and the "=" between *FJ* and *GJ* by an arrow) and one alteration to the order in (i)(b) (replacing the "=" between *BP* and *DP* by an arrow), produces orders consistent with single cancellation. Neither (i)(a) nor (i)(b) violated double cancellation. Hence, for Test 1, the order obtained from the subjects' judgments departed from that predicted on three occasions. This may be taken to indicate that at least one of the assumptions underlying this prediction is false. Which one(s), of course, it is not possible to tell from these results.

The order shown in Fig. 6.12(ii) conforms to both the single and double cancellation conditions. Remembering that Test 1 is based on Propositions 2, 5, and intradimensional subtractivity and Test 2, on these propositions together with interdimensional additivity, these results indicate that at least one of the first three of these propositions is false.

Of course, one would not reject the theory of multidimensional scaling on the basis of such a disconfirmation. In testing any theory other propositions must always be assumed to be true and, therefore, the failure of a prediction may reflect on these propositions as much as on the theory. For example, in this instance, it was assumed that all subjects attended only to the dimensions of pot/shape size and leaf shape/size, and that all subjects were able to rate their degrees of judged dissimilarity in a consistent way. Also, it was assumed that the stimuli were sufficiently distinct from one another to prevent confusions. It may have been that the pairs *E* and *F*, and *B* and *D*, were too similar to enable a reliable test of the theory of multidimensional scaling.

Whatever the cause of the disconfirmations here, evidence against the theory of multidimensional scaling has been accumulating over the last 20 years (cf. Krantz & Tversky, 1975; Tversky & Krantz, 1969; Tversky & Gati, 1982; Wender, 1971; & Wiener-Ehlich, 1978). This evidence calls into question the current attitude towards multidimensional scaling within psychology. At present it is used indiscriminately, without first testing the applicability of the theory to the stimulus domain involved. Here, it has been illustrated how the theory of conjoint measurement may be used to make such tests in certain stimulus domains.