Multidimensional scaling of individual differences for an individual person with a hand-held calculator

RICHARD G. DAVIS

Veterans Administration Medical Center, Lexington, Kentucky 40507

An extension to the INDSCAL method of individual differences scaling was developed. The method derives measures of deviation from a "standard" observer, who is defined in terms of the INDSCAL subject weights. These weights are calculated for one individual person at a time. The method is implemented in a small hand-held calculator. The calculator is used for data acquisition, and it interacts with the user in a conversational manner through an alpha-numeric display. The present implementation of the method scales 84 judgments of pair similarity among seven objects and derives subject weights on two dimensions. The method can be extended to sets of up to 16 objects and up to six dimensions. The method is a feasible, cost-effective approach to classification of individuals on complex perceptual attributes within a single 1-h session.

The judgment of pairwise similarity is a powerful and efficient tool for the assessment of perceptual structures in individuals. For the subject, a similarity judgment is relatively unconstrained. The subject is free to react to the comparative aspects of stimuli according to the idiosyncratic view of that subject. If a stimulus set is chosen as a uniform and representative sample of a domain of interest, then the perceptual character of any subject should be adequately manifested in the structure of a set of similarity judgments for paired comparisons of that stimulus set.

A rationale of perceptual assessment based on similarity leads to multidimensional scaling (MDS), which extracts relatively simple descriptive structures from the more complex networks of pairwise similarity judgments. One example of an MDS approach that honors the perceptual uniqueness of individuals is INDSCAL (Carroll & Chang, 1970). The INDSCAL algorithm computes a group stimulus space that displays stimulus similarities as geometric proximities. INDSCAL furnishes a second table of subject weights that map each subject's private stimulus space into the group stimulus space.

INDSCAL is based on a variation of the model proposed by Horan (1969). This model assumes that some set of n dimensions influences all person's similarity judgments and, further, that persons differ only on the degree of influence each dimension has on their judgments. The model allows for the case wherein a person

is uninfluenced by a dimension, which reduces the dimensionality of the model for that person. Conversely, the model admits the case for rare persons who are influenced by a dimension not apparent in the majority of personal spaces, which raises the dimensionality of the group space. Thus, the INDSCAL model can be extended to provide an assessment of the typical aspects and the idiosyncratic features of a single person's perception in some stimulus domain.

The focus of this report is on a procedure to characterize the private perceptual space of a single person within the context of a single session of contact. If the contact is in a clinical setting, then the result may be needed to guide decisions on therapeutic planning. In an experimental laboratory setting, the result may direct assignments to treatment groups to achieve homogeneity of subject characteristics.

The main result of the procedures is a set of subject weight indexes, which reflect the degree to which the subject appears to be influenced by the dimensions of the normative model. The particular microstructure of the subject's private stimulus space is not of interest here, only the indexes that relate that space to the group space. Therefore, a much smaller set of stimulus pairs can be used than would otherwise be possible. This has the benefit of greatly reducing the amount of testing time and materials needed and the computational demands in the calculation phase.

Note that the procedure to be presented here is designed to scale a single subject in relation to a group model of perception. For clarity and efficiency of discussion, it will be convenient to call the procedure 3S, for single-subject scaling.

A group stimulus space is assumed to exist, and the 3S method is initialized based on that group space.

This report was supported by the Veterans Administration Medical Center, Lexington, Kentucky, research program. Reprints are available from the author at Research Service (151A1), VAMC, Lexington, Kentucky 40507. The program and its documentation will appear in the Hewlett-Packard Users Library for the HP41C. Interim documentation is available from the author.

Therefore, the development of the group space will be discussed. The key issues are normative subjects sampling strategy, stimulus domain sampling strategy, and the method of INDSCAL solution normalization.

Subject sampling strategy is influenced by the theoretical demands of MDS and also by the intended purpose of the 3S procedure. Subjects should be chosen as an unbiased, representative sample of the population of interest. At the same time, selective constraints may be imposed. If the goal is to identify individuals with 3S who are "typical," then a highly homogeneous sample that appears to represent the majority of persons would be appropriate. A 3S procedure based on such a sample would identify "anomalous" perceivers, but it might not differentiate among them. When 3S is expected to differentiate among the anomalous perceivers, sampling strategy must be extended to include sufficient numbers of each type of anomalous perceiver expected.

