

12-1997

# Making Sense of Direction: Proximity and Order in Asymmetric Paired Comparison Data

Jonathan M. Borkum

Follow this and additional works at: <http://digitalcommons.library.umaine.edu/etd>



Part of the [Psychology Commons](#)

---

## Recommended Citation

Borkum, Jonathan M., "Making Sense of Direction: Proximity and Order in Asymmetric Paired Comparison Data" (1997). *Electronic Theses and Dissertations*. 2329.

<http://digitalcommons.library.umaine.edu/etd/2329>

This Open-Access Dissertation is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine.

MAKING SENSE OF DIRECTION:  
PROXIMITY AND ORDER IN ASYMMETRIC PAIRED COMPARISON DATA

By

Jonathan M. Borkum

B.A. State University of New York at Stony Brook, 1982

A THESIS

Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
Doctor of Philosophy  
(in Psychology)

The Graduate School  
University of Maine  
December, 1997

Advisory Committee:

Colin Martindale, Professor of Psychology, Advisor  
Joel A. Gold, Professor and Chairperson, Department of Psychology  
William Halteman, Associate Professor of Mathematics and Statistics  
Gordon E. Kulberg, Associate Professor of Psychology  
and Dean of the College of Social and Behavioral Sciences  
D. Alan Stubbs, Professor of Psychology

© 1998 by Jonathan M. Borkum  
All Rights Reserved

**MAKING SENSE OF DIRECTION:  
PROXIMITY AND ORDER IN  
ASYMMETRIC PAIRED COMPARISON DATA**

By Jonathan M. Borkum

Thesis Advisor: Dr. Colin Martindale

An Abstract of the Thesis Presented  
in Partial Fulfillment of the Requirements for the  
Degree of Doctor of Philosophy  
(in Psychology)  
December, 1997

In a square asymmetric matrix, the relationships among objects in the lower triangular half-matrix, differ from the relationships among the same objects in the upper triangular half. Square, asymmetric matrices can arise in similarity and preference data, when the direction of comparison is important. An asymmetric matrix can be rendered symmetric by averaging corresponding entries above and below the main diagonal. The difference between the original and the symmetric matrix is purely asymmetric, or skew-symmetric. The symmetric and skew-symmetric parts are orthogonal. An eigenvector-eigenvalue decomposition analyses the asymmetries into rank 2 skew-symmetric matrices, having an optimum least squares fit to the asymmetries (Gower, 1977).

In this dissertation I derive an alternating least squares, nonmetric analogue of the canonical decomposition of asymmetry, suitable for ordinal-level data. In simulation studies, the nonmetric version gives better metric and nonmetric recovery, than does the canonical decomposition, when the asymmetries have been

distorted by a range-compressing monotonic transform. The nonmetric technique appears to out-perform the canonical decomposition in detecting simplexes, and possibly in recovering multiplicative bias coefficients. However, canonical decomposition gives superior recovery after range-expanding monotonic transforms, and in the presence of error.

An eigenvalue ratio test is proposed for determining the number of eigenvectors to extract in the canonical decomposition. The test quantifies changes in the slope of the log eigenvalue plot. In simulation studies the test appears to maintain its anticipated Type I error rate. The test is "under-powered", which may help it to extract only well-identified eigenvectors.

Finally, directional similarity judgments were collected for all possible pairs of exemplars of two semantic categories. The exemplars differed in typicality. After Tversky (1977) this should produce asymmetries related to the typicality. No asymmetries were found, however. Power analysis indicated that a correlation ratio for the asymmetries of .05 could have been detected 90% of the time. An extreme groups analysis also did not indicate asymmetry. The first eigenvector underlying the symmetric data, however, was highly correlated with typicality. Hence, Tversky's model was not supported.



## Acknowledgments

My few words here can scarcely acknowledge the debts I have acquired over the years. Colin Martindale, Ph.D., has been an magnificent advisor. I would not have applied to the University of Maine if it had not been for his book, nor been accepted without his support, nor enjoyed it half as much without the conversations and projects that he shared with me. He has introduced me to the newest research paradigms in cognitive psychology, and venerable traditions in empirical aesthetics. I have never known him to be wrong about a research direction. Nonetheless, he has always allowed me to pursue the idiosyncrasies of my own curiosity, and even long-shot projects, with unfailing support. Although the first draft of this dissertation, which also served as the first draft of a book on all of scaling and measurement, was rather long and dense, Colin read it with remarkable patience and, actually, enthusiasm. Graduate school would have been a much diminished experience without the intellectual adventures to which he has introduced me over the years.

Similarly, this dissertation would not have existed if I had not taken the statistics courses taught by Joel A. Gold, Ph.D. Before these courses, I had taken for granted that statistics required a genetic talent that I simply lacked. The first of his three classes began for me on Tuesday July 10, 1984, and I still recall the discussion that morning on linear regression, for its rigor, clarity, and enthusiasm. I suppose that a significant part of my career direction changed in that class discussion, and in the many that followed. To this day, quantitative methods have a sense of majesty to me. My life would be much poorer if Joel had not shown me that majesty.

It is sometimes said that exposure to statistics can promote clear, insightful thinking about clinical matters. As I continue in the field, however, I find that also the intuitive yet precise formulation demanded in clinical work gives clarity to my thinking about quantitative psychology. It was Gordon Kulberg, Ph.D., who first challenged me to think precisely and empirically about clients.

In contacts over many years, William Halteman, Ph.D. has been a model of professionalism. With tact and clarity, he has shown me areas that I did not know in statistics, and gently challenged me to think more deeply. Alan Stubbs, Ph.D., has been unfailingly supportive, and shown me quantitative models in psychology, and the joy of new ideas.

To the faculty as a whole I owe a substantial debt for their forbearance, as I pursued the long path leading to this dissertation. It is perhaps unusual, that a student in clinical psychology would be permitted such a quantitative dissertation. I feel privileged to have been a part of a program in which such a combination was possible.

For nearly a decade now, I have had the privilege of working in a behavioral medicine program built with remarkable steadiness by Jeffrey Matranga, Ph.D., ABPP. From small beginnings in a community hospital, it has moved forward, sometimes slowly, sometimes with breathtaking leaps, and is now statewide. This would have been impossible without Jeff's good-natured focus and exceptional sense of timing and judgment. Whatever writing style as I may claim to have was undoubtedly forged in the hundreds of reports that I have dictated under Jeff's tutelage over the years. I have had the privilege, too, of learning from many coworkers, too numerous to properly name here. In particular, though, I would

mention the exuberance, wisdom, and honesty of James R. Fegan, M.D., a psychiatrist from whom I learned about life, the brilliance of Altaf Ahmed, M.D., who lifts us all to a higher level, the patience of John Ditri, M.D., the camaraderie of Trish Hildreth and Eleanor Ellis across many years, and the support and comfortable, complementary strengths of my coworkers, Willy White, M.A., Christine Gray, Psy.D., and Jeff Leonards, Ph.D.

I also thank Ken Kindya, Ph.D., for teaching me the secret of life and success, and for his friendship; Yvonne Unger, whose integrity and profound decency is a cure for any cynicism; and Ira Lipsky, for his faithfulness as a friend.

Finally, I dedicate this dissertation to Roger Borkum, and to Debi and Christopher Imhof, and to the memory of my parents, Judith and David Borkum, for more than I can put into words.

## TABLE OF CONTENTS

Acknowledgments . . . . .	ii
List of Tables . . . . .	vii
List of Figures . . . . .	x
Chapter 1. Introduction . . . . .	1
Chapter 2. Dependence of Multidimensional Techniques on Symmetric Input . . . . .	6
Chapter 3. The Relationship of Symmetric and Asymmetric Parts . . . . .	10
Chapter 4. Canonical Analysis of Skew-Symmetric Data . . . . .	13
Chapter 5. Current Techniques for Mixed Symmetric -- Skew-Symmetric Data . . . . .	23
Chapter 6. Nonmetric Analysis of Skew-Symmetries . . . . .	33
Chapter 7. Simulation Program . . . . .	45
Chapter 8. Simulation Studies: Quality Control . . . . .	59
Chapter 9. Simulation Studies: Nonmetric Analysis of Skew-Symmetry . . . . .	70
Chapter 10. Simulation Studies: Canonical Analysis of Asymmetry . . . . .	89
Chapter 11. Simulation Studies: Comparing Metric and Nonmetric Techniques . . . . .	104
Chapter 12. Eigenvalue Ratio Test . . . . .	126
Chapter 13. Analyses With Special Matrix Structures . . . . .	139
Chapter 14. Published Data Sets . . . . .	147
Chapter 15. Rationale for Experiments . . . . .	164
Chapter 16. Experiment 1: Similarity . . . . .	170
Chapter 17. Experiment 2: Typicality . . . . .	181

**TABLE OF CONTENTS**  
(Continued)

<b>Chapter 18. Discussion</b> . . . . .	<b>188</b>
<b>References</b> . . . . .	<b>202</b>
<b>Appendices</b>	
<b>I. Orthogonality of Symmetric and Skew-Symmetric Parts</b> . . . . .	<b>211</b>
<b>II. Canonical Analysis of Asymmetry: Computational Features</b> . . . . .	<b>212</b>
<b>III. Geometric Interpretation of Skew-Symmetric Matrices</b> . . . . .	<b>218</b>
<b>IV. Rescaling Factors for Wishart, Chi-Square, and Asymmetric Error</b> .	<b>221</b>
<b>V. Stimulus Words</b> . . . . .	<b>226</b>
<b>VI. Source Code of the FORTRAN 77 Program         Implementing NSKMDS</b> . . . . .	<b>227</b>
<b>VII. Sample FORTRAN 77 Source Code for Bracketing         and Bisection of Asymmetric Error Vectors</b> . . . . .	<b>241</b>
<b>Biography</b> . . . . .	<b>246</b>

## LIST OF TABLES

1.	Monotone Regression . . . . .	38
2.	Diagnostic Statistics on the Random Number Generator . . . . .	61
3.	Time of Day Effect: Analysis of Variance . . . . .	63
4.	Diagnostic Statistics for the Random Number Generator: Time of Day Effects . . . . .	66
5.	Time of Day Effect . . . . .	69
6.	Nonmetric Analysis (NSKMDS): Recovery Under Positively Accelerated Monotonic Distortion . . . . .	73
7.	Nonmetric Analysis (NSKMDS): Recovery Under Decelerating Monotonic Distortion . . . . .	76
8.	Nonmetric Analysis (NSKMDS): Recovery as a Function of Normal Error . . . . .	79
9.	Nonmetric Analysis (NSKMDS): Recovery as a Function of Asymmetric Error . . . . .	83
10.	Nonmetric Analysis (NSKMDS): Analysis of Variance . . . . .	86
11.	Nonmetric Analysis (NSKMDS): Normal Error by Asymmetric Error by Monotonic Distortion . . . . .	88
12.	Canonical Analysis of Asymmetry (CAA): Recovery Under Positively Accelerated Monotonic Distortion . . . . .	90
13.	Canonical Analysis of Asymmetry (CAA): Recovery as a Function of Normal Error . . . . .	95
14.	Canonical Analysis of Asymmetry (CAA): Recovery as a Function of Asymmetric Error . . . . .	97
15.	Canonical Analysis of Asymmetry (CAA): Analysis of Variance . . . . .	101

