

# Applications of Multidimensional Scaling in Cognitive Psychology

Edward J. Shoben  
University of Illinois

Cognitive psychology has used multidimensional scaling (and related procedures) in a wide variety of ways. This paper examines some straightforward applications, and also some applications where the explanation of the cognitive process is derived rather directly from the solution obtained through multidimensional scaling. Other applications examined include cognitive development, and the use of MDS to assess change as a function of context. Also examined is how an ideal representation is selected, whether, for example, a space or a tree is more appropriate. Finally, some inherent limitations of the method for cognitive psychologists are outlined, and some pitfalls and potential misapplications are identified.

Although there are many examples of multidimensional scaling (MDS) and related methods applied to problems in cognitive psychology, MDS has not been used to its fullest potential in this branch of psychology. Not only are there many circumstances where a straightforward application of MDS might be profitable, but there are also instances where some more subtle application of these procedures might be useful in providing answers to cognitive questions. The present paper will attempt to illustrate some of these less straightforward applications by example, as well as noting some of the more routine applications. No effort has been made to cover every application of MDS in the cognitive literature, however, since this re-

view is intended to be illustrative rather than exhaustive. Moreover, it should be admitted at the outset that the proportion of applications in comprehension and memory exceeds considerably the proportion of such examples from perception and psychophysics. Although a conscious effort has been made to draw on more recent work, older studies have been included when deemed appropriate.

This paper will examine applications of both traditional two-way MDS (e.g., Kruskal, 1964a, 1964b; Shepard, 1962a, 1962b) and three-way MDS (e.g., Carroll & Chang, 1970). Three-way analyses which can extract higher dimensionality on the same data set, are generally, but not exclusively, performed when individual differences are considered. When these individual variations are not at issue, two-way MDS is usually employed.

The most straightforward and the most common use of MDS has been descriptive. Many of these early applications (e.g., Clark, 1968; Rapaport & Fillenbaum, 1972) have used MDS to characterize the spatial structure underlying a set of stimuli. This use makes no attempt, however, to specify how the stimuli are processed. Although researchers can attempt to identify the dimension along which the stimuli vary in a multidimensional space, frequently little or no evidence has been provided that subjects are actually using these dimensions.

More recent studies have progressed beyond this descriptive use:

1. Studies suggesting that subjects are actually using the spatial representation recovered by

MDS. For example, Rumelhart and Abrahamson (1973) have used Henley's animal names and found that distance in the MDS solution would predict solutions to analogies. Similarly, Rips, Shoben, and Smith (1973) found that distance in a multidimensional space would predict reaction times (RTs) in a categorization task.

2. Studies comparing the structure of the stimulus space over time. For example, Arabie, Koslyn, and Nelson (1975) and Howard and Howard (1977) have looked at changes in the stimulus space from a developmental perspective. Voss and his colleagues (LaPorte & Voss, 1979; Bisanz, LaPorte, Vesonder, & Voss, 1978) have used MDS procedures to examine changing representations of prose with changes in context.
3. Studies using MDS analysis to evaluate various theories (e.g., Friendly, 1977; Soli & Arabie, 1979).

This paper will examine some of the descriptive uses of MDS before moving on to some of the less straightforward applications. Alternatives to MDS as a representation of data will be examined; and finally, there will be an attempt to specify the role of MDS in cognitive psychology.

### Descriptive Uses

#### Psychological Domains and Concrete Objects

Many stimulus sets have been subjected to MDS analysis. These range from consonant phonemes to color names to semantic domains. A brief and useful review has been provided by Shepard (1980).

The data set that has probably been subjected to analysis by the largest number of multidimensional scaling programs is the confusion data collected by Miller and Nicely (1955) on consonant phonemes. Shepard (1972) incorporated the fitting of an exponential decay function on the original confusion proportions in performing the MDS analysis. The resulting two-dimensional solution was readily interpretable. The first dimension distinguished between voiced (za and da) and unvoiced (fa and ka) phonemes. The second dimension separated the na-

sals (ma and na) from the other consonant phonemes. Within the remaining consonants, there was also a separation between those that are formed at the front of the mouth (fa and ba) and those formed at the back of the mouth (ga and zha). These results thus placed some order on this domain.

The color domain has long been of interest to users of MDS (Ekman, 1954). In his original paper, Shepard (1962a) provided an analysis of the judged similarity of colors. In the resulting two-dimensional solution, a *color circle* was recovered in which the various colors formed a circle with a gap between the color with the shortest wavelength (violet) and the one with the longest wavelength (red). If the other points are connected by drawing the circle, the resulting ordering is monotonic with wavelength. Moreover, the fact that red and violet are quite close to each other in the two-dimensional space (while they are maximally distant in wavelength) accords with the knowledge that red and violet are quite similar to each other.

For semantic domains, the results are generally less clear-cut. There are several possible reasons for this difference. First, semantic domains are generally of functionally infinite size (although there are exceptions, such as kin terms). Thus, it is not practicable to scale *all* animals; some selection must be made. Surprisingly, this selection is often made haphazardly. Perhaps not surprisingly, altering the sample of exemplars can affect the resulting solution. Second, semantic domains are potentially more heterogeneous than the other two domains that have been discussed. Whereas any group of colors will have some hue and some saturation, it is not easy to come up with shared dimensions for the nouns *justice*, *conservation*, and *laughter*. Third, some domains that are commonly subjected to MDS analysis (categories) may pose technical problems for computer programs that are commonly used. This last point will be discussed in greater detail in a subsequent section.

One of the most frequently scaled semantic domains is animal names. Henley (1969) scaled 30 animal names based on the dissimilarity judgments of undergraduates. In the three-dimensional solution that Henley obtained, the first dimension was clearly identifiable as *size*, with elephant, giraffe,

and camel at one extreme and rat, mouse, and chipmunk at the other. The second dimension, labeled *ferocity* by Henley, had the predators—lion, tiger, and leopard—at one extreme and the domesticated animals—cow, sheep, and donkey—at the other extreme. Finally, the third dimension, labeled *humanness* by Henley, placed the apes at one extreme and pig, dog, and wolf at the other.

Henley's last dimension is particularly difficult to interpret. It is not at all clear why giraffe, chipmunk, and camel are all "more human" than horse, dog, and leopard. Although the third dimension does separate the anthropoid apes from the rest of the animals, it is difficult to maintain that this third dimension is interpretable. Similarly, the label for the second dimension has been disputed even though the ordering of objects is quite consistent across studies (see Storm, 1980, for an exception). For example, rather than ordering the animals on ferocity, Rips, Shoben, and Smith (1973) labeled a similar dimension in their MDS solution *predacity*, arguing that this dimension separated the wild animals from the domesticated farm animals.

In addition to this dispute over labeling dimensions, there has also been an apparent controversy over how many dimensions are required. Rips et al. (1973), for example, used only two dimensions. Shoben (1976) also used two dimensions (in addition to one that separated the birds from the mammals in his solution). However, King, Gruenewald, and Lockhead (1978) have argued for a three-dimensional result, and Rumelhart and Abrahamson (1973) were unable to obtain satisfactory results in their analogies task with any dimensionality lower than three.

