A Multidimensional Scaling Stress Evaluation Table

Kenneth Sturrock and Jorge Rocha
Field Methods 2000 12: 49
DOI: 10.1177/1525822X0001200104

A Multidimensional Scaling Stress Evaluation Table

KENNETH STURROCK
JORGE ROCHA
University of Florida

This article presents a probability table for the evaluation of stress values generated by multidimensional scaling (MDS) procedures employing stress formula 1. This table is based on the probability distribution of stress values from 587,200 random similarity matrices of different sizes processed to yield results for several dimensions.

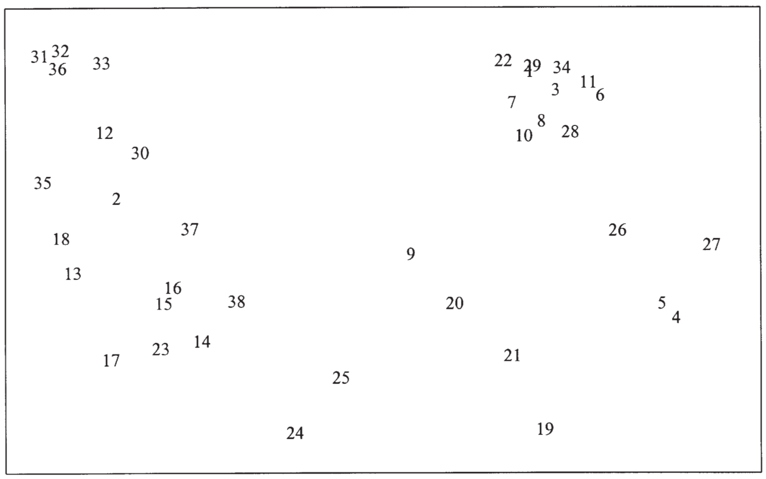
Multidimensional scaling (MDS) is widely used to assess and to visualize similarities or dissimilarities among a set of cognitive or physical objects. Similarities might be measured in percentages or in some form of correlation. Pile-sort data on a list of diseases, for example, comprise a matrix of proportions—the percentage of times that any item is placed in the same pile with any other item in a list. Distance, in miles, between all pairs of cities in a list of cities comprise dissimilarities that are appropriate for MDS analysis.

Figure 1 shows the output from an MDS analysis. Thirty-eight informants in a small village in southern Mexico were asked to pile-sort thirty-eight photographs (one for each informant) into groups of people who are in the same family. Most MDS diagrams, like the one shown in Figure 1, are two-dimensional, probably because two dimensions are easy to comprehend. Real data, however, are often best represented in more than two dimensions. To visualize three- (or four- or five-) dimensional data in two dimensions, the distance between each pair of objects in an MDS diagram may be slightly inaccurate. The sum of these inaccuracies is referred to as the stress in the MDS solution and in the diagram that represents the solution. The stress in two dimensions for the output shown in Figure 1 is 0.168.

The authors would like to thank Russ Bernard, Lance Gravlee, David Kennedy, George Mbeh, and Chris McCarty for their help and support with this project. We would especially like to thank our three anonymous reviewers for their numerous valuable suggestions.

Field Methods, Vol. 12, No. 1, February 2000 49–60

FIGURE 1
An Example of a Thirty-Eight–Object Multidimensional Scaling Diagram Representing Kinship Groups



NOTE: This diagram represents kinship groupings between a number of families in southern Mexico. The data were collected by having informants pile-sort photographs of people living in the area. The stress of the diagram nonmetrically scaled in two dimensions is 0.168.

Traditionally, stress is thought of as a goodness-of-fit measure for the diagram—but it is in fact a badness-of-fit measure because the higher the stress score, the worse the fit (Kruskal and Wish 1978:49). A high stress value usually indicates that the chosen number of dimensions is not adequate for accurately portraying the complex relationship among a set of objects. An alternative explanation could be that the objects depicted have no real relationship and therefore cannot be arrayed in the number of dimensions chosen. Regardless, a high stress value is considered undesirable.