**LIST OF TABLES**  
(Continued)

16.	<b>Canonical Analysis of Asymmetry (CAA): Normal Error by Asymmetric Error by Monotonic Distortion . . . . .</b>	102
17.	<b>Analysis Method by Positively Accelerated Monotonic Distortion: Analysis of Variance for Nonmetric Recovery . . . . .</b>	105
18.	<b>Analysis Method by Positively Accelerated Monotonic Distortion: Analysis of Variance for Metric Recovery . . . . .</b>	106
19.	<b>Analysis Method vs. Decelerating Monotonic Distortion: Analysis of Variance for Metric Recovery . . . . .</b>	110
20.	<b>Analysis Method vs. Decelerating Monotonic Distortion: Analysis of Variance for Nonmetric Recovery . . . . .</b>	112
21.	<b>Analysis Method vs. Normal Error Level: Analysis of Variance for Metric Recovery . . . . .</b>	115
22.	<b>Analysis Method vs. Normal Error Level: Analysis of Variance for Nonmetric Recovery . . . . .</b>	118
23.	<b>Analysis Method vs. Asymmetric Error Level: Analysis of Variance for Metric Recovery . . . . .</b>	120
24.	<b>Analysis Method vs. Asymmetric Error Level: Analysis of Variance for Nonmetric Recovery . . . . .</b>	121
25.	<b>Eigenvalues of Skew-Symmetries for Random Normal Matrices . . . . .</b>	128
26.	<b>Normative Values for the Eigenvalue Ratio Test . . . . .</b>	133
27.	<b>Empirically Obtained Type I and Type II Error Rates for the Eigenvalue Ratio Test . . . . .</b>	136
28.	<b>One-Dimensional Skew-Symmetries . . . . .</b>	140
29.	<b>Intercorrelations Among Cognitive Tests for Alcohol-Dependent and Control Subjects . . . . .</b>	150

**LIST OF TABLES**  
(Continued)

30.	International Trade in 1993 . . . . .	155
31.	International Trade Data: Skew-Symmetric Dimensions . . . . .	156
32.	Journal Citation Data for 1960 . . . . .	160
33.	Representation of Asymmetries in Journal Citation Data . . . . .	162
34.	Hypothetical Similarity Matrix . . . . .	165
35.	Eigenvalues of Vegetable Typicality Judgments . . . . .	184
36.	Typicality Judgments of Vegetables . . . . .	185
37.	Eigenvalues of Typicality Judgments of Tools . . . . .	186
38.	Typicality Judgments of Tools . . . . .	187



## LIST OF FIGURES

1a.	Accelerated Monotonic Transform . . . . .	56
1b.	Decelerated Monotonic Transform . . . . .	57
2.	Time of Day Effect . . . . .	64
3.	NSKMDS: Accelerated Distortion . . . . .	74
4.	NSKMDS: Decelerated Distortion . . . . .	77
5.	NSKMDS: Normal Error Level . . . . .	80
6.	NSKMDS: Asymmetric Error . . . . .	84
7.	CAA: Accelerated Distortion . . . . .	91
8.	CAA: Decelerated Distortion . . . . .	93
9.	CAA: Normal Error Level . . . . .	96
10.	CAA: Asymmetric Error . . . . .	98
11a.	Metric Recovery vs. Accelerated Distortion . . . . .	107
11b.	Nonmetric Recovery vs. Accelerated Distortion . . . . .	108
12.	Metric Recovery vs. Decelerated Distortion . . . . .	111
13.	Nonmetric Recovery vs. Decelerated Distortion . . . . .	113
14a.	Metric Recovery vs. Normal Error . . . . .	116
14b.	Nonmetric Recovery vs. Normal Error . . . . .	117
15.	Metric Recovery vs. Asymmetric Error . . . . .	122
16.	Nonmetric Recovery vs. Asymmetric Error . . . . .	124
17.	Distribution of EVR1 . . . . .	130

**LIST OF FIGURES**  
(Continued)

18.	Distribution of EVR1: High Error . . . . .	132
19.	CAA Solution for One-Dimensional Data . . . . .	141
20.	CAA Solution for Constant Error (Simplex) . . . . .	143
21.	NSKMDS Solution for Constant Error (Simplex) . . . . .	144
22.	CAA Solution for Cognitive Test Data . . . . .	151
23.	NSKMDS Solution for Cognitive Tests . . . . .	153
24.	CAA Solution for International Trade . . . . .	158
25.	CAA Solution for Journal Citations . . . . .	163
26.	Geometric Interpretation of Skew-Symmetries . . . . .	220

## Chapter 1

### Introduction .

Scaling is the application of a model to raise qualitative data to an interval or ratio level (Young, 1984). The assumptions of the model supplement the information missing in the data (Shepard, 1958). Scaling techniques are also used to display data that is already on an interval or ratio level (e.g., Torgerson, 1952). Many scaling techniques presuppose that the attribute is one-dimensional, that is, that it can be represented by a single continuum (Torgerson, 1958). The brightness of a light, loudness of a tone, and heaviness of a weight are examples of approximately one-dimensional attributes.

The goals of this dissertation, however, pertain to the scaling of multidimensional attributes. These are complex attributes that can be resolved into more than one aspect or component. For example, hue can be modeled by the color circle spanned by the two opponent process dimensions of "red-green" and "blue-yellow" (Indow, 1988; Shepard, 1980). An important feature of the multidimensional techniques is that the level of complexity of their solutions is data-driven. The multidimensional techniques will extract a single dimension when this is warranted by goodness-of-fit, and multiple dimensions when these are required to account for the data (e.g., Kruskal, 1964a; Shepard, 1962; Torgerson, 1952).

Typically, the multidimensional techniques use a similarity matrix as input. The matrix gives the degree of resemblance, proximity, or consonance of all pairs of

stimuli drawn from a set (Coombs, 1964). For example, observers may rate the similarity of musical tones (Krumhansl, 1979). In the data matrix the rows and columns correspond to the tones. The numbers  $x_{ij}$  in the body of the matrix are the average similarity rating of the tone for row "i" to the tone represented by column "j".

An advantage of the multidimensional techniques is their broad scope of application: essentially any matrix that is symmetric, and in which the rows and columns stand for the same entities, can be used as input. For example, we can use multidimensional scaling to represent the dimensions underlying the similarities among phonemes. From the standpoint of the analysis, it does not matter whether the similarities were obtained as direct ratings (e.g., Peters, 1963; Singh, Woods, & Becker, 1972), disjunctive reaction time (e.g., Münsterberg, 1894, for letters), confusion probabilities in perception (Miller, & Nicely, 1954), confusion probabilities in recall (Wickelgren, 1966), or some other behavioral process. The mathematics of the data analysis is clearly separated from the method of data collection. Thus, the multivariate analyses can range far from sensation and perception to represent cognitive or behavioral processes and, indeed, virtually any symmetric pattern of relationships among objects.

However, asymmetric relationships are far less accessible to standard multidimensional techniques. A symmetric relation is one in which order has no effect: The similarity of "a" to "b" equals the similarity of "b" to "a". Conversely, in

asymmetric relations, direction matters. The two orders of comparison give different results.

For many types of data, asymmetry seems to be the rule. The phoneme /b/ is not confused with /v/ with the same frequency as the reverse (Miller, & Nicely, 1954), nor are the rated similarities necessarily equal. Rosch (1975a) found that the psychological distance is shorter from an exemplar to a prototype than from the prototype to an exemplar. Tversky suggests that whenever two objects differ in characteristics such as "intensity, frequency, familiarity, good form, and informational content" similarity is likely to be asymmetric (Tversky, 1977, p. 332).

Tversky argues that asymmetry is inherent in directional similarity judgments: "We say that the portrait resembles the person, not that the person resembles the portrait" (Tversky, 1977, p. 328). In truth, however, the breadth of this phenomenon is not well understood. Due to the limitations of the scaling techniques, in most studies the data have been artificially made symmetric prior to analysis (Harshman, 1978).

It is likely, however, that flows through space have symmetric and asymmetric components. For example, the migration between two countries probably depends on open borders and geographical proximity (symmetric) but also on differences in population densities, economic opportunities, political stability and freedom, and in general, attractiveness (asymmetric; Tobler, 1976-77, 1979). Flows of air pollution, distribution of a species, telephone calls, postal correspondence, and vehicular traffic

probably have symmetric and asymmetric components (Harshman, 1981). So, too, would flows through time, e.g., the succession of plant species.

A complementary difficulty occurs in preference. Here the relations are supposed to be asymmetric, but they can be contaminated with symmetry. If I prefer chocolate to vanilla I should "disprefer" vanilla to chocolate. Preference reversals (Kahneman, & Tversky, 1979) violate this, and pose theoretical difficulties similar to those of asymmetric similarity.

An analogous situation arises any time we study change using more than one variable. For example, we might administer a battery of cognitive tests to two groups of children -- a group that is having difficulty learning to read, say, and a group that is not. For each group we can intercorrelate the tests, giving two symmetric matrices. Because of the symmetry, each matrix can be reduced to an upper or a lower triangular half matrix without loss of information. Thus, we can derive a matrix in which the upper half contains the correlations for one group and the lower half contains the correlations for the other group. The symmetry in this artificial matrix tells how the two groups of children resemble each other in the tests. The asymmetry tells how they differ; that is, the manner in which the correlations change with the presence or absence of reading difficulty. If we have more than two symmetric matrices that we wish to compare, we can average them in various ways to give two matrices, which can be compared as above. That is, we can use the technique for comparing two symmetric matrices as a way of conducting "single degree of freedom" contrasts, analogous to ANOVA, among more than two comparable matrices.

The problem of studying the effects of an independent variable on symmetric matrices of dependent variables is in fact exceedingly common. In MANOVA we partition the variances and covariances of the dependent variables into a between groups covariance matrix and a within groups covariance matrix. However, if we wished we could compute a separate dependent variable covariance matrix for each group, that is, for each level of the between-groups factor. Looking at the effects of the independent variable on the covariance of the dependent variables would be the same as looking at how reading difficulty status affects the correlations among cognitive tests. Thus, any time a MANOVA is conducted, techniques for jointly analyzing symmetry and asymmetry could be used to supplement our understanding of the data.

Clearly, then, a scaling method that is flexible enough to subsume symmetric and asymmetric components would have a great deal of utility. My goal in this dissertation is, first, to review a data-driven technique for analyzing asymmetric data at an absolute level of measurement. Then, by analogy to nonmetric multidimensional scaling I will develop an ordinal-level, data-driven method for analyzing asymmetries. As a third step, the behavior of the absolute-level and ordinal-level techniques will be studied using a computer program to simulate data with known characteristics. Once the techniques are better understood via simulation studies their utility will be demonstrated by applying them to asymmetric data sets already in the literature. Then the techniques will be applied to new similarity data collected to help clarify the relationship between asymmetric similarity and typicality differences.

## Chapter 2

### Dependence of Multidimensional Techniques on Symmetric Input

As context, let us first consider four interrelated topics:

- (1) The reason and way in which standard multidimensional techniques are dependent on symmetric input.
- (2) The relationship between the symmetric and purely asymmetric parts of a matrix.
- (3) A general, data-driven technique for analyzing pure asymmetries; that is, input for which there is no symmetric component.
- (4) The techniques that have been proposed for mixed symmetric-asymmetric data and a way of organizing them.
- (5) Techniques for testing the statistical significance of the asymmetries.

Standard techniques for scaling complex attributes use a similarity matrix as input. Here similarity is to be understood broadly as any indicator of consonance between the stimuli in a study. Traditional techniques require symmetric input because they use a symmetric quantity to represent similarity, and because the mathematics become unwieldy otherwise.

In metric and nonmetric multidimensional scaling, stimuli are modeled as points in a multidimensional space. The dissimilarity between two stimuli is represented in the model as the distance between the corresponding stimulus points.



