Seriation and Multidimensional Scaling: A Data Analysis Approach to Scaling Asymmetric Proximity Matrices

Joseph Lee Rodgers and Tony D. Thompson
University of Oklahoma

A number of model-based scaling methods have been developed that apply to asymmetric proximity matrices. A flexible data analysis approach is proposed that combines two psychometric procedures—seriation and multidimensional scaling (MDS). The method uses seriation to define an empirical ordering of the stimuli, and then uses MDS to scale the two separate triangles of the proximity matrix defined by this ordering. The MDS solution contains directed distances, which define an "extra" dimension that would not otherwise be portrayed, because the dimension comes from relations between the two triangles rather than within triangles. The method is particularly appropriate for the analysis of proximities containing temporal information. A major difficulty is the computational intensity of existing seriation algorithms, which is handled by defining a nonmetric seriation algorithm that requires only one complete iteration. The procedure is illustrated using a matrix of co-citations between recent presidents of the Psychometric Society. *Index terms: asymmetric data, cluster analysis, combinatorial data analysis, multidimensional scaling, order analysis, proximity data, seriation, unidimensional scaling.*

This paper is concerned with data that are asymmetric proximities indicating relationships between pairs of objects and with an analytic procedure that can reveal information about the processes that generated the data. The data analysis model is a method called seriation, and it is used in combination with multidimensional scaling (MDS).

Within the exploratory data analysis (EDA) paradigm proposed by Tukey (1977), searching for structure in a dataset is a natural early step in the development of theories. EDA includes the repeated fitting of mathematical models to data, the use of procedures not especially influenced by one or a few extreme data points, and the visual representation of data through graphical procedures. The procedure developed here is consistent with these EDA tenets.

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 16, No. 2, June 1992, pp. 105-117
© Copyright 1992 Applied Psychological Measurement Inc.
0146-6216/92/020105-13$1.90

Models of Asymmetric Relationships

Models that portray relationships between stimuli as distances or angles within metric spaces (e.g., MDS, factor analysis) represent relationships between stimuli symmetrically. But many relationships are inherently asymmetric (Tversky, 1977). A number of approaches to modeling asymmetric proximities have been proposed in the psychometric literature. These can be classified into three categories.

In the first category, the symmetric portion of the data is extracted, and the asymmetric portion is discarded as error. A common approach of this type is to average corresponding off-diagonal elements (a symmetric least-squares fit to the original asymmetric matrix) and apply symmetric models. For example, Tversky and Hutchinson (1986) analyzed 39 asymmetric proximity matrices by averaging. Alternatively, Levin and Brown (1979) defined row multipliers that produced a least squares fit between the two triangles of the asymmetric matrix.

The second approach—a more sophisticated treatment of asymmetries—involves explicitly modeling the asymmetries in addition to a sym-

Downloaded from SHAREOK. Persistent Handle URL : https://hdl.handle.net/11244/24978
Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

metric component. Young (1975) defined a Euclidean distance model modified by multiplicative asymmetric weights: A spatial representation shows the symmetric structure, and dimensional weights for each stimulus show how to expand or contract the space to account for directed distances from that stimulus. Krumhansl (1978) specified a distance-density model in which objects A and B are represented spatially, and asymmetries are accounted for through the density of other object points around A and B. Saito (1986) developed an MDS approach in which a symmetric configuration is modified by estimated "object constants." When these constants are constrained to be nonpositive, they can be considered "density constants" in relation to the distance-density model. Constantine and Gower (1978) proposed additively partitioning a square asymmetric matrix **Q** into two matrices—**S** (symmetric) and **A** (skew-symmetric; i.e., $a_{i,j} = -a_{j,i}$). They performed a singular value decomposition of **A** to obtain a least-squares fit to be plotted in two dimensions and provided an interpretation of this spatial representation of the asymmetries. Constantine and Gower also proposed fitting an unfolding model to an asymmetric matrix, which represents both row and column stimuli within the same space. Holman (1979) used a nonmetric procedure to decompose asymmetric proximities into a symmetric function and a bias function. Finally, Weeks and Bentler (1982) defined a model in which the symmetric part of a relationship is modeled as a distance, and additive constants are combined with the distance to reflect asymmetries.

The third category of approaches for modeling asymmetric proximities involves network models in which asymmetries are represented nonspatially as distances. These methods rely on directed distances within a graph-theoretic context to model asymmetries. Models in this category might be called "purely asymmetric," because no underlying symmetry need be posited. DeSarbo, Manrai, and Burke (1990) proposed a nonspatial version of the distance-density model. A number of other such approaches have been developed, including those of Cunningham (1978), DeSarbo (1982), Hutchinson (1989), and Klauer (1989).

The method developed here shares features of the second and third categories above. Visual interpretation of distances is incorporated because stimuli are represented in a Euclidean space. Distances are directed, however, and the presentation of the configuration indicates order relationships. The order relationships may be substantively meaningful, so that an added dimension of information can be represented within the space. This approach is particularly valuable if temporal relationships underlie the asymmetries in the data. Note that none of the above methods has the capability to portray both a dominance order and a proximity structure.

The conceptual basis of this approach is the following. The stimuli in an $N \times N$ asymmetric proximity matrix can be reordered to define $N!$ unique triangles of proximities. Each triangle may be fit using symmetric models; however, the proximities are directed, depending on the particular ordering of the stimuli. Thus, an exhaustive asymmetric model could be defined in which all $N!$ triangles of directed proximities are represented spatially using MDS, with the order relationships specified in the MDS solution. However, with even medium-sized proximity matrices, performing $N!$ MDS analyses would be practically impossible. The approach presented here selects an empirically interesting subset (one or two) of the $N!$ possible triangles. Seriation optimizes an objective function across the $N!$ reorderings of the N stimuli, and provides a natural empirical method to select an ordering of stimuli. Thus, this approach combines MDS, which uses information contained within triangles of a proximity matrix, and seriation, which uses information contained in the relationship between triangles. Both types of information can be displayed in the same graphical representation for substantive interpretation. Gower (1977, p. 111) anticipated this approach in his short description of his "Method 3."

