PSYCHOMETRIKA—VOL. 41, NO. 2
JUNE, 1976

EFFECTS ON INDSCAL OF NON-ORTHOGONAL PERCEPTIONS OF OBJECT SPACE DIMENSIONS

ROBERT C. MACCALLUM

THE OHIO STATE UNIVERSITY

A general question is raised concerning the possible consequences of employing the very popular INDSCAL multidimensional scaling model in cases where the assumptions of that model may be violated. Simulated data are generated which violate the INDSCAL assumption that all individuals perceive the dimensions of the common object space to be orthogonal. INDSCAL solutions for these various sets of data are found to exhibit extremely high goodness of fit, but systematically distorted object spaces and negative subject weights. The author advises use of Tucker's three-mode model for multidimensional scaling, which can account for non-orthogonal perceptions of the object space dimensions. It is shown that the INDSCAL model is a special case of the three-mode model.

A general problem in behavioral research involves the selection of a method of data analysis when several alternative methods are available. This problem is especially evident in the area of individual differences models for multidimensional scaling; given n individuals, p objects, and similarity or dissimilarity measures by each individual for all pairs of objects, a variety of such models have been proposed for describing individual differences in judgments about object relationships [Tucker and Messick, 1963; Kruskal, Note 1; McGee, 1968; Horan, 1969; Carroll and Chang, 1970; Tucker, 1972].

Two models will be the subjects of the present paper. The first model was originally proposed by Horan [1969] and was further developed by Carroll and Chang [1970]. This model is now commonly referred to as the INDSCAL model, after a widely distributed computer program which follows the computational procedure developed by Carroll and Chang. This model makes three basic assumptions about the nature of individual differences: (a) all individuals share a common r-dimensional object space; (b) observed individual differences in judgments about object relationships arise from differential weighting of the dimensions by each individual; and (c) all individuals perceive the dimensions of this object space to be independent, or orthogonal (to be developed below). The basic equation for this model can be written in scalar products form as follows:

Requests for reprints should be sent to Robert C. MacCallum, Department of Psychology, The Ohio State University, Columbus, Ohio 43210.
Key words: multidimensional scaling, individual differences models for multidimensional scaling, three-mode factor analysis, three-way scaling models.

$$x_{jj'i} = \sum_{t=1}^{r} w_{it} b_{jt} b_{j't} , \tag{1}$$

where $x_{jj'i}$ represents the scalar product between vectors from the origin to points, or objects, j and j'; b_{jt} represents the projection of point j on dimension t; and w_{it} represents the weight which individual i applies to dimension t. The INDSCAL procedure employs techniques presented by Torgerson [1958] to convert proximity judgments to scalar products, then obtains a least squares solution for the w_{it} and b_{jt} parameters. The object space configuration specified by the b_{jt}'s is not subject to rotation.

A second model, proposed by Tucker [1972], provides a slightly more general description of individual differences. The model involves an application of three-mode factor analysis [Tucker, 1966] to the three-way data array made up of n square symmetric scalar products matrices, one for each individual. This model will be referred to as three-mode multidimensional scaling. Under the assumptions of this model, individual differences are described in the following manner: (*a*) as in INDSCAL, individuals share a common r-dimensional object space; (*b*) as in INDSCAL, individuals apply differential weights to the object space dimensions; and (*c*) unlike the INDSCAL representation, individuals are allowed to perceive varying degrees of relationship among the dimensions of the object space. The three-mode model can be expressed in scalar products form as follows:

$$x_{jj'i} = \sum_{t=1}^{r} \sum_{t'=1}^{r} h_{tt'i} b_{jt} b_{j't'} , \tag{2}$$

where $x_{jj'i}$ and b_{jt} are the same as above, and $h_{tt'i}$ is determined by the perceived relation between object dimensions t and t' for individual i. When $t = t'$, then $h_{tt'i}$ is the weight which individual i applies to dimension t. Thus, the INDSCAL model is seen to be a special case of the three-mode model. If all individuals perceive the object dimensions to be independent, the parameter $h_{tt'i}$ will be non-zero only when $t = t'$. This would reduce (2) to a single summation:

$$x_{jj'i} = \sum_{t=1}^{r} h_{tti} b_{jt} b_{j't} , \tag{3}$$

which is equivalent to (1).