## Theory

MDS analysis has also been profitably applied to theories of memory (e.g., Ross 1983). Ross examined 12 well-known cognitive psychology theories: the distributed memory model (J. A. Anderson, 1977), HAM (J. R. Anderson & Bower, 1974), the short-term memory model (Atkinson & Shiffrin, 1968), levels of processing theory ( Craik & Lockhart, 1972), the propositional theory (Kintsch, 1974), the LNR model (Norman & Ru-

melhart, 1975), the dual code theory (Paivio, 1971), the SAM model (Raaijmakers & Shiffrin, 1980), the random walk model (Ratcliff, 1978), the schema model (Rumelhart & Ortony, 1977), the MOPS model (Schank, 1980), and Tulving's (1975) memory model.

Using expert subjects, Ross (1983) employed SINDSCAL (Carroll & Chang, 1970; Pruzansky, 1975) to obtain a two-dimensional solution. The resulting solution ordered the theories in terms of the unit of analysis on one dimension and the degree of rigor on the other dimension. In terms of the unit of analysis, MOPS and the schema model were at one extreme, and J. A. Anderson's distributed memory model and Ratcliff's random walk model were at the other. Theories such as MOPS are often concerned with paragraphs or stories as units, whereas Ratcliff's model and J. A. Anderson's model are usually applied to single words. On the degree of rigor dimension, theories that have strong mathematical or computer underpinnings were at one extreme, and thus MOPS, the LNR model, HAM, SAM, Ratcliff's model, and J. A. Anderson's model all scored very high on this dimension. At the other extreme, theories that are not mathematical, such as Tulving's model, levels of processing, and the dual code theory were low on this dimension.

Although Ross did not attempt to apply his scaling results to any kind of behavioral data, it is interesting to note that memory theories are organized in this way, and it may perhaps belie a general trend in that two of the theories examined by Ross (levels of processing and the dual code theory) have recently been receiving considerable criticism; interestingly, both are theories without strong mathematical or computer underpinnings and both are theories whose prominent unit of analysis is pairs of words.

## Applications of MDS Analysis

Some early studies (Clark, 1968; Fillenbaum & Rapaport, 1971) have allowed the MDS analysis to stand by itself. Although the authors typically made some comment about the dimensions re-

covered and their significance, no attempt was made to provide evidence that the derived space is actually used in processing information.

In contrast, most recent studies have attempted to obtain evidence for the general validity of the derived multidimensional space. Rips et al. (1973), for example, showed that distances in a multidimensional space would predict categorization times. Specifically, they obtained the mean reaction time (RT) to verify statements such as "A robin is a bird." For the 12 birds and 12 mammals they employed, Rips et al. found that semantic distance correlated with categorization time quite highly.

Similarly, Shoben (1976) used semantic distance to predict RTs in a Same-Different task. The stimulus domain in this study was a set of six mammals and six birds. Scaled together (in contrast to most earlier studies), these 12 exemplars and the two superordinate terms *bird* and *mammal* yielded a three-dimensional solution that was readily interpretable. One dimension, as usual, was *size* with bear at one extreme and robin, sparrow, and mouse at the other. A second dimension was labeled *predacity* with bear and lion at one extreme and goose and chicken at the other. A third dimension simply clustered the birds at one extreme and the mammals at the other. For Same category judgments, such as *hawk-eagle*, Shoben was able to show first that the distance between hawk and bird and the distance between eagle and bird predicted Same RT quite well. Moreover, the distance between hawk and eagle was not predictive of Same RT, contrary to some semantic memory theories (such as Schaeffer & Wallace, 1969, 1970). In addition, Shoben was also able to use the distances obtained from MDS to predict the ease with which people can determine that the exemplars are from different categories. For the first item, the distance between the exemplar and its category was predictive of RT. For the pair bear-goose, for example, the bear-mammal distance was predictive of RT. Moreover, the distance between the second exemplar and the category of the first was also predictive, in this example, the goose-mammal distance. Here, small distances led to long RTs. Once again, the distance between exemplars was a relatively poor predictor.

### Task Processing Account

The MDS solution could be used here to provide a processing account of the task. If subjects compared exemplars to each other, then it might have been expected that the distance between the exemplars would be the best predictors of RT. Distances in the multidimensional space and regression analyses were able to place severe constraints on any viable processing account of this Same-Different task.

Rumelhart and Abrahamson (1973) also used MDS to provide a processing account of an analogies task. Although more sophisticated accounts of analogical reasoning are now available (Sternberg, 1977), Rumelhart and Abrahamson's work is instructive in that it shows another way in which MDS analysis can be applied to provide some theoretical constraint on a cognitive task.

Rumelhart and Abrahamson were concerned with four-term analogies such as "fox:horse::chipmunk:\_\_\_\_\_." Subjects were asked to select the best completion (in one experiment) from among alternatives—antelope, donkey, elephant, and wolf—where elephant is the best answer. Spatially, Rumelhart and Abrahamson noted that the ideal point or solution to the analogy in the multidimensional space can be located by first determining the relationship (in all three dimensions in this case) between the first two terms and then by applying those differences to the third term. Thus, people have solved the analogies by constructing a parallelogram whose vertices are the first three terms of the analogy. In the present example, fox is smaller than horse (Dimension 1), somewhat more ferocious than horse (Dimension 2), and slightly less human than horse (Dimension 3). The ideal point is thus an animal that is larger than chipmunk, less ferocious than chipmunk, and slightly more "human" than chipmunk. Elephant comes closest to this ideal point. Interestingly, the alternatives can also be rank-ordered (as done by Rumelhart and Abrahamson), and subjects will judge that antelope is the second-best completion to the problem, donkey is third, and wolf is poorest.

In addition to predicting subjects' first choices by distances from an ideal point, Rumelhart and Abrahamson also developed a theory about the distribution of subjects' responses. Specifically, they assumed that subjects' choices would be in proportion to their distances from the ideal point. More formally, they proposed that the distribution of responses would follow Luce's (1959) choice rule,

$$Pr(X_i | X_1, \dots, X_n) = \frac{v(d_i)}{\sum_j v(d_j)} \quad [1]$$

Here,  $d_i = X_i - I$ : the distance between alternative  $X_i$  and the ideal point, and  $v(\ )$  is a monotonically decreasing function, and  $P(X_i | X_1, \dots, X_n)$  is the probability of selecting the  $i$ th item from the  $n$  alternatives. Following Shepard's (1972) original work, Rumelhart and Abrahamson (1973) assumed  $v(x) = \exp(-ax)$ , where  $a$  is constrained to be positive, because Shepard obtained good fits to recall data using this kind of function. In addition, this function has the virtue of requiring only one parameter.

The results corresponded quite well to the theoretical predictions, both qualitatively and quantitatively. Examination of even the data for third and fourth choices showed a close correspondence between theory and data. Moreover, this correspondence held up even when the distance among the possible alternatives was systematically varied.

This exemplifies a case in which MDS analysis appears to have actually aided in the formulation of a psychological processing model. Because it was relatively easy to determine the distance of each of the alternatives from the ideal point, given the availability of the multidimensional space, Rumelhart and Abrahamson (1973) were led to formulate the expanded theory based on Luce's (1959) choice rule.

