Nonmetric Interactive Multidimensional Scaling with Multiple Subjects

Robert M. Hamer
Medical College of Virginia, Virginia Commonwealth University

A method for nonmetric interactive multidimensional scaling (MDS) of similarity judgments is described which is also capable of using responses from previous judges to supplement the judgments of a current subject. The method combines recent advances in interactive MDS with recent advances in numerical methods in MDS to produce a program capable (1) of performing nonmetric interactive MDS and (2) of fitting a wide variety of models, such as the individual differences model. The empirical investigation compared three versions of the system: (1) a metric simple Euclidean model-fitting version (similar to previous interactive scaling programs); (2) a metric individual differences version; and (3) a nonmetric individual differences version. There were no statistically significant differences among the three versions.

This paper describes a method for nonmetric interactive multidimensional scaling (MDS) of judgments capable of using responses from previous judges to supplement the judgments of a current subject. The computer algorithm uses two mathematical models: (1) the interactive scaling model (Young & Cliff, 1972), which derives an r-dimensional space containing the stimuli in which $r + 1$ of the points lie on the axes of the space and (2) a Euclidean distance model used to represent the subject's judgments about successive subsets of the stimuli.

Young and Cliff (1972) contains a complete description of their system and the mathematical model; consequently, only a brief summary will be presented here.

Given that there are n stimuli, assumed to correspond to n points in an $r < n$ dimensional space. The exact dimensionality of this space is unknown. There exists a subject who will make similarity judgments about these points.

1. Select a "focus" of p points and collect judgments about all possible pairs from the subject.
2. Select the two points farthest apart and pass the first dimension through them. Find the point with the highest residual and pass the next dimension through it. Continue until the residual is less than a preset minimum. These points define a "frame."
3. Select new points to make a new focus and obtain the judgments needed to have all the judgments in this focus. If there are no more stimuli left unjudged, go to 4. Otherwise go back to 2.
4. Collect all uncollected judgments between the final "frame" stimuli and the remaining stimuli. This gives all the information to determine the coordinates of all the points in a r-dimensional space.

Method

The stimuli were a set of 30 animal names, which had been used previously in MDS research by Henley (1969) and Baker (1974). This set of stimuli was used because both of the previous studies had produced similar structures for this set; thus, there appeared to be a fairly firm idea of what the structure ought to look like. Having a well-established idea of the structure of the data was useful, as it made it possible to establish how successful this procedure was in recovering the structure.

The design used was a two by three randomized block design. Thus, there were six cells, each with eight subjects. The between-blocks factor, which had two levels, was a factor consisting of the set of stimuli used in the initial focus, and the factor with three levels which incorporated the MDS model used and the transformation method used. The within-blocks factor, which had three levels, was a model-transformation factor using subjects who were matched (assigned) to each other. The design was analyzed by a multivariate analysis that will be explained below.

The two sets of starting stimuli were selected to test the hypothesis that starting with basis stimuli would reduce the number of judgments required. If the final basis was composed entirely of stimuli which came from the first focus, a minimum number of judgments would be required. Conversely, if the final basis was composed entirely of stimuli that came solely from the last focus, a maximum number of judgments would be required (Girard & Cliff, 1976, p. 60).

Baker and Young (1975) and Baker (1974) attempted to decrease the number of judgments their procedure required by making it increasingly difficult to replace frame stimuli as the session advanced. Green (1977) and Green and Bentler (1979) also addressed this issue. They proposed using information on the stimuli, if available, to try to include in the first focus stimuli that were likely to be final basis stimuli.

Because the stimuli used in this study had been used in previous studies (Baker, 1974; Hen-ley, 1969), some information on their structure was available. For half the subjects, stimuli were chosen that were at the periphery of the spaces previously found, as these stimuli were likely to be the basis stimuli (Young & Cliff, 1972). For the remaining subjects the stimuli used in the initial focus were chosen randomly.

Previous interactive scaling systems mentioned (Baker & Young, 1975; Young & Cliff, 1972) assumed judgments to be ratio (that is, ratio distances or, minimally, linear transformations of ratio distances). This assumption was necessary because until relatively recently, computer technology did not exist to allow an interactive program running on a minicomputer dedicated to human experimentation to use nonmetric transformations. This technology now exists.