Generally, distance is defined to be a symmetric quantity (Fréchet, 1906): the distance from New York to Boston equals the distance from Boston to New York. Although we can invent a space in which symmetry does not hold (e.g., Wilson, 1931), the space would lose much of the intuitive accessibility needed for exploratory modelling.

In principal components analysis and factor analysis the stimuli are represented, not as points in a multidimensional space, but as vectors drawn from the origin. In principal components and factor analysis, the similarity between two stimuli is modeled as the cosine of the angle between the stimulus vectors (Hotelling, 1933; Jackson, 1924). Of course the angle between vector "i" and vector "j" equals the angle between "j" and "i", and hence the representation of similarity is symmetric.

In the various forms of cluster analysis, similarity is indicated by membership in the same clusters. If object "i" is in the same cluster as "j", then "j" is in the same cluster as "i", and again the relation is modeled with an inherently symmetrical process (Lorr, 1983).

In some cases symmetry is inherent in the mathematics. Principal components and factor analysis separate a matrix into a set of eigenvectors. Each stimulus object has a coefficient, or loading, on each vector. The similarity between "i" and "j" is predicted from the product of their loadings on an eigenvector (Hotelling, 1933):

$$\hat{s}_{ij} = a_{iI} * a_{jI}$$

where  $\hat{s}_{ij}$  is the similarity between i and j predicted from component "I",  $a_{iI}$  is the loading of object i on component I, and  $a_{jI}$  is the loading of j on component I.

Because  $a_{iI} * a_{jI} = a_{jI} * a_{iI}$ ,  $\hat{s}_{ij} = \hat{s}_{ji}$ . Here, symmetry is imposed by the fact that

multiplication is commutative. This stricture can be avoided by allowing the loadings to take on imaginary values (Basilevsky, 1983). However, it is unclear how a representation in terms of imaginary numbers should be related to data comprised of real numbers. The mathematics do not prohibit modelling asymmetric similarity but do complicate it.

Classical MDS is computed on an interval-level similarity matrix by (1) squaring all entries; (2) subtracting the row and column mean from each entry; (3) adding the grand mean to each entry; and (4) conducting an eigenvector decomposition as is done in principal components (Torgerson, 1952). Thus, except in the unusual case that step (2) brings an asymmetric matrix to symmetry, asymmetric input will lead to an eigenvector decomposition of an asymmetric matrix. Hence, the mathematics discourage asymmetric analyses with metric MDS, in precisely the same way that they discourage the analysis of asymmetries with principal components.

Nonmetric MDS and most cluster analytic solutions are obtained with iterative, numerical algorithms. Here there is no computational stricture, and the requirement of symmetry is imposed only by how distance and category membership are usually defined. This allows us some freedom later in the dissertation for defining a nonmetric analysis of asymmetry.

The arguments presented in this section do not impede multidimensional techniques from accepting asymmetric input, so much as from providing asymmetric output. However, the two constraints are essentially interchangeable. As will be shown in the next chapter, the symmetric and asymmetric parts of a square matrix are

orthogonal. Therefore, a symmetric solution will be unable to account for the asymmetric aspects of the data. While principal components, cluster analysis and multidimensional scaling programs could be rewritten to allow asymmetries in the data, the asymmetries would be left unexplained.

### Chapter 3

#### The Relationship of Symmetric and Asymmetric Parts

In a square, asymmetric data matrix, the similarity of "i" to "j" does not in general equal the similarity of "j" to "i". That is, the value in row i, column j, is not constrained to equal the value in row j, column i:

$$x_{ij} \neq x_{ji} \text{ for at least one } i, j \text{ pair.}$$

We can impose symmetry on the matrix by replacing  $x_{ij}$  and  $x_{ji}$  with the average of the two:

$$x_{ij} \rightarrow x_{ij \text{ symm}} = .5*(x_{ij} + x_{ji})$$

and

$$x_{ji} \rightarrow x_{ji \text{ symm}} = .5*(x_{ij} + x_{ji}) = x_{ij \text{ symm}}$$

The difference between the original and "symmetrized" matrices is the information lost in the conversion to symmetry:

$$x_{ij \text{ resid}} = x_{ij} - .5*(x_{ij} + x_{ji}) = .5*(x_{ij} - x_{ji})$$

$$x_{ji \text{ resid}} = x_{ji} - .5*(x_{ij} + x_{ji}) = .5*(x_{ji} - x_{ij})$$

Note that  $x_{ij \text{ resid}} = -x_{ji \text{ resid}}$ . In the derived symmetric matrix, the values in cells i,j and j,i are equal. In the residual matrix, the values in corresponding cells are equal in magnitude but opposite in sign. The residual matrix would therefore be called "anti-symmetric" or "skew-symmetric" (Basilevsky, 1983). The skew-symmetric matrix is purely asymmetric; it contains no symmetric part, as  $x_{ij \text{ resid}} + x_{ji \text{ resid}} = x_{ij \text{ resid}} - x_{ij \text{ resid}} = 0$  for all i, j. The values in the skew-symmetric matrix are orthogonal to the

entries in the derived symmetric matrix. This can be seen by computing the sums of cross products of the numbers in the symmetric part with the numbers in the skew-symmetric part. For the off diagonal cells  $ij$  and  $ji$  we have

$$\begin{aligned}
 & x_{ij \text{ symm}} * x_{ij \text{ resid}} + x_{ji \text{ symm}} * x_{ji \text{ resid}} \\
 &= x_{ij \text{ symm}} * x_{ij \text{ resid}} + x_{ij \text{ symm}} * (-x_{ij \text{ resid}}) \\
 &= 0.
 \end{aligned} \tag{1}$$

This is true for all  $i \neq j$ . For the diagonal cells,  $x_{ii \text{ resid}} = 0$ , and therefore

$$\begin{aligned}
 & x_{ij \text{ symm}} * x_{ij \text{ resid}} + x_{ji \text{ symm}} * x_{ji \text{ resid}} \\
 &= x_{ij \text{ symm}} * 0 + x_{ji \text{ symm}} * 0 \\
 &= 0.
 \end{aligned} \tag{2}$$

Hence the sum of cross products between the symmetric and skew-symmetric parts, that is, the sum of Equations 1 and 2 across all  $i, j$ , is zero. Thus, an asymmetric matrix can be additively decomposed into orthogonal symmetric and skew-symmetric parts.

Preference and other "dominance" matrices are often regarded as ideally skew-symmetric. Consider, for example, ratings of ice cream flavors on a scale from -10 (strongly dislike) through 0 (neutral) to +10 (strongly like). If I prefer chocolate to vanilla by +5 units then presumably I would "prefer" vanilla to chocolate by -5 units. As another example, if tone "i" is +5 units louder than tone "j", then tone "j" should be -5 units "louder" than tone "i".

Because of the prevalence of preference and other dominance data, there are well-developed techniques for analyzing skew-symmetric matrices (e.g., Thurstone,

1927). Specifically, a one-dimensional scale having a least squares fit to the input matrix is given simply by the matrix's column averages (Mosteller, 1951).

Unfortunately, traditional approaches to purely asymmetric data have been one-dimensional, and hence not useful for our present purposes.

## Chapter 4

Canonical Analysis of Skew-Symmetric Data <sup>1</sup>

As noted, a square, asymmetric matrix can be additively decomposed into symmetric and skew-symmetric parts. We can apply standard techniques in scaling to the symmetric part, extracting components, clusters, dimensions, etc. Thus our focus will be on identifying a method for analyzing the skew-symmetries. In fact, it is relatively straightforward to derive a principal components-like solution for skew-symmetric matrices.

Principal components is a subset of a more general form of analysis called the singular value decomposition (Leon, 1990), Eckart-Young decomposition (Eckart, & Young, 1936), or basic structure of a matrix (Horst, 1965). Principal components is essentially restricted to symmetric input matrices, but all matrices have singular value decompositions (Leon, 1990).

In principal components, a correlation matrix is analyzed into a set of eigenvectors. The correlations are approximated from the component loadings by

$$\hat{x}_{ij} = a_{iI} * a_{jI} + a_{iII} * a_{jII} + a_{iIII} * a_{jIII} \dots$$

Similarly,

$$\hat{x}_{ji} = a_{jI} * a_{iI} + a_{jII} * a_{iII} + a_{jIII} * a_{iIII} \dots = \hat{x}_{ij} .$$

---

<sup>1</sup> Parts of this chapter were presented at the American Psychological Association annual convention in San Francisco, August, 1991, (Borkum, 1991b).

In the singular value decomposition, a matrix is analyzed into separate row and column eigenvectors:

$$\hat{x}_{ij} = (a_{iI} * b_{jI}) + a_{iII} * b_{jII} + a_{iIII} * b_{jIII} \dots$$

where  $a_{iI}$  is the loading of object  $i$  on row component  $I$ ,

$b_{jI}$  is the loading of object  $j$  on column component  $I$ , and the remaining terms are defined analogously. Note that

$$\hat{x}_{ji} = (a_{jI} * b_{iI}) + a_{jII} * b_{iII} + a_{jIII} * b_{iIII} \dots$$

which need not equal  $\hat{x}_{ij}$ .

The terms in parentheses show the reconstruction of  $\hat{x}_{ij}$  and  $\hat{x}_{ji}$  from the first row and column eigenvector. To predict  $\hat{x}_{ij}$ , the entry in row  $i$ , column  $j$  of the input matrix, we use the loading of  $i$  on the row eigenvector,  $a_{iI}$ , and the loading of  $j$  on the column eigenvector,  $b_{jI}$ . To predict  $x_{ji}$ , the entry in row  $j$ , column  $i$ , we use the loading of  $j$  on the row eigenvector,  $a_{jI}$ , and the loading of  $i$  on the column eigenvector,  $b_{iI}$ .

Because  $a_{iI}$  and  $b_{iI}$  can differ, and  $a_{jI}$  and  $b_{jI}$  can differ,  $\hat{x}_{ij}$  need not equal  $\hat{x}_{ji}$ . The predicted matrices are not constrained to symmetry.

We can think of principal components as the singular value decomposition of a symmetric correlation matrix. In a symmetric matrix the rows and columns are the same, so the two sets of eigenvectors are identical.

When a matrix is skew-symmetric each column is the negative of the corresponding row. Therefore, in the singular value decomposition, the row eigenvectors of a skew-symmetric matrix will be identical to the column eigenvectors, except for an alternation in sign and the order in which they are extracted (Basilevsky,



1983; Gower, 1977). For example, the first column eigenvector equals the second row eigenvector:

$$b_{ii} = a_{iii} .$$

The second column eigenvector equals the negative of the first row eigenvector:

$$b_{iii} = -a_{ii} .$$

The same relationships hold between the third and fourth eigenvectors, the fifth and sixth eigenvectors, etc. This simplifies the results of the singular value decomposition and makes it quite analogous to the principal components decomposition of the symmetric part of the data. The identity between row and column eigenvectors enables the singular value decomposition to reconstruct the skew-symmetry:

$$\hat{x}_{ij} = a_{ii} * b_{ji} + a_{iii} * b_{jii} \dots$$

But as noted above,  $b_{ji} = a_{jii}$  and  $b_{jii} = -a_{ji}$ , so

$$\hat{x}_{ij} = a_{ii} * a_{jii} - a_{iii} * a_{ji} \dots$$

Analogously,

$$\hat{x}_{ji} = a_{jii} * a_{iii} - a_{ii} * a_{jii} \dots = -\hat{x}_{ij} .$$

Taken together, the first two eigenvectors define the elementary skew-symmetric matrix that best approximates, in a least-squares sense, the skew-symmetric input (Basilevsky, 1983). Similarly the third and fourth eigenvectors define an elementary skew-symmetric matrix that best approximates the residuals left by the first two eigenvectors. This property holds for all pairs of eigenvectors.

