014523-3-T

HUMAN PERFORMANCE CENTER
DEPARTMENT OF PSYCHOLOGY

The University of Michigan, Ann Arbor

Constraining Nonmetric Multidimensional Scaling Configurations

ELLIOT NOMA and
JANICE JOHNSON

D D C

RECEIVED

OCT 20 1977

D

Technical Report No. 60

August 1977

THE HUMAN PERFORMANCE CENTER

DEPARTMENT OF PSYCHOLOGY

The Human Performance Center is a federation of research programs whose emphasis is on man as a processor of information. Topics under study include perception, attention, verbal learning and behavior, short- and long-term memory, choice and decision processes, and learning and performance in simple and complex skills. The integrating concept is the quantitative description, and theory, of man's performance capabilities and limitations and the ways in which these may be modified by learning, by instruction, and by task design.

The Center issues two series of reports. A Technical Report series includes original reports of experimental or theoretical studies, and integrative reviews of the scientific literature. A Memorandum Report series includes printed versions of papers presented orally at scientific or professional meetings or symposia, methodological notes and documentary materials, apparatus notes, and exploratory studies.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 014523-3-T, TR-60	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CONSTRAINING NONMETRIC MULTIDIMENSIONAL SCALING CONFIGURATIONS.		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER Technical Report No. 60
7. AUTHOR(s) Elliot Noma and Janice Johnson		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0648
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Performance Center University of Michigan Ann Arbor, Michigan 48109		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Engineering Psychology Program Office of Naval Research 800 N. Quincy Street, Arlington, Virginia 22217		12. REPORT DATE August, 1977
		13. NUMBER OF PAGES 56
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

multidimensional scaling
interpretation of scaling outputs
goodness-of-fit measures
model testing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The interpretation of multidimensional scaling outputs is usually based on the identification and labeling of geometric structures in space. Some of the most commonly used structures are reviewed. Interpretation of the scaling outputs requires many psychological and mathematical assumptions including the assumption that the configuration with the lowest stress is the output desired. Unfortunately, little is known about the uniqueness of a configuration generated from fallible data and this non-uniqueness also affects the interpretation of the spatial outputs. A scaling method incorporating information in addition to the dissimilarities is

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

Unclassified

Block 20 - con't.

proposed and the implications of this approach for the interpretation of a configuration are discussed.

Constraining Nonmetric Multidimensional

Scaling Configurations [1]

Elliot Noma

Janice Johnson

The University of Michigan

August, 1977

Abstract

The interpretation of multidimensional scaling outputs is
usually based on the identification and labeling of geometric
structures in the space. Some of the most commonly used structures
are reviewed. Interpretation of the scaling outputs requires many
psychological and mathematical assumptions including the assumption
that the configuration with the lowest stress is the output desired.
Unfortunately, little is known about the uniqueness of a configuration
generated from fallible data and this non-uniqueness also affects
the interpretation of the spatial outputs. A scaling method incor-
porating information in addition to the dissimilarities is proposed
and the implications of this approach for the interpretation of
a configuration are discussed.

Constraining Nonmetric Multidimensional

Scaling Configurations [1]

Elliot Noma and Janice Johnson

The University of Michigan

Introduction

People organize their experience of the world. In some fashion,
they build a cognitive structure allowing them to see the similarities
and differences among events. This facility is crucial, since people
can then use knowledge derived from past experiences to deal with
present or anticipated future situations. Similarities and dif-
ferences among events may be modeled as aggregates of similarities
(or differences) along psychological continua or between psychological
categories. In addition, these psychological continua and categories
are assumed to correspond in some way to identifiable stimulus
properties. For example, they may correspond to the physical dimen-
sions or semantic categories of the stimuli.

Many scaling techniques attempt to define correspondences among
three measures: similarity judgments, subjective measures, and ob-
jective measures. In most methods, however, the experimenters make
a priori assumptions about certain aspects of cognitive structure.
For example, in magnitude estimation a subject judges the extent
to which a stimulus manifests a prespecified characteristic. This
estimate is assumed to be function of objectively measured parameters.

By contrast, structures in a multidimensional scaling need
not be identified beforehand. In the model used by Shepard-Kruskal
nonmetric multidimensional scaling, each stimulus is represented

as a point in a coordinate space. The interpoint distances are measured using a Minkowski metric, which means that if points j and k have coordinates x_{jk}, x_{j2}, ..., x_{jd} and x_{k1}, x_{k2}, x_{kd} respectively in a d-dimensional space, then their interpoint distance is

$$d_{jk} = [\sum_{i=1}^{d} |x_{ji} - x_{ki}|^r]^{1/r}.$$

The exponent r, which can range from 1 to ∞, determines the type of Minkowski metric. When r = 2, the above equation yields the Pythagorean Theorem, which of course defines the Euclidean distance function. The Euclidean metric is by far the most commonly used, but sometimes it is fruitful to consider other metrics. Another common metric, in which r = 1, is often referred to as the City-Block metric.

Starting with a measure of interpoint similarity, dissimilarity, distance, or other measure of interstimulus association, the nonmetric multidimensional scaling algorithm attempts to place stimuli in the space such that stimuli which have been judged to be very similar, not dissimilar, close together, etc. are represented by points that are close to each other in the space. Conversely, very dissimilar stimuli should be represented by points that are far apart. This relationship between similarity and distance is called the monotonicity requirement, because the interpoint distances should be monotonically related to the input dissimilarity measures.

In most cases, however, no set of points in a space with a fixed metric, r, and dimensionality, d, can satisfy the monotonicity requirement. It is assumed, however, that deviations

4

from a perfect fit are due to measurement errors, and an attempt
is made to find the constellation of points that best satisfies
the monotonicity requirement. That is, a scaling algorithm tries
to find a set of point coordinates that minimizes some loss func-
tion in the same way that a linear regression finds parameters
that minimize the sum of squared deviations of expected and ob-
served values. Kruskal (1964a) defines a loss function called
stress:

$$\text{STRESS}_1 = \sqrt{\frac{\sum_{jk}(d_{jk} - \hat{d}_{jk})^2}{\sum_{jk} d_{jk}^2}}$$

where the d_{jk}'s are the distances derived by the scaling algorithm,
and the $\hat{d}_{jk}$'s are the values of the interpoint distances satisfying
the monotonicity requirement. Note that a stress of zero indicates
a perfect fit in the sense of satisfying this requirement. That is
to say, nonmetric multidimensional scaling attempts to define a
configuration that minimizes stress, and then uses this measure
as an indicator of goodness-of-fit (in the same sense as the mean
square due to error in a linear regression).

To minimize stress, an algorithm is employed that starts with
an arbitrary initial configuration and iteratively steps the point
coordinates toward a lower stress configuration. After a pre-
specified number of iterations, or after several iterations that
do not appreciably decrease the stress, the algorithm terminates
and a final configuration is printed. This configuration is often
called the local minimum solution, since any movement of one or
more of the points away from the current location in space will

increase the stress. This does not mean, however, that this is
the absolutely lowest stress configuration. There may exist many
local minimum solutions for a given set of dissimilarities. If
all these solutions could be located, at least one would have the
lowest stress associated with it. This configuration, or group of
configurations, is called the global minimum solution. In most
cases it is assumed that the local minimum found by the multi-
dimensional scaling algorithm is actually the global minimum.
The configuration is then interpreted in terms of correspondences
between groupings of points and psychological, physical, or
semantic dimensions or categories.

On the surface, it seems that multidimensional scaling
methods render a final configuration without reference to any
external constraints or a priori hypotheses. Therefore, multi-
dimensional scaling can be very helpful if a priori hypotheses
are vague or nonexistent, since post hoc interpretations are
often possible (their validity depending, of course, on the repli-
cability of the orderings or groupings). It can also be helpful
in suggesting alternative interpretations by alerting the experi-
menter to qualitative as well as quantitative deviations from
the expected results. Two reasons for the ease of interpretation
of a multidimensional scaling output are the (usually) low dimen-
sionality and Minkowski metric of the output configuration.

However, the unconstrained nature of multidimensional scaling
also has its drawbacks. Aside from the many hidden psychological
assumptions, little is known about the uniqueness of configurations

generated from fallible data, but the programs blindly search out the configuration with the lowest possible stress, and give no information about any other possible configurations. Thus, there is no way to predict how perturbations from the local minimum solution will affect either the stress or the interpretability of the spatial output. It might well be that the configuration with the lowest stress is very difficult to interpret, or does not manifest hypothesized structures, whereas a configuration with only slightly higher stress is dramatically more interpretable or in accord with previous hypotheses.

In this paper we will review some of the theoretical assumptions underlying multidimensional scaling, as well as some of the more traditional methods of interpretation. We also survey the research on the uniqueness of the scaling solution. Finally, we propose an alternative approach to the interpretation problem - CONSCAL, a multidimensional scaling program which allows a user to constrain the scaling solution in accordance with certain hypothesized structures.

In order to understand the questions involved in interpretations, we must review the basic assumptions underlying multidimensional scaling. These assumptions are both mathematical and psychological. We first examine the psychological assumptions, and their implications for deriving and interpreting configurations.

Psychological Assumptions of Multidimensional Scaling

The most crucial psychological assumption of multidimensional scaling is that people's internal stimulus representations may be meaningfully modeled in spatial terms. In other words, it is assumed that a person internally organizes representations of stimuli in a form functionally analogous to a "psychological map." Within this map, the stimuli are points and the interstimulus similarities (or subjective distances) are an increasing function of the distances between points in the map. The interpoint distances in the psychological map are generally taken to be fixed and to be accessible to a subject in a judgment task. This psychological map is the "underlying configuration" which multidimensional scaling methods seek to recover. This assumption has led directly to two lines of research that consider (1) the conditions that perfectly scalable data must satisfy and (2) why deviations from perfectly scalable data, or errors, occur. No satisfactory axiomatization of spatial models has been developed, but some necessary conditions are outlined by Beals, Krantz and Tversky (1968). The testability of such axioms is, however, in doubt. A possible explanation for errors has been explored by Ramsay (1969). He models the errors by accepting the validity of the spatial model and assumes the distances to be constantly varying according to some probability distribution, with any individual judgments being based on a sample taken from this distribution.