Stimulus sampling strategy will directly influence the results of a 3S procedure. If a set of stimuli that represent one dimension of perception is omitted, then the 3S method will be insensitive to that source of differences. Moreover, stimuli that span perceptually significant ranges on each of the dimensions must appear in the group space. Therefore, the stimulus set should be designed to uniformly span the stimulus domain of interest. However, when a particular application of 3S is focused on a local region of a perceptual domain, there may be a workable subset of stimuli that can determine the group space.

Adequate stability of the group stimulus space depends upon the number of stimuli included and the dimensional size of the space. Generally, authorities encourage use of the largest number of stimuli possible. However, experimental logistics encourage minimum stimulus numbers. In the context of two-way MDS, there are numerous suggestions about stimulus numbers. Kruskal and Wish (1978, p.34) recommend "a rough rule of thumb . . . [of] . . . at least twice as many pairs as parameters to be estimated." Hence for n stimuli and d dimensions, $n(n-1)/2 \geq 2nd$. Sherman (1972) gives a ratio of degrees of freedom as an index of adequate stability, $df(D)/df(X)$, where $df(D) = n(n-1)/2$ and $df(X) = d(n-1) - d(d-1)/2$. Sherman proposes a ratio greater than 2.5. Others (Spence & Domoney, 1974; Young, 1970) suggest a ratio of three for unique judgments to estimated parameters. Ramsay (1980) studied the MULTISCALE maximum likelihood estimation algorithm and found two replications of 15 stimuli or five replications of 10 stimuli were adequate.

Beyond statistical reliability in the group space, a large number of stimuli is worthwhile for practical reasons. Given a large and densely populated group stimulus space, numerous advantages follow. Such a space will be more likely to exhaustively represent the dimensions that influence the perceptual process represented by the space. Particular dimensions of interest can be evaluated by small stimulus lists abstracted from

the group space using the present 3S procedure. In a large space, the overdetermination of loci should promote a more reliable model of the microstructure in the space. Finally, in clinical applications, a 3S user may wish to construct a stimulus set to evaluate a particular aspect in one individual, and the more populous group space is more likely to meet the needs of such special test attempts. Recognizing that the largest possible stimulus space should be constructed as a basis for 3S, the recommended minimums for two, three, four, five, and six dimensions are 9, 13, 17, 21, and 24 stimuli, respectively.

The group stimulus space determined by INDSCAL may be normalized along each dimension by any scalar constant. The choice of normalization method is arbitrary, and the essential structure in the group space is invariant over all possible normalization methods. Carroll and Chang (1970) selected the method that sets the variance on each group stimulus space dimension equal to 1. This method has the advantage of giving vector lengths in the subject weight space that are approximately equal to the proportion of variance in common between the subject's private perceptual space and the group stimulus space. This property of the subject weight vector is convenient. Moreover, the associated group stimulus space depicts the structure for that hypothetical observer who weights all dimensions equally.

However, the typical person in a population of interest may not weight each dimension equally. It is quite reasonable to expect that one or two dimensions may be dominant in similarity judgments for most persons. Hence, the group stimulus space can be misleading, as Carroll and Chang (1970) have normalized their solution. In fact, it can be shown that the unity-variance normalized model may display extremely unrepresentative structure. Moreover, the 3S procedure seeks to scale persons vis-à-vis the group space. The interpretation of such a scaling would be more convenient if the hypothetical observer represented by the group space were characteristic of the most frequently observed individual spaces. Therefore, some other normalization method is desirable for the present purpose. What is the consequence of variations in normalization of the INDSCAL solution?

It can be shown that while the orientation of the INDSCAL solution is unique, a particular pair of stimulus/subject weight spaces is not unique. In fact, any particular solution given by INDSCAL can be rewritten using the following equations. In the group stimulus space, for every coordinate x in the original solution and x_n in the renormalized solution, $x_n = x/\sqrt{a}$, and in the subject weight space for every weight w in the original solution and w_n in the renormalized solution, $w_n = aw$, where a is the constant of renormalization. This result follows from the fact that the stimulus and subject weight arrays are interdependent arrays that summarize the several private stimulus spaces used to calculate

the INDSCAL solution. Also, the essential structure in the group stimulus space is invariant over a multiplicative transform.