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Seriation of Asymmetric Data

Background

Seriation is a procedure that orders stimuli according to structure within a proximity matrix. The method was developed within electrical engineering, graph theory, and archeology. Seriation was first presented in the psychometric literature by Hubert (1974, 1976; Baker & Hubert, 1977). He asserted: "One of the basic problems of data analysis that has concerned applied researchers for many years deals with the sequencing of objects along a continuum" (Hubert, 1974, p. 9). Relationships have been noted between seriation and other psychometric methods, including unfolding (Defays, 1978; Heiser, 1981; Hubert & Golledge, 1981), MDS and correspondence analysis (Mardia, Kent, & Bibby, 1979), and cluster analysis and Thurstone's Law of Comparative Judgment (Hubert & Golledge, 1981). Seriation methods can order stimuli in both symmetric and asymmetric matrices—the focus here is on the latter. The method presented applies specifically to square (one-mode) asymmetric proximity matrices. For seriation treatments of rectangular data matrices, see Cliff, Collins, Zatkin, Gallipeau, and McCormick (1988) or Coombs and Smith (1973).

A formal mathematical statement of the seriation problem can be found in Hubert (1976). The description here is conceptual. Begin with a square asymmetric matrix of proximities between stimuli. Assume that the stimuli are arbitrarily ordered and that order relationships are transitive. Table 1 shows a 6 × 6 matrix of proximities. If the stimulus order is arbitrary, any other ordering produces a different matrix with the same information. Table 2 shows the asymmetric matrix associated with another of the 6! (720) different orderings of the six stimuli. Note, for example, that the 2,1 element of the matrix in Table 2 is the same element as the 4,5 element in Table 1. Each ordering produces a different upper and lower triangle. Reverse orderings (e.g., 6-5-4-3-2-1 versus 1-2-3-4-5-6) produce identical triangular structures on opposite sides of the main diagonal. Clearly, the structure contained in Table 2 is relatively hidden in Table 1.

Table 1
Proximity Matrix Indicating Proximity Between All Pairs of Six Stimuli

	Stimulus					
Stimulus	1	2	3	4	5	6
1		0	0	2	6	0
2	12		9	14	15	5
3	3	0		7	10	0
4	0	0	0		1	0
5	0	0	0	0		0
6	8	0	4	11	13	

Table 2
The Same Proximity Matrix as in Table 1, With Rows and Columns Permuted to Show Greater Structure

	Stimulus					
Stimulus	5	4	1	3	6	2
5		0	0	0	0	0
4	1		0	0	0	0
1	6	2		0	0	0
3	10	7	3		0	0
6	13	11	8	4		0
2	15	14	12	9	5	

Table 3 uses real data—the win/loss records of the top ten women tennis players in 1989 (Flink, 1990). Players are ordered alphabetically. Table 4 depicts one of the 10! reorderings of the same data, obtained with the seriation algorithm developed here. Table 4 shows (from bottom to top) an empirically-based rank order of the players based on their pairwise win-loss record. The empirical ranking closely matches the one actually defined by a panel of *World Tennis* experts (Flink, 1990).

In general, the seriation procedure orders stimuli according to some objective. The objective is defined within a mathematical function that is optimized using one of several different procedures. Once the seriation solution is found, the ordering of the stimuli is no longer arbitrary. With asymmetric data, the most popular seriation objective function has been to find the ordering of the stimuli maximizing the sum of the elements within one triangle (say, the lower triangle) which simultaneously minimizes the

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Table 3
1989 Record of the Top Ten Women Tennis Players, With Wins on the Rows and Losses on the Columns

Player and Order	Player									
	1	2	3	4	5	6	7	8	9	10
1. Evert		1	0	0	0	2	0	0	1	0
2. Garrison	1		0	1	0	1	1	0	0	0
3. Graf	2	2		0	3	4	4	3	3	6
4. Maleeva	0	0	0		0	0	0	0	0	1
5. Navratilova	0	4	0	2		1	1	3	2	1
6. Novotna	0	0	0	0	0		0	1	1	0
7. Sabatini	1	2	1	0	1	1		3	0	1
8. Sanchez	0	0	1	1	0	1	1		0	0
9. Seles	1	1	0	3	0	0	0	1		0
10. Sukova	0	0	0	0	1	1	0	0	0	

sum in the other triangle (De Soete, Hubert, & Arabie, 1988; Hubert, 1976). Although this goal is easy to state, it is difficult to achieve except with an exhaustive search of the $N!$ possible orderings (actually, a search of $N!/2$ orderings suffices, because reverse orderings can count as a single search). With even medium-sized matrices, such a search is computationally intractable (e.g., a 10×10 matrix requires a search of $10!/2 = 1{,}814{,}400$ orderings; a 20×20 matrix requires 1.2×10^{18} orderings).

Therefore, computational procedures, such as the locally optimal pairwise exchange (LOPI) procedure (Baker & Hubert, 1977), dynamic programming (Hubert & Arabie, 1986; Hubert & Golledge, 1981), and simulated annealing (De Soete et al., 1988), have been applied to reduce the size of the search. Although these procedures substantially reduce computational labor compared to an exhaustive search, they are still computationally-intense methods. For example, dynamic programming requires at most 2^N evaluations of the objective function (which is smaller than $N!/2$ for all $N > 4$). Nevertheless, 2^N still increases exponentially, and precludes analysis of very large datasets.