### Free Recall Analysis

MDS analysis can also play a useful role in the analysis of free recall. The distances derived from the multidimensional space can be used as predictors of output order in free recall. Shepard (1972) did this analysis by fitting an exponential transfor-

mation to the recall data. A different and promising approach is one developed by Friendly (1977). He derived the average distance between any pair of objects  $i$  and  $j$  by the following equation:

$$D_{ij} = \frac{\sum_{t=1}^T \phi_{ijt} |l_{it} - l_{jt}|}{\sum_{t=1}^T \phi_{ijt}} \quad [2]$$

Here,  $l_{it}$  is the position of item  $i$  on trial  $t$  and  $\phi_{ijt}$  is an indicator variable which is 1 if items  $i$  and  $j$  are both recalled on trial  $t$  and zero otherwise. If giving greater weight to items recalled in close contiguity to each other is desired, then this expression can be modified by extracting the square root of the squared difference in recall position, i.e.,

$$D_{ij}^* = \left[ \frac{\sum \phi_{ijt} |l_{it} - l_{jt}|^2}{\sum \phi_{ijt}} \right]^{1/2} \quad [3]$$

With this proximity matrix, Friendly was able to obtain a highly interpretable two-dimensional solution that had a very low stress value.

Whereas Friendly derived his similarity matrix from free recall protocols, a related question is whether distances obtained from an MDS solution can predict recall data. For example, Caramazza, Hersh, and Torgerson (1976) found little correlation between distance and recall proximity on early trials, but positive and usually significant correlations on later trials. For categories whose exemplars varied the most in the multidimensional space, such as mammals, these correlations were quite high; for categories whose exemplars showed little variation, such as fish, the correlations were much lower. Obviously, there must be variation in the space in order to predict patterns of free recall. Generally, similarity as measured by direct ratings or by distance in a multidimensional space is a reasonably good prediction of results in free recall experiments (Friendly, 1979; Schwarz & Humphreys, 1973).

### Music Perception

Perhaps the area in which MDS has contributed most to cognitive psychology in the past 5 years is in the area of music perception. Early psychological investigations focused solely on the dimen-

sion of pitch height (Stevens & Volkman, 1940), and subsequent work has provided evidence of more complicated structure largely in its emphasis on the octave (Shepard, 1964). According to this formulation, the notes of the scale could be thought of in terms of a chroma circle, analogous to the circle observed for colors.

Recently, the investigation has proceeded in a musical context. That is, many studies now ask subjects, not for their judgment of similarity of two tones in the absence of explicit context, but for the similarity of two tones in the context of a diatonic scale, or the similarity of two passages in the context of a melody. This practice has led to the recovery (through MDS) of much more complicated representations.

Perhaps the seminal work along these lines was done by Krumhansl (1979). She presented subjects with a simple musical context: a major triad chord, an ascending major scale, or a descending major scale. Differences among these contexts were slight. Subjects were then presented with a pair of tones whose similarity they were to judge in the context.

The raw similarity results alone are interesting. For stimuli in the major triad, other major triad stimuli were judged to be most similar, followed by diatonic tones. Nondiatonics were judged as least similar. This same pattern was observed for diatonic tones. For nondiatonic tones, there was little effect of this categorical variable. Instead, the similarity in pitch height was the primary determinant of the similarity between a nondiatonic tone and another one.

The MDS representation that Krumhansl obtained is a more complicated version of the chroma circle. Of particular interest is the three-dimensional solution that resembles an inverted cone. At the base of the cone are the components of the major triad. For the C major scale she employed, these items are middle C, E, G, and high C, reading clockwise around the circle. The diatonic tones cluster at the next level. These, too, can be viewed as a circle. Reading clockwise, the tones here are D, F, A, and B. Finally, at the base of the inverted cone are the nondiatonic tones, which, like the other clusters, are arranged in ascending order when the circular base is read in a clockwise direction.

Thus, it appears that more than a chroma circle is present when the tones are presented in a scale context. Regardless of musical training, people make the classification of tones as suggested by music theory; this representation can be recovered through MDS.

Krumhansl and her colleagues (Krumhansl, Bharucha, & Kessler, 1982) have generalized this result from tones to chords. In a scale context, people judged the major chords (for C major: CEG, FAC, and GBD) as central unless the scale context was a minor key, in which case the corresponding chords occupied the center of the MDS representation. More specifically, for the A minor scale Krumhansl et al. employed, the corresponding chords were A minor, D minor, and E major, respectively. Chords that were not part of the scale sequence (such as the E major for the C major scale) were at the periphery of the space. This pattern held across scale types.

Recently, Pollard-Gott (1983) has extended this line of work to passages from classical music. More specifically, her subjects listened to a Liszt sonata over three sessions. In each, the listeners were encouraged to take notes and to think about the relations among the passages. After each session, subjects judged the similarity of the 28 pairs of stimuli that were generated from the eight passages (4 to 16 measures in length) Pollard-Gott studied.

These similarity data were analyzed using SINDSCAL (Carroll & Chang, 1970; Prozansky, 1975). The data are particularly interesting when examined across sessions. After the first session, the dimensions recovered by the program represent fairly gross physical features of the passage; the dimensions are most adequately labeled, for example, as happy-sad, high-low, simple-complex, and loud-soft. However, by the second session, a different dimension emerges: a thematic dimension. By the second session, this dimension separated without overlap passages that deal with Theme A from those that deal with Theme B. This separation is even more pronounced after the third listening session. Thus, it seems that a thematic dimension emerges as the subject becomes more knowledgeable about the composition.

Strong support for this interpretation is provided

by the results obtained in an expert condition conducted by Pollard-Gott (1983). In this condition, a one-dimensional MDS solution was obtained; that single dimension accounted for 84% of the variance. As with the fourth dimension obtained with the naive subjects, this single thematic dimension clustered all the passages that dealt with Theme A at one extreme and all the passages that dealt with Theme B at the other.

### Measurement of Change in Structure

Some of the most impressive applications of MDS have been in demonstrating a change in the representation of stimuli. For example, by examining a set of stimuli in several conditions, it may be possible either to show an effect of context and/or to extract a more complicated multidimensional solution. Less dramatically, it may be demonstrated that context does have an effect.

One straightforward application of this strategy was performed by LaPorte and Voss (1979), who presented subjects with 20 nouns taken from one of two simple stories. For one story, these nouns included *fields, clouds, vegetation, train, approach, decade, troops, plague, eggs, and food*. Initially, subjects all performed a pairwise rating of similarity for all pairs of nouns. The resulting matrices were subjected to a three-way MDS analysis, and a two-dimensional solution was obtained where one dimension separated man-made objects (train, approach, and troops) from natural ones (fields, vegetation, and food). The second dimension was interpreted by the authors as separating animals from nonanimals, with eggs and food at one extreme and fields and clouds at the other.

Subsequent to this initial rating, subjects either repeated the task (a control condition) or read a story about how grasshoppers become a pest every 10 years. These experimental subjects then performed the rating task on the same words once again. Interestingly, the resulting space had both similarities and dissimilarities to the original solution.

The first dimension was virtually the same as before and could be characterized as a man-made vs. natural dimension. However, the second di-

mension now reflected the temporal ordering of events in the story such that nouns that occurred early in the story appeared at one extreme and nouns that appeared at the end of the story appeared at the other end of the dimension. Thus, it seems that a story context can alter the relationship among terms. The results for a second story used by LaPorte and Voss showed similar, although more complicated, results.

A more extensive framework for studying prose representations has been developed by Bisanz, LaPorte, Vesonder, and Voss (1978). They presented subjects with stories involving animal names that had a clear thematic structure. In one condition, for example, the animals varied along a *leadership* dimension and along a *helpfulness* dimension. Bisanz et al. compared the MDS solution obtained prior to reading the story with one derived after reading the story. In the former case, they obtained a typical animals solution, similar to the one obtained by Rips et al. (1973) in which the two dimensions could be characterized as *size* and *ferocity*. After reading the story, and being instructed to rate the animals in terms of their similarity to each other in the story, the themes (such as helpfulness and leadership) were recovered. Interestingly, these two dimensions were not recovered equally well.