The issue of normalization therefore becomes: among the infinite set of scalars a , which values would provide a more desirable model for the 3S procedure? The present 3S procedure calculates the subject weights relative to a standard observer (SO), which is a hypothetical person defined by averages computed in the native INDSCAL subject space. This renormalized INDSCAL model has the following desirable properties: (1) The SO lies on the diagonal in the subject weight space. (2) The entries in the group stimulus space are the root mean square values of the respective entries in the several private subject spaces. (3) The group space is the model for a hypothetical person whose scalar products array is the average of the several individual scalar products arrays for the subjects included in the analysis. (4) The general structure of the renormalized group stimulus space does not present a distorted impression of the individual subject spaces.

THE STATICAL MODEL

The 3S procedure is initialized from an INDSCAL group stimulus space that has n stimuli and r dimensions. This stimulus space array can be written as the Kronecker product of the column vectors to produce an n^2 by r matrix, G . This G array is derived from an optimum INDSCAL data acquisition procedure in terms of stimulus set definition and subject sampling procedures, as discussed above.

The set of dissimilarity judgments for a single subject can be converted into a table of cross products by the method given by Torgerson (1958), which is presented below (Equation 3). This array, Z , can be postmultiplied by the right pseudoinverse of G to yield a vector W with r elements. The vector W is the set of subject weights for the individual who produces the Z array, and W relates that subject's private perceptual space to the group stimulus space.

This procedure follows from the equation given by Carroll and Chang (1970, p. 286),

$$Z^* \cong WG, \quad (1)$$

where Z^* is the set of n by n cross-product arrays obtained from m subjects, W is the set of m subject weight vectors with r elements each, and G is the n^2 by r Kronecker product form of the group stimulus space.

When Equation 1 is reduced to the single-subject case and solved for W by postmultiplication with the right pseudoinverse of G , the result is:

$$\hat{W} = Z[G(G^T G)^{-1}], \quad (2)$$

which is the equation solved by the 3S procedure. For simplicity, the term $G(G^T G)^{-1}$ is denoted G^* in the discussions below.

If the dissimilarity data used for a single subject in Equation 2 are originally used as one of the subjects on the INDSCAL analysis that produced the G^* in Equation 2, then the weight vector will be exactly that given by INDSCAL for that subject. If the subject's data are not included in the original analysis, then the weights vector given by Equation 2 is an approximate estimate of that subject's weights if the data had been included in the INDSCAL analysis.

Note that 3S does not derive a stimulus space for the subject; rather, 3S determines a vector of subject weights, which could then be used to operate on the group stimulus space to derive a model for that subject's private space. However, the weights vector is a summary index that is sufficient to classify the individual with respect to the SO in the subject weight space and the group stimulus space. Hence, computation of the private stimulus space for the subject is not necessary.

INITIALIZATION OF 3S

The 3S program must be initialized for a particular application. A G^* array must be computed from the preexisting INDSCAL group stimulus space, and the renormalization constants vector must be derived from the subject weights space.

The G^* array is calculated as follows: (1) Formation of the Kronecker products array G , given a group stimulus array, S , of n stimuli by r dimension, is given by the formula for each element in G , $G_{ij} = s_{pj} s_{qj}$, for all p and q from 1 to n , for all j from 1 to n^2 . (2) Next, the G array is multiplied by its transpose, which gives an r by r array. (3) Next, the matrix inverse of the previous result is determined. (4) Finally, the previous r by r array is multiplied into the array G to give the G^* array. The G^* entries are rounded to the number of significant digits to be used, and the array is then loaded into the G^* registers.

The renormalization constants are calculated from the INDSCAL subjects space. The mean subject weight for each dimension is the constant for that dimension. These constants are used by the 3S program to divide the weights initially calculated. The 3S procedure applies these renormalization constants after calculating the subject weights. The renormalization step could be performed in the preliminary initialization phase. However, the present method allows 3S to accept a subject's raw data from the original INDSCAL solution to check the result from 3S against the result given by INDSCAL itself. This is useful to verify the integrity of the 3S program and the correctness of initialization computations.