Most multidimensional scaling programs also assume that a subject's psychological map can be modeled successfully, or at

least satisfactorily, by a Minkowski r-metric. This is despite the fact that Minkowski r-metrics are only a small subset of the metrics which might be considered (see Shepard, 1974).

It is furthermore assumed that a person can meaningfully organize any set of stimuli. Specifically, the stimuli may be ordered or classified with respect to some psychological referent or referents. These referents are usually assumed to correspond to geometric shapes superimposed on the scaled output. Relating psychological referents to the output configuration is often done by labeling the axes of the coordinate space. After examining the first coordinate of all points, attempts are made to define a psychologically meaningful criterion for distinguishing points with large positive first coordinate values from those large negative values. This process is repeated for all d coordinate axes in a d-dimensional space. Another method attempts to assign psychologically meaningful labels to groups of points that are scaled near each other in the multidimensional space. Such groups of points are referred to as clusters.

As Cliff and Young (1968) point out, a subject's responses depend only indirectly on physical characteristics of the stimuli, and more on how the subject has personally organized the items. A subject's personal organization, moreover, will depend on which similarities or differences between stimuli are most salient or significant to him. It follows from the assumption of personalized psychological referents that, unless there is a high degree of agreement across subjects as to which types of interstimulus relationships are most significant, multidimensional scaling outputs might be expected to vary across subjects. Not only can relative interpoint distances vary, but in many instances the dimensions of variation, the

dimensionality, and even the metric of the space have been shown to be different for different individuals (Hyman & Well, 1967; Spector, Rivizzigno & Golledge, 1976). This means that care must be taken when pooling subject responses. Even though the pooled configuration may be easy to analyze, it also may lead to an incorrect or unrepresentative interpretation. It is better to analyze and interpret scaling plots separately for each subject, or to use a scaling method that creates a group space and indicates how each of the individual spaces may be mapped onto this space. A program which does the latter is INDSCAL (Carroll & Wish, 1974), which assumes that all subject spaces are characterized by the same dimensions, dimensionality, and metric, but which allows for differential weightings of the common dimensions in producing individual configurations. The INDSCAL model is usually interpreted to mean that all subjects evaluate stimuli with respect to the same psychological aspects. They then use a common method of amalgamating these many aspects, but differentially weight them according to individual saliency or attention factors.

Deciding on a Configuration

Although multidimensional scaling programs yield a single
"best" configuration in terms of minimizing stress (assuming that
a global minimum is obtained), this does not imply that they auto-
matically yield the "most correct" or "most acceptable" configuration.
Nor is there any algorithm which does. This is because stress is
only one of several factors used in evaluating a configuration.
Others include metric, dimensionality, degree of degeneracy, and
meaningfulness or interpretability.

Stress and other goodness-of-fit measures can only be rough
guidelines which help one to decide among alternative configurations.
One reason for this is that there are no concrete statistical guide-
lines for determining whether certain stress values imply "signifi-
cantly" good fit, or a "significant difference" in goodness-of-fit
between configurations. Kruskal (1964a) has published some rough
guidelines indicating what he considers to be "excellent", "good",
"fair", etc. stress values. These guidelines, however, are strictly
rule-of-thumb, since stress values are a function of the metric of
the space and the number of points to be scaled, as well as of the
scalability of the data. Monte Carlo studies (see Klahr, 1969)
scaling random distances, and producing stress distributions, have
presented tables of stress values at the .05 levels of the distri-
butions. However, these cannot be taken too seriously because,
in scaling noise, they provide only an irrelevant comparison. Any
sort of redundancy or consistency in the data will increase the
chance of a good fit. Therefore, most subjects in almost any

11

reasonable task will be more consistent than the 95% least consistent
random orderings. The question most experimenters wish to answer
is not whether there is structure in the data - that is assumed.
The question is the nature of the structure, and whether it has been
reasonably represented in a given multidimensional scaling configuration.

Other Monte Carlo studies, such as that of Young (1970), have
randomly picked configurations and added noise to the interpoint dis-
tance matrices. Unfortunately, these analyses confound the measures
of recovery of the true interpoint distances (metric determinacy)
and the stress values.

Since there is no statistic for determining the "correct" con-
figuration, the usual approach is to obtain several different solutions
using a variety of metrics and dimensionalities. Choice of an appro-
priate Minkowski metric is often based upon comparison of stress
values across metrics. As noted by Shepard (1974), this is invalid
because degeneracies, with resultant lower stress values, are more
prevalent in the City-Block ($r=1$) than in the Euclidean metric ($r=2$,
see Arnold, 1971), and most prevalent of all in the dominance metric
($r=\infty$). This means that one should primarily rely on interpretability
or theoretical considerations, such as the hypothesized integrality or
non-integrality of stimulus dimensions (see Hyman & Well, 1967, 1968;
Garner, 1974; Somers & Pachella, 1977) in determining an appropriate
metric.

Goodness-of-fit measures, interpretability, and theoretical
considerations are also used to determine the appropriate number
of dimensions. One frequently used rule-of-thumb is "looking for

the elbow" in the stress curve. Since stress decreases with increasing dimensionality, one must look for something besides mere increase in goodness-of-fit (in fact, there will always be a perfect solution for N points in N-2 dimensions - see Lingoes, 1971). Obtaining solutions in a number of different dimensions (for a given metric) and plotting stress versus dimensionality yields a curve which often shows that stress decreases dramatically for increasing dimensionality up to a certain point. After this point, adding dimensions decreases stress minimally. The usual interpretation of this phenomenon is that the added dimensions are needed to accurately fit the distances up to the correct dimensionality, and extra dimensions only fit noise. The point offering the "most for one's money" - minimizing dimensionality while maximizing goodness-of-fit - is called the "elbow" of the curve (see Figure 1). Once again however, interpretability and theoretical considerations must weigh heavily. If the elbow indicates that the appropriate dimensionality is three, but only two of the dimensions can be identified, then the third dimension is of little theoretical value (see Torgerson, 1965). In addition, there are distinct advantages to configurations of low dimensionality, since they greatly facilitate the visualization of structure. Shepard (1974) reports that in spite of the advantages of low dimensionality, most people tend to err on the side of deciding on too many dimensions, rather than too few.

A further consideration in determining the appropriateness of a configuration is the possibility of obtaining a degenerate solution.
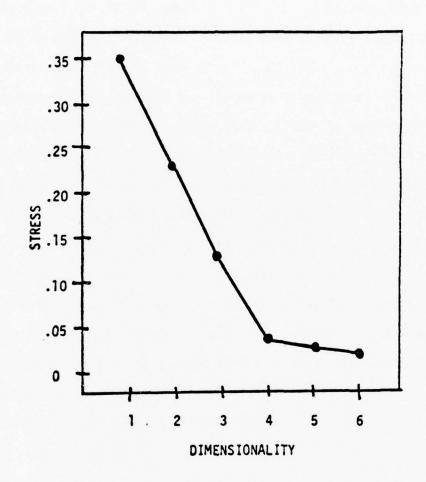
Figure 1. Diagram of stress plotted as a function of dimensionality.
The break or "elbow" in the curve at four dimensions indicates
that this is the proper number of dimensions needed to adequately
describe the psychological space.

14

This means that the scaling program attempts to increase goodness-of-fit by collapsing points upon one another. This can often mean that the dimensionality is too low. Lingoes (1977b) has specified a number of characteristics of degeneracy, including: (1) many points are lying atop one another ($d_{ij} = 0$ for many i,j); (2) the stress approaches zero for a large number of points and low dimensionality; (3) the maximal number of iterations is used to obtain the configuration; (4) stress improves with decreasing dimensionality; (5) there are outliers - points only weakly related to the rest of the configuration; and (6) the function relating input and derived interpoint distances (called the Shepard diagram) is step-like in nature. This indicates a large number of tied derived distances. In addition, Lingoes and Guttman (1967) have developed a coefficient of deformation, which indicates the degree of degeneracy in a configuration. In order to minimize or eliminate degeneracy, one can use one of several procedures. (1) Remove outliers from analysis, since they obscure structure among the ramaining points by pushing them together into one section of the configuration (Lingoes, 1977b). (2) Analyze subsets, or clusters, separately (Lingoes, 1977b; Shepard, 1974). This may clarify structure within groups. On the other hand, relationships between clusters are not accounted for, and therefore intercluster comparisons cannot be made. (3) Increase the dimensionality (Lingoes, 1977b). This may prevent groupings from collapsing to a point. This method, however, is not always helpful, since increasing the dimensionality may obscure rather than elucidate inter- and intra-cluster structures by making visualization more difficult. (4) Use metric methods, such as restricting the shape of the function relating

input and derived interpoint distances to eliminate step patterns in the Shepard diagram (Shepard, 1974; Shepard & Crawford, 1975).

We must reiterate our emphasis on two points. First, meaningfulness and ease of interpretation of a configuration are crucial, and interact with all other guidelines for deciding upon a "most correct" configuration. That is, interpretability is very important for deciding on a "best" configuration.