METHODS FOR MEASURING STRESS

Working under the premise that there is, in fact, structure to the relationship between the objects, Kruskal and Wish (1978) outlined alternative methods for determining the best dimensionality for an MDS diagram. One method involves scaling the diagram in several different dimensions, graphing the values of the resulting stress, and choosing the dimensionality with the highest drop in stress compared to the next lower dimensionality. The second method involves applying a series of rules of thumb to the same stress plots and interpreting the shape of the curve with knowledge gained through "experience and intuition" (Kruskal and Wish 1978:54–56).

An additional approach involves identifying boundaries on the upper and lower limits of stress and comparing the diagram's stress to these limits. The lower limit of stress, or a perfect dimensional fit, would result in a stress value of zero. This level is typical when dealing with very small numbers of objects, because n objects will always fit perfectly in $n - 1$ dimensions (Kruskal and Wish 1978:52). For example, three objects can be scaled perfectly in two dimensions.

The upper bound of stress is more difficult to determine, as it depends on the number of objects and the chosen dimensionality. One way to determine the upper bound of stress is to scale randomly generated matrices. The idea is that random matrices should have no real structure that connects the objects represented; these matrices would, therefore, produce a worst-case stress value when scaled. The lowest observed stress value from such unstructured matrices could then be used to place an upper limit on the level of stress generated by structured matrices. Since we expect some variation in stress for randomly generated matrices of a given size, large numbers of random matrices must be scaled to create a distribution of stress values for comparison.

Klahr (1969) laid the groundwork for this approach by showing that there was a difference between stress values for randomly generated matrices and matrices with a real structural arrangement of objects. Klahr posed the following problem: If a large number of unstructured, randomly generated matrices were produced and scaled in various dimensions, how frequently would reasonable or structured-looking stress values be generated? Klahr scaled 3,000 randomly generated matrices with different combinations of objects and dimensions. He found that a randomly generated eight-object matrix would virtually never scale to produce a stress value that was similar to an actual, structurally related matrix of eight objects in two dimensions. When the number of objects in the random matrices was increased to 16, then structured-looking stress values would not be generated until the matrix was scaled in more than five dimensions (Klahr 1969:326–27).

Spence and Ogilvie (1973) reported an experiment in which they repeatedly created and scaled a few configurations of randomly generated matrices. They used a curve-fitting procedure to extrapolate the stress values for the MDS configurations that they did not run, and they produced a table depicting mean stress values for MDS configurations of 12–48 objects in from 1–5 dimensions (Spence and Ogilvie 1973:515; see also Spence and

FIGURE 2
Stress Formulas 1 and 2

$$Stress_1^{\,2} = \frac{\sum_{i,j}(d_{ij} - \tilde{d}_{ij})^2}{\sum_{i,j} d_{ij}^2}$$

$$Stress_2^{\,2} = \frac{\sum_{i,j}(d_{ij} - \tilde{d}_{ij})^2}{\sum_{i,j}(d_{ij} - d)^2}$$

NOTE: d_{ij} represents the raw dissimilarity values for each pair of objects. $\tilde{d}_{ij}$ represents the scaled distances in a particular number of dimensions for each pair of objects. d represents the mean of all d_{ij}.