SUBJECT WEIGHTS INTERPRETATION

Meaningful interpretation of the INDSCAL subject weights must be limited to ratios of weights on any two dimensions or to sets of such ratios in higher order solutions. This follows from the fact that the essential

information in the subject weight space is in the directions of the subject weight vectors from the origin to each subject locus.

The renormalization process used by 3S places the SO on the diagonal. If that direction is taken as a zero reference vector, then the subject weights resulting from 3S can be expressed as a signed angle between the subject's own vector and that of the SO. Then, in the two-dimensional case, a single number fully describes the locus of the subject being scaled in relation to the SO. Use of the angle index for subject weights discourages errors of interpretation by persons who are not sensitive to the vector nature of the INDSCAL subject weight space.

PROGRAM DESCRIPTION

Phase 1

Data acquisition is mediated by an interactive program. The calculator has a 12-character alphanumeric display for a 24-character display register; thus, there are verbal prompts to the user by the program. The program prompts with the next item number, displays the response, invites a change, and either asks the current item again or proceeds to the next item.

Phase 2

The raw data may require that some value be added to all observations to place the data onto a ratio scale. The method used is one of several given by Torgerson (1958). All possible triples of data are considered. The three possible values for the constant that are computable in each triple are determined as follows: $C_i = x_i - x_j - x_k$, $C_j = x_j - x_i - x_k$, and $C_k = x_k - x_i - x_j$. The largest C over all possible triples is taken as the additive constant. This method is the same one used in INDSCAL. Calculation of the additive constant is optional and may be excluded with some minor reduction in precision.

Phase 3

The data matrix is converted into a cross-products matrix, denoted B^* , which is determined by the method given by Torgerson (1958), as follows:

$$b_{ij} = 1/2n(1/n \sum_i^n d^2_{ij} + 1/n \sum_j^n d^2_{ij} - 1/n^2 \sum_i^n \sum_j^n d^2_{ij} - d^2_{ij}), \quad (3)$$

where b_{ij} is a cross product and d_{ij} is a distance in Row i , Column j of the B^* and distance matrices, respectively. The B^* matrix is then normalized so that the total sum of squared entries equals one.

Phase 4

The weights are calculated by operating on B^* with the G^* matrix. The matrix B^* is treated as a row vector of the matrix rows. The G^* array is cross multiplied,

column by column, with B^* ; these products are summed in each column to yield the INDSCAL weights. The subject weights define a subject vector, the length of which is approximately equal to the proportion of variance in common between the standard observer object space and the individual's object space. This is true only if the scale used at this point is in terms of the original INDSCAL object space. Next, renormalization factors are applied to the subject weights. Then, the angles in the coordinate axis planes between the bisector and the individual are calculated.

Phase 5

The entire results are printed in this phase. A set of response summary data is included, which relates the mean, standard deviation, and frequency distribution of the responses. These parameters may be evaluated to determine if the individual's response patterns are not radically atypical. The response frequency data are printed graphically as an interval histogram. The subject weight angle and proportion of common variance are then printed.

HARDWARE CONFIGURATION

A Hewlett-Packard Model 41C was used, with accessories as follows: three memory modules, a printer, and a magnetic card reader. There are 256 memory registers available, of which 147 were allocated to program and 109 to data storage. The program space is equivalent to 1,000 bytes, or approximately 300 to 400 instructions.

Memory storage requirements were met with five partitions: the response sorting key list, 16 registers; the response data and cross-products array, 28 registers; the standard observer G^* array, 28 registers; a response frequency vector, 9 registers; and a set of miscellaneous data, 28 registers. Memory for data was carefully allocated. Memory requirements were minimized by using special techniques of data storage and accepting constraints in the calculations.