A simpler optimization algorithm is presented here that works even with large datasets. Many data analytic settings support sacrificing the single globally optimal solution in exchange for faster convergence, as long as the resulting ordering is close to optimal. The seriation method presented here requires only simple programming and a single complete iteration through the stimuli. Thus, it provides quick solutions with even very large proximity matrices. If more

Table 4
Seriated Ordering of the 1989 Top Ten Women Tennis Players

Seriated Order	Seriated Order									
	6	4	8	2	10	1	9	5	7	3
6. Novotna		0	1	0	0	0	1	0	0	0
4. Maleeva	0		0	0	1	0	0	0	0	0
8. Sanchez	1	1		0	0	0	0	0	1	1
2. Garrison	1	1	0		0	1	0	0	1	0
10. Sukova	1	0	0	0		0	0	1	0	0
1. Evert	2	0	0	1	0		1	0	0	0
9. Seles	0	3	1	1	0	1		0	0	0
5. Navratilova	1	2	3	4	1	0	2		1	0
7. Sabatini	1	0	3	2	1	1	0	1		1
3. Graf	4	0	3	2	6	2	3	3	4	

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

precision is required and/or if the proximity matrix is small, then LOPI, dynamic programming, or simulated annealing would be preferred.

A Simple Nonmetric Seriation Procedure

The seriation algorithm presented here is a slight adaptation of a nonmetric counting rule suggested by Hubert and Golledge (1981) and has relationships to early procedures proposed by Kendall (1955) and Flueck and Korsh (1974). From all possible $N!$ orderings of the N stimuli defining a proximity matrix $\mathbf{Q}$, the objective is to order the stimuli according to "number of victories" over the other stimuli. A "victory" for stimulus I over stimulus J occurs when $Q_{i,j} > Q_{j,i}$. To find this ordering, construct a competition matrix from the original $\mathbf{Q}$ matrix by comparing corresponding off-diagonal elements. The competition matrix $\mathbf{T}$ is defined as:

$$T_{ij} = \begin{cases} 1 \text{ if } Q_{ij} > Q_{ji} \\ .5 \text{ if } Q_{ij} = Q_{ji} \\ 0 \text{ if } Q_{ij} < Q_{ji} \end{cases} . \qquad (1)$$

Next, the stimuli are ordered by sorting on the row averages—the r_is—(or, equivalently, the sums) in $\mathbf{T}$. Thus, stimuli that consistently dominate other stimuli are placed lower in the ordering. This is a nonmetric procedure for maximizing the sum in the lower triangle of $\mathbf{T}$. Because only the size comparisons and not the size of the elements are accounted for, one or a few extreme proximities do not have an undue effect on the solution.

The remaining problem involves resolving ties in the row averages. Baker and Hubert (1977) recommended using the Younger A criterion (Younger, 1963) in which the sizes of the original elements in $\mathbf{Q}$ are used (metrically) to resolve ties. Three-way and four-way ties are resolved iteratively by passing through the tied elements several times until larger entries in $\mathbf{Q}$ are in the lower triangle.

Some Seriation Statistics

This formulation supports the development of the following "seriation statistics" designed to provide diagnostic information to extend the seriation results:

1. ASY is a measure of the asymmetry of a proximity matrix (cf. Hubert & Baker, 1979);
2. TIES counts the number of tied row averages in $\mathbf{T}$;
3. The d_is are domination statistics, one for each stimulus; and
4. FIT is a measure of how close the ordering comes to finding perfect seriation structure.

Because the information the procedure uses to order stimuli is contained within the asymmetries of the matrix, ASY can suggest whether using the procedure is appropriate. (If not, the stimuli can be ordered on other grounds; e.g., Defays, 1978; Hubert, 1974.) ASY is a special type of variance based on absolute deviations:

$$\text{ASY} = \frac{\sum\sum_{i<j} |T_{ij} - .5|}{.25N(N-1)} . \qquad (2)$$

First, note that the average of off-diagonal entries in the indicator matrix $\mathbf{T}$ must equal .5. Further, ASY is invariant with respect to whether it is computed in the upper or lower triangle. In a symmetric matrix, all entries in $\mathbf{T}$ will be .5 (because each competition ends up tied), and ASY = 0. In a completely asymmetric matrix (where none of the corresponding off-diagonal entries are the same), all $[N(N-1)/2]$ entries in each triangle will be 1s or 0s. In such a matrix, ASY $= .5N(N-1).5/[N(N-1).25] = 1.0$. Thus, ASY varies between 0 for perfectly symmetric and 1.0 for perfectly asymmetric matrices. ASY is equal to the proportion of nontied indicators (0s or 1s) in either triangle of $\mathbf{T}$. ASY uses a nonmetric definition of asymmetry—corresponding off-diagonal elements are simply counted as equal or not. The size of the difference could be accounted for through a similar approach, but the statistic would have indeterminate bounds.

TIES counts the proportion of ties in the row means. This is a measure of how much metric information is used in the optimization criterion. Because the first criterion used orders the stimuli

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

on increasing row means, tied r_is suggest ambiguity in this ordering (which is, if possible, resolved by the Younger A criterion discussed above). Thus, TIES is defined as follows:

$$\text{TIES} = (\Sigma K_i)/(N - 1) \quad , \tag{3}$$

where

$$K_i = \begin{cases} 1, \text{ if } r_i = r_{i-1} \\ 0, \text{ otherwise,} \end{cases} \tag{4}$$

for each row r, $r > 1$. When TIES = 0, there are no tied row means, and the ordering of the stimuli is unambiguous. When TIES = 1, all the row means are tied, and the resolution by the Younger A criterion makes the seriation procedure a completely metric procedure. If the data are binary, and TIES is close to 1.0, there is no information in the matrix on which to order the objects, and the ordering will be arbitrary.