In addition to recovering the story structure, Bisanz et al. were also able to demonstrate that the recovered structure could predict performance in a memory task. In particular, they asked subjects to determine if pairs of animals were both helpful or both not helpful. At least for affirmative responses, it was clear that distance in the multidimensional space was predictive of RT, and the authors were able to formulate a regression model (where distance in the multidimensional space was one of the independent variables) that predicted the latencies of all the trials quite well.

Actually, it is possible that Bisanz et al. might have been able to obtain greater predictability by looking at another distance in the space they obtained. The only distance that Bisanz et al. examined is the distance between the two stimuli. For example, in the pair lion-tiger, they looked only at the distance between the two animals. Assume

for the moment that lion is helpful and that tiger is not, and further assume that the subject processes lion first. In this case it might be profitable to examine the distance between the point for lion and the point for helpful and between tiger and helpful, if the decision required is whether both are helpful. Logically, these are the distances that should be involved in the model that Bisanz et al. proposed, and it is surprising that they appear not to have looked at these distances. In this case, a small lion-helpful distance should be facilitory and a small tiger-helpful distance should be inhibitory. This kind of analysis (used earlier by Shoben, 1976, in a Same-Different task with semantic categories) would provide additional support for the model espoused by Bisanz et al.

### Cognitive Development

One obvious place to look for examples of MDS as an index of change is in work with children. Particularly in cognitive development, it should be possible to use MDS to assess changes in people's subjective representations over time. It would be particularly desirable, for example, if it were possible to show that young subjects tended to organize concepts along perceptual dimensions, whereas adults tended to employ more abstract dimensions.

One fairly straightforward illustration of this approach is a study by Howard and Howard (1977). They obtained similarity judgments for a set of 10 animals that were chosen from Henley's (1969) original set. The subjects were first-graders, third-graders, sixth-graders, and college students. The resulting matrices obtained from all subjects were analyzed with Carroll and Chang's (1970) INDSCAL program. A three-dimensional space was obtained in which the three dimensions extracted were *size*, *domesticity*, and *predativity*. These last two dimensions are usually thought of as equivalent (see Rips et al., 1973), but Howard and Howard (1977) make a good case that these dimensions are distinguishable although not uncorrelated. For the domesticity dimension, mouse, rabbit, lion, deer, and bear are on one side of the midpoint, and horse, cow, sheep, dog, and pig are on the other side. For the predativity dimension, lion and bear are at

one extreme; mouse, rabbit, and deer all cluster near the other pole of this dimension.

Howard and Howard assessed age differences by looking at the weight assigned to each dimension in the subject space. Averaging over subjects within age groups, they found that younger subjects tended to emphasize the size dimension, while older subjects emphasized the more abstract domesticity and predativity dimensions. More specifically, first- and third-grade subjects placed much greater weight on the size dimension than on either of the other two. Sixth-graders placed equal weight on the size and predativity dimensions and less weight on the domesticity dimensions. College students were essentially the opposite of the young children; they placed greatest weight on the predativity dimension and least weight on the size dimension. Thus, it does indeed seem that increasing age is accompanied by an increasing reliance on abstract dimensions, at least for this stimulus domain.

A more complex analysis along similar lines has been performed by Miller and Gelman (1983). They used techniques developed by Arabie, Kosslyn, and Nelson (1975) to investigate the development of the concept of number. As in Howard and Howard (1977), four age groups were used: kindergartners, third-graders, sixth-graders, and graduate students. Using a modification of the method of triads, Miller and Gelman (1983) had subjects judge which of three digits was most similar and which was least similar. To cut down on the number of triads to be judged by very young children, Miller and Gelman used a balanced incomplete sampling procedure developed by Arabie et al. (1975). The resulting similarity matrices were analyzed separately for each age group using KYST2A (Kruskal, Young, & Seery, 1977). Additional analyses were also performed; these will be discussed briefly later.

For the 10 digits, two-dimensional solutions were obtained for all four age groups. For the young children, the solution can be described as a semicircle in which the digits are ordered by magnitude. The kindergartners appear to have some difficulty with zero, but they and the third-graders otherwise provide a perfect ordering in terms of magnitude along the semicircle. Older children and adults, in contrast, reveal more than magnitude information

in their solutions. For sixth-graders, the odd-even dimension is clearly present in addition to a magnitude dimension. Particularly for adults, other relationships are also present; and their solution, in some cases, does not order the digits perfectly along the magnitude dimension. For example, the powers of two (2, 4, 8) are closer together in the two-dimensional solution than they should be according to the magnitude dimension.

Thus, Miller and Gelman (1983) were able to demonstrate rather elegantly that there is increasing complexity in children's conception of number as age increases. For young children, the concept of number is strongly associated with counting. With age, the concept becomes richer, as other relations (oddness, or powers) become part of the concept.

Supporting these conclusions, Miller and Gelman (1983) employed the INDCLUS<sup>1</sup> program (Arabie & Carroll, 1983) as a further means of investigating the change with age. In this analysis, all age groups must have the same clusters; what differed (in addition to the additive constant) were the weights assigned to these clusters by each group of subjects. Five of the seven clusters pertained to counting, and the other two were the odd numbers excluding 1 (3, 5, 7, 9) and the powers of 2 (2, 4, 8). For the kindergartners and the third-graders, the five counting clusters were all assigned higher weights than these last two clusters. For the adults, on the other hand, the powers-of-2 cluster had the highest weight, and the weight for the odd-numbers-excluding-1 cluster was fourth, only slightly behind the large-numbers (6, 7, 8, 9) cluster. Sixth-graders were between these two extremes. The weight for the powers-of-2 cluster ranked third, and the weight for the odd-numbers-excluding-1 cluster ranked fifth.

Thus, this clustering analysis very nicely complements the MDS analysis in that it shows an increasing complexity in the concept of number with increasing age. Young children's conception of number is almost entirely dependent on counting; greater complexity occurs with increasing

age. Methodologically, it is interesting to note that the clustering analysis performed by Miller and Gelman (1983) parallels the scaling work done by Howard and Howard (1977). In both cases, one representation was obtained: a three-dimensional solution in the former and a set of seven clusters in the latter. In both cases, the argument for developmental change relies on an orderly change in the weights assigned to each dimension (in the former case) or each cluster (in the latter case). The Miller and Gelman (1983) paper is particularly convincing because, in addition to this logic, these authors also examined another representational measure (in this case MDS solutions) independently, in which the representation was obtained independently for each age group.

#### Increased Dimensionality Resulting from Examination of Different Contexts

Previous examples have described a few of the studies that have used MDS to index change. Phrased more generally, these studies have looked at the representation of a stimulus domain in several different contexts. One of the interesting byproducts of this approach is that it may be possible to extract more dimensions in this circumstance than if the same stimulus domain is examined in only one context. It is suggestive, for example, that Howard and Howard (1977) were able to extract three dimensions from a set of 10 animal names, where Rips et al. (1973) were able to extract only two from a set of 14 animal names, with both using the INDSCAL model.