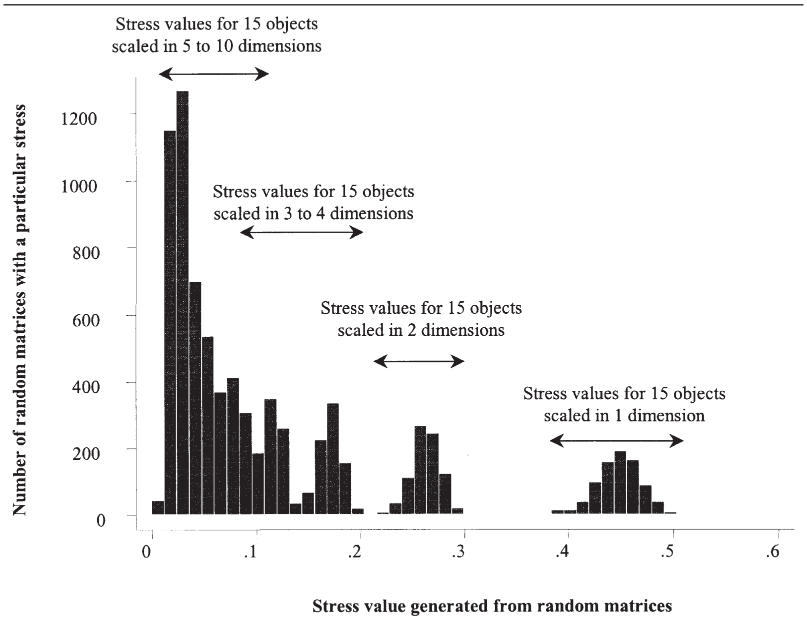
Graef 1974:331). Levine (1978) continued this line of research by creating and scaling 14,000 MDS configurations and producing a table showing the percentage chance of encountering a particular stress value for a given MDS configuration. Levine's table displays stress values for MDS configurations in 1–5 dimensions with 6, 8, 10, 12, 16, 20, and 24 objects in the matrix.

Unfortunately, from a practical standpoint, Levine's (1978) work was based on the less commonly used stress formula 2. Stress formula 2 is a slightly different way of calculating ideal versus depicted discrepancies in the MDS diagram by using the variance of the distances instead of the sum of squared distances in the denominator (see Figure 2). Although stress formulas 1 and 2 results can be converted to each other with a bit of work, Levine's table is not as useful as it might be to researchers who use the much more common stress formula 1, since formula 2's results are always larger (Shepard, Kimball, and Nerlove 1972:80–83; De Leeuw and Stoop 1984:391–93).

APPLYING MORE COMPUTING POWER TO THE PROBLEM

To make our results as useful as possible, we produce a table based on a brute force approach to determining the upper stress boundaries using stress formula 1. A UNIX shell script was written to invoke the Statistical Analysis

FIGURE 3
A Composite Histogram Depicting Frequencies of Stress from 8,000 Randomly Generated Matrices Containing Fifteen Objects at Various Dimensionalities



System (SAS) statistical package, which in turn created and scaled random matrices nonmetrically using a TORSCA initial configuration. The matrices were created by filling the cells with values chosen from a uniform random distribution of numbers from 0 to 1. The shell script then extracted the stress values from the SAS output files and placed them into a comma-delimited text file.

A second shell script was used to repeatedly run the first script. The scripts were written so that several copies could be run simultaneously on the University of Florida's Digital Equipment Corporation (now Compaq) Alpha cluster. The end result was the creation and scaling of 587,200 random matrices. The matrix configurations covered were 4–100 objects and 1–10 dimensions. Each configuration below 61 objects was run 800 times, while each configuration with 61 or more objects was run only 400 times. We did not bother running configurations that were expected to resolve perfectly—that

TABLE 1
Stress Values for MDS-Scaled Randomly Generated Matricies of Fifteen Objects (800 matrices per dimension)

Dimensions	*Minimum Stress*	*Mean Stress*	*Maximum Stress*
1	0.382	0.449	0.51
2	0.212	0.263	0.3
3	0.141	0.172	0.203
4	0.083	0.116	0.149
5	0.058	0.082	0.113
6	0.033	0.058	0.08
7	0.027	0.042	0.058
8	0.012	0.029	0.046
9	0.009	0.025	0.042
10	0.007	0.019	0.034

NOTE: MDS = multidimensional scaling.