The domination statistics, d_is, are derived from the r_is. The average of all row means r_i must equal .5. Thus, $r_i > .5$ indicates stimuli that have dominated more than an expected number of other stimuli under an assumption of uniform domination; stimuli with $r_i < .5$ have dominated fewer than expected. The domination statistic, d_i, is defined as $d_i = 2(r_i - .5)$ for each of the i rows. If the ith stimulus dominated none of the other stimuli, $d_i = -1$; if it dominated all other stimuli, $d_i = 1$. If it dominated exactly half of the other stimuli, $d_i = 0$. Thus, d_i can range from -1 to +1: Like the correlation coefficient, the sign of d_i indicates the direction of domination, and the size of d_i indicates the strength of domination.

Finally, FIT measures how close the seriation algorithm comes to finding perfect seriation structure in relation to the objective function. Such structure occurs when lower objects in the ordering have victories over those higher in the ordering, but not over those even lower; or, equivalently, when a redefined indicator matrix corresponding to the final solution has all the 1s below the diagonal and only 0s above. FIT can be defined by using the seriation-defined order of the stimuli, and then slightly redefining the indicator matrix **T**. In TFIN (the final competition matrix), if $P_{ij} = P_{ji}$, then instead of assigning a value of .5, assign a value of 0; thus, only entries in **Q** that literally dominate are given a value of 1.0. Then, FIT is defined as the proportion of total competitions that produce winners in the lower triangle of TFIN:

$$\text{FIT} = \frac{\sum\sum_{i>j} \text{TFIN}_{ij}}{\sum_i \sum_j \text{TFIN}_{ij}} \quad . \tag{5}$$

FIT = 1.0 if every 1.0 element in TFIN is in the lower triangle; FIT = .5 if exactly half of the 1.0 entries are in the lower triangle and half in the upper (as expected by chance in a randomly ordered matrix). Values of FIT should be substantially larger than .5 before concluding that the seriation procedure found a useful and defensible ordering. A low value of FIT for a matrix suggests that seriation is not a proper analytic model to use. Ties do not show up in the computation of FIT, because they are represented neither in the numerator nor the denominator. FIT does not indicate whether a better solution could have been found with a better optimization procedure (e.g., dynamic programming), or whether some other objective criterion could have produced a "better" ordering of the stimuli.

TRIAGE

This seriation procedure and associated seriation statistics are available within a larger system of preprocessing and diagnostic routines called TRIAGE (Rodgers, Thompson, & Thompson, 1989), which is a set of SAS MACROS written in PROC MATRIX in SAS (SAS Institute, Inc., 1985). The seriation procedure is also available in PROC IML coding. (Copies of the coding, documentation, and test runs are available on request.) This procedure has been used successfully to order stimuli from proximity matrices as large as 41 × 41, which is much too large for the other methods listed above to treat successfully. The results obtained from the simple procedure presented here have been compared to those obtained from dynamic programming in the 15 × 15

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

matrix analyzed by Hubert and Golledge (1981, p. 435). The procedure ordered the stimuli exactly the same as dynamic programming. The results also were compared to those obtained by applying LOPI to several constructed datasets (Hubert, 1976, pp. 35, 44). Results were consistently close, although not always identical.

The output of a seriation procedure is an ordering of the stimuli. The ordering of the stimuli may be interpretable and, if so, the procedure has successfully found a unidimensional dominance scale along which the stimuli fall. Also, the proximity matrix then can be ordered, and an MDS model can be fit to one or both triangles.

MDS of the Triangles of the Seriated Proximity Matrix

If the stimuli are ordered according to some interpretable dimension, then the two triangles may contain information that is separately meaningful. For example, suppose information is available about the membership overlap between 41 divisions of the American Psychological Association (APA) (see Rodgers, 1988). The proximity matrix contains $Q_{i,j}$s indicating the proportion of total members of Division i who are also members of Division j. Note that the numerators of these proportions are equal for $Q_{i,j}$ and $Q_{j,i}$, but the denominators are different and reflect the total size of each division. Thus, $Q_{i,j}$ and $Q_{j,i}$ will differ to the extent that the sizes of the divisions are different, and the seriation procedure (which uses the difference between $Q_{i,j}$ and $Q_{j,i}$) should therefore order the divisions by size. Once ordered, proximities in the upper triangle are directed and indicate the specific relationships of large divisions to small divisions. Numbers in the lower triangle are ordered to specify directed relationships of small divisions to large divisions. Separate MDS analyses of the two triangles have these different interpretations. Note that the assumption of transitivity that the seriation procedure imposes on the ordering is reasonable when size (or time) is the dimension on which the stimuli are ordered.

Distances in the MDS solution are directed according to the particular triangle being scaled. This feature can be portrayed within the MDS space by drawing arrows between points indicating the direction of the relationship. In practice, the $N(N-1)/2$ arrows that are drawn considerably confuse the plot in which the stimuli are scaled. To compromise, only the $N-1$ arrows indicating the ordering of the stimuli produced by the seriation can be drawn; then, "visual transitivity" is relied on to fill in other arrows. Alternatively, the stimuli can be represented with numbers or letters indicating the seriated ordering, so that the unidimensional seriation structure—obtained by analysis of the asymmetries—is captured in the MDS representation along with the symmetric proximity representation. This sacrifices the visual impact of arrows forcing the directed distance interpretation of the solution, but produces a more understandable configuration to study.