In nonmetric MDS, the number of parameters to be estimated from the judgments is greater than that in metric MDS. In interactive scaling, with a typically "small" focus size (often 5 or less) the dimensionality is "large" in relation to the focus size (often 3 or more). Thus, the ratio of parameters estimated (in a nonmetric procedure) to number of data points (unique judgments) is usually too large to produce stable estimates of these parameters. Spence (undated) stated that a good rule of thumb for minimum number of judgments to use in incomplete data collection procedures is $3(n)(m)$, where n is the number of dimensions and m is the number of stimuli, in order to obtain stable estimates.

In MDS, one way to increase the ratio of data points collected to parameters estimated is to use multiple subjects and an individual differences (ID) model. In an ID model, multiple subjects can be used to estimate the same set of parameters, with the only parameters unique to each subject being a vector of weights used to express the salience of each dimension for each subject. Thus, the ratio of data points collected to parameters estimated is large.

It may seem inappropriate to consider using an ID model in a single-subject interactive scaling situation. The procedure was as follows: First, a subject was run through the interactive

scaling procedure, using a metric transformation. After the procedure found a set of basis stimuli and collected all judgments required for a configuration, it collected all remaining judgments, both as a check and for further use. A second subject was assigned to (matched with) the subject just mentioned and was run through the interactive scaling procedure, using a metric transformation and a two-subject individual differences model. The two subjects were the current subject and the one to whom he/she had been assigned. The previously collected judgments from the first subject were used to supplement the judgments of the current subject; each focus was analyzed using a two-subjects ID model metric transformation analysis.

A third subject was assigned to (matched with) the first subject and was run through the procedure using a two-subject model (using the first subject's judgments as before) and a nonmetric (Kruskal's least squares monotonic) transformation. Thus, the contrast between the first and second levels of this factor addressed the ID model issue, and the contrast between the second and third addressed the metric-nonmetric issue.

Each subject was handed a set of instructions, escorted to a cubicle, shown how to use the terminal, and then left alone to interact with the program. The program presented pairs of stimuli to the subjects and requested judgments on a 1 to 9 similarity scale. After the program had obtained all possible within-focus judgments, it would analyze them, selecting those frame points that appeared to define a subspace, and then randomly selecting new points to make up the next focus. This process continued until all points had been used in a focus.

The program then selected a final frame and collected all judgments between the final frame stimuli and the remaining stimuli. It also collected, *as a check only,* all remaining judgments. These were used as criteria against which to evaluate the judgments predicted by the program.

Results

There were a variety of dependent variables, which were so highly correlated that only one of them will be discussed in detail. All were measures of fit (goodness of fit or badness of fit) between the distances resulting from the interactive scaling procedure and the distances resulting from a nonmetric MDS of the complete judgment matrix. The one which will be discussed was the Fisher Z transformation ($\tanh^{-1}r$) of the correlation between corresponding elements of the distance and judgment matrices. The measure was defined as follows:

Let $\mathbf{D}$ be the 30×30 matrix of distances from the interactive scaling procedure and let $\mathbf{D}^*$ be the 30×30 matrix of distances resulting from the MDS of the complete set of judgments. Let $\mathbf{v}$ and $\mathbf{v}^*$ be formed by reshaping $\mathbf{D}$ and $\mathbf{D}^*$, respectively, into vectors. The dependent measure was $\tanh^{-1}(r_{vv})$. This is a measure of the linear relationship between corresponding elements of the two matrices. The distances, rather than the coordinates, were used because (1) the distances are invariant to the particular orientation of the axes (the coordinates are not) and (2) to allow for differences in dimensionality.

Table 1 contains means, standard deviations, and the ANOVA table for the measure. There was no significant main effect for starting method; a carefully selected starting subset did not appear to improve the fit measures or to reduce the number of judgments required (all methods and both starting sets required about 200 judgments). This might have been attributable to (1) true lack of effect, (2) wrong subset chosen, or (3) lack of power.

The multivariate test of transformation method addressed the issues of whether supplementation improved the quality of the solutions and whether fitting a nonmetric model at each frame improved the solutions. Neither did.