Second, determination of the metric, dimensionality, and "acceptable" stress level, as well as the interpretation, are almost entirely post hoc, unless there are strong theoretical limits on the metric and dimensionality. This fact has two major implications. First, the criteria for "acceptability" are highly subjective. That is, there are no statistical tests with significance levels for the goodness-of-fit measures. Second, even though interpretability is crucial in deciding on a configuration, the dangers of over-interpretation are particularly acute. This means that some interpretation can be found for almost any configuration if one has enough creativity and persistence. Thus, the validity of any interpretation must rely heavily on the replicability of the basic structures in the configuration. We now discuss the problems and methods of interpretation.

Interpreting Configurations

The interpretation of multidimensional scaling outputs is based on the identification and labeling of different types of structures, several of which will be reviewed here.

1. <u>Vectors</u>. The order of points projected onto a vector through the space may suggest interpretations of the point constellation. Vectors are generally found by searching for orderings in the configuration which correspond to objective orderings of physical continua. One can also use subjective orderings of physical continua obtained by methods such as magnitude estimation or unidimensional similarity judgments. Usually, the vectors of most interest are the axes of each of the dimensions, but one may also plot other vectors through the space by regressing external variables onto the coordinates of each point. Chipman and Carey (1975) use a similar technique to locate vectors corresponding to loudness, pitch, volume and density in a space of noise bands.

An alternative method for locating vectors through a Euclidean output space is to apply principal components analysis or factor analysis to the scaled interpoint distances (see Napior, 1972). In the case of principal components analysis, the results correspond to rotating the coordinate system to maximize the variance of the first coordinates of all points. The second axis maximizes the variance of the second coordinates of all points with the restriction that this axis is orthogonal to the first. The process is repeated for all d dimensions in the space. Factor analysis yields similar results. In neither case, however, is a substantive interpretation offered. The procedures merely attempt to elucidate any structure that may be hidden in the scaling output. Degerman (1970) has introduced an interesting variation of this approach which divides the space into discrete, qualitative clusters and continuous, Minkowski space.

2. <u>Polar coordinate patterns</u>. These patterns are interpretations of the point coordinates projected onto a two-dimensional plane. Instead of relating external variables to the orderings of the points projected onto vectors, the points are located in reference to their polar coordinates. That is, external values are associated with distance from an origin and the angular separation of points. The days of the week and the color circle are two examples of item sets that may be treated in this fashion.

One can also use an "ideal points" model to characterize a configuration (Shepard, 1972; Cliff & Young, 1968). In this model, it is assumed that the important parameters are the distances of each point in the configuration from some hypothetical "origin" or "ideal point" in the space, from which one or more relevant vectors might emanate. This model is a form of the polar coordinate pattern interpreting only the distances from the origin.

3. <u>Clusters</u>. Looking for groupings of items in the space is another method of interpreting multidimensional scaling outputs. Such regions or groupings might simply be areas of the space partitioned from the rest of the space with no specific restrictions on intergroup or intragroup relationships between points. In most cases, however, one is interested in groupings in which the points are related to each other in some meaningful fashion, often called clusters. These clusters may be overlapping or non-overlapping subsets of the items in a given configuration. There is no rigorous or universally accepted definition of a cluster. Intuitively, however, an item in a cluster is more similar to other items in the cluster than to items

18

outside the cluster. In addition, the items in a cluster should share some attribute aside from their proximal locations in the multi-dimensional space.
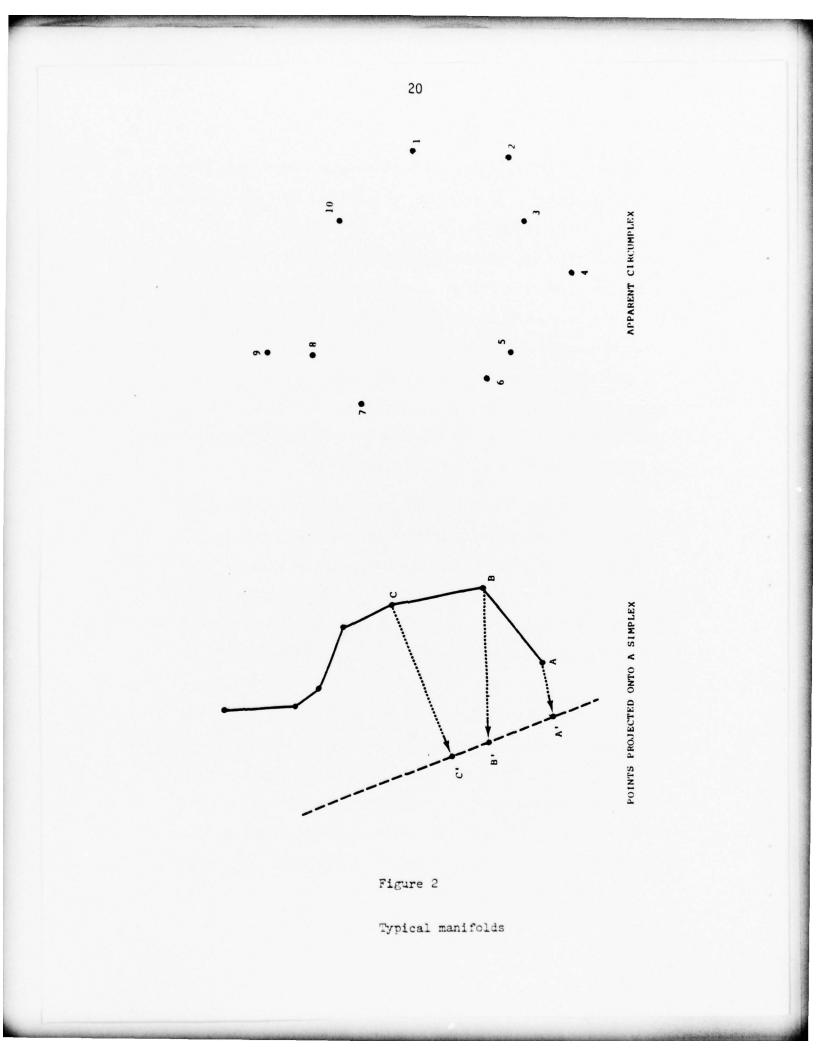
Two major methods are used for locating clusters. One method is clustering algorithms (Jardine & Sibson, 1971; Sneath & Sokal, 1973), which usually determine the membership and boundaries of clusters solely on the basis of the original dissimilarities or the derived interpoint distances. The second is a visual examination of the configuration to locate items which are grouped together in the space and which also share common features.

Clustering algorithms, which mechanically group items into clusters, can be very helpful in elucidating structure for high-dimensional, not-easily-visualized configurations where structure is often not apparent. There are algorithms for finding clusters or partitions with non-overlapping boundaries, such as Johnson's (1967) hierarchical clustering and Lingoes' (1977b) PEP-U. One non-hierarchical clustering algorithm which permits overlapping boundaries is the additive clustering technique of Shepard and Arabie (1975).

Visually examining the configuration for clusterings of items is a widely-used method, although there is disagreement concerning what characterizes a "good" cluster (cf. Lingoes, 1977b; Shepard, 1972; Shepard & Chipman, 1970). It is important to note that different clusterings based on common attributes may be derived from different theoretical considerations. One should also be apprised of the danger in concluding that one has found "clusters"

19

in two-dimensional projections of a higher-dimensional space, since points may cluster in a particular projection of the space, yet be far apart in the actual space.

Statistical and substantive clustering methods need not, of course, be mutually exclusive. Lingoes (1977b) points out that statistically significant but uninterpretable clusters (and structures in general) will usually be useless. One must have a theory or explanation to render a cluster meaningful or useful. On the other hand, clusters which do not appear upon replication must be regarded with suspicion. This is also the case when other tasks, such as stimulus sorting, yield different clusterings.

4. <u>Manifolds</u>. Manifolds, like clusters, are subsets of the scaled items. However, they differ from clusters in that specific relationships among members of the subset, such as orderings, are hypothesized. That is, manifolds are subsets of items which have a particular structure in and of themselves. Generally, manifolds are structures of a dimensionality of d-1 or less, embedded in a d-dimensional space. Some such structures (see Figure 2) are the simplex (points that may be placed on a vector in the space), the circumplex (points that may be placed on a circle around an arbitrary origin), and the radex (a polar coordinate pattern which is a combination of simplices and nested circumplices). Often the structure of a set of items which should be scaled in d-1 or fewer dimensions is distorted by the addition of extraneous dimensions.

20

APPARENT CIRCUMPLEX

POINTS PROJECTED ONTO A SIMPLEX

Figure 2

Typical manifolds

One example of this phenomenon is the "horseshoe" pattern which may result when a simplex is scaled in two-dimensional space, and uses the extra degrees of freedom to bend back upon itself (Kendall, 1971). In 3-space, a simplex may form a helical pattern.

In order to identify candidate manifolds, one may first use the scaling algorithm to suggest the existence of "interesting" structures. Such structures are then analyzed separately to see if they can indeed fit into a space of lower dimensionality. The I items composing the structure can be extracted and the IxI matrix of dissimilarities can be examined to "confirm" the structure. Lingoes and Borg (1977) detail a number of methods for identifying spatial manifolds.

5. _Isovalue contours._ Plotting the "isovalue contours" for a given external variable is another way to interpret a configuration (Abelson, 1954). In this procedure, a function of the external rating of each scaled item such as rated preference, assigns a rating to each point in the space. The rating of an item close to a given point strongly affects the rating of the point. The further away the item is, the lower its effect on the rating. The rating of a point is the sum of the effects of all items upon that point. To compute R_p, the rating of a given point p, Abelson uses the formula:

$$R_p = \sum_{i=1}^{N} \frac{r_i}{1+d_{pi}^2}$$

where r_i is the rating for the i-th item and d_{pi} is the distance between the point p and item i. Using this formula, each point on the plane is assigned a value, and contours are drawn connecting the points with equal value.