The most striking example of this higher dimensionality is an analysis of consonant phonemes performed by Soli and Arabie (1979). In this study, they used the data from Miller and Nicely (1955), which they first transformed to conform more closely to the INDSCAL model (see Arabie & Soli, 1982, for the details and justification of this procedure). In contrast to earlier studies, Soli and Arabie (1979) also used the full set of confusion matrices (including those with relatively severe distortion) in their analysis. They obtained a four-dimensional solution that accounted for 69% of the variance.

<sup>1</sup>INDCLUS (for INDividual Differences CLUstering) is the three-way version of MAPCLUS, the Arabie and Carroll (1980) algorithm for fitting the Shepard-Arabie (1979) ADCLUS model.

Their first dimension ordered the phonemes in terms of periodicity/burst order with /n/ and /m/ at one end of the dimension and /p/, /t/, /k/, /f/, and /s/ at the other extreme. The second dimension ordered the phonemes in terms of their first formants, with the voiced consonants at one extreme and the nasals and voiceless stops at the other. The third dimension was ordered in terms of the second formants, with phonemes with falling second formants at one end and phonemes with rising second formants at the other extreme. Finally, the fourth dimension was ordered in terms of spectral dispersion. This dimension separates two groups of fricatives from the other phonemes and is largely relevant to a particular listening condition.

Soli and Arabie (1979) have also been able to show that the salience of these dimensions varied markedly with the kind of listening condition. For example, both the periodicity/burst dimension and the first formant dimension were salient under extreme degradation. For the second formant and spectral dispersion dimensions, increasing degradation decreased the weight given to these dimensions.

Thus, by their ability to make the Miller-Nicely data conform to the INDSCAL model, Soli and Arabie have been able to extract higher dimensionality from this data set than had been previously done. Moreover, this higher dimensionality has enabled them to make some arguments concerning the relative importance of acoustic as opposed to phonemic properties in the underlying representation. By examining a wide variety of listening contexts, Soli and Arabie seem to have obtained additional information out of an old, and often analyzed, set of data.

#### MDS and the Underlying Representation

There have been some efforts to use MDS and related techniques to determine the appropriate underlying representation. Friendly (1977) for example, examined three possible representations of information in memory: dimensional representations, tree-structure representations, and network representations. According to his definitions, di-

dimensional representations are structures where each stimulus is represented by a continuous value on a number of dimensions. Tree structures are hierarchical representations, and networks are undirected graphs. Friendly has argued that there is a uniquely appropriate method for analyzing data for each of these three types of memory models. Specifically, MDS is the proper procedure for dimensional representations, and hierarchical clustering is the appropriate representation for tree structures. For purposes of this discussion, Friendly has restricted himself to minimum-spanning trees (thereby imposing a hierarchical requirement) as the appropriate tool of analysis for network representations.

Unfortunately, the most straightforward conclusion does not appear to be warranted. That is, it is not the case that data are analyzed with all three procedures and then the correct representation determined on the basis of which procedure best fits the data. In fact, Friendly was careful to make this point. However, it does appear that such a procedure can provide some evidence in favor of one method or another and that the methods are reasonably sensitive to variations in underlying structure.

Friendly's approach can be characterized as theory driven. Instead of examining the data in an effort to infer the appropriate underlying representation, he starts with a theoretical structure and then looks to see if that hypothesized theoretical organization was evident in the data. In his examples, he assumes a particular structure and then analyzes the data by an appropriate method to ascertain if the results are consistent with the hypothesized structure. In his examples, the data are always free recall, where the proximity matrix is derived by the method summarized earlier in this paper. However, this methodology is suitable for any proximity matrix.

This general approach has been carried a step further by Reitman and Rueter (1980). They held a particular theory of free recall and wished to obtain some confirming evidence. In their theory, chunks are an important component, and the order in which the stimuli should be recalled is also prescribed. Their technique, only peripherally related to MDS, first identifies the chunks, and then par-

titions these chunks into a lattice. Lastly, this lattice is converted into an ordered tree. Although the details of this procedure are not important here, Reitman and Reuter have been able to show that their method can recover the structure of a known stimulus set. In a subsequent example, they were able to illustrate nicely the differences in the organization of ALGOL terms for novice and expert users. Hirtle (1982) has recently extended this work.

Thus, it appears that one useful application of MDS and related techniques is in evaluating theories. If the underlying representation is sufficiently well specified, then it should be possible to obtain data from which the underlying representation could be recovered. At a minimum, a successful outcome argues for the plausibility of the theory.

### The Nature of the Underlying Representations

Although the examples cited in the preceding section suggest that these techniques can profitably be used to assess the viability of a particular theory, they explicitly do not attempt to compare competing theories directly. However, a recent paper by Pruzansky, Tversky, and Carroll (1982) has suggested that such comparisons may be both possible and profitable. Pruzansky et al. compared MDS (as exemplified by KYST2A) with tree structures (as exemplified by Sattath & Tversky's, 1979, ADDTREE). Pruzansky et al. first demonstrated that each provided a better fit to artificial data generated from its appropriate representation than the other method did. Specifically, KYST provided a better fit when the data set was generated from a plane, and ADDTREE provided a better fit when the data set was generated by a tree. Moreover, this relationship held over varying numbers of stimuli and levels of noise.

Subsequently, Pruzansky et al. found two empirical measures that predicted which procedure would provide the better fit (as measured by either the product-moment correlation or by Stress Formula 2). The first measure is the skewness of the data, as determined by the standard measure of skewness: the third central moment divided by the cubed standard deviation. The second measure,

elongation, is more complicated. This statistic is motivated theoretically by the nature of binary rooted trees. For any triple of terminal nodes, it is usually the case that two of the terms form a subcluster that does not include the third. Thus,  $\phi_{ij} \leq \phi_{jk} \leq \phi_{ik}$  if  $i$  and  $j$  are the two terms that form the subcluster. It would thus also be expected that the differences would have the relationship  $\phi_{jk} - \phi_{jk} \leq \phi_{ki} - \phi_{ij}$ . Looking at the triangle of distances, the middle side is closer in length to the long side than it is to the short side. The elongation measure used by Pruzansky et al. is simply the proportion of triangles in which this difference relation holds.

Pruzansky et al. (1982) examined the utility of these measures by calculating them for 20 data sets. Generally, when the proportion of elongated triangles exceeded .65, ADDTREE provided a better fit to the data; otherwise, KYST did. Similarly, when the skewness was less than  $-.4$ , ADDTREE provided a better fit. Generally, these two measures tended to be negatively correlated, and there were no cases where the two measures truly conflicted.

The result of greatest interest is that the data sets for which ADDTREE provided the better fit could generally be described as conceptual stimuli, whereas the data sets that could be classified as perceptual were almost always better fit by KYST. The data sets that involved colors, forms, or letters were usually fit better by KYST, but data sets that contained names of category exemplars were usually better fit by ADDTREE. Although Pruzansky et al. noted that factorial designs (as commonly used in perceptual data sets) tended to favor KYST and that natural selection (as commonly used in conceptual data sets) tended to favor ADDTREE, they offered no other explanation for this dichotomy.