Given that the eigenvectors work in pairs, it is perhaps unsurprising that the eigenvalues also occur in pairs. That is, the first two eigenvalues are equal in

magnitude, opposite in sign, and tell the amount of variation in the input matrix that is explained by the first elementary skew-symmetric matrix. Similarly the third and fourth eigenvalues differ from each other only in their signs, and indicate the match between the input and the second elementary skew-symmetric pattern, and so on. If we extract an odd number of eigenvectors, the last eigenvalue will be zero, as the last, unpaired eigenvector cannot improve the representation of skew-symmetry.

Aside from this pair-wise property, the eigenvalues of a skew-symmetric matrix are comparable to those of a symmetric matrix. Moreover, successive pairs of eigenvectors provide an optimal approximation to the input matrix, in the sense of minimizing the sum of squared residuals.

To understand the eigenvectors, consider the formulas for predicting entry  $x_{ij}$ :

$$\hat{x}_{ij} = a_{jI}a_{II} - a_{jII}a_{iI} + \dots$$

and entry  $x_{ji}$ :

$$\hat{x}_{ji} = a_{jII}a_{iI} - a_{jI}a_{II} + \dots$$

Asymmetries will occur when  $a_{jII}a_{iI}$  is much greater than  $a_{jI}a_{II}$ , i.e., when  $a_{iI}$  is larger than  $a_{II}$  and when  $a_{jII}$  is larger than  $a_{jI}$ . We can imagine that the eigenvectors reflect complementary properties, say, an ability to transmit information (component I) and an ability to receive information (component II). Asymmetry will occur between an object that is specialized as a "transmitter" ( $a_{iI}$  is high,  $a_{II}$  is low) and an object specialized as a "receiver" ( $a_{jII}$  is high,  $a_{jI}$  is low). If both objects are transmitters, or if both are receivers, or if both transmit and receive with equal facility, the asymmetries between them will be mild or absent altogether. For example, if object i

could transmit and receive equally well, that is, if  $a_{iI} = a_{iII}$ , and if object  $j$  were similarly nonspecialized,  $a_{jI} = a_{jII}$ , then

$$\begin{aligned}\hat{x}_{ij} &= a_{jI}a_{iII} - a_{jII}a_{iI} \\ &= a_{jI}a_{iI} - a_{jI}a_{iI} \\ &= 0\end{aligned}$$

and similarly

$$\begin{aligned}\hat{x}_{ji} &= a_{jII}a_{iI} - a_{jI}a_{iII} \\ &= a_{jI}a_{iI} - a_{jI}a_{iI} \\ &= 0\end{aligned}$$

(Harshman, 1981).

The two formulas

$$\hat{x}_{ij} = a_{jI}a_{iII} - a_{jII}a_{iI} + \dots$$

and

$$\hat{x}_{ji} = a_{jII}a_{iI} - a_{jI}a_{iII} + \dots$$

have a geometric as well as a conceptual interpretation (Constantine, & Gower, 1978; Gower, 1977). Suppose we represent components I and II as orthogonal axes in two-dimensional space. Then object  $i$  can be plotted as a point whose x-coordinate is  $a_{iI}$  and whose y-coordinate is  $a_{iII}$ . The skew-symmetric relationship between  $i$  and  $j$  cannot be given by the distance between points  $i$  and  $j$ , nor by the angle between line segments joining these points with the origin, because distance and angle are symmetric relations. Note, however, that if  $a_{iI}$  is much greater than  $a_{iII}$ , then the object point for  $i$  will be near the 3 o'clock position in the graph. If  $a_{jII}$  is much

greater than  $a_{ji}$  then the object point for  $j$  will be near the 12 o'clock position. Thus, the area of the triangle formed by the two object points and the origin will be relatively large. It is this area that represents the skew-symmetries. To represent skew-symmetry, however, the area must be signed: if, in a clockwise "radar-sweep" around the origin,  $j$  follows  $i$  by less than 180 degrees, then we say that object " $i$ " dominates object " $j$ " in the model, and  $\hat{x}_{ij} > \hat{x}_{ji}$ . If  $j$  precedes  $i$  in the clockwise sweep or, equivalently,  $j$  follows  $i$  by more than 180 degrees, then we say that object  $j$  dominates object  $i$  in the model, and  $\hat{x}_{ji} > \hat{x}_{ij}$ . (Because one of the dimensions may be reflected, it is always necessary to determine empirically which direction in an analysis, clockwise or counterclockwise, corresponds to dominance.)

In fact, it turns out that the signed triangular area is proportional to the skew-symmetries, and that the constant of proportionality is two. That is, the skew-symmetric relationship between  $i$  and  $j$  is equal to twice the signed triangular area bounded by the two object points and the origin. A demonstration of this is given in Appendix 3. If object  $i$  were not specialized as a transmitter or a receiver, that is, if  $a_{ji} = a_{ij}$ , and if object  $j$  were similarly nonspecialized,  $a_{ji} = a_{ij}$ , then the object points for  $i$  and  $j$  would fall along a straight line, 45° from the  $x$  and  $y$  axes. The "triangle" formed by the two object points and the origin would actually be a straight line, that is, a "triangle" whose area is zero.

Triangles, of course, are inherently two-dimensional. As we add more dimensions to the solution, we increase the number of triangles used to approximate the relationship between any two objects. When eigenvectors are first extracted from

the skew-symmetric matrix, the only possible triangles are on the plane defined by dimensions I and II, the plane defined by dimensions III and IV, the plane defined by V and VI, and so on (Harshman, 1981). Dimension I, for example, does not interact with any dimensions other than II. Dimension II does not interact with any dimensions other than I. Thus, if  $T$  dimensions are extracted, there are at most  $T/2$  triangles between points "i", "j", and the origin. The sum of the signed areas of these triangles gives the approximation of the dominance relation between objects "i" and "j". We may wish to rotate the dimensions to simple structure, so that each dimension has as many zero loadings on it as possible. However, there is a tradeoff. Although the dimensions are simpler, the number of interactions between dimensions may increase. Triangles may now be defined, for example, on the planes spanned by dimensions I and III, I and IV, II and III, II and IV, etc. Harshman (1981) describes in greater detail the issues involved in the rotation of skew-symmetric representations.

The geometric interpretation of the eigenvectors of a skew-symmetric matrix was first discussed by Gower (1977), and the graphical display is called a "Gower diagram" by Harshman (1981). The geometric interpretation will be useful to us later, when we consider nonmetric techniques.

To scale a square, asymmetric matrix, then, we can separate it into symmetric and skew-symmetric parts. Each part can then be analyzed into its the singular value decomposition. We can then compute a canonical redundancy analysis (Stewart, & Love, 1968) to determine the degree to which the components underlying the symmetric and skew-symmetric parts are the same. Because the symmetric and skew-

symmetric parts of the input are orthogonal sources of information, significant overlap would be a substantive property of the data.

A great advantage of singular value decomposition is its ease of computation with standard statistical software packages. Begin with a skew-symmetric matrix. From this, derive the raw score sum of squares, sum of cross products (SCP) matrix. In this derived matrix, the entry in row 1, column 1, is the raw score sum of squares of row 1 from the original, skew-symmetric matrix. The entry in row 1, column 2 of the SCP matrix is the raw score sum of cross products between row 1 and row 2. The other entries in the SCP matrix are defined analogously. The SCP matrix is symmetric, because the sum of cross products between rows "i" and "j" equals the sum of cross products between rows "j" and "i". The SCP matrix, then, could be analyzed using a standard principal components procedure. In fact, this is precisely its advantage because, as shown in Appendix 2, the eigenvectors of the SCP matrix are identical to the eigenvectors of the original skew-symmetric matrix (Basilevsky, 1983).

Computation, then, involves just four steps:

- (1) The asymmetric input matrix is transposed so that column 1 becomes row 1, etc. The transposed matrix is added to the original asymmetric matrix, and all entries are divided by 2. The resulting matrix is the symmetric part of the data. It can be analyzed into its eigenvectors using a principal components routine.

(2) The transposed matrix is subtracted from the original matrix. The resulting matrix is the skew-symmetric part of the data.

(3) A sum of squares, sum of cross products (SCP) matrix is computed from the derived skew-symmetric matrix.

(4) The eigenvectors of the SCP matrix are extracted using a principal components routine.

One caution is in order: the eigenvalues of the SCP matrix are equal to the squares of the eigenvalues of the original matrix. Thus, the analysis of the symmetric part of the data is not strictly comparable to the analysis of the skew-symmetric part. In the analyses that follow this does not seem to pose a difficulty, but it can be corrected by taking square roots of the skew-symmetric eigenvalues.

The singular value decomposition of a skew-symmetric matrix is also known as the canonical decomposition of asymmetry (Gower, 1977). It is a variant of one of the DEDICOM models (Harshman, 1978). Although it is general purpose, data-driven, analogous to principal components, and enjoys least squares properties, it has rarely if ever been used.

Rather, there is a scattered literature of numerous special purpose techniques for handling asymmetries. In the next chapter I will review this literature, and group the special purpose techniques by the assumptions they make about the data. In addition to considerable usefulness in its own right, the canonical analysis of

**asymmetry is a convenient way of examining whether the data meets some of the assumptions underlying special purpose analyses.**



## Chapter 5

### Current Techniques for Mixed Symmetric - Skew-Symmetric Data

In the vast majority of studies employing MDS, principal components, or factor analysis for scaling purposes, similarity is simply assumed to be symmetric (Harshman, 1978). Generally, each pair is presented in only one direction. Similarity values for both directions of comparison are set equal to the single value obtained.

Nonetheless, methods of analysis for data that are neither purely symmetric nor purely skew-symmetric have been proposed sporadically for about 40 years. To my knowledge they have not been collected until the present work. Here they are listed in approximately decreasing order of the strength of assumptions they make about the asymmetric part of the data. Methods early in the list tend to be theory-driven; those later in the list are more data-driven.

(1) Imposed symmetry. A symmetric matrix can be derived from the initial input by replacing all  $x_{ij}$  and  $x_{ji}$  with their average,  $.5*(x_{ij} + x_{ji})$ . This technique, and the even simpler approach of collecting only  $x_{ij}$  and assuming  $x_{ji} = x_{ij}$ , has been used in virtually all multidimensional approaches to scaling (Harshman, 1978). It yields a symmetric matrix on which principal components, cluster analysis, and multidimensional scaling can be computed. Aside from convenience, there are two possible rationales for its use.

First, the asymmetries may be random error. This is rarely if ever subjected to statistical test. However, tests exist, and are presented at the end of this chapter.

Second, the asymmetries may be due to "constant error" in which the entry for every pair is displaced up or down by a fixed increment caused by the experimental design. For example, suppose that two tones are presented in sequence, and that the first tone is always rated two units higher than it normally would be, simply because it is presented first. Suppose further that the true value of "i" is +5 and that the true value of "j" is +3. When "i" is presented first, the difference between the two tones,  $x_{ij}$ , is  $(5+2) - 3 = +4$ . When "j" is presented first the difference is  $x_{ji} = 5 - (3+2) = 0$ . If we isolate the skew-symmetric part of the data we find

$$x_{ij \text{ resid}} = .5*(4 - 0) = 2$$

and

$$x_{ji \text{ resid}} = .5*(0 - 4) = -2.$$

The same amount of deviation from symmetry, +2 vs. -2, would be found for all pairs of stimuli, and could be removed by imposing symmetry on the input matrix. One test of the constant error hypothesis would be to isolate the skew-symmetric part of the data and see if in fact the deviations were all of the same amount. The mean of the above-diagonal entries, and the negative of the mean of the below-diagonal entries, would be computed. Then the average of these two numbers would be subtracted from the above-diagonal entries, and added to the below-diagonal entries. The resulting off-diagonal deviation scores should be close to zero, and should have a mean square that reflects chance only. If an independent measure of chance is available, an  $F$  test may be feasible, as described at the end of this chapter. Alternatively, one could examine the row or column sums. In a constant error matrix

the sums should decline or increase in a strictly linear fashion. The results of the canonical analysis also take on a characteristic form for constant-error data: The sum of cross products matrix for constant error is a perfect simplex. Hence, the object points lie along the circumference of a half circle.