6. <u>Relating different representations</u>. Yet another interpretation method is exemplified by procrustean analysis algorithms (Gower, 1975; Cliff, 1966; for applications see Shepard & Chipman, 1970). These methods rotate, translate, and reflect points to obtain a "best fit" of one configuration onto another. The basic rationale for such methods is that comparing a scaled output to a pre-interpreted configuration gives some indication of the validity of placing the same interpretation on the scaled configuration. The customary statistics are the sum of squares error of the fit and the product-moment correlation of the point coordinates given the best fit up to a rigid translation, rotation, and reflection of the coordinates. Unfortunately, these statistics, at present, do not have significance levels indicating true goodness-of-fit. Therefore the interpretation of these statistics is as subjective as the interpretation of stress.

A variation upon the fitting-of-one-configuration-to-a-target approach is the use of several scaled outputs to generate a group space - a space which represents the "important" features of the aggregate of all individual outputs. An example of this approach, PINDIS (Borg, 1977), uses a model similar to that employed by INDSCAL (Carroll & Wish, 1974). Basically, the model postulates a group space with fixed axes. Each individual, when assigning values to stimuli, uses this space to generate the inter-item dissimilarities after stretching or shrinking each axis. This means that each individual evaluates the coordinates of each point according to the group space coordinates, but then multiplies this coordinate by a stretching factor to indicate the dimension's

relative saliency. Using a multi-configuration extension of the procrustean analysis algorithm, PINDIS generates a group space and dimensional weights for all dimensions for all individuals. The argument is that all individuals are using a common set of dimensions, so the axes are fixed and can therefore be interpreted.

8. <u>The configuration itself</u>. Last, but not least, it is possible to regard the configuration as meaningful in and of itself, without reference to any physical dimensions or categories (Bailey, 1974; Chipman & Carey, 1975). This approach allows one to characterize an internal arrangement of a stimulus domain, perhaps defining new "dimensions" or groupings, on the basis of a multidimensional scaling output. One example of this is the color space (Shepard, 1962).

Several different interpretation methods can, of course, be used in analyzing a single data set. Some analyses may be more appropriate for some configurations and others for other configurations. The different types of structure are not mutually exclusive. Comparisons among structure-recovery methods are, however, not meaningful for two reasons. First, many methods are difficult, if not impossible, to compare due to differences in representations and ways of calculating goodness-of-fit. Second, the heuristic value of a given method depends upon the problem being addressed.

We now discuss the stress, the goodness-of-fit measure, and some of the mathematical considerations when applying the nonmetric multidimensional scaling algorithm.

Stress

Since the introduction of the stress parameter (see Kruskal, 1964a, b), much attention has been paid to stress as (a) a goodness-of-fit measure (Kruskal & Carroll, 1969; Stenson & Knoll, 1969; Klahr, 1969; Wagenaar & Padmos, 1971; Spence & Ogilvie, 1973), (b) an index of recovery of the "true" configuration obscured by noise (Young, 1970; Sherman, 1972), (c) an indicator of the appropriate dimensionality of the representation (Isaac & Poor, 1974; Spence & Graef, 1974), and (d) a measure of the underlying metric (Arnold, 1971). All these papers are based on the observation that, in a Euclidean space, ordered data on interstimulus proximities sufficiently constrain the solution to an interval scale (see Abelson & Tukey, 1959, 1963; Shepard, 1966). In a Monte Carlo study to validate this claim, Shepard (1966) reports correlations in excess of .99 between "true" and reconstructed distances for all test configurations of 10 or more points. This means that an excellent reconstruction of the original point configuration is made from error-free data. Young (1970) and Sherman (1972) evaluated the reconstructive powers of multidimensional scaling for fallible data. They arbitrarily placed points in a space and generated a dissimilarities matrix by randomly moving the points before calculating the interpoint distances. They also report good recovery of the "true" configuration under conditions of moderate perturbations of many points. These studies, however, obscure one important point. The scaled outputs may be solutions to very ill-conditioned functions. That is, large deviations from the

local minimum configuration may produce only small increases in stress. One of these non-optimal solutions, moreover, may be more interpretable than the local minimum solution. This means that small rank reversals in the ordinal dissimilarity measures due to experimental error could easily have obscured key structures by leading to the best-fitting rather than the "true" configuration. By the same token, there are situations in which the local minimum is in a deep valley of the stress function, so that even slight changes in any of the point coordinates lead to large increments in stress. In this case, the function is well-conditioned.

A method of interpretation should distinguish between these two cases, since a perfect fit to one of the basic structures (e.g., vector, polar coordinate pattern, cluster) usually requires movement of points away from their local minimum locations. If it were possible to constrain the spatial configuration to perfectly fit a prespecified structure, observing the amount of increase in stress and thereby determining how well- or ill-conditioned the stress function is for any particular set of input proximities, one might gain some indication as to the validity of an interpretation. For this reason we propose just such a method for constrained scaling and interpretation called CONSCAL.

Confirmatory Multidimensional Scaling (CONSCAL)

Placing constraints on either the form of the monotone distance function (the function relating input proximities and derived distances) or the locations of the points is not a new concept in multidimensional scaling. Shepard and Crawford (1975) add penalty functions to the

standard stress measure to specify the shape of the Shepard diagram
(the monotone distance function). One option in McGee's (1968) "common
elastic multidimensional scaling" (CEMD) program permits the simul-
taneous scaling of several individual proximity matrices into different
configurations with a penalty function constraining all configurations
to be "somewhat" alike.

In our basic model for confirmatory scaling, the interpoint
distances are a function satisfying the following:

(a) DECOMPOSABILITY. The distance between points is a function
of componentwise contributions.

(b) INTRADIMENSIONAL SUBTRACTIVITY. Each componentwise con-
tribution is the absolute value of the scale difference.

These are two of the three assumptions used by Tversky and Krantz
(1970) to characterize the Minkowski metric.

Assumption (a) means that the distance between two points,
x and y, is:
$$d(x,y) = F\left[\phi_1(x_1,y_1), \ldots ,\phi_n(x_n, y_n)\right] \qquad (1)$$

where F is an increasing function in each of its n arguments. The
ϕ's are symmetric on both arguments and nonzero except when $x_i = y_i$
in which case $\phi_i(x_i,y_i) = 0$. Psychologically, one interpretation
of this assumption is that the proximity of two stimuli is determined
using a two-stage evaluation. First, an aspect is picked and the
subject judges the relative difference of the two stimuli with
respect to the particular aspect. This process is repeated until
all relevant aspect differences have been evaluated, at which time

these judgments are amalgamated into a single index of overall proximity.

Assumption (b) means that the proximity measure may be written as follows:

$$d(x,y) = F\left[|x_1-y_1|, \ldots |x_n-y_n|\right] \qquad (2)$$
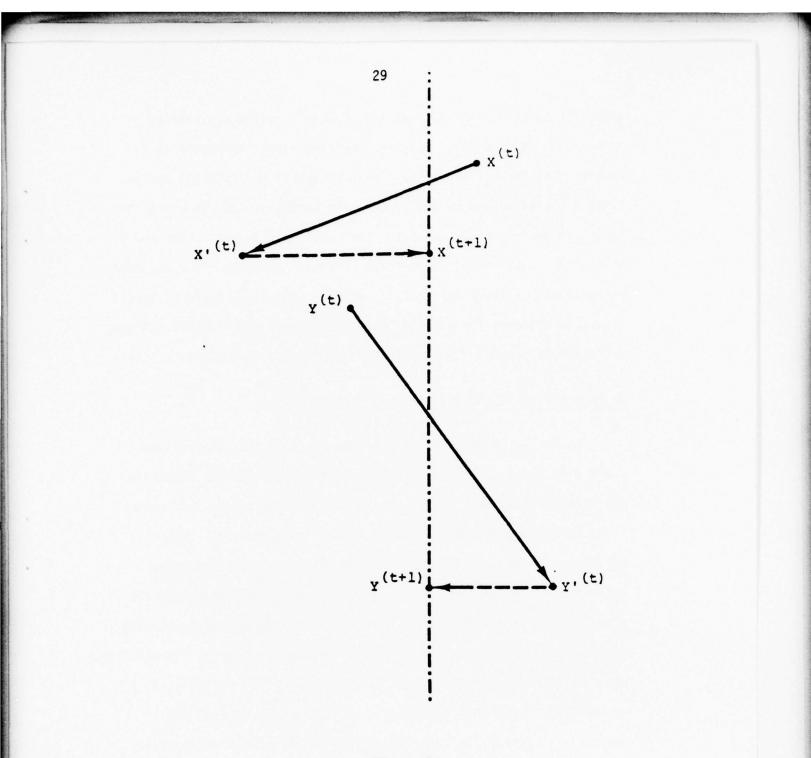
This differs from Eq. 1 primarily in that the x_i and y_i values are points along an axis.

In addition, we make the assumption that the x_i values along each aspect axis may be evaluated, at an ordinal scale, using other physical or psychological measurement methods. For example, assume the dissimilarities of pairs of rectangles are scaled in two dimensions and the axes are identified with psychological area and shape. One way to test this model is to use magnitude estimates of area and shape for each of the scaled rectangles to establish an order of x_1's for area and x_2's for shape. These orders are then viewed as constraints on the coordinates of each point. So, if $x_1 < y_1$, then the points x and y must be located so that the first coordinate of point x is smaller than the first coordinate of point y. That is, the magnitude estimates establish a weak order of projections of all points onto the axes. Operationally this means that a set of numbers from some ordering task, such as a magnitude estimation or a Thurstone scaling of paired comparisons (Thurstone, 1927), is given as supplementary information to the scaling program. As the program iterates toward a configuration of lowest stress, the configuration is forced to satisfy the ordinal constraints from the ordering data. This process may be implemented by stepping the

28

configuration at each iteration and then moving the points to conform
to the ordering constraints. Figure 3 illustrates the procedure
for one hypothetical iteration of two points in a two-dimensional
space. Furing the t-th iteration, points located at $x^{(t)}$ and $y^{(t)}$
are moved to $x'^{(t)}$ and $y'^{(t)}$. Since there is now a violation of
the ordering requirement along one dimension, the points are moved
to $x^{(t+1)}$ and $y^{(t+1)}$ to satisfy this requirement. This process
is repeated at each step until a local minimum is reached.