A matrix formulation of the models will clarify their relationships. Both models could be represented by the matrix equation:

$$X_i = B\,H_i B'. \tag{4}$$

X_i is a p-by-p scalar products matrix for individual i; B is a p-by-r matrix specifying the common object space configuration; and H_i is an r-by-r symmetric matrix specifying the nature of individual i's perception of the object space dimensions. In the INDSCAL model, H_i is restricted to being diagonal for each individual, thus indicating that the individuals perceive the object space dimensions to be independent. In the three-mode model, there is no such restriction on H_i, thus allowing individual differences in perceived relations among the object space dimensions. Note especially that the relations among the object space dimensions are not defined in matrix B, but rather in the individual H_i matrices.

The INDSCAL model and the three-mode model are clearly the most advanced and most valuable techniques available for analyzing individual differences in multidimensional scaling. A researcher faces a very real problem in attempting to choose between these two models. The use of either model implies adoption of the assumptions discussed above with respect to the representation of individual differences. If these assumptions do not in fact hold, the model in use may provide an inappropriate representation of the phenomena being investigated. Shepard [1972] discusses this problem in a more general sense. The present study is concerned with the consequences of employing the INDSCAL model when one of its assumptions is known to be violated. The particular assumption under study is the notion that all individuals perceive the object space dimensions to be independent. To investigate this matter, data have been generated which systematically violate this assumption; these data have then been analyzed via the INDSCAL procedure. Results will provide some information about the consequences of using the INDSCAL model when it is known that the three-mode model is more appropriate.

Method

Since both models assume the presence of a common object space and differential weighting of dimensions, these aspects of the data were held constant across all sets of simulated data. A two-dimensional common object space was devised by placing nine points in the form of a square. This configuration is shown in Figure 1. Using a pseudo-random number generator, subject weights for the two object dimensions were created for 100 hypothetical individuals. These weights were generated so as to be uniformly distributed on the interval zero-one on each dimension. This set of weights and object space configuration were employed in the generation of every set of simulated data.

The manipulation of the perceived relations between object dimensions was accomplished by controlling the value of θ_i, the angle which individual i perceives between the two object dimensions. If all θ_i are 90° for the 100 individuals, the data will fit both models perfectly. The present study involves the behavior of the INDSCAL model when we define a distribution of θ_i which cannot be directly accounted for by that model.