(2) Categorical judgment. Harris (1957) applied a version of the Method of Successive Intervals to asymmetric, square matrices. This provides a one-dimensional scaling of the row objects, and a one-dimensional scaling of the columns. The row and column scales are constrained to be linear transformations of each other, differing at most by an arbitrary unit of measurement, and by an additive constant. The additive constant is equivalent to a single skew-symmetry dimension that should emerge in a CAA solution. The differences in unit of measurement between rows and columns would be analogous to a multiplicative bias coefficient, and should emerge as a second skew-symmetry dimension.

(3) Multidimensional unfolding (MDU). In multidimensional unfolding, the row objects and column objects are plotted as separate points in a joint space. If, as is usually the case in skew-symmetric data, the row and column objects are the same, each object will be represented twice in the solution. Multidimensional unfolding was suggested as one option by Gower (1977), but it has several disadvantages.

(a) Representing the same object by two different points can be confusing.

(b) If the diagonal entries are included in the analysis, then the distance between an object's two points would reflect both the degree to which it enters into asymmetric relations with other objects, and its degree of self-similarity. A purely

skew-symmetric matrix could not properly be represented: all relationships between objects would be asymmetric, requiring that the two points for each object be widely spaced. However, the diagonal entries in a skew-symmetric matrix are zero, suggesting that each object's self-similarity is very high.

(c) In multidimensional unfolding, the symmetric and asymmetric aspects of the data are represented on the same "map". If the asymmetries are numerically small but theoretically interesting, they may participate little in the solution.

(d) In multidimensional unfolding most of the information in the data matrix is ignored, leading readily to degenerate solutions (Zielman & Heiser, 1993).

In multidimensional unfolding the symmetries and skew-symmetries are accounted for with the same dimensions. In a canonical analysis of asymmetry, we would compute the eigenvectors of the symmetric and skew-symmetric parts separately and relate them with a canonical redundancy analysis. If the redundancies approached 100%, the multidimensional unfolding model would receive support.

(4) In the "wind" (Wish, 1967) and gradient (Tobler, 1976-77, 1979) models, the symmetric part of the data provides a multidimensional scaling solution or similar map, with the objects represented as points on the map. The symmetric relations are modeled as the distances between object points. The asymmetric relations are pictured as vectors, directions, or "lines of force" within the map. This implies that the number of dimensions underlying the asymmetries is less than or equal to the number of dimensions for the symmetric part of the data. In practice, the maps are usually two-dimensional, so that the asymmetries must be one- or two-dimensional. This can

be checked by computing a canonical analysis of asymmetry to determine the actual number of dimensions.

(5) In the "drift-" or "slide-vector" model, asymmetries are presumed to be one-dimensional (Zielman & Heiser, 1993). The symmetric part of the data is represented by the distance between object points in a multidimensional scaling configuration. The skew-symmetric part of the data is modeled as a single vector, of a given direction within the configuration. If the separation between two object points is orthogonal to the drift vector, the relationship between the objects is symmetric in the model. If two object points are separated from each other along the length of the vector, the objects will have an asymmetric relationship in the model. The more two objects are separated from each other, and the more their direction of separation coincides with the direction of the drift vector, the greater their asymmetry.

The drift vector can be thought of as a way of introducing asymmetry into the distances between points: It is easier to travel in the direction of the vector than to oppose it. In the drift vector model, the asymmetries are one-dimensional. Moreover, this dimension can be completely specified as a linear combination of the dimensions underlying the symmetric part of the data. This can be tested by determining whether (a) the skew-symmetric part of the data yields a single eigenvector; and (b) whether the canonical redundancy of this eigenvector on the eigenvectors underlying the symmetric part of the data approaches 100%.

(6) In Young's (1975) ASYMSCAL the similarities are analyzed into a spatial MDS configuration. The asymmetries are attributed to a different weighting of the

symmetric dimensions induced by each row stimulus. An implication of the model is that the same number of dimensions underlie the symmetric and skew-symmetric parts of the data. This can be tested with the canonical analysis of asymmetry.

(7) Spatial density and contrast models. The spatial density (Krumhansl, 1978) and feature contrast (Tversky, 1977) models resemble the drift vector approach by providing one-dimensional accounts of the asymmetries. However, the drift vector was a direction in the space underlying the symmetric part of the data. The vector underlying the asymmetries could in theory be perfectly specified by a linear combination of the dimensions for the symmetric part. In the feature contrast and spatial density models there is no necessary relationship between the dimension underlying the asymmetries and those underlying the symmetric part. Rather, the asymmetry dimension is hypothesized to be related to an external characteristic of the objects. In the feature contrast model this is the salience or typicality of the objects. Higher similarity judgments are expected when comparing a less salient object to a more salient object than for the reverse comparison. In the spatial density model higher similarity judgments are expected when comparing an object with few close neighbors in the MDS solution, to an object with many close neighbors.

The canonical analysis of asymmetry can provide the start of a test of the assumptions, by determining whether a single dimension underlies the asymmetries. Thereafter, however, it is necessary to test the models more precisely, by correlating the asymmetric dimension with an independent measure of stimulus salience or with the number of near neighbors of a stimulus in the MDS space.

(8) **Bias coefficients.** The general bias coefficient model, of which the spatial density and additive feature contrast models are special cases (Nosofsky, 1991) posits simply that the rows are weighted differently than the columns. This implies that the asymmetry is one-dimensional. The asymmetry dimension is otherwise unconstrained by hypothesis. The assertion of one-dimensional asymmetries can be assessed by using the canonical analysis to determine the actual number of vectors underlying the skew-symmetric part of the matrix.

(9) **Canonical analysis of asymmetry.** This procedure was discussed above. An asymmetric matrix is additively decomposed into orthogonal symmetric and skew-symmetric parts. A sum of squares, sum of cross products matrix is derived from each part and the eigenvectors are extracted. This can be thought of as an extension of a principal components-like approach to an asymmetric matrix. The canonical analysis of asymmetry is data-driven and does not make assumptions about the asymmetries.

(10) **DEDICOM, or the "Decomposition into Directional Components"** (Harshman, 1978, 1981) closely resembles the canonical analysis of asymmetry. In DEDICOM, a matrix of asymmetric relations between objects is divided into a set of factors underlying the objects and an asymmetric matrix of interrelationships among the factors. There may be one set of factors (single-domain model) or two, depending on whether we expect the same set of factors to underlie the objects in their row and column roles. The number of factors is usually small compared with the number of objects in the study. As a result, the matrix of factor interrelationships is generally

much smaller and easier to interpret than the interrelationships among the objects. For a symmetric input matrix, DEDICOM reduces to principal components analysis. For skew-symmetric input, DEDICOM yields the singular value decomposition or canonical analysis of asymmetry. DEDICOM is data-driven and does not require assumptions about the asymmetries.

(11) Separate row and column solutions. In an asymmetric matrix the rows do not equal the columns. In the general case the dimensions underlying the rows would not be the same as the dimensions underlying the columns. Therefore, one approach to fully representing the input is to compute separate MDS configurations for the rows and the columns. This is the approach taken in Smallest Space Analysis for asymmetric data (Lingoes, 1972). As in DEDICOM and CAA, no assumptions are made about the asymmetries.

Although this approach is data-driven and exhaustive, it may not be the most useful way of representing the data. First, the row objects are treated as different than the column objects, even though in most of our matrices the row and column objects will be the same. This is not a difficulty if we hypothesize that different dimensions underlie the objects in their row and column roles. However, there is no equivalent of the single-domain DEDICOM model. Second, the partition into row and column solutions may not be as meaningful as the partition into symmetric and skew-symmetric components.



### Significance testing of the asymmetries

In similarity data, the asymmetries are often numerically small, and hence compete with the hypothesis that they are random error. Several statistical tests are possible:

(1) For a matrix whose entries are frequency counts or probabilities: We can use a chi-square test to compare the expected entries under the hypothesis of symmetry with the actual, observed cell entries (Bishop, Fienberg, & Holland, 1975). The expected entries are simply the entries in the symmetrized matrix:  $E_{ij} = E_{ji} = .5*(X_{ij} + X_{ji})$ . Cells on the main diagonal are left out of the test because they are symmetrical by definition. The chi-square is evaluated at  $.5*(K^2 - K)$  degrees of freedom, where K is the number of objects in the study.

(2) For parametric data, in which a number of subjects have each given similarity ratings for all possible pairs of K objects, significance testing can be developed by analogy to ANOVA. Five sources of variance may be identified: (1)  $AB_{\text{symm}}$ : the variation among the symmetries in the data, after averaging across subjects; (2)  $AB_{\text{asym}}$ : the variation in the data's skew-symmetries, after averaging across subjects; (3) S: the variation among subject means, after averaging across the stimulus pairs; (4)  $AB_{\text{symm}} \times S$ : how the symmetries interact with (vary across) subjects; and (5)  $AB_{\text{asym}} \times S$ : how the asymmetries interact with (vary across) subjects. Using formulas given in Keppel (1982, pp. 635-642) we can derive the expected mean square for each of the five variance components in the design:

$$\begin{array}{ll}
 \text{AB}_{\text{symm}} & \sigma_e^2 + \sigma_{\text{ABsymm} \times \text{S}}^2 + S\sigma_{\text{ABsymm}}^2 \\
 \text{AB}_{\text{asym}} & \sigma_e^2 + \sigma_{\text{ABasym} \times \text{S}}^2 + S\sigma_{\text{ABasym}}^2 \\
 \text{S} & \sigma_e^2 + \sigma_{\text{ABsymm} \times \text{S}}^2 + \sigma_{\text{ABasym} \times \text{S}}^2 + k^2\sigma_s^2 \\
 \text{AB}_{\text{symm}} \times \text{S} & \sigma_e^2 + \sigma_{\text{ABsymm} \times \text{S}}^2 \\
 \text{AB}_{\text{asym}} \times \text{S} & \sigma_e^2 + \sigma_{\text{ABasym} \times \text{S}}^2
 \end{array}$$

As may be seen, the ratio

$$\text{AB}_{\text{asym}} / (\text{AB}_{\text{asym}} \times \text{S}) = (\sigma_e^2 + \sigma_{\text{ABasym} \times \text{S}}^2 + S\sigma_{\text{ABasym}}^2) / (\sigma_e^2 + \sigma_{\text{ABasym} \times \text{S}}^2)$$

has an expected value of 1, if  $\sigma_{\text{ABasym}}^2$ , the true average asymmetric variance in the population, is zero. Hence, this ratio would be a reasonable test statistic for the asymmetries. As a ratio of two sample variances, it should be distributed as F. The degrees of freedom would be

$$\text{numerator: } .5(k^2 - k)$$

$$\text{denominator: } .5(s-1)(k^2 - k)$$

where  $s$  is the number of subjects, and  $k$  is the number of objects or stimuli, in the study.

The asymmetric part of the matrix could be extracted for each subject. The asymmetry by subject interaction would be a logical error term for most purposes. If we wished to include level differences among subjects as a source of error, the pooled within-asymmetries variance would be a reasonable choice for an error term.