For two or more points the monotone order of projection onto
each axis is determined using Kruskal's (1964a, b) monotone regression
on the coordinates in place of its more usual application to the
interpoint distances. Within the permissible range of values for
a coordinate, the movement toward a minimum stress configuration
proceeds as in regular nonmetric multidimensional scaling, but
the specified order of any two points along any one dimension can
never be reversed.

Kruskal's monotone regression, as applied to interpoint distances,
has two options known as the primary and secondary approaches. In
the primary approach, tied interitem dissimilarities need not be
mapped into equal interpoint distances, while in the secondary
approach the mapping must be onto equal interpoint distances.
In both cases, violations of the general monotonicity requirement
mean contributions to the stress. In CONSCAL, these two options
are also available when specifying the monotone order of projection
onto the axes. Here, the primary approach is called weak dimensional

Figure 3

During each iteration Confirmatory Scaling moves all points twice. First, points $x^{(t)}$ and $y^{(t)}$ are moved to locations $x'^{(t)}$ and $y'^{(t)}$ using the method of steepest descent. Second, these points are moved to locations $x^{(t+1)}$ and $y^{(t+1)}$ to satisfy the monotonicity requirement of the point coordinates with respect to the horizontal coordinate axis.

30

monotonicity, while the secondary approach is called semi-strong
dimensional monotonicity. Semi-strong dimensional monotonicity is
usually used for scaling stimuli in a factorial experimental design,
since all tied values of the independent variables used to create the
factorial design should map into tied coordinate values. When psycho-
logical variables specify the order of projection onto the axes, weak
dimensional monotonicity should usually be used since there is little
reason to believe, for example, that two stimuli which elicit category
estimates of 6 on a 1 to 10 scale are precisely equal psychologically.

An Example: Multidimensional Scaling of Ellipses

The following examples come from a study of the interactions
among dimensions of stimulus variation (for a theoretical discussion
see Somers & Pachella, 1977). One way of studying these interactions
is to investigate the relationship between unidimensional judgments
and interstimulus dissimilarity ratings. Previous studies using
rectangles as stimuli in scaling tasks indicate that two dimensions -
size (width times height = area) and shape (width divided by height)
were the relevant psychological dimensions for predicting dissimilarity
ratings (Krantz & Tversky, 1975; Noma, Note 1). For ellipses, it
was therefore hypothesized that the two dimensions of area and
eccentricity (the "size" and "shape" dimensions of ellipses) would
be the relevant dimensions.

We were interested, basically, in three questions. First, how
relevant are the dimensions of area and eccentricity as predictors
of judged dissimilarity? Second, how do "physical area" and "physical

shape" (derived from physical measurements of the stimuli) compare with "judged area" and "judged shape" (derived from magnitude estimation data) as dimensions characterizing the configuration? Lastly, since area-eccentricity and and major axis-minor axis are physically equivalent pairs of dimensions, are they also psychologically equivalent? In other words, an ellipse can be uniquely specified by noting either its area and eccentricity or the lengths of its major and minor axes. Therefore, one might hypothesize that the area-eccentricity and the major axis-minor axis pairs of dimensions would equally well characterize the two-dimensional scaling configurations. For rectangles, however, area and shape seem to be better descriptors of the two-dimensional configuration than are height and width (Krantz & Tversky, 1975).

A factorial design with four equally spaced levels of area crossed with four equally spaced levels of eccentricity was employed in constructing the stimuli. The largest ellipse was in a 3:1 ratio to the smallest, and the most eccentric was in a 1.66:1 ratio to the least eccentric. These sixteen ellipses, drawn by a CALCOMP plotter, were photographed, and black-on-white slides were made.

Four subjects made global dissimilarity judgments for all possible pairs of ellipses (excluding identical pairs). The entire set was presented three times, in a different random order each time, and the results were averaged for each subject. The subjects received the following instructions verbally:

We are interested in how people perceive complex figures. In this experiment you will be asked to judge how similar ellipses are to each other. You will be shown pairs of drawings like this. [sample pair of slides shown] Rate the similarity of the pair on a scale of 1 to 10 (integers only), 1 being the most similar, 10 the least similar (or most different). If the two drawings of a pair were identical, for example, you would rate the pair zero. (This will never be the case, though - the two stimuli in a pair will always be different.)

Base your judgment on the overall similarity of the figures. To give you an idea of what the whole set of figures is like, I'll run through the slides briefly. [the 16 ellipses shown singly, in random order.]

You will have as much time as you need to judge each pair. Mark your judgment in the appropriate space on the answer sheet. Be sure to use the full range of ratings.

Do you have any questions?
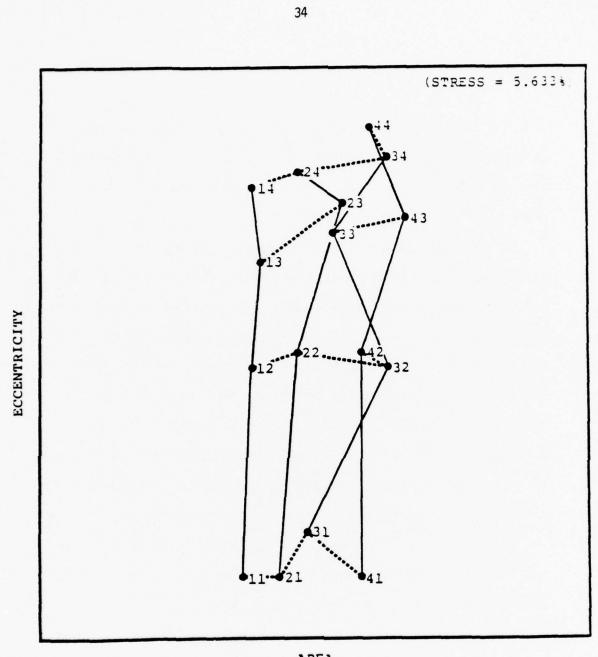
In another session, subjects made magnitude estimates, on the same 1-10 scale, of four properties of the ellipses (which are presented individually) - area, eccentricity, length of major axis, and length of minor axis. The order of these four tasks varied among subjects. Six judgments were made of each of the four properties for each of the 16 ellipses (384 judgments total), and the results were averaged for each subject in each task.

Unconstrained multidimensional scaling of the global dissimilarity judgments showed generally good fits for all four of the subjects in two-dimensional Euclidean space. One configuration (DT) had a stress of 13.1% and the other subjects' configurations ranged between 5.6% and 8.7% stress. We were reasonably confident that the local minimum problems were being avoided because starts from either random or "hypothesized best fit" (area by eccentricity

factorial design) configurations resulted in virtually identical
stress values and configurations. For two subjects, a third dimen-
sion was added, but this made little difference in stress, and the
extra dimension was uninterpretable.

In all four cases, clearly interpretable dimensions of area
and eccentricity were present. There were a few minor deviations
from the hypothesized orderings along the dimensions, as can be seen
in Figures 4-7, and one major reversal of area levels within the
smallest eccentricity level in the highest-stress configuration
(DT, Figure 7). One question that cannot be answered using tradi-
tional interpretation techniques is, how meaningful are such reversals?
Are they merely noise, or does the subject actually have some anomaly
in his or her cognitive structure? One way we can try to answer this
is to use a constrained multidimensional scaling analysis.

As can be seen in Table 1, constraining the configuration to fit
the factorial design according to which the stimuli were constructed
causes increases in stress from about 2-4% for each subject, indicating
that this model does reasonably well for all four subjects. In fact,
the configuration with the major reversal (DT, Figure 7) shows the
second-lowest increase in stress - only 2.4%. This increases our
confidence in the factorial design with respect to DT, since even
though her deviations from the model appeared to be more systematic
than those of the other subjects, they seem to be no more important.