Two aspects of the distribution of θ_i were controlled. First, the range

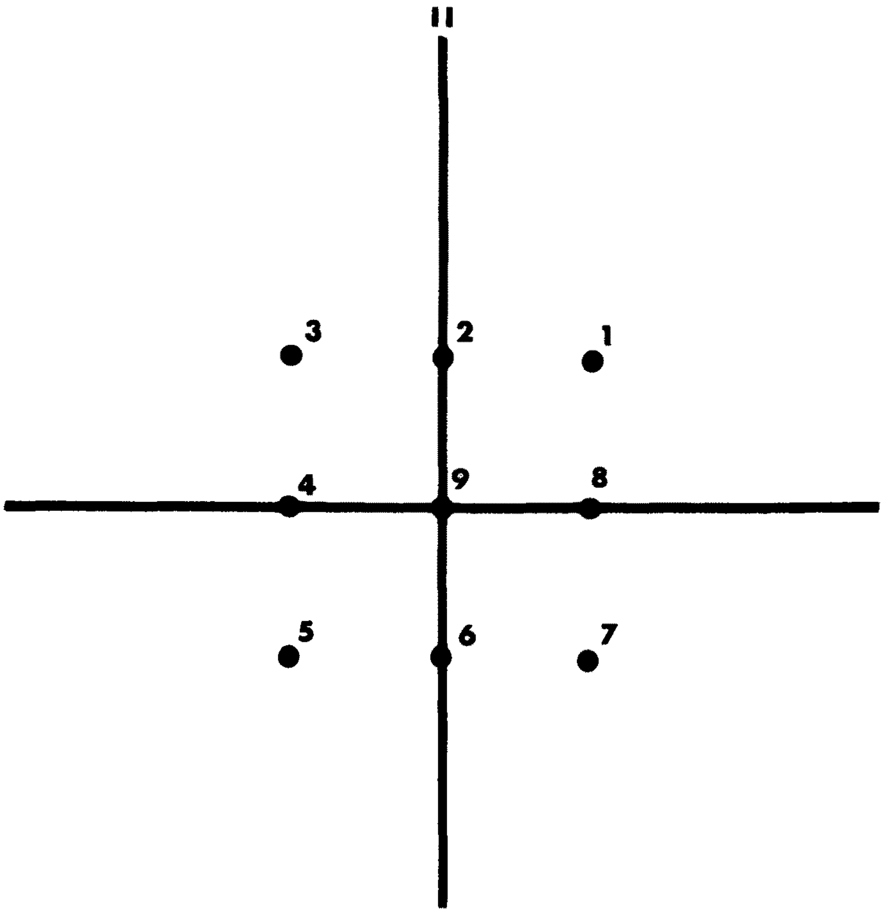

FIGURE 1.
Original Underlying Object Space for Simulated Data.

of the distribution of θ_i was allowed to take on one of three possible values: $\pm$ 0°, $\pm$ 15°, or $\pm$ 30°. This allowed control of the degree of individual differences in perception of the relation between the two dimensions. Second, the midpoint of the range of θ_i was allowed to take on one of three values: 30°, 60°, or 90°. This allowed control of the "mean" perception of the relation between the object dimensions. Combining these two factors gives rise to nine different conditions with respect to the distribution of θ_i. The design is presented in Table 1. When the range of θ_i was $\pm$ 0°, all individuals had the same θ_i, and only one set of simulated data could be generated for each midpoint of θ_i. Three sets of simulated data were generated for the cells in which the range of θ_i was $\pm$ 15° or $\pm$ 30°. Under these conditions, angles

TABLE 1

Design of Simulated Data Conditions

Range of Θ_i	Midpoint of Range of Θ_i	Interval of Θ_i	Number of Sets of Data
± 0°	30°	$30° \leq \Theta_i \leq 30°$	1
± 0°	60°	$60° \leq \Theta_i \leq 60°$	1
± 0°	90°	$90° \leq \Theta_i \leq 90°$	1
± 15°	30°	$15° \leq \Theta_i \leq 45°$	3
± 15°	60°	$45° \leq \Theta_i \leq 75°$	3
± 15°	90°	$75° \leq \Theta_i \leq 105°$	3
± 30°	30°	$0° \leq \Theta_i \leq 60°$	3
± 30°	60°	$30° \leq \Theta_i \leq 90°$	3
± 30°	90°	$60° \leq \Theta_i \leq 120°$	3

were generated for the 100 individuals so as to be uniformly distributed across the specified range. It was then possible to combine the object space matrix, the individual weight matrix, and the information about angles between dimensions to construct scalar products matrices for each of the 100 individuals. The design gives rise to 21 separate sets of simulated data. Each fits the three-mode model perfectly, and each is characterized by a specified distribution of θ_i. Each of these 21 sets of data was then analyzed in two dimensions by the INDSCAL procedure.

Results and Discussion

The INDSCAL program provides a goodness of fit measure which is the product moment correlation between all the elements of the observed scalar products matrices and all the elements of the scalar products matrices as reconstructed by the INDSCAL solution. The goodness of fit of INDSCAL to the 21 sets of simulated data is shown in Table 2. The procedure obviously fits all 21 sets of data extremely well. It is verified that INDSCAL fits perfectly when all θ_i are 90°.

The fact that INDSCAL fits so well seems rather surprising, considering the extensive violation of the INDSCAL assumptions in many of the conditions. This indicates that individual differences in perceived relationships among dimensions do not have great effect on the fit of the INDSCAL procedure. Thus, when poor fit is observed in practice, it cannot necessarily be attributed to non-orthogonal perceptions of object space dimensions. It would more likely be attributable to other factors; *e.g.*, other systematic effects occurring in the perception of the objects, or error of measurement. Since this measure of fit is not very sensitive to controlled variation in the distribution of θ_i, one must wonder whether this index provides in any sense

TABLE 2

Goodness of Fit of INDSCAL Procedure to 21 Sets of Simulated Data

Midpoint of Range of Θ_i	Range of Θ_i		
	$\pm 0^\circ$	$\pm 15^\circ$	$\pm 30^\circ$
30°	.9961	.9959 .9956 .9956	.9922 .9925 .9943
60°	.9981	.9939 .9938 .9949	.9850 .9828 .9820
90°	1.0000	.9935 .9930 .9921	.9734 .9758 .9759

a measure of the appropriateness of the INDSCAL model to a given set of data. This is an important question, and will be discussed after the remaining results are presented.