One problem in fitting the conceptual data sets could be in the nature of the stimuli. Although most of the conceptual data sets examined by Pruzansky et al. are unpublished, there is sufficient familiarity with the procedural details of the first eight they studied to offer some speculation. The first seven data sets (referenced by Pruzansky et al. as the Mervis et al. data sets) involved 19 exemplar names and one category label. The eighth data set (Hendley, 1969) consisted of 30 animal names with no category label. For the first seven data sets, the

superiority of ADDTREE, as measured by the difference in both of Pruzansky et al.'s indices of goodness of fit averaged .15 for  $r^2_i$  and .07 for  $r^2_M$ . The superiority of ADDTREE for the eighth data set was about half these means, .08 and .03, respectively. Only one of the seven data sets examined by Mervis et al. (1980) had smaller differences between the two procedures than Henley's (1969) data set.

### Problems with Categories

Because of the way in which people judge categories, it is difficult for MDS programs to accommodate categorical data sets. In particular, people seem to regard all exemplars as similar to their category name. For example, Shoben (1976) found that the similarity between goose and bird was virtually equivalent to the similarity between hawk and eagle, which in turn was less than the similarity between hawk and bird. More abstractly, even atypical exemplars of a category are regarded as highly similar to their category name. This general finding causes problems for scaling algorithms. For example, the similarity between robin and goose is quite low, but each is highly similar to the concept of bird. In terms of distances in a multidimensional space, robin and goose should be quite far from each other, but both must be quite close to bird. Obviously, it is impossible to satisfy both conditions. This kind of conflict is not present if superordinate terms are not among the test stimuli.

It is not clear how this problem can be solved. One possibility is to adapt a proposal of Krumhansl's (1983) and try to measure typicality of terms separately. Krumhansl has argued that similarity is a function, not only of the distance between the two objects, but also of the distance (typicality) of each of the objects to the superordinate. In the context of the musical scale, Krumhansl was able to vary the context (which scale is used) and thereby to show that the vertical structure (typicality) has an effect on the similarity judgment independent of the effect of the distance between the two objects (horizontal structure). In fitting the function that described the contribution of both horizontal and vertical structure, Krumhansl used the chroma cir-

cle of tones as the horizontal structure. For other domains, there may be no comparable underlying structure available; the applicability of this approach is thus not immediately obvious.

Another approach is to ask the purpose for which the MDS solution is obtained. If, for example, one does not wish to claim that the underlying representation is spatial but rather to use the distances in the multidimensional space to predict some dependent variable (such as RT) in another experiment, then other options may be available. For example, Shoben (1976) was particularly interested in the distances obtained by MDS between category exemplars and their superordinate. His first scaling solution used three dimensions to depict six mammal exemplars, six bird exemplars, and two superordinate terms. An examination of the Shepard diagram, which plots fitted distances against the original data, indicated that the greatest disparity between the actual data and the fitted distances occurred for pairs that involved superordinate terms. To use precisely these distances in subsequent analyses, Shoben rescaled the data and used the weights option to ensure that the disparity between actual and fitted distances would be minimal for pairs involving a superordinate term. In this case, this manipulation was successful in that the exemplar-superordinate distances obtained from the weighted MDS solution predicted RT in a Same-Different task quite well.

### Choosing an Appropriate Representation

There are no simple rules for choosing the appropriate representation. Instead, there are a number of criteria whose utility will depend on the particular circumstance. One obvious benchmark is the interpretability of the solution. Clearly, an uninterpretable solution is not convincing evidence for the plausibility of the underlying representation. More often, there are competing arguments concerning which of two representations is more interpretable. For example, Sattath and Tversky (1979) have argued that the ADDTREE representation of Henley's animal data is more understandable than the one provided by MDS.

Sattath and Tversky have also used another criterion for deciding the appropriate representation: goodness of fit. Pruzansky et al. (1982) have also employed this criterion in their comparison of ADDTREE and KYST. From the two measures they used, it is clear that different measures of goodness of fit may give different results, and, moreover, that a particular index of goodness of fit may favor one method over another. For example, one of the indices used by Pruzansky et al. is directly related to stress, which is minimized by KYST. Not surprisingly, KYST fares better when this particular index is used.

Pruzansky et al. have also suggested that there are statistical features of the data that may also help in making a decision. As noted earlier, they suggested that the skewness and the elongation of the data will enable determining in advance whether ADDTREE or KYST will yield a representation that is more faithful to the original data. These two values can therefore be measured with the expectation of having some idea of the procedure that will give a better fit.

All of these procedures are useful, but it can be argued that they do not provide a definitive answer, either individually or collectively. As Krumhansl (1983) has noted, there are often theoretical reasons for adopting a particular approach. She has given an excellent example of this theoretical motivation in her work with music. As noted earlier, in the absence of specific context, the notes of a scale may be conceived as a chroma circle, much like the results for colors. With context, however, she was able to show that the similarity of a pair of tones depended not only on the distance in the chroma circle, but also on the distance to the superordinate, which in this case was the scale. In a C major scale, for example, G is closer to C than might be expected from the chroma circle alone. Similarly, in this scale, C# is further from C than would be expected in the absence of explicit context.

This approach has proven very fruitful. Not only was Krumhansl able to provide a satisfactory spatial representation of her own data, but with some additional assumptions related to those in her spatial density model (Krumhansl, 1979), she was able

to account for her data at more microcosmic levels as well. For example, she could account for the similarity of  $i$  to  $j$  being greater than the similarity of  $j$  to  $i$  when  $j$  was closer to the superordinate term.

Thus, it seems that there is no easy answer to the question of the single most appropriate representation. All of the methods individually have problems, and there is no magic rule for combining approaches. Moreover, slight variations in technique can have fairly substantial effects on the outcome. Perhaps the most dramatic example of the large improvement in the scaling solution as a function of a relatively small (to a cognitive psychologist) change in procedure is the work of Soli and Arabie (1979) on the Miller-Nicely data. By using INDSCAL rather than MDSCAL and employing a transformation of the original data to make them better conform to the assumptions of the INDSCAL model, they were able to recover many more dimensions than Shepard (1972) did in his seminal work on this data set. Consequently, it is premature to specify any rule or rules for determining the preferred underlying representation for any particular data set.

### Problems and Pitfalls in MDS Analysis

It might be speculated that one of the reasons why MDS analysis is applied so infrequently to cognitive problems is that it is a complicated set of techniques with which few cognitive psychologists are familiar. Because of this complexity, there are a number of pitfalls into which cognitive psychologists are prone to fall and there are a number of interesting cognitive problems to which the power and potential applicability of MDS analysis have not been applied.

#### Number of Stimuli

The simplest problem to avoid is using too few stimuli for the number of dimensions. In two dimensions, for example, it is trivially easy to accommodate the similarities among six stimuli. It is very likely that a seemingly low value of stress will be obtained even if the similarities are gen-

erated randomly. Klahr (1969), for example, has performed monte carlo simulations that provide users with stress values beyond which there can be rejection of the hypothesis that the data are random. These values are provided, of course, for a given number of stimuli and a given number of dimensions.

As a general rule of thumb, it would appear that it is very unwise to use fewer than 9 or 10 stimuli in a two-dimensional solution and fewer than 14 or 15 when three dimensions are employed. Moreover, it should be noted here that the stress values given in Klahr are artificially high (Arabie & Boorman, 1973) in that they used a now obsolete and clearly inferior initial configuration. Nevertheless, these tables (or their equivalent) are the best available and they should be consulted. The major import of this discussion is that there should be wariness of MDS solutions performed with too few stimuli. Thus, Shafto's (1973) argument for case grammar, which depends heavily on a two-dimensional representation of eight stimuli, should be viewed with a high degree of skepticism.