Comparing judged area and judged eccentricity with physical area
and physical eccentricity produced little difference in either stress
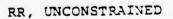
34

AREA

RR, UNCONSTRAINED

Figure 4

AREA

JL, UNCONSTRAINED

Figure 5

36

(STRESS = 7.048%)

ECCENTRICITY

AREA

TM, UNCONSTRAINED

Figure 6

AREA

DT, UNCONSTRAINED

Figure 7
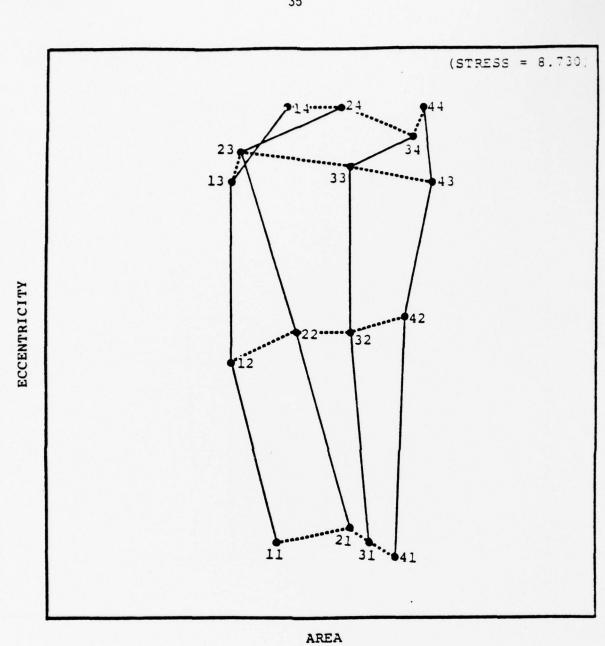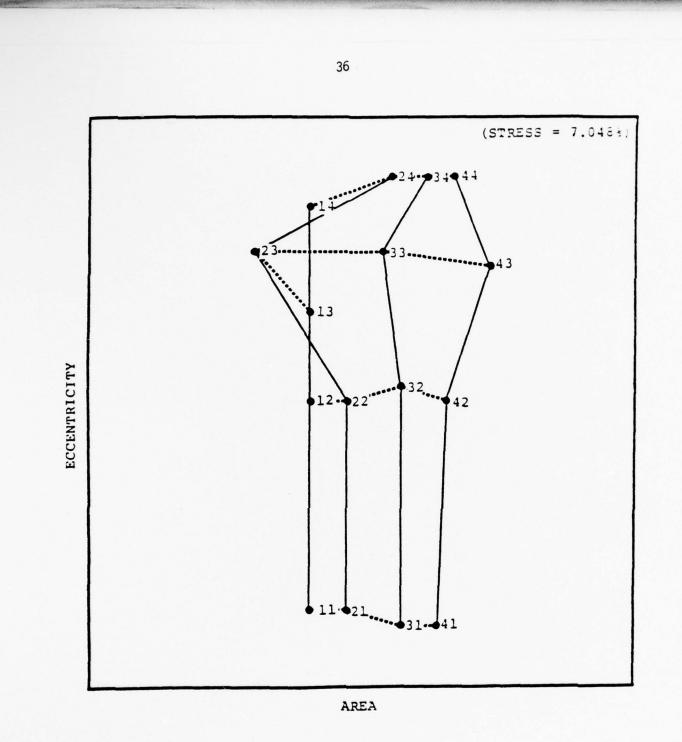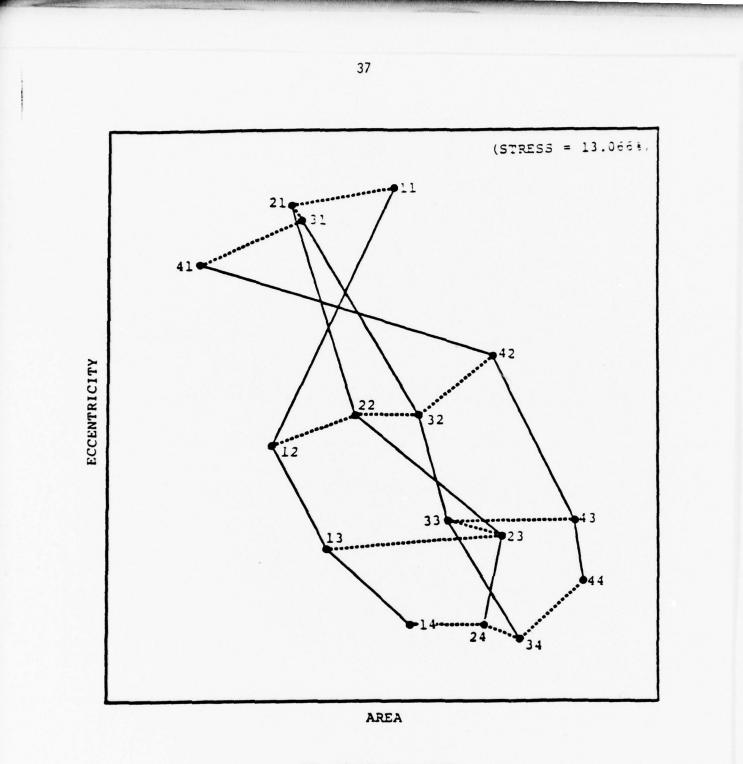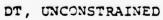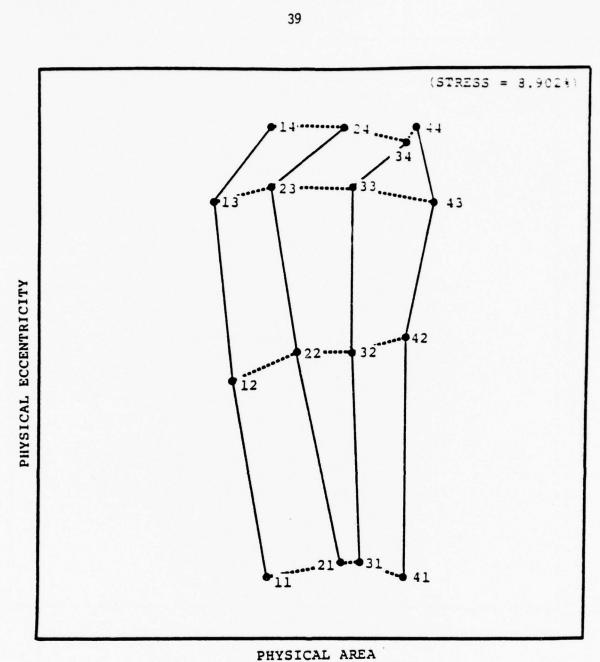
JUDGED ECCENTRICITY

RR, CONFIRMATORY

JUDGED AREA AND ECCENTRICITY, STRONG MONOTONICITY

Figure 8

PHYSICAL AREA

JL, CONFIRMATORY
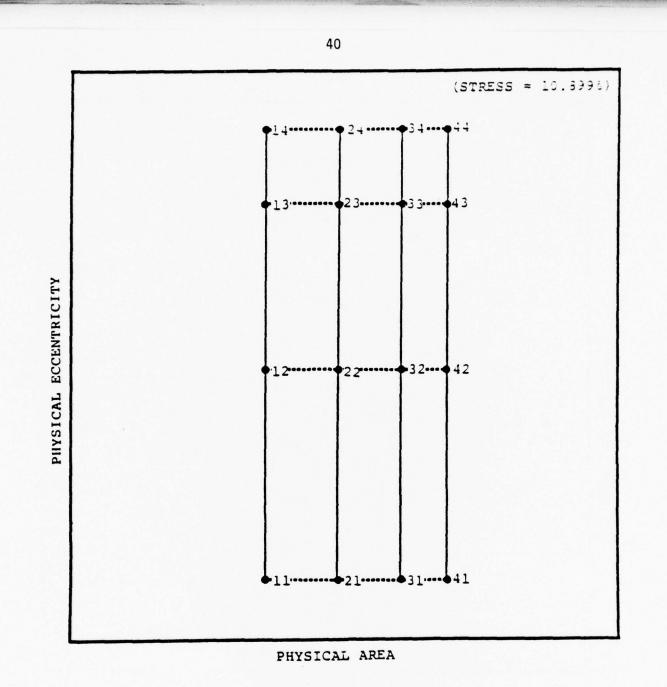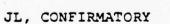
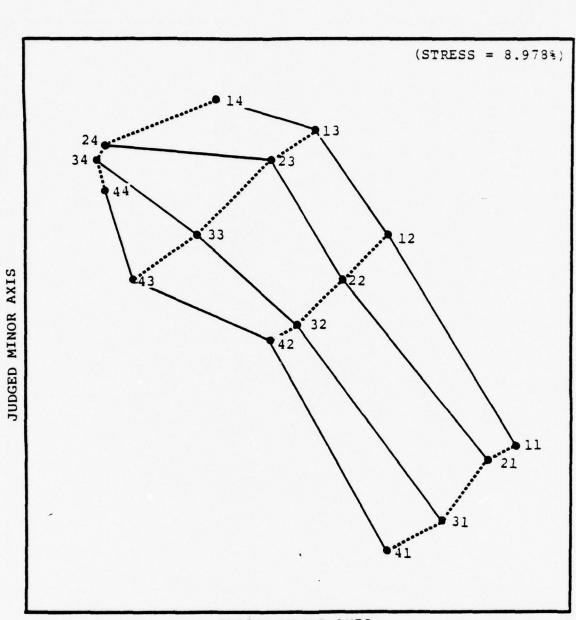PHYSICAL AREA AND ECCENTRICITY, WEAK MONOTONICITY

Figure 9

PHYSICAL AREA

JL, CONFIRMATORY

PHYSICAL AREA AND ECCENTRICITY, STRONG MONOTONICITY

(FACTORIAL DESIGN)

Figure 10
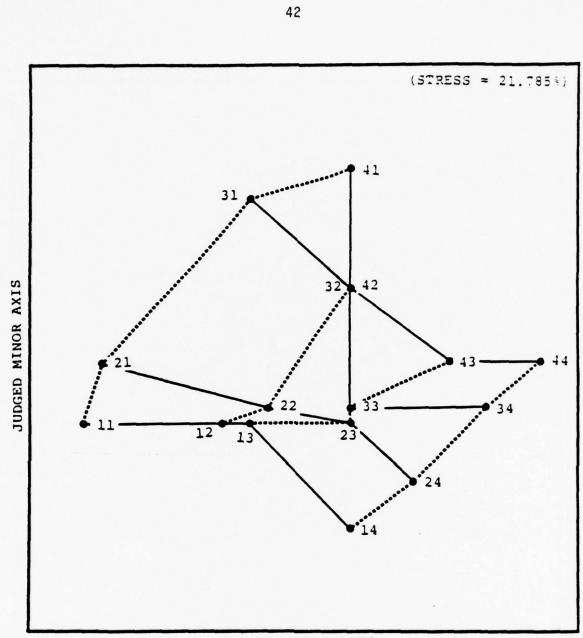
TM, CONFIRMATORY

JUDGED MAJOR AND MINOR AXES, STRONG MONOTONICITY

Figure 11

42

(STRESS = 21.785%)