Let us now consider the object space configurations offered by the INDSCAL solutions for these simulated data. These solutions were found to be virtually invariant with respect to changes in the range of θ_i. However, systematic effects were observed with respect to the changes in midpoint θ_i. Mean configurations for the seven separate INDSCAL object space solutions representing each level of midpoint θ_i are presented in Figures 2, 3, and 4. As evidence for the striking consistency of the object configurations across all levels of range of θ_i, the author computed mean squared deviations of the coordinates of the seven object space solutions from the mean object space for each level of midpoint θ_i. These mean squared deviations were .00005 when midpoint θ_i was 30°, .00007 when midpoint θ_i was 60°, and .00024 when midpoint θ_i was 90°. Thus, Figures 2, 3, and 4 can be taken to be representative of INDSCAL object space solutions for all 21 sets of simulated data.

These figures reveal very obvious distortions of the original square configuration into a parallelogram. The degree of compression is directly related to the deviation of midpoint θ_i from 90°. The dimensions recovered

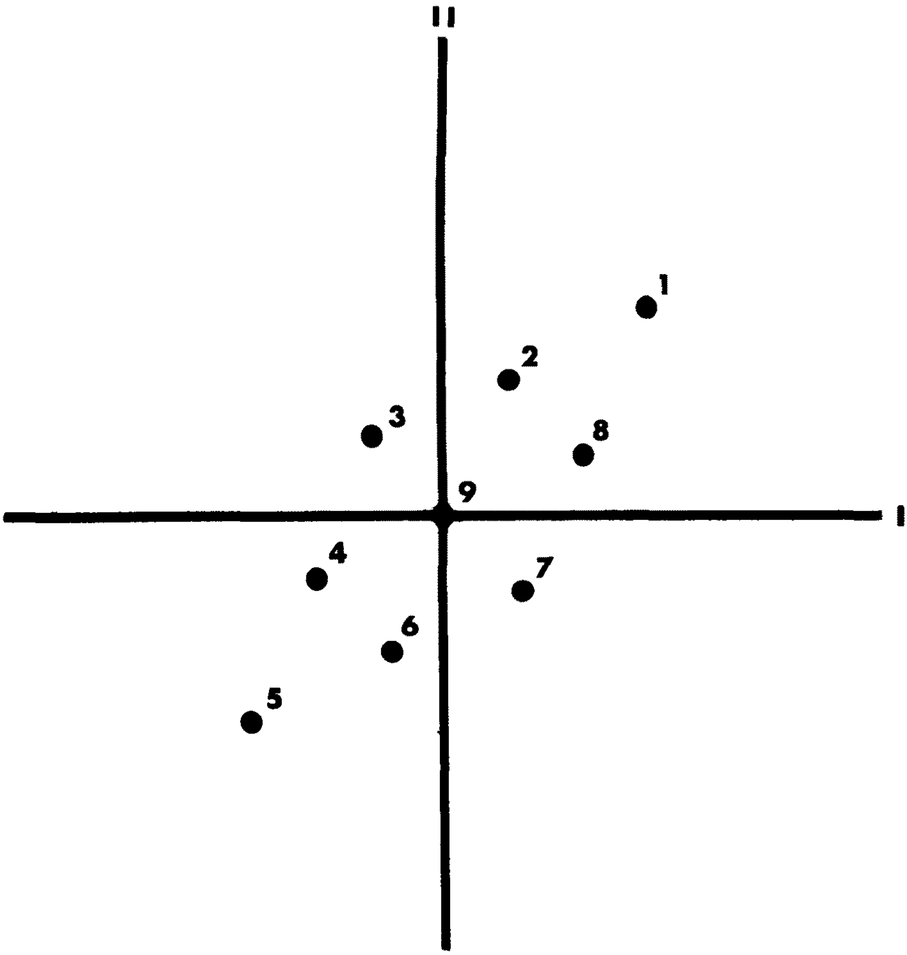

FIGURE 2.
Mean INDSCAL Object Space Solution for Midpoint $\theta_i = 30°$.