## Chapter 6

### Nonmetric Analysis of Skew-Symmetries

In the canonical analysis of asymmetry, the asymmetric relation between two objects,  $i$  and  $j$ , say, is modeled as the difference between two cross-products:

$$\hat{x}_{ij} = r_i s_j - r_j s_i$$

The canonical analysis of asymmetry gives a least squares representation of the asymmetry on what is essentially an absolute scale of measurement.  $\hat{x}_{ij}$  approximates the actual asymmetry values,  $x_{ij}$ , without changing their mean, standard deviation, rank order, etc. The geometric interpretation is on a ratio level. If  $r$  and  $s$  are plotted as orthogonal vectors, the area of the triangle formed by the origin and the object points  $(r_i, s_i)$  and  $(r_j, s_j)$  is equal to one-half of  $\hat{x}_{ij}$ , the predicted asymmetry between  $i$  and  $j$ .

Most data in psychology, however, is thought to correspond to an ordinal level of measurement (Stevens, 1946). In attempting an absolute or a ratio level of representation, we would be allowing the scale mean and/or the scale mean and standard deviation, plus any unknown monotonic transforms that have altered the data, to influence the representation. However, the scale mean, standard deviation, and order-preserving transforms are arbitrary in ordinal-level data. Therefore the representation would include incidental features of the data and be less parsimonious.

It seems reasonable, then, to generalize the canonical analysis of asymmetry, so that it is making use of only the rank order of the asymmetries in the data. We can readily do so, by analogy to the nonmetric multidimensional scaling of symmetric

data. In what follows, the nonmetric generalization will be called nonmetric, skew-symmetric multidimensional scaling, or NSKMDS. Although the feasibility of such an undertaking was noted by Gower (1977), the actual development of a nonmetric, skew-symmetric technique appears to be new in this dissertation.<sup>2</sup>

In NSKMDS, as in the canonical analysis of asymmetry, the predicted asymmetry values are given as the differences between sums of cross products

$$\hat{x}_{ij} = r_i s_j - r_j s_i$$

and interpreted geometrically as twice the area of the triangle formed by the object points  $i$  and  $j$ , and the origin. As in the canonical analysis of asymmetry we will seek a least squares representation. The difference from the canonical analysis of asymmetry, is that in NSKMDS we will require only that the predicted asymmetry values  $\hat{x}_{ij}$  have the same rank order as the actual asymmetries  $x_{ij}$ .

To accomplish this we must first define what we mean by a least squares, ordinal level relationship. This is provided by Kruskal's Stress Formula 2 (Borg & Lingoes, 1987; Kruskal & Wish, 1977) modified for skew-symmetric input:

$$S_2 = \sum_{ij} [\hat{x}_{ij} - \delta_{ij}]^2. \quad (3)$$

$S_2$  is a "badness of fit" measure, as it gives the sum of squared departures of the predicted asymmetries,  $\hat{x}_{ij}$ , from the quantities they are designed to represent,  $\delta_{ij}$ . Note that if this were the canonical analysis of asymmetry the formula would show  $x_{ij}$  in place of  $\delta_{ij}$ : we would be trying to minimize the sum of squared departures of the

---

<sup>2</sup> Parts of this chapter were presented at the American Psychological Association annual convention in New York City in August, 1995, (Borkum, 1995).

predicted from the actual asymmetry values. Instead, however, we seek only a least squares correspondence between the predicted asymmetries and a new quantity,  $\delta_{ij}$ . The  $\delta_{ij}$  values are called "disparities" in nonmetric MDS, and we will adopt the same terminology here. Disparities are numbers chosen to be as close as possible to the predicted asymmetries  $\hat{x}_{ij}$ , with the proviso that they maintain the same rank order as the actual asymmetries  $x_{ij}$ . Thus the disparities are a device for representing the rank order of the actual asymmetries in the badness of fit formula. If we can find  $\hat{x}_{ij}$  values that minimize  $S_2$ , the  $\hat{x}_{ij}$ 's will approximate the rank order among the actual asymmetries, to a least squares criterion.

It should be noted that, as written, the  $S_2$  badness of fit formula is incomplete. The value of  $S_2$  could be infinitely reduced simply by shrinking the configuration to a single point at the origin. Then all disparities  $\delta_{ij}$  and all predicted asymmetries  $\hat{x}_{ij}$  would be zero, mimicking perfect goodness of fit. In the actual Stress Formula 2 this is prevented through normalization, that is, by dividing Equation 3 by the sum of squared deviation scores of the  $\delta_{ij}$ 's.

$$S_2 = \frac{\sum_{ij} [\hat{x}_{ij} - \delta_{ij}]^2}{\sum_{ij} [\delta_{ij} - \delta_{\text{mean}}]^2}$$

Then, any shrinkage in the configuration would reduce the denominator as well as the numerator, preventing a decrease in  $S_2$ . The normalization can be carried out as a separate step, however, and does not affect the derivation that follows.

It should also be remembered that the  $\hat{x}_{ij}$  values in Equation 3 are not infinitely free to vary. In NSKMDS, as in the canonical analysis of asymmetry, they are derived from the representation of the objects  $i$  and  $j$  in terms of the vectors  $r$  and  $s$ :

$$\hat{x}_{ij} = r_i s_j - r_j s_i$$

The  $\hat{x}_{ij}$  values derive from the position of objects  $i$  and  $j$  in a two-dimensional space. To reduce badness of fit, we must adjust the objects' coordinates  $(r_i, s_i)$ ,  $(r_j, s_j)$  in this spatial model so that the predicted asymmetries  $\hat{x}_{ij}$  will approach the disparities  $\delta_{ij}$  in Equation 3.

Because the  $\hat{x}_{ij}$  values must derive from the location of object points  $(r_i, s_i)$  and  $(r_j, s_j)$  in a two-dimensional space, the  $\hat{x}_{ij}$ 's will generally not give a perfect match to the  $\delta_{ij}$ 's. And because the  $\delta_{ij}$ 's must maintain the same rank order as the actual asymmetries in the data, the  $\delta_{ij}$  values will generally not match the  $\hat{x}_{ij}$ 's exactly. Therefore,  $S_2$  will rarely equal 0. Our goal is simply to minimize  $S_2$ , to obtain a good approximate solution.

Our task, then, is two-fold. We must find  $\delta_{ij}$  values that are as close as possible to the predicted asymmetries  $\hat{x}_{ij}$  while maintaining the same rank order as the actual asymmetries  $x_{ij}$ . Then, we must find predicted asymmetries that are as close as possible to the  $\delta_{ij}$ 's, by adjusting the coordinates of objects  $i$  and  $j$  on the vectors  $r$  and  $s$ . In each of these two steps, "as close as possible" will mean "in a least squares sense".

#### Finding the $\delta_{ij}$ values

The  $\delta_{ij}$ 's can be obtained through Kruskal's (1964b) block averaging algorithm. First, arrange the actual asymmetries  $x_{ij}$  in ascending order. There are usually  $.5*(K^2 - K)$  asymmetries, one for each pair of objects. Thus, sorting the asymmetries also

sorts the object pairs. Now, next to each actual asymmetry value  $x_{ij}$  place the predicted asymmetry  $\hat{x}_{ij}$ , computed from

$$\hat{x}_{ij} = r_i s_j - r_j s_i ,$$

for the same pair of points. Thus, if the observed asymmetry between  $i$  and  $j$  is 4.3, say, then next to 4.3 place the predicted asymmetry between the same object points  $i$  and  $j$ . An example is shown in Table 1. If the predicted asymmetries (column 2 in Table 1) had a perfect monotonic relationship to the actual asymmetries (column 1), then the  $\hat{x}_{ij}$ 's would increase or hold steady as we went down the table, just as the  $x_{ij}$ 's do.

To make this happen, start at the top of the table, and read down until coming to the first pair where  $\hat{x}_{ij}$  drops instead of increasing or holding steady. If this is the ninth pair, say, average the  $\hat{x}_{ij}$  for the eighth and ninth pairs, and use this mean value in place of the eighth and ninth  $\hat{x}_{ij}$ 's (column 3). If this mean value is less than the seventh  $\hat{x}_{ij}$ , average the seventh, eighth, and ninth  $\hat{x}_{ij}$ 's, and use this value instead of the three original  $\hat{x}_{ij}$ 's. This mean  $\hat{x}_{ij}$  is compared with the sixth  $\hat{x}_{ij}$ , etc. Eventually, even if we have to go back and average the first nine  $\hat{x}_{ij}$ 's, we will have a string of  $\hat{x}_{ij}$ 's that does not decline (column 5).

Proceeding in this way through the list we can be sure that the predicted asymmetries  $\hat{x}_{ij}$  will increase or stay the same, and thus have a perfect monotonic relationship to the observed asymmetries  $x_{ij}$ . Any  $\hat{x}_{ij}$  that we do not need to average is left unchanged. Thus it is indeed as close as possible (i.e., identical) to the original

$\hat{x}_{ij}$ . When we have to average two or more  $\hat{x}_{ij}$ 's to achieve monotonicity, the mean  $\hat{x}_{ij}$  is as close as possible to the original  $\hat{x}_{ij}$ 's in a least squares sense. Thus, the  $\hat{x}_{ij}$ 's,

Table 1

Monotone Regression<sup>1</sup>

Skew-Symm Data	Skew-Symm Computed From NSKMDS Model	1st Pass	2nd Pass	Final Result
0.7	0	0	0	0
0.8	3	3	3	3
1.9	3	3	3	3
2.0	5	5	5	5
2.3	8	8	8	8
2.4	9	9	9	9
3.8	12	12	10.67	10.67
5.5	13	10.5	10.67	10.67
6.2	8	10.5	10.67	10.67
7.1	14	14	14	14

<sup>1</sup> Note. Adapted from Borg and Lingoes, 1987, p. 35.

after smoothing with the averaging technique, are the  $\delta_{ij}$ 's. They have a perfect monotonic relationship to the observed asymmetries, and they are as close as possible to the  $\hat{x}_{ij}$  predicted asymmetries.



### Optimizing the $\hat{x}_{ij}$ values

Although our goal is to find values of  $\hat{x}_{ij}$  that minimize  $S_2$ , we must do so indirectly, by adjusting the coordinates of  $i$  and  $j$  on vectors  $r$  and  $s$  in

$$\hat{x}_{ij} = r_i s_j - r_j s_i . \quad (4)$$

Therefore let us substitute Equation 4 into the formula for  $S_2$ :

$$S_2 = \sum_{ij} [\hat{x}_{ij} - \delta_{ij}]^2.$$

$$S_2 = \sum_{ij} [r_i s_j - r_j s_i - \delta_{ij}]^2.$$

Squaring gives

$$S_2 = \sum_{ij} [r_i^2 s_j^2 - 2r_i r_j s_i s_j - 2r_i s_j \delta_{ij} + r_j^2 s_i^2 + 2r_j s_i \delta_{ij} + \delta_{ij}^2]$$

Let us focus at first on adjusting the coordinate of a single object,  $i$ , on vector  $r$ :

$$S_{2 \text{ for } i} = r_i^2 \sum_j s_j^2 - 2r_i s_i \sum_j r_j s_j \\ - 2r_i \sum_j s_j \delta_{ij} + s_i^2 \sum_j r_j^2 + 2s_i \sum_j r_j \delta_{ij} + \sum_j \delta_{ij}^2$$

We want the value of  $r_i$  that will minimize  $S_2$ . We can find this by setting the first partial derivative of  $S_2$  with respect to  $r_i$  equal to zero, and solving for  $r_i$ :

$$0 = 2r_i \sum_j s_j^2 - 2s_i \sum_j r_j s_j - 2\sum_j s_j \delta_{ij} \\ r_i = [s_i \sum_j r_j s_j + \sum_j s_j \delta_{ij}] / \sum_j s_j^2 \quad (5)$$

Note that because the derivative is linear with respect to  $r_i$  there is only one solution, that is, there are no multiple roots. Moreover, we can assume that this solution is not a maximum because badness of fit, as assessed by Equation 4, can be made worse without limit.