Data packing is a technique for reducing memory utilization that involves placing multiple values into a single data register. Savings of 50% to 80% are possible, but the unpacking routines use additional program space and increase calculation time. The response key list was packed to a size of 16% of the unpacked size, and the G^* matrix was packed to 50%. The method was restricted to the symmetrical case so that the data and G^* matrices could be additionally reduced. The present algorithm accepts 84 judgments on seven objects judged twice in both orders of presentation. The G^* array allows only two weights on two dimensions to be computed.

The G^* array was packed two values per register with four significant digits and sign per value; thus each weight is based on G^* elements resolved to 1 part in 10,000. If three values of a G^* array are stored in each register, the present program can be easily extended to three dimensions. However, the precision of the G^*

elements would be reduced to 1 part in 100, since each element would have only two significant digits. The standard deviation of subject weights determined by 3S has been .05 to .2; therefore, the reduced precision for a three-dimensional G^* may be negligible.

The present program capabilities can be substantially enlarged at the expense of increased operational complexity. The present hardware configuration could be increased by one 64-register memory module. Non-essential features of the program could be removed, and the data acquisition and weight calculation program segments could be loaded and executed separately. If these steps are included, then the 3S procedure may be extended up to six dimensions and up to 16 stimuli.

SOLUTION PRECISION

The 3S method in the HP41C was used to calculate the subject weights for subjects who were in the initial analysis that was used to generate the G^* matrix. The 3S values were exactly equal to the original INDSCAL values, which were output to three significant digits. This degree of precision was expected, since the G^* matrix entries were expressed to four significant digits. If the G^* matrix is packed three values per register, then the entries will contain only two significant digits. However, due to the relatively large coefficient of variation for INDSCAL weights, this reduced precision should not degrade the interpretability of the 3S results.

DISCUSSION

The Human-Machine Interface

Calculator keyboards can be intimidating to the unsophisticated user. Typically, there are many keys, which tend to be small. The HP41C will accept keyboard overlays that are intended to relabel keys that have function assignments that are different from those printed on the permanent keyboard labels. In addition, the keyboard overlay could be fashioned to physically cover all unused keys in a particular application. Effective key size can be increased by assigning the same function to two or more adjacent keys. These hardware and software features allow a highly customized interface for a wide range of applications that should lead to excellent human-machine interactions.

The HP41C provides features that permit the user to tailor the keyboard to the specific task at hand. Any key can be assigned to any function, either an existing hardware function in the machine architecture or a function that has been programmed. The function that will be performed by a key when it is pressed is displayed when the key is pressed. That function is not performed if the key is released more than .5 sec after the keypress begins. Therefore, each key can be previewed by the user to determine its current function assignment.

Significance of the Method

The INDSCAL algorithm provides an object space that contains coordinate axis locations with psychological significance. The model assumes that the observers are using one or more orthogonal dimensions for discrimination of the objects (or may be described as though they were). Further, the model assumes that a major source of variance can be described by the apparent weight each observer places on each dimension. The resultant object space is descriptive only of the data upon which it is based. Nonetheless, there is no reason to reject the notion that INDSCAL solutions are good first approximations to some generalizable perceptual model. Moreover, comparisons to maximum likelihood derived models support this assumption (Ramsay, 1977). Also, comparison of INDSCAL object spaces with perceptual models known a priori illustrates the generality of INDSCAL models (Wish & Carroll, 1974). In summary, while the method described here may be regarded as unorthodox, it can be safely used, if applied with caution. The propriety of the technique may be decided on an empirical basis for each application, in spite of the tenuous theoretical basis of this method.

In psychological laboratories, subjects are generally assigned to experimental conditions at random, and within conditions, subjects are assumed to be essentially equivalent. However, individuals may interact differently with an experimental variable due to systematic individual differences. The result could be inflated within-groups variance and attendant loss of statistical power. Assessment of these individual differences and assignment to experimental conditions based on these differences are often neglected due to the excessive expense of such procedures. The present method can be included in research designs with a modest increase in time and expense.

This method may be of exceptional value when major sources of variance arise from a priori subject properties, such as in studies of semantics, values, opinions, and other social psychology phenomena. In studies of learning and memory, the similarity structure of the stimuli used may be an important covariable, and individual differences assessment of perception of the stimulus set can lead to more precise views of the phenomena under study.