when midpoint $\theta_i \neq 90°$ do not correspond to the original underlying dimensions as constructed. Those original dimensions could be obtained in the distorted INDSCAL spaces via oblique rotation, but this is not permissible in INDSCAL solutions. These results indicate that when there are consistently non-orthogonal perceptions of the object space dimensions, INDSCAL gives rise to a systematically distorted view of the true underlying object space. Furthermore, the extent of this distortion is directly related to the deviation of the median perception from orthogonality.

It is of utmost importance to recognize that this distortion gives rise to considerable complications in the *interpretation* of the object space dimen-

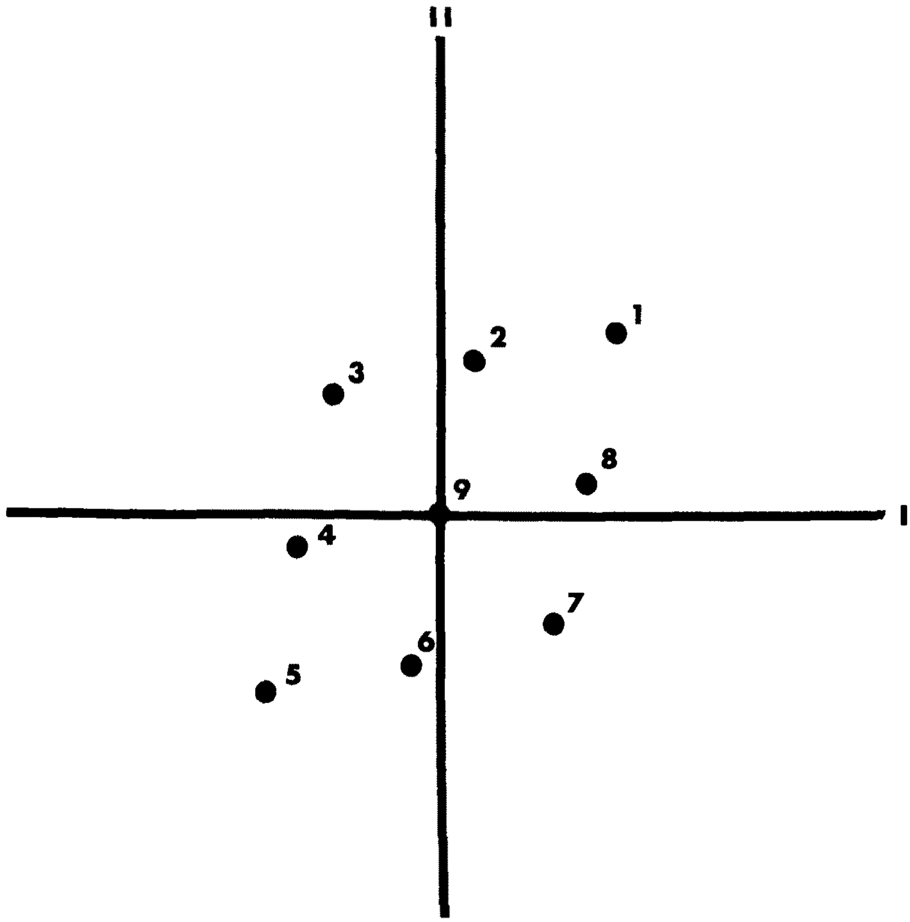

FIGURE 3.
Mean INDSCAL Object Space Solution for Midpoint $\theta_i = 60°$.

sions. There is a high degree of simplicity inherent in the dimensions of Figure 1. These dimensions might be considered as being analogous to simple perceptual attributes of the set of objects; *i.e.*, the attributes to which the individuals attend when judging similarity of objects. Considering now the INDSCAL solutions, note the considerable complexity inherent in the dimensions of Figures 2 and 3. These dimensions *do not* correspond to the very simple underlying dimensions of the original object space. The implication for practical applications is critical: when the *relevant object dimensions* are not, on the average, perceived to be *mutually independent,* then those dimensions *will not be recovered* by INDSCAL, and psychological interpretation will suffer.