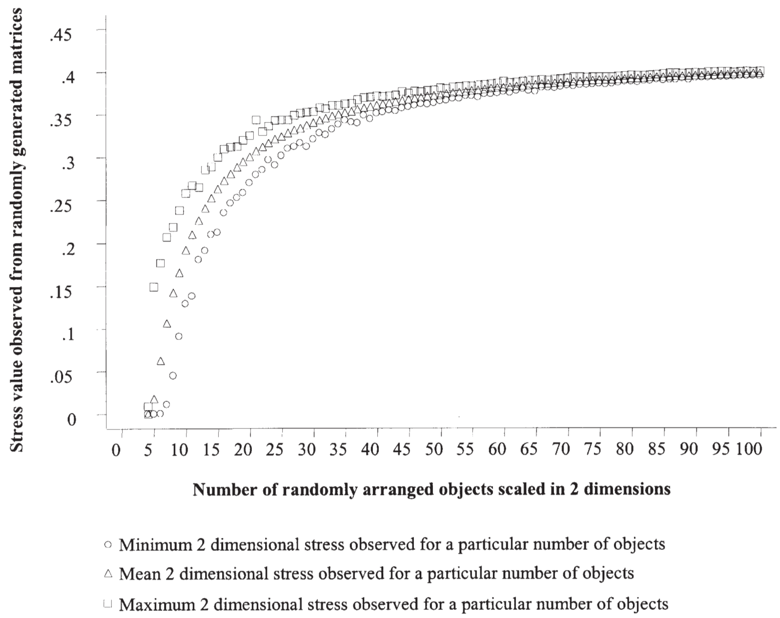
is, when the number of dimensions was one less than the number of objects, or higher.

RESULTS

The next step was to examine the distributions of each configuration and to determine the percentile values for use in constructing a table. For example, Table 1 and Figure 3 show the results of 8,000 randomly generated matrices containing fifteen objects scaled in 1–10 dimensions. Figure 3 shows how the distributions cluster around definite average points according to the number of dimensions in which the matrices were scaled. The fifteen-object matrices scaled in one dimension have considerably higher stress than those scaled in two dimensions.

As the dimensionality increases the stress decreases, as expected, since it becomes easier for the MDS algorithm to fit the objects into the allowed space. Note that as the number of objects scaled increases there is less variation in the stress results obtained. Figure 4 shows that as the number of objects scaled in two dimensions increases, the minimum and maximum stress values observed converge and begin to level off. It appears that the stress generated from scaling 100 random objects in two dimensions is not much worse than the stress generated from scaling 40 random objects.

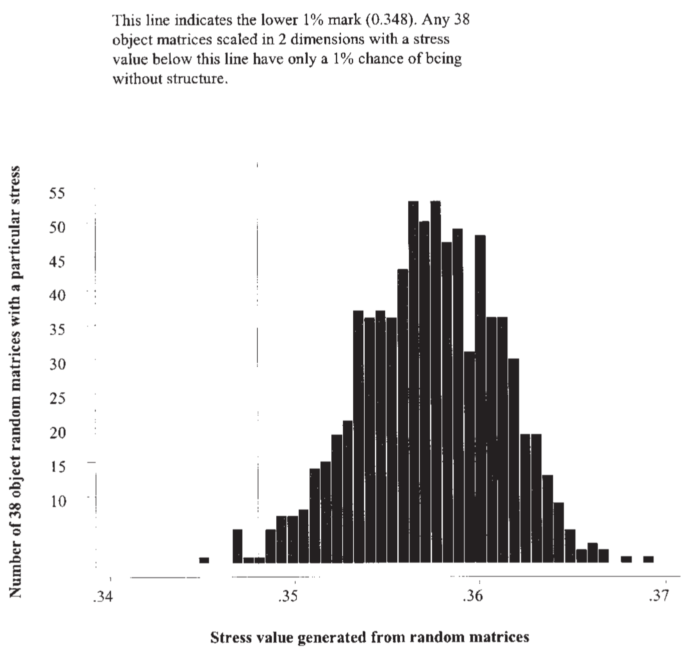
FIGURE 4
The Increase of Stress
with Number of Objects Scaled