The means tended to reflect patterns in a direction opposed to the predictions. In some respects, it appeared that nonmetric analyses pro-

Table 1
Means and Standard Deviations for r_{vv*} and $\tanh^{-1}(r_{vv*})$

Starting Method	Transformation Method					
	Unsupplemented Metric		Supplemented Metric		Supplemented Nonmetric	
	r	$\tanh^{-1}(r)$	r	$\tanh^{-1}(r)$	r	$\tanh^{-1}(r)$
Periphery	.5159	.5737	.5728	.6724	.4590	.5052
	.0612	.0820	.1328	.2046	.1129	.1510
Random	.5041	.5727	.4662	.5327	.4364	.4850
	.1471	.2020	.1824	.2576	.1646	.2122

Analysis of Variance

Source[a]	df	F	p
S	1,14	0.77	.40
T	2,13	1.13	.35
T1	1,13	0.28	.60
T2	1,13	2.43	.14
S × T	2,13	0.80	.47
S × T1	1,13	1.58	.23
S × T2	1,13	0.75	.40

[a]S is Starting Method; T is Transformation; T1 is Supplemented vs. Unsupplemented Contrast; T2 is Metric vs. Nonmetric Contrast.
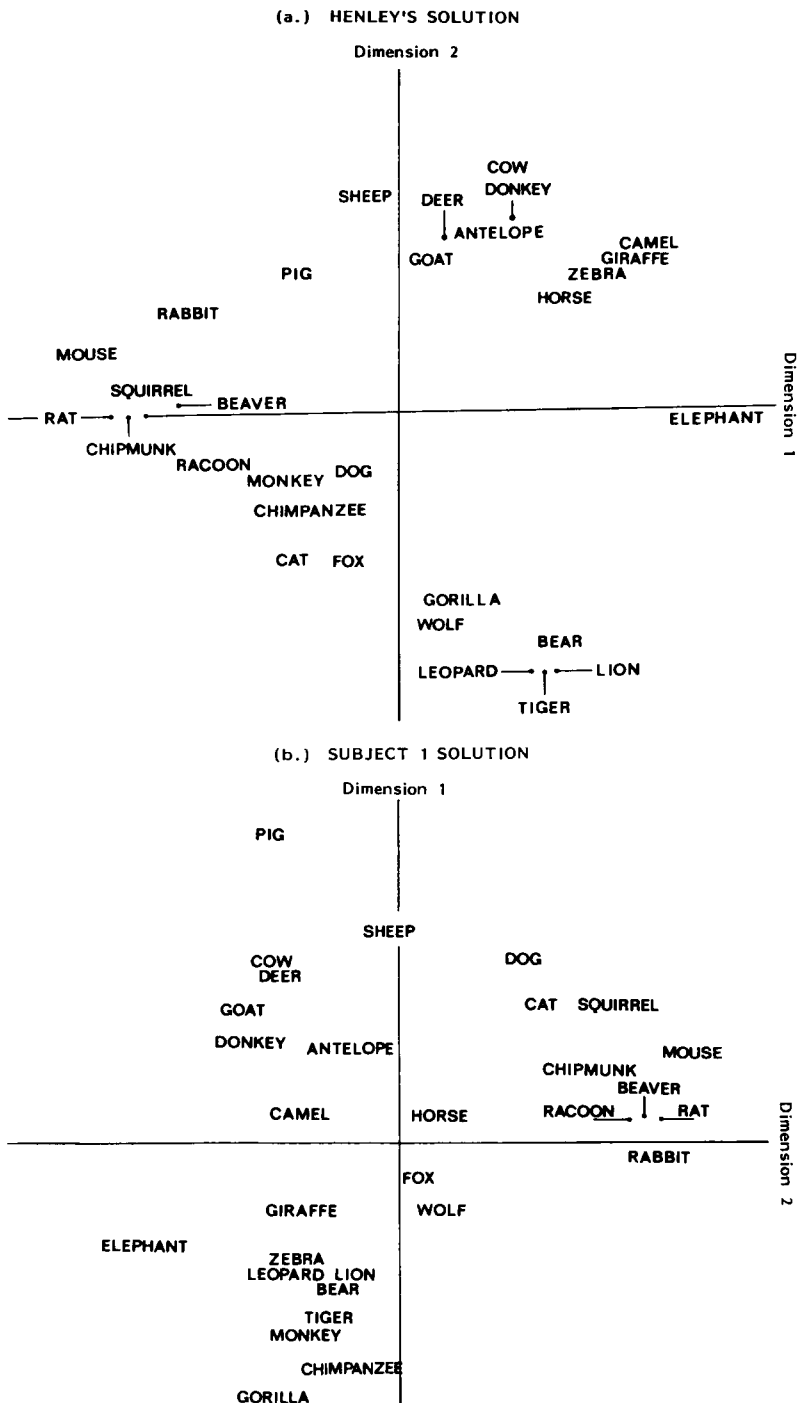
duced poorer solutions than did metric ones. Additionally, introducing supplementary judgments appeared to introduce more variation into the process than it accounted for. There were also hints that the treatments had some sort of complex effect on the variances, although it is impossible to say that with any certainty. Randomization tests were used to test the same hypotheses as the linear models analyses did, because the dependent variables (Fisher Z transformations of correlations of dependent observations) did not meet the assumptions necessary for linear models analysis. The p values obtained from the randomization tests were nearly identical to those obtained from the t distribution.