JUDGED MINOR AXIS
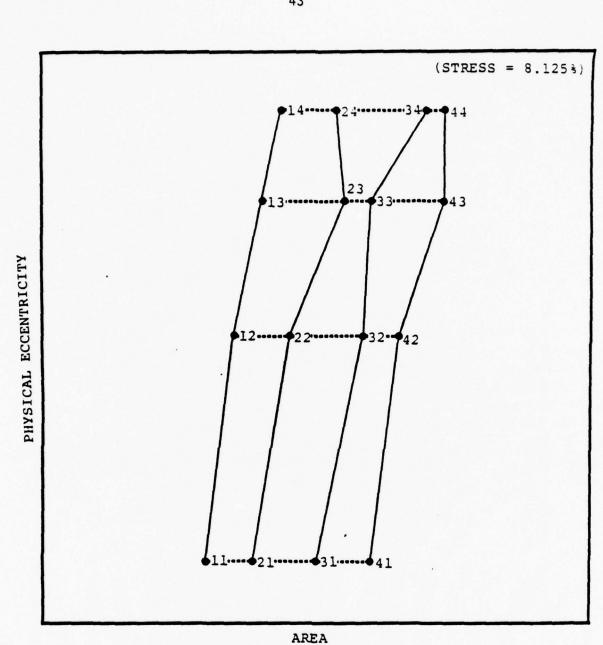
JUDGED MAJOR AXIS

RR, CONFIRMATORY

JUDGED MAJOR AND MINOR AXES, STRONG MONOTONICITY

Figure 12

RR, CONFIRMATORY

ECCENTRICITY ONLY, STRONG MONOTONICITY

Figure 13

Table 1

Stress values for configurations with and without constraints, for four subjects

	DT	RR	JL	TM
A. Unconstrained	13.066	5.633	8.730	7.048
B. Weak Physical Area Physical Eccentricity	13.618	5.793	8.902	7.393
C. Strong Physical Area Physical Eccentricity	15.505	9.512	10.899	9.529
D. Weak Psychological Area Psychological Eccentricity	16.314	7.191	9.281	8.163
E. Strong Psychological Area Psychological Eccentricity	17.396	7.606	9.954	8.453
F. Weak Psychological Major Axis Psychological Minor Axis	16.741	18.593	10.719	8.905
G. Strong Psychological Major Axis Psychological Minor Axis	17.396	21.779	11.254	8.978
H. Strong Physical Eccentricity	14.490	8.125	9.516	7.882
I. Strong Physical Area, Eccentricity minus Unconstrained	2.439	3.879	2.169	2.481

values (see Table 1) or configurations (see Figures 8 and 9 for
sample configurations). One would naturally expect the factorial
design, with strong monotonicity (see Figure 10), to produce higher
stress than any of the other models because of the large number of
ties which must be satisfied. These results indicate two things:
(1) that subjects' scalings of area and eccentricity are reasonably
veridical (which is not particularly surprising), and (2) that models
using physical versus judged area and eccentricity are for the most
part interchangeable, with preference perhaps going for the factorial
design model, because of its greater simplicity.

Comparing the "scaled area-eccentricity" to the "scaled major
axis-minor axis" models proved more interesting. For three of the
subjects, the area-eccentricity and major axis-minor axis models
were approximately equivalent in terms of stress, and produced
highly similar configurations (compare Figures 6 and 11, for example).
However, for one subject, there was a dramatic difference in stress
between area-eccentricity and major axis-minor axis configurations.
For RR, at least, even though the two models are physically equivalent,
they are not psychologically equivalent (compare Figures 8 and 12).
This comparison also shows that there can be dramatic individual
differences between subjects regarding the applicability of certain
models even though the configurations may appear quite similar.

Using confirmatory multidimensional scaling, it is also possible
to constrain only a subset of the dimensions. This might be especially
helpful if one has strong hypotheses only about some of the dimensions
a subject is expected to use, but not about all of them. For example,

see Figure 13, in which eccentricity, but not area, is constrained.

Discussion

Most traditional interpretation techniques assume the uniqueness
of the configuration. Structures are imposed on the resultant con-
figuration to yield an interpretation (an exception being the inter-
pretation of manifolds by extracting subsets of points and analyzing
these submatrices of the original dissimilarity matrix). As we
have seen, however, there may be a wide variety of possible config-
urations with almost identical stress values. What this means in
terms of interpretations is not clear. It does mean that a measure-
ment theoretic or confirmatory analysis method is needed. The measure-
ment theoretic approach could be an extension of the Krantz and
Tversky (1975) tests, incorporating an error theory that measures
the degree to which their measurement axioms are violated. (This
would also allow a non-parametric goodness-of-fit test.) Another
extension of the tests would cover nonfactorial experimental designs.

Until such necessary and sufficient conditions are defined, con-
firmatory multidimensional scaling could fill the void. In fact,
confirmatory scaling may have an advantage over the Krantz and Tversky
tests in that estimates can be made of the "importance" of violations
of the conditions. For instance, one axiom of the Tversky and
Krantz (1970) axiom system is

(c) INTERDIMENSIONAL ADDITIVITY. The distance is a function of
the sum of componentwise contributions.

One testable condition that may be derived from this is that the dissimilarity judgments, δ, must satisfy:

$$\delta(A_1 S_1, A_2 S_2) = \delta(A_1 S_2, A_2 S_1)$$

where A_1 and A_2 are two levels on dimension one and S_1 and S_2 are two levels on dimension two with stimulus $A_i S_j$ being on the i-th level of dimension one and the j-th level of dimension two. If the subject is asked to rank the dissimilarities of all stimulus pairs, systematic violations of this condition could occur. What remains unanswered, however, is whether these violations are psychologically important. That is, is this systematic bias an artifact of the requirement that the subject must report some ordering even if he is indifferent with respect to several possible orderings? In this case, the subject will probably establish a unique ordering using a simple rule even if this rule is of no importance in his actual handling of the stimulus representation. By scaling this data using a confirmatory multidimensional scaling, these anomalies are shown to be unimportant, as they contribute little to the stress.

Even though the analysis is confirmatory only in the sense that fixing certain axes in a factor analysis is a confirmatory factor analysis, the results provide stronger support for a potential interpretation than that provided by most traditional methods. Perhaps the appropriate approach is to use one of the more traditional methods or a theoretical analysis to formulate a family of possible interpretations. By applying confirmatory multidimensional scaling, a decision can then be made as to the relative validity of each interpretation.

A situation in which the nonuniqueness of the solution is important is in analyses in which one or more scaled outputs are compared among themselves or against pre-interpreted configurations. The validity of these approaches may be questioned since there may be slight modifications of each configuration that would produce vastly different goodness-of-fit measures across configurations and change the group space in an approach such as PINDIS.

One of the major unsolved problems of confirmatory scaling is the interpretation of increases in stress with added constraints to the configuration. One method for evaluating stress increases proposed by David Krantz (personal communication) is a pseudo-F-test. In the unconstrained multidimensional scaling of N points in a d-dimensional space, using Young's (1970) terminology, there are $N(N-1)/2$ degrees of freedom of the dissimilarities and $d(N-1) - d(d-1)/2$ degrees of freedom of the coordinates. Not surprisingly, Young (1970) has demonstrated that, in general, the stress increases with either increases in the degrees of freedom of the dissimilarities (number of points) or decreases in the number of degrees of freedom of the coordinates. In certain cases, such as that of strong dimensional mono-tonicity with a factorial design, the degrees of freedom of the coordinates are drastically decreased. For instance, in a four-by-four factorial experimental design, there are 120, or $16(16-1)/2$, degrees of freedom of the dissimilarities. In an unconstrained multidimensional scaling of the points in two dimensions there are 29, or $2(16-1) - [2(2-1)/2]$, degrees of freedom of the coordinates. By contrast, a confirmatory multidimensional scaling in a two-dimensional

four-by-four design using strong dimensional monotonicity has 5, or
$2(4-1) - [2(2-1)/2]$, degrees of freedom of the coordinates. This would
seem to imply, extending Young's analysis, that the stress in the
confirmatory analysis should be much higher than that of the unconstrained
solution. However, in our analysis of three of the four subjects we
found no large differences in stress when comparing unconstrained
and confirmatory analyses. Using the degrees-of-freedom point of
view, this seems to imply that the confirmatory factorial design
is the best representation of the data. Currently it is unclear how
weak dimensional monotonicity and non-factorial designs could be
interpreted in light of a degrees-of-freedom analysis.

Finally, we reiterate that the scaling solution which is optimal
in terms of some algorithm or goodness-of-fit measure should not
automatically be taken to be the optimal solution for other purposes,
such as interpretation of model-testing. Even though goodness-of-fit
measures tell us something about the appropriateness of the scaling
model, they indicate little in and of themselves about the interpre-
tability of the scaled results. As we have shown in the case of
multidimensional scaling, possible interpretations may be rejected
unwarrantedly, and "interesting" distortions in the unconstrained
scaled outputs may be only mathematical anomalies.

This new conceptualization of optimality can be extended to
other types of scaling. The correct question in most scaling situations
may be how the goodness-of-fit measure is affected if a given model is
satisfied, rather than how closely the scaled output resembles our pre-
interpreted hypothetical space.

Reference Note

1. Noma, E. Analysis of the psychological dimensions of rectangles.
University of Michigan, unpublished manuscript, 1976.

References

Abelson, R.P. A technique and a model for multi-dimensional attitude scaling.
Public Opinion Quarterly, 1954, 18, 405-418.

Abelson, R.P., & Tukey, J.W. Efficient conversion of nonmetric information into
metric information. Proceedings American Statistical Association Meetings,
Social Statistics Section, 1959, 226-230.