The percentiles of the distributions for one, two, and three dimensions were calculated and used to create Table 2. Researchers can use this table to compare the stress in actual MDS solutions to the 1% value of a distribution of stress values generated from random matrices with the same number of objects and scaled in the same number of dimensions. For example, Figure 1 comes from a data set with thirty-eight objects that was scaled in two dimensions. The stress for that Figure 1 is 0.168. According to Table 2, if that thirty-eight–object matrix happened to scale in two dimensions with a stress value of 0.348, we expect a 1% chance that the objects in the matrix are randomly arranged (see Figure 5). Therefore, we can claim with some certainty that the thirty-eight–object matrix with a two-dimensional stress of 0.168 is not random, or without structure.

It may be difficult to pick out the relevant factors from a two-dimensional diagram, but as Kruskal and Wish (1978:49) pointed out, the number of

FIGURE 5
Depiction of the 1% Cutoff for a Thirty-Eight–Object Matrix in Two Dimensions



explanatory factors is not necessarily the same as the number of axes used to depict the diagram. As always, it is a good idea to use a number of different methods to triangulate outcomes. Additionally, it must be kept in mind that there is no substitute for quality of data. Poorly collected data about a structurally related set of objects may be no better than random noise. Just because the stress from a particular matrix is less than the boundaries shown in Table 2 does not mean that the results of the scaling will always be useful to the researcher.

CONCLUSION

Our approach to building this table has precedent, but previous studies were hampered by the low computational power available to researchers. Our

TABLE 2
Nonmetric MDS Stress 1% Left-Hand-Tail Cutoff Values

Number of Objects	*Stress in One Dimension*	*Stress in Two Dimensions*	*Stress in Three Dimensions*
5	0.002	0.000	0.000
6	0.019	0.002	0.000
7	0.161	0.024	0.002
8	0.227	0.071	0.012
9	0.256	0.104	0.038
10	0.286	0.133	0.058
11	0.328	0.160	0.084
12	0.347	0.183	0.103
13	0.368	0.199	0.117
14	0.387	0.217	0.134
15	0.393	0.228	0.147
16	0.411	0.242	0.154
17	0.421	0.254	0.162
18	0.431	0.263	0.172
19	0.441	0.269	0.182
20	0.446	0.279	0.189
21	0.459	0.284	0.196
22	0.466	0.293	0.200
23	0.471	0.301	0.205
24	0.475	0.302	0.212
25	0.480	0.308	0.216
26	0.487	0.313	0.221
27	0.488	0.317	0.223
28	0.490	0.319	0.229
29	0.497	0.324	0.232
30	0.499	0.328	0.235
31	0.503	0.330	0.240
32	0.505	0.333	0.241
33	0.507	0.337	0.243
34	0.510	0.339	0.247
35	0.512	0.342	0.250
36	0.514	0.343	0.252
37	0.517	0.347	0.255
38	0.519	0.348	0.255
39	0.519	0.349	0.258
40	0.522	0.352	0.260
41	0.525	0.354	0.262
42	0.526	0.356	0.264
43	0.528	0.357	0.265
44	0.530	0.358	0.267
45	0.530	0.360	0.269
46	0.532	0.362	0.269

(continued)

TABLE 2 Continued

Number of Objects	*Stress in One Dimension*	*Stress in Two Dimensions*	*Stress in Three Dimensions*
47	0.534	0.363	0.271
48	0.534	0.365	0.272
49	0.534	0.365	0.274
50	0.536	0.366	0.275
51	0.537	0.368	0.276
52	0.539	0.369	0.277
53	0.539	0.370	0.278
54	0.540	0.371	0.280
55	0.542	0.372	0.281
56	0.542	0.373	0.281
57	0.543	0.374	0.282
58	0.543	0.375	0.284
59	0.544	0.375	0.284
60	0.544	0.376	0.285
61	0.546	0.377	0.286
62	0.547	0.378	0.287
63	0.546	0.380	0.287
64	0.548	0.379	0.288
65	0.548	0.380	0.289
66	0.548	0.381	0.290
67	0.549	0.381	0.290
68	0.550	0.382	0.291
69	0.550	0.382	0.291
70	0.550	0.384	0.292
71	0.551	0.384	0.292
72	0.551	0.384	0.293
73	0.552	0.385	0.294
74	0.552	0.386	0.294
75	0.553	0.386	0.294
76	0.553	0.386	0.296
77	0.554	0.386	0.296
78	0.553	0.387	0.296
79	0.553	0.388	0.297
80	0.555	0.388	0.297
81	0.555	0.389	0.298
82	0.556	0.390	0.299
83	0.555	0.390	0.299
84	0.556	0.390	0.299
85	0.556	0.390	0.300
86	0.556	0.391	0.300
87	0.557	0.391	0.300
88	0.557	0.391	0.300