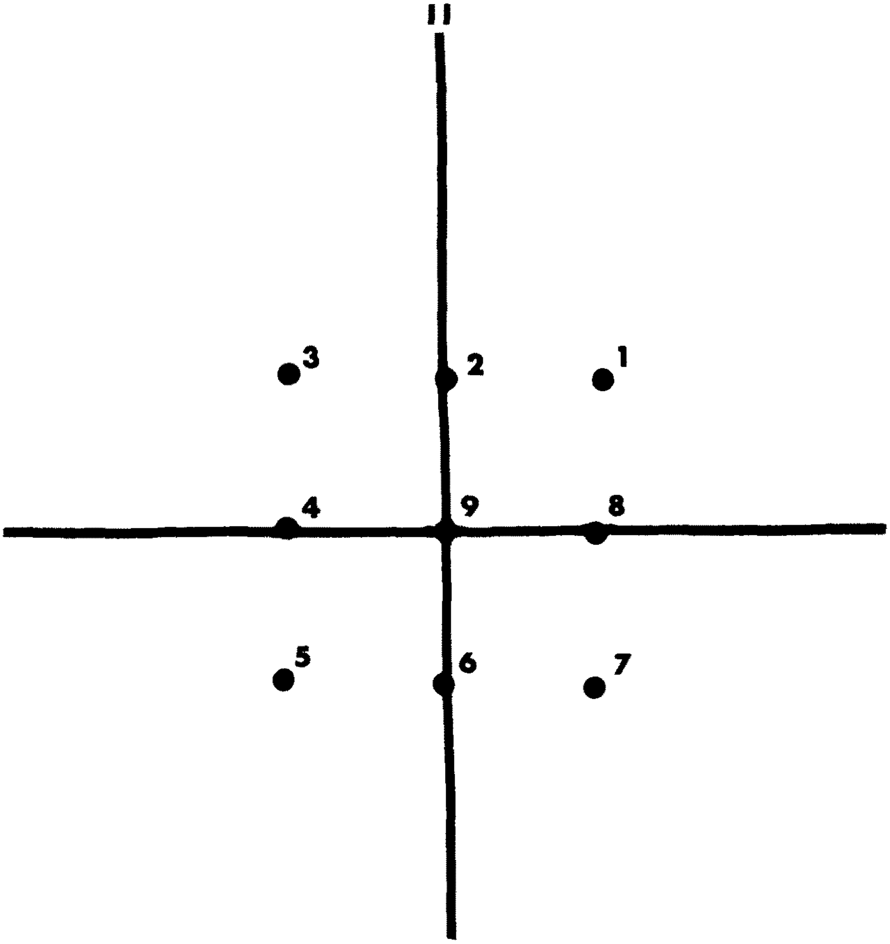

FIGURE 4.
Mean INDSCAL Object Space Solution for Midpoint $\theta_i = 90°$.

The fact that the INDSCAL object space solutions for the various sets of simulated data reveal configurations which may be thought of as accurate representations of perceived configurations is of little consequence. The INDSCAL procedure seeks to recover not only a configuration of points, but, more importantly, a set of unique dimensions which hopefully correspond to relevant perceptual attributes of the objects. Obviously, the latter goal is not accomplished under the violations being studied here.

Let us now examine the subject weights provided by the INDSCAL analyses of the simulated data. The most interesting information in these results involved the occurrence of negative weights. Table 3 shows the number

TABLE 3

Number of Negative Weights Occurring in INDSCAL Solutions for 21 Sets of Simulated Data

Midpoint of Range of Θ_i	Range of Θ_i		
	$\pm 0^o$	$\pm 15^o$	$\pm 30^o$
30^o	23	22 22 22	21 21 21
60^o	8	8 8 7	7 7 6
90^o	0	0 0 0	0 0 0

of negative weights occurring in the INDSCAL solutions for the 21 sets of simulated data. It is apparent that the effect of the range of θ_i is negligible, while the midpoint of θ_i has a clear effect on the results. The presence of consistent non-orthogonal perceptions clearly causes negative weights to occur in the INDSCAL solutions. Recall that all of the original simulated weights were generated to lie between zero and one. Thus, the perceptual spaces of all hypothetical persons could be presented in terms of a real Euclidean space by the three-mode scaling model. The occurrence of negative weights in INDSCAL solutions for data in which the midpoint of θ_i is not 90° reveals the inability of INDSCAL to account for the judgments of all individuals in such a group in Euclidean terms.