Therefore, Equation 5 tells us the coordinate to assign to object  $i$  on vector  $r$ , to minimize  $S_2$ .

That is, it tells how to adjust the  $r_i$  values so as to minimize the sum of squared discrepancies

between the disparities  $\delta_i$  and the predicted asymmetries  $\hat{x}_i$ . By using Equation 5 with each object  $i$  in turn, we can adjust all of the coordinates on vector  $r$ .

The steps for vector  $s$  are almost identical:

$$S_{2 \text{ for } i} = r_i^2 \sum_j s_j^2 - 2r_i s_j \sum_j r_j s_j - 2r_i \sum_j s_j \delta_{ij} + s_i^2 \sum_j r_j^2 + 2s_i \sum_j r_j \delta_{ij} + \sum_j \delta_{ij}^2$$

We then set the first partial derivative of  $S_2$  with respect to  $s_i$  equal to zero, and solve for  $s_i$ :

$$0 = -2r_i \sum_j r_j s_j + 2s_i \sum_j r_j^2 + 2 \sum_j r_j \delta_{ij}$$

$$s_i = [r_i \sum_j r_j s_j - \sum_j r_j \delta_{ij}] / \sum_j r_j^2 \quad (6)$$

Equation 6, applied to each object  $i$  in turn, tells us the coordinates to assign on vector  $s$  so as to minimize  $S_2$ .

In practice, we will replace each  $r_i$  and  $s_i$  element as soon as the new value is computed. As a result, the values of the coordinates depend on the order in which they are updated. A preferable approach might be to store the new values of  $r_i$  and  $s_i$  in separate vectors, say  $r'_i$  and  $s'_i$ , and to replace  $r_i$  with  $r'_i$  and  $s_i$  with  $s'_i$  after all of the elements had been recomputed. This would correspond to adjusting the configuration as a whole to a least squares criterion. In experience with the algorithm, however, it appears that the two approaches lead to essentially identical solutions. Replacing the  $r_i$  and  $s_i$  values as they are computed provides some savings in time and memory overhead, and hence was the approach adopted.

Locality parameter. In practice, we use a modified formula for  $S_2$ :

$$\begin{aligned} S_{2 \text{ for } i} = & [r_i^2 \sum_j s_j^2 \exp[K * \text{abs}(\delta_{ij})] - 2r_i s_j \sum_j r_j s_j \exp[K * \text{abs}(\delta_{ij})] \\ & - 2r_i \sum_j s_j \delta_{ij} \exp[K * \text{abs}(\delta_{ij})] + s_i^2 \sum_j r_j^2 \exp[K * \text{abs}(\delta_{ij})] \\ & + 2s_i \sum_j r_j \delta_{ij} \exp[K * \text{abs}(\delta_{ij})] + \sum_j \delta_{ij}^2 \exp[K * \text{abs}(\delta_{ij})]. \end{aligned}$$

The modified formula differs from the original only in weighting each term on the right hand side by

$$\exp[K*\text{abs}(\delta_{ij})] ,$$

where K is a negative integer. This gives greater weight, when computing  $S_2$ , to  $\delta_{ij}$ 's that are smaller in absolute value. Thus, the weighting should help encourage solutions that accurately reproduce the smaller asymmetries, that is, that preserve fine-grained aspects of the data. The weighting is useful because a common feature of degenerate solutions is that small skew-symmetries tend towards zero, and effectively drop out of the solution.

In NSKMDS, "K" is referred to as a locality parameter. Unless otherwise specified, K is set to zero in the analyses, so that

$$\exp[K*\text{abs}(\delta_{ij})] = \exp[0] = 1$$

and there is no differential weighting of smaller asymmetries. When NSKMDS has appeared to produce a degenerate solution, negative values of K will sometimes be used in an attempt to eliminate the degeneracy.

The modifications in  $S_2$  lead to a similar weighting in the equations for determining the coordinates of each point on each iteration. Thus, Equation 5 is modified slightly, from

$$r_i = [s_i \sum_j r_j s_j + \sum_j s_j \delta_{ij}] / \sum_j s_j^2 \quad (5)$$

to

$$r_i = [s_i \sum_j r_j s_j e^{K\text{abs}(\delta_{ij})} + \sum_j s_j \delta_{ij} e^{K\text{abs}(\delta_{ij})}] / \sum_j s_j^2 e^{K\text{abs}(\delta_{ij})} \quad (5a)$$

and,

$$r_i = [s_i \sum_j r_j s_j e^{K\text{abs}(\delta_{ij})} + \sum_j s_j \delta_{ij} e^{K\text{abs}(\delta_{ij})}] / \sum_j s_j^2 e^{K\text{abs}(\delta_{ij})}$$

and, similarly, Equation 6 is modified from

$$s_i = [r_i \sum_j r_j s_j - \sum_j r_j \delta_{ij}] / \sum_j r_j^2 \quad (6)$$

to

$$s_i = [r_i \sum_j r_j s_j e^{K_{abs}(\delta)} - \sum_j r_j \delta_{ij} e^{K_{abs}(\delta)}] / \sum_j r_j^2 e^{K_{abs}(\delta)}. \quad (6a)$$

### NSKMDS algorithm

The NSKMDS algorithm proceeds iteratively. First the disparity values  $\delta_{ij}$  are calculated to approach the predicted asymmetries,  $\dot{x}_{ij}$ . Then the coordinate of each object on each of the two dimensions is adjusted to give new predicted asymmetries that approach the disparities. These steps, adjusting the disparities and then adjusting the coordinates, comprise one iteration. On the next iteration the disparities are recomputed from the new  $\dot{x}_{ij}$ 's, and the object coordinates are readjusted to approach these new disparities. The algorithm proceeds in this fashion until  $S_2$ , the badness of fit between the disparities and the predicted asymmetries, stops improving.

Thus, NSKMDS belongs to the alternating least squares class of algorithms. It interleaves two steps, each of which improves goodness of fit in a least squares sense. There are two advantages in particular associated with alternating least squares techniques. Because the least squares solutions on each iteration will usually improve, and certainly cannot worsen, badness of fit, and because badness of fit cannot improve beyond  $S_2 = 0$ , the algorithm must converge. Moreover, because the improvement on each iteration is optimal in a least squares sense, the algorithm finds "the conditional global minimum, since it is conditional on the values used to start the entire process" (Takane, Young, & de Leeuw, 1977, p.63). That is, the algorithm finds the best possible solution, given the starting configuration of object points.

Two difficulties can arise in NSKMDS, however, or indeed in alternating least squares algorithms more generally. First, there may be a local minimum "between" the

starting configuration and the globally optimal solution. In that event the algorithm will converge on the local solution, which may not be adequate, and which may not be replicable given a different starting configuration. To help avoid this, NSKMDS uses the canonical analysis of asymmetry to generate the starting configuration. The results of the canonical analysis should generally be close enough to the optimal ordinal-level solution to avoid local minimum problems. There is no guarantee that this will be the case, however. Moreover, if the optimal metric and nonmetric solutions differ markedly, the risk of a local minimum solution will increase. The second potential difficulty is a degenerate solution. When discussing the unnormalized  $S_2$  formula,

$$S_2 = \sum_{ij} [\dot{x}_{ij} - \delta_{ij}]^2. \quad (7)$$

we noted that it could be minimized by shrinking the configuration to a single point, located at the origin. Then all coordinates and all disparities would be zero, reducing  $S_2$  to zero. This problem was solved by dividing Formula 3 by the sum of squared deviation scores of the disparities,

$$S_2 = \sum_{ij} [\dot{x}_{ij} - \delta_{ij}]^2 / \sum_{ij} [\delta_{ij} - \delta_{\max}]^2$$

so that any reduction in the overall size of the configuration would be penalized. The normalization, however, leaves a "loophole":  $S_2$  can be minimized by partitioning the object points into two sets, and shrinking each set to a single point. Degenerate solutions are most likely to occur when relatively few observed asymmetries are being used to estimate a large number of model parameters, the  $(r, s)$  coordinates (Borg & Lingoes, 1987). Because the number of vector coordinates and the number of degrees of freedom in the data is the same in

NSKMDS as in symmetric multidimensional scaling, NSKMDS should be no more prone to degenerate solutions than is MDS generally.

Appendix 6 gives the source code for a FORTRAN 77 program implementing the NSKMDS algorithm.

In the chapters that follow, I will review simulation studies testing the behavior of the NSKMDS algorithm and comparing it with the canonical analysis of asymmetry. I will then apply the two techniques to representative asymmetric data sets in the literature to demonstrate the potential utility of the analyses. Then data collected for this dissertation will be analyzed to shed light on a dominant theory about asymmetries in similarity ratings.

## Chapter 7

### Simulation Program

Neither the canonical analysis of asymmetry, which is rarely if ever used, nor nonmetric skew-symmetric MDS, which is new in this dissertation, have been subject to systematic investigation. Simulation studies can be helpful in this regard, by showing the behavior of the algorithms under known conditions. In this chapter I will discuss the program used to conduct the simulation studies.

#### Overview.

The simulation studies were conducted using a FORTRAN 77 program, written by the present author, compiled for personal computers, and designed for studying the performance of analysis techniques under a wide range of user-specified conditions.

The program proceeds in a series of stages. First, a symmetric and skew-symmetric matrix of the same order are generated, and added together to produce a generally asymmetric matrix. (Either part can be set to zero, however, to give a purely symmetric or purely skew-symmetric matrix.) In the simulations, the resulting matrix is the "true" configuration.

The program then generates a number of error perturbed versions of the true matrix to imitate data collected from a number of subjects. The type and amount of error are specified in advance by the user. Also, if indicated, the matrix elements are subjected to a monotonic distortion to simulate ordinal-level data.

An average matrix is then derived from the error perturbed versions. The degree to which the average matrix represents the error-filled versions is quantified with descriptive

measures such as the correlation ratio. Because the true scores are not used in these indices, they are measures that would ordinarily be available to a researcher with actual data.

The average matrix is then analyzed using the technique under investigation, and regenerated using the results of the analysis. The regenerated matrix is compared with the "true" configuration to determine the efficacy of the analysis under the given conditions.

This constitutes a single trial in the simulation. For a particular set of conditions, numerous trials would be run, each beginning with a different "true" matrix.

Let us examine several of these steps a bit more closely.

Type of error.

1. Normal error. Luce (1989, p. 260), in reviewing his own psychophysical studies, noted that the primary source of error appeared to be a strong influence on responses by the stimulus from the preceding trial. If stimuli are presented in a different random order for each subject, this type of error would be equivalent to adding random deviates to each cell of the asymmetric matrix. Hence, for one error condition available in the simulation program, random, normally-distributed values are added on a cell-by-cell basis to the asymmetric matrix for each "subject", or replication, within a trial. Each entry in the error-perturbed matrix is thus represented as

$$x_{ij}' = x_{ij} + z_{ijk}$$

where  $z_{ijk}$  is the error added to cell  $ij$  for subject, or replication,  $k$ .

Normal error enters at the level of the similarity judgment. on a given trial, a person comparing two stimuli may assign a value that is higher or lower than it ordinarily would be,



due to such uncontrolled factors as the similarity of the immediately preceding trial. In this model, normal error pertains to the trial and is added directly to the similarity judgment.