Although some general guidelines for nonmetric MDS can be formulated, there are no such easy rules for other commonly used algorithms, such as INDSCAL or SINDSCAL. Here, the degree of constraint provided by the data depends, not only on the number of stimuli, but also on the number of subjects. For example, it would be incorrect to fault Soli and Arabie (1979) for extracting four dimensions from the Miller-Nicely data when there were 16 stimuli and 16 "subjects" (actually listening conditions). Clearly, some kind of simulation work must be done before treating this issue definitively. At a minimum, however, when the number of subjects is very small, it would seem prudent to follow the strictures outlined above for nonmetric MDS analysis. Moreover, when the number of subjects is small, the orientation of the axes may not be well defined.

#### Minkowski $r$ Metrics

Nearly all of the applications of MDS in cognitive psychology have assumed Euclidean distance. Consistent with Shepard (1964) and with

Garner (1972), many researchers have gone on to assume that their stimuli were wholistic or integral, whereas others never considered any other metric.

There are at least two other metrics that are of inherent interest to cognitive psychologists. In most general terms, the Minkowski family of metrics is described by Equation 4 where  $r \geq 1$ :

$$d_{ij} = [\sum_k |x_{ik} - x_{jk}|^r]^{1/r} . \quad [4]$$

If the exponent  $r$  is equal to 2, Equation 4 gives the familiar Euclidean case. Two other cases are of particular interest: when  $r$  equals 1 and when  $r$  approaches infinity. In the latter case, the distance between two objects reduces to the maximum difference between the objects on any dimension. Thus, two objects that differ from each other by a moderate amount on each of three dimensions are actually closer to each other than two objects that are identical on two dimensions but very different on a third. Although there are few applications of this *dominance* metric, Arnold (1971) has argued for this particular metric in the semantic terms he examined.

When  $r$  equals 1, the distance between two points is determined by adding the differences on all dimensions. In two dimensions, this procedure resembles distances in a city where a person must travel three blocks north and two blocks east, for example, to get from 43rd Street and 7th Avenue to 46th Street and 5th Avenue in New York City. Because of this analogy, this value of  $r$  is often referred to as the city-block metric. This metric is particularly interesting to cognitive psychologists because it implies that the dimensions recovered by MDS are separable, rather than integral.

Despite the inherent appeal of the city-block metric, there are understandable, though not excusable, reasons for its lack of prominence. The first reason is that numerical problems are more likely to be encountered in non-Euclidean metrics. For the Euclidean distance model, either a small number of random initial configurations or a rational starting configuration can be selected and there will be reasonable confidence that a near optimum solution will be obtained. This procedure will rarely yield satisfactory results for the city-block metric. Local minimum problems are far more

common; it appears that many more random starting configurations must be used (Arabie, 1973) when  $r$  equals 1 than when the Euclidean metric is used.

Fortunately, Arnold (1971) has developed a procedure that seems to surmount most of these problems. In a little noticed paper, Arnold approached both the city-block metric and the dominance metric (he set  $r = 32$  to approximate the dominance metric) by a successive approximation procedure. He first obtained a solution in Euclidean space; this solution was then the initial configuration for  $r = 1.5$  and  $r = 2.5$ . Approaching  $r = 1$ , he used the resulting configuration when  $r = 1.5$  as the starting configuration for a solution with  $r$  set at 1.25, and so on, until he obtained a solution with  $r = 1$ . In a similar way, Arnold obtained a representation of his stimuli using the dominance metric. Using this procedure, Arnold found that the Euclidean solution gave the poorest fit (as measured by stress) to the data. Stress declined monotonically as  $r$  decreased from 2 to 1, and it also declined monotonically as  $r$  increased from 2 to 32. Interestingly, the best fit was obtained with the dominance metric.

Although this procedure appears very successful and well worth using, it does appear that it has one important limitation: It does not seem to work well with two-dimensional solutions (Carroll & Arabie, 1980). The reasons for this failure are not presently known.

### Using the Options

Finally, it appears that many users of MDS are not aware of all the options that most computer programs permit. For example, although INDSCAL is characterized as three-way (stimuli by stimuli by subjects), the last way need not be actual individual subjects. Soli and Arabie (1979), for example, have used conditions in lieu of subjects very successfully. Miller and Gelman (1983) used age as their individual difference variable in their INDCLUS analysis. It would seem that INDSCAL might be particularly helpful in examining the effects of context in cognition.

There are a number of more technical options within the often used programs that are seldom used but are potentially useful. For example, KYST permits the similarities to be weighted so that, for example, similarities could be weighted according to their reliability or consistency across subjects. Friendly (1977) has shown that this kind of weighting procedure can be used in obtaining a satisfactory MDS representation of free recall data. If tied data are viewed as particularly meaningful, then it might be desirable to elect the secondary approach to ties instead of the primary approach. There are other examples, as a careful reading of the program documentation will reveal; the important point is that there are options available that can be usefully applied to problems in cognitive psychology only if users are aware of them.

### Conclusions

MDS has been usefully and frequently applied to problems in cognitive psychology. In many areas, such as the work in music perception and the development of models of analogical reasoning, the applications have been particularly inventive. Some of the examples described above illustrate that there are many more problems in cognitive psychology to which MDS and related procedures might profitably be applied. Particularly with the advent of new methods for determining the underlying structural representation, such as the work of Pruzansky et al. (1982) and Krumhansl (1983), it appears that MDS may become less a method of exploratory analysis and more a method of theoretical analysis. In addition, developments in individual difference models, such as the introduction of INDCLUS, suggest that it may be possible to examine the structure of a set of stimuli as a function of condition or experimental context by using these algorithms.

From the perspective of cognitive psychology, it appears that the limitation on MDS is that it tells little about processing. Cognitive psychology has many examples of problems where two very different structures coupled with two different sets of processing assumptions can predict the same results. Two prominent examples are the propositional imagery debate and semantic memory models.

Anderson (1978) has even argued that some of these fundamental structural questions can never be resolved. At a minimum, it appears that choices can certainly not be made without a consideration of the accompanying processing assumptions.

Thus, if the goal of the cognitive psychologist is to construct sensible theories of cognition, then it is clear that MDS analysis can be quite helpful in half of the battle. Although it may not help directly in determining the underlying process, it does provide evidence about the structural properties of the data and how that structure might change under various conditions. Moreover, MDS analysis might actually lead to the development of a processing model, as it did for Rumelhart and Abrahamson (1973). With the current interest of cognitive psychology in the nature of the underlying representation and the current interest in the effects of context, the potential applicability of MDS and related techniques in this field appears to be large.