Experimental psychologists regularly manipulate the physical parameters of a stimulus and study the corresponding variance in behavior. It is assumed that persons who are exposed to the same physical stimulus also respond to the same functionally effective internal stimulus. Likewise, it is assumed that different physical stimuli evoke functional stimuli that are different. If these assumptions do not hold, then the variances within and between treatments may increase due to these systematic errors. A 3S pretest can establish the relationship between the physical stimulus domain and that of stimuli as "perceived" by each subject. Then the

experimenter may utilize this scaling in several ways. First, the weight index on each dimension of perception could be used to adjust physical stimulus parameters so that subjects in each treatment share a similar functional stimulus. Alternatively, the experimenter may use the weight index as a covariate and gain control by statistical adjustment. Another approach would be to admit subjects into the study based on the 3S weights so that all subjects in all conditions share a common perceptual structure. The choice of which of these to use would depend on the particulars of a given experiment. Whatever the choice, the 3S pretest offers the opportunity to remove some systematic variance from the error terms. This variance may be excluded or accounted for, but the primary benefit is that previously undetectable effects may be exposed.

In clinical settings, important judgments must be made within the time span of a particular client/clinician interaction. The more sophisticated techniques of MDS probably have not penetrated the clinical scene due to the dependence on computer centers that is traditionally associated with these methods. This method illustrates a step toward the use of MDS in clinical settings that is largely free of these constraints.

It is feasible for a clinician to assess the structure of a client's perception of that client's significant social network in multidimensional terms and to relate that client to some normative group in quantitative terms, all within the time of a single office session. The client could be asked to judge similarity of pairs composed of members of the immediate family, close personal friends, and other significant figures such as superiors at work, leaders of institutions such as church, and the like. These similarities on, say, 68 pairs of 17 persons can be processed by an extended version of 3S to yield weights on four dimensions, which describe how the client perceives his immediate social environment.¹ Note, that the clinician may choose the person-stimuli to suit the particular client. Also, the client is given an unconstrained task that does not impose potentially distorting task demands. Such an assessment could be done in 1 h from onset of instructions to interpretable results.

MDS and related techniques are being raised to a high level of sophistication. However, these methods tend to

remain in the hands of specialists with access to large computer center resources. The increasing power of small hand-held calculators and the decreasing cost of such machines remove an important barrier that can lead to diffusion of advance MDS techniques into the clinical scene.

REFERENCES

- CARROLL, J. D., & CHANG, J. J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 1970, **35**, 283-319.
- HORAN, C. B. Multidimensional scaling: Combining observations when individuals have different perceptual structures. *Psychometrika*, 1969, **34**, 139-165.
- KRUSKAL, J. B., & WISH, M. *Multidimensional scaling*. Sage University Series on Quantitative Applications in Social Sciences, 07-011, Beverly Hills, Calif: Sage Publications, 1978.
- RAMSAY, J. O. Maximum likelihood estimation in multidimensional scaling. *Psychometrika*, 1977, **42**, 241-266.
- RAMSAY, J. O. Some small sample results for maximum likelihood estimation in multidimensional scaling. *Psychometrika*, 1980, **45**, 139-144.
- SHERMAN, C. R. Nonmetric multidimensional scaling: A Monte Carlo study of the basic parameters. *Psychometrika*, 1972, **37**, 323-355.
- SPENCE, I., & DOMONEY, D. W. Single subject incomplete designs for nonmetric multidimensional scaling. *Psychometrika*, 1974, **39**, 469-490.
- TORGERSON, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.
- WISH, M., & CARROLL, J. D. Applications of individual differences scaling to studies of human perception and judgment. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2), *Psychophysical judgment and measurement*. New York: Academic Press, 1974.

NOTE

1. The dimensions that explain the typical person's similarity judgments would have been previously identified by a scaling study designed to construct the stimulus space model for person perception. The stimulus structure for a particular test would be derived from this model. If particular dimensions of person perception are of interest, then a specific subset of stimuli that are most salient on that dimension could be selected and a G* for that subset could be used to initialize the 3S method.

(Received for publication September 13, 1980;
revision accepted February 17, 1981.)