In certain types of proximity matrices with a large number of zero entries, most of the nonzero entries may be seriated into the lower triangle. Stimuli with time dependencies—in which relationships only exist for stimuli in one temporal direction, for example—can produce such data. A "misordered" proximity matrix can hide such triangular structure (see Figure 1, below). When seriated, the directed nature of the time dependencies becomes apparent. In this case, only one MDS application may be necessary.

Any type of asymmetric proximities (e.g., proportions, frequencies, rates, overlap measures) can be treated with this procedure. Often, theoretical grounds may suggest the dimensions on which a seriation procedure should order the stimuli (as above with the APA data). In cases without theoretical guidance, the seriation solution can be inspected to discover an interpretable structure. Because seriation extracts a unidimensional ordering from the stimuli, multiple dimensions in the asymmetric information can be confounded within the unidimensional seriation solution. Note that there may be theoretically interesting triangles of information other than the two specified by the seriation algorithm presented here. Certainly MDS (or other) analyses may be used on these triangles as well, although optimality properties for which

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

the seriation procedure searches are not features of such an analysis.

The approach presented here is more of an exploratory procedure and less of a specific asymmetric model than the methods described above. This method sacrifices some mathematical elegance, but gains flexibility and tractability, which is justified to the extent that substantive information is more easily obtained by the data analyst. Thissen and Steinberg (1988) demonstrated data analysis applications of item response theory, arguing that "the goal of the analysis of psychological data, however mathematical, is psychology, not mathematics" (p. 385).

Illustration—A Co-Citation Matrix

The Data

Presumably, information about substantive relationships between researchers is contained in the frequency with which they cite one another. An 18 × 18 asymmetric similarity matrix was constructed of co-citation rates of Psychometric Society presidents between 1970 and 1987. Rates were defined by examining all *Psychometrika* articles through 1985 authored or coauthored by these 18 presidents, and dividing the number of times President i cited President j by the number of articles authored by President i (these numbers are therefore interpreted as citations per article). The citation rate matrix is presented in Table 5.

On theoretical grounds, the seriation procedure was expected to order the presidents on a time-related dimension, because asymmetries at least partially reflect the fact that younger presidents had more opportunity to cite older presidents than vice versa. (Other differences between pairs of presidents that might be confounded within the ordering include differential tendencies to cite many papers or important papers; these are expected to be less systematically reflected than the time structure, however, because there is not an obvious unidimensional scale along which these types of differences should fall). Obviously, the particular ordering in Table 5—the chronological ordering of the presidents—already contains time-related structure. Thus, this order is not an arbitrary order of the 18! possible orderings; the matrix already has achieved some of the goal of seriation, in that many larger entries are in the lower triangle and smaller entries are in the upper triangle.

Results

Table 6 presents the matrix from Table 5 after seriation was applied. ASY = .72, indicating substantial asymmetry. The value of TIES was .24; the Younger A criterion had to resolve two, two-way ties and one, three-way tie. From the 151 informative (nonzero) entries, 109 (71%) are in the lower triangle. FIT = .85, indicating that 85% of the active "competitions" were reflected in the derived ordering. Domination statistics d_i and row averages r_i also are shown in columns of Table 6.

The seriation of the presidents appears to give an empirical time order, one related to "psychometric age" of the president. The seriation ordering and the chronological order had a Spearman correlation of $r = .73$. Of those who moved at least five places between the two orderings, Luce, McDonald, and Novick "became older" in the empirical ordering, and Messick "became younger." Luce may have moved for substantive rather than time-related citation behavior, however. Because his work treated areas other than scaling, test theory, factor analysis, and structural equations modeling, he would be expected to cite others less often in this group, almost all of whom are identified with one or more of these areas. Such substantive structure can be extracted from the information in Table 6 by explicit analysis of the proximities.

Each triangle in Table 6 was submitted to an MDS, with an empirical justification for analysis of these particular triangles (out of the 18! possible orderings). If arrows are drawn on top of the MDS solution indicating "who is citing whom," the two solutions give a natural asymmetric representation of the citations data. The stimuli also may be identified by numbers representing the seriated ordering, which is a preferable technique for graphical purposes.

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Table 5
Co-Citation Rates of Psychometric Society Presidents From 1970–1986, With Citations "By" on Rows and Citations "Of" on Columns