Abelson, R.P., & Tukey, J.W. Efficient utilization of non-numerical information
in quantitative analysis: General theory and the case of simple order. Annals
of Mathematical Statistics, 1963, 34, 1347-1369.

Arnold, J.B. A multidimensional scaling study of semantic distance. Journal
of Experimental Psychology Monograph, 1971, 90, 349-372.

Bailey, K.D. Interpreting smallest space analysis. Sociological Methods and
Research, 1974, 3, 3-29.

Beals, R., Krantz, D.H., & Tversky, A. Foundations of multidimensional scaling.
Psychological Review, 1968, 75, 127-142.

Borg, I. Geometric representation of individual differences. In Geometric
Representations of Social Science Data, J.C. Lingoes, L. Guttman, & E. Roskam,
Eds. Ann Arbor, Michigan: Mathesis, 1977 (in press).

Carroll, J.D., & Wish, M. Models and methods for three-way multidimensional scaling.
In D.H. Krantz, R.C. Atkinson, R.D. Luce, & P. Suppes (Eds.), Contemporary
Developments in Mathematical Psychology, vol. 2. San Francisco: Freeman, 1974.

Chipman, S.F., & Carey, S. Anatomy of a stimulus domain: The relation between
multidimensional and unidimensional scaling of noise bands. Perception and
Psychophysics, 1975, 17, 417-424.

Cliff, N. Orthogonal rotation to congruence. Psychometrika, 1966, 31, 33-42.

Cliff, N. & Young, F.W. On the relation between unidimensional judgments and
multidimensional scaling. Organizational Behavior and Human Performance,
1968, 3, 269-285.

References - con't

Degerman, R. Multidimensional analysis of complex structure: Mixtures of
    class and quantitative variation. Psychometrika, 1970, 35, 475-491.

Garner, W.R. The Processing of Information and Structure. Potomac, MD: Erlbaum
    Associates, 1974.

Gower, J.C. Generalized procrustes analysis. Psychometrika, 1975, 40, 33-51.

Hyman, R., & Well, A. Judgments of similarity and spatial models. Perception
    and Psychophysics, 1967, 2, 233-248.

Hyman, R., & Well, A. Perceptual separability and spatial models. Perception
    and Psychophysics, 1968, 3, 161-165.

Isaac, P.D., & Poor, D.D.S. On the determination of appropriate dimensionality
    in data with error. Psychometrika, 1974, 39, 91-109.

Jardine, N., & Sibson, R. Mathematical Taxonomy. London: Wiley, 1971.

Johnson, S.C. Hierarchical clustering schemes. Psychometrika, 1967, 32, 241-254.

Kendall, D.G. Seriation from abundance matrices. In F.R. Hodson, D.G. Kendall,
    & T. Tautu, Mathematics in the Archeological and Historical Sciences.
    Edinburgh University Press: 1971.

Klahr, D. A Monte Carlo investigation of the statistical significance of Kruskal's
    nonmetric scaling procedure. Psychometrika, 1969, 34, 319-330.

Krantz, D.H., & Tversky, A. Similarity of rectangles: An analysis of subjective
    dimensions. Journal of Mathematical Psychology, 1975, 12, 4-34.

Kruskal, J.B. Multidimensional scaling by optimizing goodness of fit to a non-
    metric hypothesis. Psychometrika, 1964a, 29, 1-27.

Kruskal, J.B. Nonmetric multidimensional scaling: A numerical method. Psychometrika,
    1964b, 29, 115-129.

Kruskal, J.B., & Carroll, J.D. Geometrical models and badness of fit functions.
    Multivariate Analysis, 1969, 2, 639-671.

References - con't.

Lingoes, J.C. Some boundary conditions for a monotone analysis of symmetric matrices. Psychometrika, 1971, 36, 195-203.

Lingoes, J.C. Identifying directions in the space for interpretation. In J.C. Lingoes, L. Guttman, & E. Roskam (Eds.), Geometric Representations of Social Science Data. Ann Arbor, Michigan: Mathesis, 1977b (in press).

Lingoes, J.C., & Borg, I. Identifying spatial manifolds for interpretation. In J.C. Lingoes, L. Guttman, & E. Roskam (Eds.), Geometric Representations of Social Science Data. Ann Arbor, Michigan: Mathesis, 1977 (in press).

McGee, V.E. Multidimensional scaling of N sets of similarity measures: A nonmetric individual differences approach. Multivariate Behavioral Research, 1968, 3, 233-248.

Napior, D. Nonmetric multidimensional techniques for summated ratings. In R.N. Shepard, A.K. Romney, & S.B. Nerlove (Eds.), Multidimensional Scaling, vol. 1. New York: Seminar, 1972.

Ramsay, J.O. Some statistical considerations in multidimensional scaling. Psychometrika, 1969, 34, 167-182.

Sherman, C.R. Nonmetric multidimensional scaling: A Monte Carlo study of the basic parameters. Psychometrika, 1972, 37, 323-355.

Shepard, R.N. The analysis of proximities: Multidimensional scaling with an unknown distance function. II. Psychometrika, 1962, 27, 219-246.

Shepard, R.N. Metric structures in ordinal data. Journal of Mathematical Psychology, 1966, 3, 287-315.

Shepard, R.N. A taxonomy of some principal types of data and of multidimensional methods for their analysis. In R.N. Shepard, A.K. Romney, & S.B. Nerlove (Eds.), Multidimensional Scaling, vol. 1. New York: Seminar, 1972.

References - con't.

Shepard, R.N. Representation of structure in similarity data: Problems and prospects. Psychometrika, 1974, 39, 373-421.

Shepard, R.N., & Arabie, P. Additive cluster analysis of similarity data. Paper presented at U.S. - Japan Seminar, Theory, Methods and Applications of Multidimensional Scaling and Related Techniques, 1975.

Shepard, R.N. & Chipman, S.F. Second-order isomorphism of internal representations: Shapes of states. Cognitive Psychology, 1970, 1, 1-17.

Shepard, R.N., & Crawford, G. Multidimensional scaling based on the fitting of constrained difference functions. Paper presented at U.S. - Japan Seminar, Theory, Methods and Applications of Multidimensional Scaling and Related Techniques, 1975.

Somers, P. & Pachella, R.G. Interference among sources of information in complex integrated displays. Human Performance Center Technical Report No. 58 (014523), University of Michigan, February, 1977.

Spector, A.N., Rivizzigno, V.L., & Golledge, R.G. The analysis of individual cognitive configurations of the city using multidimensional scaling. In R.G. Golledge, & A.N. Spector (Eds.), Cognitive configurations of the city: Vol. II. Columbus, Ohio: Department of Geography, Ohio State University, 1976.

Sneath, P.H., & Sokal, R.R. Numerical Taxonomy. San Francisco: Freeman, 1973.

Spence, I., & Graef, J. The determination of the underlying dimensionality of an empirically obtained matrix of proximities. Multivariate Behavioral Research, 1974, 9, 331-341.

Spence, I., & Ogilvie, J.C. A table of expected stress values for random rankings in nonmetric multidimensional scaling. Multivariate Behavioral Research, 1973, 8, 511-517.

References - con't.

Stenson, H.H., & Knoll, R.L. Goodness of fit for random rankings in Kruskal's nonmetric scaling procedure. Psychological Bulletin, 1969, 71, 122-126.

Thurstone, L.L. A law of comparative judgment. Psychological Review, 1927, 34, 273-286.

Torgerson, W.S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Tversky, A., & Krantz, D.H. The dimensional representation and the metric structure of similarity data. Journal of Mathematical Psychology, 1970, 7, 572-596.

Wagenaar, W.A., & Padmos, P. Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. British Journal of Mathematical and Statistical Psychology, 1971, 24, 101-110.

Young, F.W. Nonmetric multidimensional scaling: Recovery of metric information. Psychometrika, 1970, 35, 455-473.

Footnote

OFFICE OF NAVAL RESEARCH, CODE 455
TECHNICAL REPORTS DISTRIBUTION LIST

Director, Engineering Psychology
Programs, Code 455
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217 (5 cys)

Defense Documentation Center
Cameron Station
Alexandria, VA 22314 (12 cys)

Dr. Robert Young
Director, Cybernetics Technology
    Office
Advance Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Col. Henry L. Taylor, USAF
OAD(E&LS) ODDR&E
Pentagon, Room 3D129
Washington, D.C. 20301

Office of Naval Research
International Programs
Code 102IP
800 North Quincy Street
Arlington, VA 22217 (6 cys)

Director, Statistics and
    Probability Program, Code 436
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director, Physiology Program
Code 441
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Branch Office
ATTN: Dr. J. Lester
495 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. Charles Davis
536 South Clark Street
Chicago, IL 60605

Commanding Officer
ONR Branch Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Dr. M. Bertin
Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco 96503

Director, Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6 cys)

Mr. John Hill
Naval Research Laboratory
Code 5707.40
Washington, D.C. 20375

Director
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research
    Laboratory
Naval Submarine Base
Groton, CT 06340

Chief, Aerospace Psychology
    Division
Naval Aerospace Medical Institute
Pensacola, FL 32512

Navy Personnel Research and
    Development Center
Management Support Department
Code 210
San Diego, CA 92512

Dr. James H. Howard
Catholic University
Department of Psychology
Washington, D.C. 20064

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Journal Supplement Abstract
    Service
American Psychological Association
1200 17th Street, NW
Washington, D.C. 20036 (3 cys)

Dr. Ward Edwards
Director, Social Science
    Research Institute
University of S. California
Los Angeles, CA 90007

Director, Human Factors Wing
Defence & Civil Institute of
    Environmental Medicine
Post Office Box 2000
Downsview, Ont. CANADA

Dr. A.D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF ENGLAND