### References

- Anderson, J. A. Neural models with cognitive applications. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale NJ: Erlbaum, 1977.
- Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington DC: Winston, 1974.
- Arabie, P., & Boorman, S. A. Multidimensional scaling of measures of distance between partitions. *Journal of Mathematical Psychology*, 1973, *10*, 148–203.
- Arabie, P., & Carroll, J. D. MAPCLUS: A mathematical programming approach to fitting the ADCLUS model. *Psychometrika*, 1980, *45*, 211–235.
- Arabie, P., & Carroll, J. D. INDCLUS: An individual differences generalization of the ADCLUS model and the MAPCLUS algorithm. *Psychometrika*, 1983, *48*, 157–169.
- Arabie, P., Kosslyn, S. M., & Nelson, K. E. A multidimensional scaling study of visual memory of 5-year-olds and adults. *Journal of Experimental Child Psychology*, 1975, *19*, 327–345.
- Arabie, P., & Soli, S. D. The interface between the types of regression and methods of collecting proximity data. In R. G. Golledge & J. N. Rayner (Eds.), *Proximity and preference: Problems in the multidimensional analysis of large data sets*. Minneapolis: University of Minnesota Press, 1982.
- Arnold, J. B. A multidimensional scaling study of semantic distance. *Journal of Experimental Psychology Monograph*, 1971, *90*, 349–372.
- Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 1968.
- Bisanz, G. L., LaPorte, R. E., Vesonder, G. T., & Voss, J. F. On the representation of prose: New dimensions. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 357–358.
- Caramazza, A., Hersh, H., & Torgerson, W. Subjective structures and operations in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 103–117.
- Carroll, J. D., & Arabie, P. Multidimensional scaling. *Annual review of psychology*, 1980, *31*, 607–649.
- Carroll, J. D., & Chang, J. J. Analysis of individual differences in multidimensional scaling via an *N*-way generalization of 'Eckart-Young' decomposition. *Psychometrika*, 1970, *35*, 283–319.
- Clark, H. H. On the meaning and use of prepositions. *Journal of Verbal Learning and Verbal Behavior*, 1968, *7*, 421–431.
- Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671–684.
- Ekman, G. Dimensions of color vision. *Journal of Psychology*, 1954, *38*, 467–474.
- Fillenbaum, S., & Rapaport, A. *Structures in the subjective lexicon*. New York: Academic Press, 1971.
- Friendly, M. L. In search of the *m*-gram: The structure of organization in free recall. *Cognitive Psychology*, 1977, *9*, 188–249.
- Friendly, M. Methods for finding graphic representations of associative memory structures. In C. R. Puff (Ed.), *Memory organization and structure*. New York: Academic Press, 1970.
- Garner, W. R. Information integration and form of encoding. In A. Melton & E. Martin (Eds.), *Coding processes in human memory*. New York: Wiley, 1972.
- Henley, N. M. A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 176–184.
- Hirtle, S. C. Lattice-based similarity measures between ordered trees. *Journal of Mathematical Psychology*, 1982, *25*, 206–225.
- Howard, D. V., & Howard, J. H. A multidimensional scaling analysis of the development of animal names. *Developmental Psychology*, 1977, *13*, 108–113.
- King, M. C., Gruenewald, P., & Lockhead, G. R. Classifying related stimuli. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, *4*, 417–427.

- Kintsch, W. *The representation of meaning in memory*. Hillsdale NJ: Erlbaum, 1974.
- Klahr, D. Monte Carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure. *Psychometrika*, 1969, 34, 319–330.
- Krumhansl, C. L. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 1979, 11, 346–374.
- Krumhansl, C. L. *Set-theoretic and spatial models of similarity: Some considerations in application*. Paper presented at the Mathematical Psychology meetings, Boulder CO, August, 1983.
- Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, 1982, 8, 24–36.
- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, 29, 1–27. (a)
- Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, 29, 115–129. (b)
- Kruskal, J. B., Young, F. W., & Seery, J. B. *How to use KYST 2, a very flexible program to do multidimensional scaling and unfolding*. Murray Hill NJ: Bell Telephone Laboratories, 1977.
- LaPorte, R. E., & Voss, J. F. Prose representations: A multidimensional scaling approach. *Multivariate Behavioral Research*, 1979, 14, 39–56.
- Luce, R. D. *Individual choice behavior*. New York: Wiley, 1959.
- Miller, G. A., & Nicely, P. E. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 1955, 27, 338–352.
- Miller, K., & Gelman, R. The child's representation of number: A multidimensional scaling analysis. *Child Development*, 1983, 54, 1470–1479.
- Norman, D. A., Rumelhart, D. E., & The LNR Research Group. *Explorations in cognition*. San Francisco: Freeman, 1975.
- Paivio, A. *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston, 1971.
- Pollard-Gott, L. Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology*, 1983, 15, 66–94.
- Pruzansky, S. How to use SINDSCAL: A computer program for individual differences in multidimensional scaling. Murray Hill NJ: Bell Telephone Laboratories, 1975.
- Pruzansky, S., Tversky, A., & Carroll, J. D. Spatial versus tree representations of proximity data. *Psychometrika*, 1982, 47, 3–24.
- Raaijmakers, J. G. W., & Shiffrin, R. M. SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14). New York: Academic Press, 1980.
- Ratcliff, R. A theory of memory retrieval. *Psychological Review*, 1978, 85, 59–108.
- Rapaport, A., & Fillenbaum, S. An experimental study of semantic structures. In A. K. Romney, R. N. Shepard, & S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences* (Vol. 2). New York: Seminar Press, 1972.
- Reitman, J. S., & Reuter, H. H. Organization revealed by recall orders and confirmed by pauses. *Cognitive Psychology*, 1980, 12, 554–581.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1–20.
- Ross, B. H. *A note on scaling memory theories*. Unpublished manuscript, 1983.
- Rumelhart, D. E., & Abrahamson, A. A. A model for analogical reasoning. *Cognitive Psychology*, 1973, 5, 1–28.
- Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale NJ: Erlbaum, 1977.
- Sattath, S., & Tversky, A. Additive similarity trees. *Psychometrika*, 1979, 42, 319–345.
- Schaeffer, B., & Wallace, R. Semantic similarity and the comparison of word meanings. *Journal of Experimental Psychology*, 1969, 82, 343–346.
- Schaeffer, B., & Wallace, R. The comparison of word meanings. *Journal of Experimental Psychology*, 1970, 86, 144–152.
- Schank, R. C. Language and memory. *Cognitive Science*, 1980, 4, 243–284.
- Schwartz, R. M., & Humphreys, M. S. Similarity judgments and free recall of unrelated words. *Journal of Experimental Psychology*, 1973, 101, 10–15.
- Shafto, M. The space for case. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 551–562.
- Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 1962, 27, 125–140. (a)
- Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika*, 1962, 27, 219–246. (b)
- Shepard, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, 1, 54–87.
- Shepard, R. N. Psychological representations of speech sounds. In E. E. David & P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill, 1972.

- Shepard, R. N. Multidimensional scaling, tree-fitting, and clustering. *Science*, 1980, *210*, 390–398.
- Shepard, R. N., & Arabie, P. Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review*, 1979, *86*, 87–123.
- Shoben, E. J. The verification of semantic relations in a same-different paradigm: An asymmetry in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 365–379.
- Soli, S. D., & Arabie, P. Auditory versus phonetic accounts of observed confusions between consonant phonemes. *Journal of the Acoustical Society of America*, 1979, *66*, 46–59.
- Sternberg, R. J. Component processes in analogical reasoning. *Psychological Review*, 1977, *84*, 353–378.
- Stevens, S. S., & Volkman, J. The relation of pitch to frequency: A relative scale. *American Journal of Psychology*, 1940, *53*, 329–353.
- Tulving, E. Ecphoric processes in recall and recognition. In J. Brown (Ed.), *Recall and recognition*. London: Wiley, 1975.

#### Acknowledgments

*Preparation of this paper was supported in part by Grant No. BNS 82-17674 from the National Science Foundation. The author thanks Lawrence E. Jones for his careful reading of an earlier draft.*

#### Author's Address

Send requests for reprints or further information to Edward J. Shoben, Department of Psychology, University of Illinois, 603 E. Daniel Street, Champaign IL 61820, U.S.A.