President and Order	Order																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Guttman		0.0	0.0	.1	.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2. Messick	.1		.2	.5	0.0	0.0	.1	0.0	.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3. Bock	.7	.1		.1	0.0	0.0	.3	.1	0.0	.1	0.0	.1	0.0	0.0	0.0	0.0	0.0	0.0
4. Shepard	.7	.7	0.0		1.7	1.5	.2	0.0	0.0	0.0	.3	0.0	0.0	0.0	0.0	.3	0.0	0.0
5. Kruskal	.2	0.0	0.0	1.8		1.7	0.0	0.0	0.0	0.0	.2	0.0	.2	0.0	.2	.3	0.0	0.0
6. Carroll	.1	0.0	.2	1.1	1.5		.1	0.0	.1	0.0	.5	1.2	.1	.2	.2	0.0	.8	.3
7. Luce	0.0	0.0	.2	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8. Jöreskog	1.0	0.0	.2	0.0	0.0	0.0	0.0		0.0	.1	0.0	0.0	0.0	0.0	0.0	.2	0.0	0.0
9. Cliff	.4	0.0	0.0	.3	.3	.1	0.0	.2		.1	.6	.1	0.0	.1	0.0	0.0	0.0	0.0
10. Novick	.2	0.0	0.0	0.0	0.0	0.0	.5	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11. Young	.7	.1	.3	.9	1.3	.9	.1	.2	.2	0.0		.1	0.0	0.0	.1	.1	1.8	2.1
12. Ramsay	.1	.1	.1	.2	.5	.4	.1	.2	.1	.2	.2		0.0	0.0	.1	0.0	.2	.2
13. Bentler	.3	0.0	.1	0.0	0.0	.1	0.0	1.9	.3	.1	.2	.1		0.0	.3	.9	.1	.1
14. Hubert	0.0	0.0	0.0	.4	.2	.1	.3	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0
15. Bloxom	.2	0.0	.2	0.0	0.0	.2	.6	.5	.1	.1	0.0	.1	.1	0.0		0.0	.1	0.0
16. McDonald	1.3	0.0	.1	.1	.1	.1	0.0	.6	0.0	.1	.1	0.0	.1	0.0	0.0		0.0	.1
17. Takane	.3	.2	.2	1.3	1.8	1.5	.4	.2	.3	.2	3.5	1.7	.2	0.0	.2	0.0		2.0
18. De Leeuw	.1	.1	0.0	.5	1.4	.9	.1	.2	.1	0.0	2.6	.1	.1	.1	.1	.1	1.4	

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Table 6
Seriation Results From Psychometric Society Presidents' Co-Citations

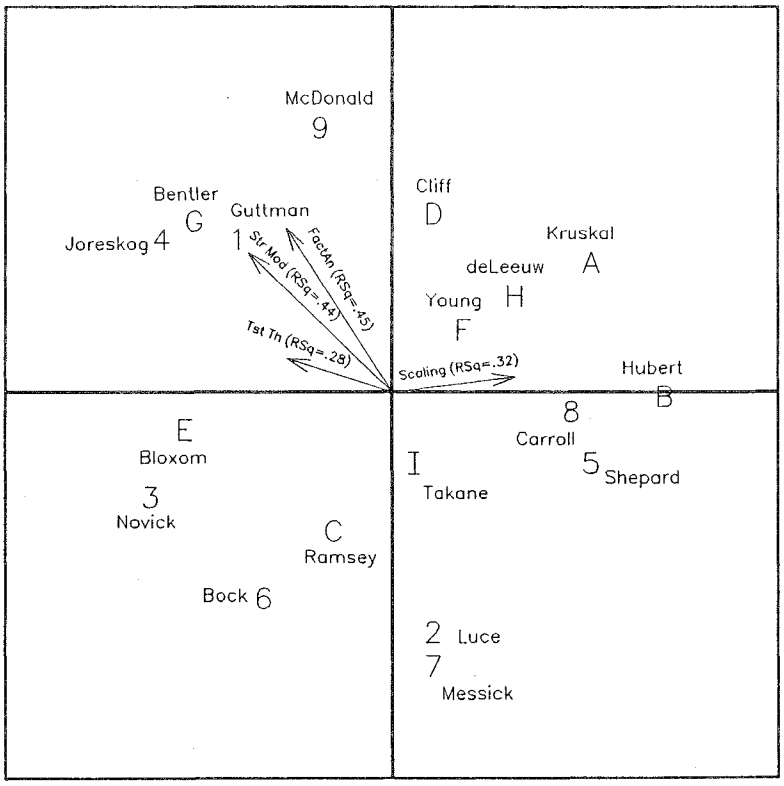
Seriated Order	Seriated Order																		r_i	d_i
	1	7	10	8	4	3	2	6	16	5	14	9	15	12	11	13	18	17		
1. Guttman		0.0	0.0	0.0	.1	0.0	0.0	0.0	0.0	.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.06	-.88
7. Luce	0.0		0.0	0.0	0.0	.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.17	-.66
10. Novick	.2	.5		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.31	-.38
8. Jöreskog	1.0	0.0	.1		0.0	.2	0.0	0.0	.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.33	-.34
4. Shepard	.7	.2	0.0	0.0		0.0	.7	1.5	.3	1.7	0.0	0.0	0.0	0.0	.3	0.0	0.0	0.0	.39	-.22
3. Bock	.7	.3	.1	.1	.1		.1	0.0	0.0	0.0	0.0	0.0	0.0	.1	0.0	0.0	0.0	0.0	.39	-.22
2. Messick	.1	.1	0.0	0.0	.5	.2		0.0	0.0	0.0	0.0	.1	0.0	0.0	0.0	0.0	0.0	0.0	.44	-.12
6. Carroll	.1	.1	0.0	0.0	1.1	.2	0.0		0.0	1.5	.2	.1	.2	1.2	.5	.1	.3	.8	.44	-.12
16. McDonald	1.3	0.0	.1	.6	.1	.1	0.0	.1		.1	0.0	0.0	0.0	0.0	.1	.1	.1	0.0	.47	-.06
5. Kruskal	.2	0.0	0.0	0.0	1.8	0.0	0.0	1.7	.3		0.0	0.0	.2	0.0	.2	.2	0.0	0.0	.47	-.06
14. Hubert	0.0	.3	0.0	0.0	.4	0.0	0.0	.1	0.0	.2		0.0	0.0	0.0	0.0	0.0	0.0	0.0	.47	-.06
9. Cliff	.4	0.0	.1	.2	.3	0.0	0.0	.1	0.0	.3	.1		0.0	.1	.6	0.0	0.0	0.0	.53	.06
15. Bloxom	.2	.6	.1	.5	0.0	.2	0.0	.2	0.0	0.0	0.0	.1		.1	0.0	.1	0.0	.1	.56	.11
12. Ramsay	.1	.1	.2	.2	.2	.1	.1	.4	0.0	.5	0.0	.1	.1		.2	0.0	.2	.2	.58	.17
11. Young	.7	.1	0.0	.2	.9	.3	.1	.9	.1	1.3	0.0	.2	.1	.1		0.0	2.1	1.8	.61	.22
13. Bentler	.3	0.0	.1	1.9	0.0	.1	0.0	.1	.9	0.0	0.0	.3	.3	.1	.2		.1	.1	.67	.33
18. De Leeuw	.1	.1	0.0	.2	.5	0.0	.1	.9	.1	1.4	.1	.1	.1	.1	2.6	.1		1.4	.72	.44
17. Takane	.3	.4	.2	.2	1.3	.2	.2	1.5	0.0	1.8	0.0	.3	.2	1.7	3.5	.2	2.0		.89	.78