TABLE 2 Continued

Number of Objects	*Stress in One Dimension*	*Stress in Two Dimensions*	*Stress in Three Dimensions*
89	0.557	0.392	0.301
90	0.558	0.392	0.302
91	0.558	0.393	0.302
92	0.557	0.393	0.302
93	0.558	0.393	0.303
94	0.559	0.394	0.303
95	0.559	0.394	0.303
96	0.559	0.394	0.304
97	0.559	0.395	0.304
98	0.559	0.395	0.304
99	0.560	0.396	0.304
100	0.560	0.396	0.305

NOTE: MDS = multidimensional scaling.

results match those of Spence and Ogilvie (1973) and correlate well with Levine's (1978). There remain some unanswered questions involving the shapes of the random distributions, the possibility of using permutation statistics to generate the comparison random matrices, and the possibility of creating accurate formulae to describe stress levels across matrix configurations. Nevertheless, we feel that Table 2 is a useful addition to the methods described by Kruskal and Wish (1978) and by others. We invite those interested to examine the data set stored at http://www.bebr.ufl.edu/ks/MDSstress/.

REFERENCES

De Leeuw, Jan, and Ineke Stoop. 1984. Upper bounds for Kruskal's stress. *Psychometrika* 49 (3): 391–402.

Klahr, David. 1969. A Monte Carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure. *Psychometrika* 34 (3): 319–30.

Kruskal, Joseph B., and Myron Wish. 1978. *Multidimensional scaling.* Sage University paper series on Quantitative Applications in the Social Sciences #11. Beverly Hills, CA: Sage.

Levine, David M. 1978. A Monte Carlo study of Kruskal's variance based measure of stress. *Psychometrika* 43 (3): 307–15.

Shepard, Roger N., Romney A. Kimball, and Sara Beth Nerlove. 1972. *Multidimensional scaling: Theory and applications in the behavioral sciences. Volume I—Theory.* New York: Seminar Press.

Spence, Ian, and John C. Ogilvie. 1973. A table of expected stress values for random rankings in nonmetric multidimensional scaling. *Multivariate Behavioral Research* 8 (4): 511–17.
Spence, Ian, and Jed Graef. 1974. The determination of the underlying dimensionality of an empirically obtained matrix of proximities. *Multivariate Behavioral Research* 9 (3): 331–41.

KENNETH STURROCK is a Ph.D. anthropology student at the University of Florida and an MPH student at the University of South Florida. His past work has included preservation of previously unwritten languages and the study of the transmission of knowledge through social networks. Currently, his research focuses on the behavioral and cultural factors that influence the selection and use of drugs for the treatment of infectious disease.

*JORGE ROCHA is a Ph.D. candidate at the University of Florida. His interests are centered on materialist theories of sociocultural change and methodologies for operationalizing research strategies derived from those theories. His dissertation research focused on understanding the processes that drive rural household agricultural producers in Peru to intensify the use of their natural resources and the cognitive changes in their conceptualization that accompanied those processes. His recent publications include "Culture, Rationality and Decision Making" (*Research in Economic Anthropology *1996) and "Estructura Economica Regional de America Latina" (*Revista de Ciencias Sociales de la Universidad Autonoma de Nuevo Leon *1999).*