Discussion

The key issue in this study was whether a nonmetric interactive MDS procedure would produce better results than the metric procedures previously used. It did not. In fact, although there were no significant differences, the patterns of means and standard deviations showed a slight tendency for the nonmetric procedure to produce poorer results. Supplementing a subject's judgments with the judgments of an earlier subject did not help, either.

Two sets of starting stimuli were used: a random set and a set chosen to be at the periphery of the space defined by previous studies of the

Figure 1
Dimensions 1 and 2 of Henley's Solution and
the Solution of a Typical Subject (Subject 1)

(a.) HENLEY'S SOLUTION

Dimension 2

COW
DONKEY
SHEEP   DEER
ANTELOPE
CAMEL
GIRAFFE
GOAT   ZEBRA
HORSE
PIG
RABBIT
MOUSE
SQUIRREL
BEAVER
RAT
CHIPMUNK
RACOON
MONKEY   DOG
CHIMPANZEE
CAT   FOX
GORILLA
WOLF
BEAR
LEOPARD   LION
TIGER

Dimension 1

ELEPHANT

(b.) SUBJECT 1 SOLUTION

Dimension 1

PIG

SHEEP

COW
DEER   DOG
GOAT   CAT   SQUIRREL
DONKEY   ANTELOPE
MOUSE
CHIPMUNK
BEAVER
CAMEL   HORSE   RACOON   RAT
RABBIT

Dimension 2

FOX
GIRAFFE   WOLF
ELEPHANT
ZEBRA
LEOPARD LION
BEAR
TIGER
MONKEY
CHIMPANZEE
GORILLA

same stimuli. Which set of starting stimuli were used did not appear to make any difference.

Previous studies of these stimuli produced three-dimensional solutions (Baker, 1974; Baker & Young, 1975; Henley, 1969), of which only the first two dimensions were usually interpretable. One dimension was usually a size dimension, and the second was usually interpreted as ferocity. Figure 1 contains plots of Henley's group solution and the solution from the first subject in this study. The solutions look comparable: One dimension is clearly a size dimension, and the other can be interpreted as a bipolar dimension ranging from domestic to ferocious.

The problem of finding the proper way to assess the fit between two MDS solutions has not yet been completely solved. Young and Cliff (1972), Baker and Young (1975), Baker (1974), and Green (1974) used correlations of various types to assess fit. Baker and Young (1975) also used canonical correlation to compare the coordinates of two solutions. Because of possible differences in dimensionality, as well as difficulties in interpreting the exact meaning of the canonical correlations, that was not done here.

The correlations tended to be lower than those found with previous procedures, averaging around .50, accounting for only about one quarter of the variation. Thus, this nonmetric individual differences MDS procedure appears to result in different solutions than other procedures applied to the same data.

References

Baker, R. F. *Interactive scaling with individual subjects: Two evaluations.* Unpublished doctoral dissertation, The University of North Carolina, Chapel Hill, 1974.

Baker, R. F., & Young, F. W. A note on empirical evaluation of the ISIS procedure. *Psychometrika,* 1975, *40,* 413–415.

Girard, R. A., & Cliff, N. A monte carlo evaluation of interactive multidimensional scaling. *Psychometrika,* 1976, *41,* 43–64.

Green, R. S. *A monte carlo evaluation of a computer-interactive extended transitivity dominance scaling model.* Unpublished doctoral dissertation, University of Southern California, Los Angeles, 1974.

Green, R. S. *Evaluating some improvements in interactive MDS.* Paper presented at the meeting of the Psychometric Society, Chapel Hill NC, April 1977.

Green, R. S., & Bentler, P. M. Improving the efficiency and effectiveness of interactively selected MDS data designs, *Psychometrika,* 1979, *44,* 115–119.

Henley, N. M. A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior,* 1969, *8,* 176–184.

Spence, I. *Incomplete experimental designs for multidimensional scaling.* Unpublished manuscript, undated.

Young, F. W., & Cliff, N. Interactive scaling with individual subjects. *Psychometrika,* 1972, *37,* 385–415.

Acknowledgments

Author's Address

Send requests for reprints or further information to Robert M. Hamer, Medical College of Virginia, Virginia Commonwealth University, MCV Station, Box 710, Richmond VA 23298.