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Using ALSCAL (SAS Institute, Inc., 1983, p. 1), the classical nonmetric MDS model was fit to both the upper and lower triangle of Table 6. The MDS fit statistics obtained from these two procedures were S-STRESS = .38 and R^2 = .58, and S-STRESS = .51 and R^2 = .23 for the lower triangle and upper triangle, respectively. Only the results from the lower triangle were interpretable and interesting (because in the seriated solution, most of the nonzero entries are in the lower triangle, which explains the poorer fit values for the upper triangle). The two-dimensional solution is shown in Figure 1. Presidents are numbered/lettered according to their seriation-derived order (numerals are continued with letters), increasing with time. Thus, low numbers/letters (e.g., Guttman and Luce) are placed into this space according to how others cited them. Information about their citations (most of which were 0) is contained in the upper triangle, which is not represented in this space. High numbers/letters (De Leeuw and Takane) are placed into the space according to how they cited others. Thus, the higher the stimulus number/letter in Figure 1, the more actively did that president's citation behavior contribute to this space.

Consistent with the EDA tenets, the directions in this space were substantively validated using external data (see, e.g., Kruskal & Wish, 1978, p. 35). The March, 1986, issue of *Psychometrika* published annotated bibliographies for test theory (Lewis, 1986), factor analysis (Mulaik, 1986), structural modeling (Bentler, 1986), and scaling (Torgerson, 1986). In each area, the number of references to each president was tabulated and regressed into the space portrayed in Figure 1. Direction cosines defined the angular orientations, which are represented in the space

Figure 1
Derived Stimulus Configuration From Co-Citation Data for Psychometric Society Presidents, 1970–1986



Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

as arrows. Inspection of the space shows considerable structure. The right (and especially upper) side contains presidents who were primarily devoted to studying scaling. The upper left quadrant contains developers of factor analysis and structural equations modeling. Test theorists are on the left-hand side. Three presidents were "unstable" when slightly different versions of the MDS model (e.g., metric models) were fit. Ramsey and Cliff moved between the upper right quadrant and the lower left quadrant across these different solutions, and Bloxom also moved toward the scaling direction. These shifts properly reflect the attention of these presidents to both scaling and test theory issues. The other presidents were relatively stable across different implementations of the MDS algorithm.

To illustrate further the flexibility of this method—and still within the EDA framework—the lower triangle of Table 6 was submitted to a cluster analysis (using Ward's procedure in SAS). The results contained additional information not obvious in the MDS solution, particularly about the scaling specialists. The two major clusters clearly separated those who focused on scaling from the rest. The four subclusters contained developers of structural equations modeling (Guttman, McDonald, Jöreskog, and Bentler), the developers of ALSCAL (Takane, Young, and De Leeuw), scaling experts who worked at Bell Laboratories (Shepard, Kruskal, and Carroll), and all others.

Discussion

Asymmetric data are both informative and common. Approaches derived previously that explicitly apply to asymmetries (as opposed to treating them as error on top of symmetric processes) have been strongly model-based. This paper defined an exploratory procedure that supports a different approach to the analysis of asymmetric proximity matrices. Conceptually, all $N!$ orderings and the resulting (say) lower triangles could be analyzed with MDS, and directed distances indicated. Practically, seriation provides a mechanism to select one or two triangles that are of particular empirical interest. The resulting MDS portrayal represents symmetric information from within the relevant triangle, and the ordering of the stimuli reflects asymmetric information from relationships between the two triangles. These two procedures—separately and in combination—can provide substantive information to the data analyst, as illustrated by the analysis of the Psychometric Society presidents' data.

References

Baker, F. B., & Hubert, L. J. (1977). Applications of combinatorial programming to data analysis: Seriation using asymmetric proximity measures. *British Journal of Mathematical and Statistical Psychology, 30,* 154–164.

Bentler, P. M. (1986). Structural modeling and *Psychometrika*: An historical perspective on growth and achievements. *Psychometrika, 51,* 35–52.

Cliff, N., Collins, L., Zatkin, J., Gallipeau, D., & McCormick, D. (1988). An ordinal scaling method for questionnaire and other ordinal data. *Applied Psychological Measurement, 12,* 83–97.

Constantine, A. G., & Gower, J. C. (1978). Graphical representation of asymmetric matrices. *Applied Statistics, 27,* 297–304.

Coombs, C., & Smith, J. E. K. (1973). On the detection of structure in attitudes and developmental processes. *Psychological Review, 80,* 337–351.

Cunningham, J. P. (1978). Free trees and bidirectional trees as representations of psychological distance. *Journal of Mathematical Psychology, 17,* 165–188.