Interpretation of these results requires that one draw a distinction between the INDSCAL *model* and the INDSCAL *computational procedure*. While the INDSCAL model is Euclidean in nature, the computational procedure presented by Carroll and Chang does not constrain the individual weights to being non-negative. As a result, the procedure can produce non-Euclidean solutions; *i.e.*, solutions exhibiting some negative weights. However, this does not necessarily represent an undesirable quality in the computational procedure. The occurrence of negative weights constitutes potentially important information. In discussing this matter, Carroll and Chang point out that it is generally unnecessary to constrain the weights to be non-negative, stating that if the *model* is *systematically* violated, this may be reflected in negative weights. Thus, the occurrence of negative weights may indicate that the *model* is not psychologically meaningful or appropriate for the given set of data.

Present results offer additional information with respect to such an interpretation. The occurrence of negative weights in a practical application of INDSCAL might be taken as evidence for the presence of consistent non-orthogonal perceptions of the true object space dimensions. It is now clear that the presence of such perceptions is one characteristic which can give rise to negative INDSCAL weights. In such cases, the INDSCAL model is, of course, not appropriate.

Given these general results, let us reconsider the question raised earlier: that is, does excellent fit indicate that the INDSCAL model is appropriate for the given set of data? The results indicate that this is definitely not so. It is possible to observe extremely high fit, yet a badly distorted object space, and corresponding poor recovery of underlying dimensions, as well as large numbers of negative weights. In this light, it is important to draw a distinction between *fitting* the data and *representing the underlying structure* of the data. It is seen that the INDSCAL procedure may fit the data very well, while the actual solution may not provide an appropriate representation of either the stimuli or the individuals. Clearly, such solutions are of little value. The problem could perhaps be somewhat alleviated by the development of a new measure of fit, which might take into account the nature of the weights. An alternative approach might be to employ a computational procedure which would restrict the INDSCAL weights to being non-negative. This approach is currently being investigated by the author.

Conclusions

Results indicate that the INDSCAL procedure is susceptible to a variety of serious difficulties when the assumption that the individuals perceive the underlying dimensions to be orthogonal is violated. It is certainly conceivable that such problems could be occurring in practical applications and going unnoticed. In such situations, the INDSCAL model would be providing unreasonable solutions due to violation of this basic assumption. In such cases, Tucker's three-mode multidimensional scaling model would be more appropriate. Since a researcher does not know whether this assumption is violated in his data, the author suggests that it would generally be safer to employ Tucker's three-mode model as an individual differences model for multidimensional scaling. If orthogonal perceptions of the object space dimensions would provide an appropriate representation of the data, such a representation could be achieved by proper transformation of the three-mode solution. The author is presently developing a formal procedure for acomplishing such a transformation.

These problems are clearly relevant to empirical applications of these scaling techniques, and therefore merit concern and further investigation.

REFERENCE NOTE

1. Kruskal, J. B. *How to use MDSCAL, a program to do multidimensional scaling and multidimensional unfolding.* Unpublished report, Bell Telephone Laboratories, 1968.

REFERENCES

Carroll, J. D. and Chang, J. J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika,* 1970, **35**, 283–319.

Horan, C. B. Multidimensional scaling: Combining observations when individuals have different perceptual structures. *Psychometrika,* 1969, **34**, 139–165.

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

Shepard, R. N. Introduction to Volume I. In R. N. Shepard, A. K. Romney, and S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences.* Vol. I: Theory. New York: Seminar Press, 1972.

Torgerson, W. W. *Theory and methods of scaling.* New York: Wiley, 1958.

Tucker, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika,* 1966, **31**, 279–311.

Tucker, L. R. Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika,* 1972, **37**, 3–27.

Tucker, L. R and Messick, S. An individual differences model for multidimensional scaling. *Psychometrika,* 1963, **28**, 333–367.

Manuscript received 10/14/74
Revised manuscript received 6/26/75
Final version received 8/25/75