Defays, D. (1978). A short note on a method of seriation. *The British Journal of Mathematical and Statistical Psychology, 31,* 49–53.

DeSarbo, W. S. (1982). GENCLUS: New models for general non-hierarchical clustering analysis. *Psychometrika, 47,* 449–475.

DeSarbo, W. S., Manrai, A., & Burke, R. (1990). A nonspatial methodology for the analysis of two-way proximity data incorporating the distance-density hypothesis. *Psychometrika, 55,* 229–254.

De Soete, G., Hubert, L., & Arabie, P. (1988). On the use of simulated annealing for combinatorial data analysis. In W. Gaul & M. Schader (Eds.), *Data, expert knowledge, and decisions* (pp. 328–340). Berlin: Springer-Verlag.

Flink, S. (1990, February). On top of the world. *World Tennis,* 31–34.

Flueck, J. A., & Korsh, J. R. (1974). A branch search algorithm for maximum likelihood paired comparison ranking. *Biometrika, 61,* 621–626.

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Gower, J. C. (1977). The analysis of asymmetry and orthogonality. In J. R. Barra, F. Brodeau, G. Romer, & B. van Cutsem (Eds.), *Recent developments in statistics* (pp. 109–123). New York/Amsterdam: North Holland.

Heiser, W. J. (1981). *Unfolding analysis of proximity data.* Unpublished doctoral dissertation, University of Leiden, The Netherlands.

Holman, E. W. (1979). Monotonic models for asymmetric proximities. *Journal of Mathematical Psychology, 20,* 1–15.

Hubert, L. J. (1974). Some applications of graph theory and related non-metric techniques to problems of approximate seriation: The case for symmetric proximity measures. *The British Journal of Mathematical and Statistical Psychology, 27,* 133–153.

Hubert, L. J. (1976). Seriation using asymmetric proximity measures. *The British Journal of Mathematical and Statistical Psychology, 29,* 32–52.

Hubert, L. J., & Arabie, P. (1987). Unidimensional scaling and combinatorial optimization. In J. De Leeuw, W. J. Heiser, J. Meulman, & F. Critchley (Eds.), *Multidimensional data analysis* (pp. 181–196). Leiden, The Netherlands: DSWO Press.

Hubert, L. J., & Baker, F. B. (1979). Evaluating the symmetry of a proximity matrix. *Quality and Quantity, 13,* 77–84.

Hubert, L. J., & Golledge, R. G. (1981). Matrix reorganization and dynamic programming: Applications to paired comparisons and unidimensional seriation. *Psychometrika, 46,* 429–442.

Hutchinson, J. W. (1989). NETSCAL: A network scaling algorithm for nonsymmetric proximity data. *Psychometrika, 54,* 25–52.

Kendall, M. G. (1955). Further contributions to the theory of paired comparisons. *Biometrics, 11,* 43–62.

Klauer, K. C. (1989). Ordinal network representation: Representing proximities by graphs. *Psychometrika, 54,* 737–750.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review, 85,* 445–463.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills: Sage.

Levin, J., & Brown, M. (1979). Scaling a conditional proximity matrix to symmetry. *Psychometrika, 44,* 239–244.

Lewis, C. (1986). Test theory and *Psychometrika*: The past twenty-five years. *Psychometrika, 51,* 11–22.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis.* New York: Academic Press.

Mulaik, S. A. (1986). Factor analysis and *Psychometrika*: Major developments. *Psychometrika, 51,* 23–34.

Rodgers, J. L. (1988). Structural models of the American Psychological Association in 1986: A taxonomy for reorganization. *American Psychologist, 43,* 372–382.

Rodgers, J. L., Thompson, T. D., & Thompson, P. A. (1989). TRIAGE: A preprocessor for proximity data. *Psychometrika, 54,* 352–353.

Saito, T. (1986). Multidimensional scaling to explore complex aspects in dissimilarity judgment. *Behaviormetrika, 20,* 35–62.

SAS Institute, Inc. (1983). *SUGI supplemental library user's guide* [Computer program manual]. Cary NC: Author.

SAS Institute, Inc. (1985). *SAS user's guide: Basics* [Computer program manual]. Cary NC: Author.

Thissen, D., & Steinberg, L. (1988). Data analysis using item response theory. *Psychological Bulletin, 104,* 385–395.

Torgerson, W. S. (1986). Scaling and *Psychometrika*: Spatial and alternative representations of similarity data. *Psychometrika, 51,* 57–64.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading MA: Addison-Wesley.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Tversky, A., & Hutchinson, J. (1986). Nearest neighbor analysis of psychological spaces. *Psychological Review, 93,* 3–22.

Weeks, D. G., & Bentler, P. M. (1982). Restricted multidimensional scaling models for asymmetric proximities. *Psychometrika, 47,* 201–208.

Young, F. W. (1975, August). *An asymmetric Euclidean model for multiprocess asymmetric data.* Paper presented at the US–Japan Seminar on Multidimensional Scaling, University of California at San Diego, La Jolla.

Younger, D. H. (1963). Minimum feedback arc sets for a directed graph. *Institute of Electrical and Electronics Engineers Transactions on Circuit Theory, 10,* 238–245.

Acknowledgments

Earlier versions of this work were presented at the meetings of the Psychometric Society in Montreal in 1987 and in Los Angeles in 1988. The authors appreciate the assistance of Paul Thompson, who was involved in programming the seriation algorithm.

Author's Address

Send requests for reprints or further information to Joseph Lee Rodgers, Department of Psychology, University of Oklahoma, Norman OK 73019, U.S.A.

Downloaded from apm.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016