In MDS and principal components analysis we depict stimuli as points or directions in a multidimensional space. That is, each stimulus is represented as a vector of coordinates on the dimensions of the space. The coordinates might simply be a convenient way of picturing the data – a graphical technique. Alternatively, we might assume that the dimensions are in some senses true, for example, that they are aspects of how people represent the stimuli internally. If the dimensions have a reality separate from the similarity judgments, then the dimension coordinates themselves may be subject to random error – a person may misperceive a given aspect of a given stimulus object. This will affect similarity judgments involving the stimulus; if the stimulus vectors are perturbed, the similarity values derived from the vectors will also be perturbed. However, adding normally-distributed error to the stimulus vectors will not, in general, cause normally-distributed error in the similarity values. Rather, the error distribution of the similarities will depend on how the stimulus vectors are compared to give similarity values. Three types of comparisons are considered here, giving three types of error distributions: (1) a squared Euclidean distance model – similarity may be modeled as the squared distance between stimulus points; (2) a principal components (scalar products) model – similarity may be given as the angle between stimulus vectors; and (3) an asymmetry model – similarity may be given as the signed area of the triangle formed by the two stimulus points and the origin. Each of these cases gives rise to a different distribution of error in the derived similarity values, as described in the following paragraphs.

2. Chi-square error. Error may enter in at the level of the stimulus vectors, before the asymmetric matrix is constructed. If the error were added to the vectors underlying the symmetric part of the matrix, and if the symmetric part were constructed using a squared Euclidean distance model, each cell in the average matrix would be given as

$$x_{ij}' = (1/k) \sum_k [x_i + z_k - x_j - z_k]^2$$

which, after squaring equals

$$\begin{aligned} x_{ij}' &= (1/k) \sum_k [x_i^2 + 2x_i z_k - 2x_j x_j - 2x_j z_k + z_k^2 \\ &\quad - 2z_k x_j - 2z_k z_k + x_j^2 - 2x_j z_k + z_k^2] \end{aligned}$$

which

$$\begin{aligned} &= (1/k) [x_i^2 + \{2x_i \sum_k z_k\} - 2x_j x_j - \{2x_j \sum_k z_k\} + \sum_k z_k^2 \\ &\quad - \{2x_j \sum_k z_k\} - (2\sum_k z_k z_k) + x_j^2 - \{2x_j \sum_k z_k\} + \sum_k z_k^2] \end{aligned} \quad (8)$$

If the error has a mean of zero the terms in braces will tend towards zero. If the error is uncorrelated with itself, then the term in parentheses will be smaller than the remaining terms. Therefore, Equation 8 will tend towards

$$x_{ij}' = (1/k) \sum_k [(x_i - x_j)^2 + z_k^2 + z_k^2]$$

which is distributed as chi-square with two degrees of freedom and a noncentrality parameter equal to  $(x_i - x_j)^2$ . Ramsay (1969) notes that this is preferable to normal error in a distance model as the  $x_{ij}'$  entries here are constrained to be positive.

3. Wishart error. If error enters into the model at the level of the symmetry vectors, and if these vectors are combined using a principal components (scalar products) model, the entries in the average matrix would be given by

$$x_{ij}' = (1/k) \sum_k [(x_i + z_k) * (x_j + z_k)]$$

$$= (1/k) (x_i x_j + \{x_i \sum_k z_{ik}\} + \{x_j \sum_k z_{jk}\} + \sum_k z_{ik} z_{jk}) \quad (9)$$

If the error has a mean of zero, the terms in braces will tend towards zero, and Equation 9 will be approximately equal to

$$= (1/k) (x_i x_j + \sum_k z_{ik} z_{jk})$$

which has a product normal distribution with mean equal to  $x_i x_j$ . Product normal distributions resemble a normal distribution. If  $z_{ik}$  and  $z_{jk}$  each have a mean of 0 and a standard deviation of 1, then  $x_{ij}'$  in Equation 9 will have a mean of  $x_i x_j$ , a standard deviation of 1, and will be symmetric. It will be both heavier-tailed and more peaked than a normal distribution, with a kurtosis of 6, versus 3 for a normal distribution (Craig, 1936, pp. 2 and 3; Meeker, Cornwell, & Aroian, 1981). Note that the diagonal entries are given as

$$x_{ii}' = (1/k) (x_i^2 + \sum_k z_{ik}^2)$$

which is distributed as chi square with one degree of freedom and a noncentrality parameter equal to  $x_i^2$ . This pattern, of product normal distributions for the off-diagonal entries of a matrix, and chi-square distributions on the diagonal, matches that of a Wishart matrix (Seber, 1984).

4. Asymmetric error. If error is added to the vectors underlying the asymmetric part of the matrix, and if these vectors are combined to give elementary skew-symmetries, the error-perturbed skew-symmetries in the average matrix will be

$$\begin{aligned} x_{ij}' &= (1/k) \sum_k [(r_i + z_{rik}) * (s_j + z_{sjk}) - (s_i + z_{sik}) * (r_j + z_{rjk})] \\ &= (1/k) (r_i s_j + \{r_i \sum_k z_{sjk}\} + \{s_j \sum_k z_{rik}\} + \sum_k z_{rik} z_{sjk} \\ &\quad - s_i r_j - \{s_i \sum_k z_{rjk}\} - \{r_j \sum_k z_{sik}\} - \sum_k z_{sik} z_{rjk}) \end{aligned} \quad (10)$$

If the error vectors have a mean of zero the terms in braces will drop out of the equation, leaving

$$x_{ij}' = (1/k) (r_i s_j - s_i r_j + \sum_k z_{rik} z_{sjk} - \sum_k z_{sik} z_{rjk}).$$

This is distributed as the difference of two product normal distributions. For the diagonal entries  $j=i$ , and

$$\begin{aligned} x_{ii}' &= (1/k) (r_i s_i - s_i r_i + \sum_k z_{rik} z_{sik} - \sum_k z_{sik} z_{rik}), \\ &= 0 \end{aligned}$$

In the simulation program this is referred to as asymmetric error, as the error only contributes to the skew-symmetric part of the matrix.

The simulation program permits incongruent combinations of error and true score models. For example, the user can specify a chi-square error distribution, and a principal components-like model for the true scores. In that case the error will be constructed as

$$\text{error}_{ij} = (1/k) (\sum_k [z_{ik} - z_{jk}]^2)$$

and the symmetric part of the true scores will be constructed as

$$x_{ij} = x_i * x_j .$$

In this way the effects of type of error can be distinguished from the effects of the model used to construct the true scores.

Amount of error. In general, the amount of error introduced by a vector will depend on the type of error, unless we take steps to remove the confounding.

First, we need to specify an index of the amount of error, that can be used across error types. In the simulation program, the index is the sum of squared differences between cells in the original and error-perturbed matrices.

Second, for each type of error we need to rescale the error vectors to produce the desired residual sum of squares.

For normal error this is straightforward. Because a random deviate is added to each cell of the matrix, the residual sum of squares is simply the sum of squares of the deviates. We can rescale the deviates by simply normalizing them to a sum of squares of one

$$z_{ijk} / \text{sqrt} (\sum_{ij} z_{ijk}^2)$$

and then multiplying each  $z_{ijk}$  by the square root of the desired amount of residual sum of squares. This same process is used in the two artificial cases: adding chi-square error, generally appropriate only for squared distance models, to symmetries created through scalar products, and adding Wishart error to distance models.

When we add chi-square error to a distance model, however, the situation becomes more complicated. In this case, instead of adding error, cell by cell, to the already-constructed symmetric matrix, the error is added to the symmetry vectors before they are used to construct the matrix. Thus, error enters into the matrix as

$$x_{ij}' = (1/k) (\sum_k [x_i + z_{ik} - x_j - z_{jk}]^2)$$

And the sum of squared residuals for a given matrix  $k$  is

$$\sum_{ij} (x_{ij}' - x_{ij})^2 = \sum_{ij} [(x_i + z_{ik} - x_j - z_{jk})^2 - (x_i - x_j)^2]^2$$

To determine how the error vectors should be rescaled, we introduce a scaling coefficient,  $C$ , and set the equation equal to the desired amount of residual variation (DARV)

$$\text{DARV} = \sum_{ij} [(x_i + Cz_{ik} - x_j - Cz_{jk})^2 - (x_i - x_j)^2]^2$$

or

$$0 = \sum_{ij} [(x_i + Cz_{ik} - x_j - Cz_{jk})^2 - (x_i - x_j)^2]^2 - \text{DARV}$$

When expanded (see Appendix 4), this equation is a fourth degree polynomial in the unknown rescaling coefficient:

$$\begin{aligned}
0 = & C^4(n\Sigma_i z_{ik}^4 - 4\Sigma_i z_{ik}^3 \Sigma_j z_{jk} + 6\Sigma_i z_{ik}^2 \Sigma_j z_{jk}^2 - 4\Sigma_i z_{ik} \Sigma_j z_{jk}^3 + n\Sigma_j z_{jk}^4) \\
& + C^3(4n\Sigma_i x_i z_{ik}^3 - 12\Sigma_i x_i z_{ik}^2 \Sigma_j z_{jk} + 12\Sigma_i x_i z_{ik} \Sigma_j z_{jk}^2 - 4\Sigma_i x_i \Sigma_j z_{jk}^3 \\
& - 4\Sigma_j x_j \Sigma_i z_{ik}^3 + 12\Sigma_j x_j z_{jk} \Sigma_i z_{ik}^2 - 12\Sigma_i z_{ik} \Sigma_j x_j z_{jk}^2 + 4n\Sigma_j x_j z_{jk}^3) \\
& + C^2(4n\Sigma_i x_i^2 z_{ik}^2 - 8\Sigma_i x_i^2 z_{ik} \Sigma_j z_{jk} - 8\Sigma_i x_i z_{ik}^2 \Sigma_j x_j \\
& + 16\Sigma_i x_i z_{ik} \Sigma_j x_j z_{jk} + 4\Sigma_i x_i^2 \Sigma_j z_{jk}^2 - 8\Sigma_i x_i \Sigma_j x_j z_{jk}^2 + 4\Sigma_i z_{ik}^2 \Sigma_j x_j^2 \\
& - 8\Sigma_i z_{ik} \Sigma_j x_j^2 z_{jk} + 4n\Sigma_j x_j^2 z_{jk}^2) - \text{DARV}
\end{aligned} \tag{11}$$

where DARV is the desired amount of residual variation. All of the terms in Equation 11 are known except for the scaling coefficient, C, and thus the equation is of the form

$$0 = aC^4 + bC^3 + dC^2 - \text{DARV} . \tag{12}$$

In general, quartic equations such as this will have four roots, some of which may be negative or even imaginary. However, for our purposes we need only a single, positive root.

In the simulation program this root is obtained through the methods of bracketing and bisection. That is, consider Equation 12. It will be satisfied for those values of C for which the right hand side equals zero. If we initially set C equal to zero, the right hand side will equal a negative number, that is, -DARV. In the program, C is then increased by increments of .1, until the right hand side of Equation 12 first equals a positive number. Thus we have identified an interval, .1 in length, for which Equation 12 is negative at the lower border, and positive at the upper border. Presumably the root, the point at which Equation 12 crosses zero, is bracketed within this interval.

To refine our estimate of the root we bisect the interval, keeping track of the half interval in which the sign change occurs. We then bisect the half interval, and so on, until we have converged on the place where the sign change occurred. This is very likely to be a root, and hence the desired rescaling factor.

Because Equation 12 is a continuous function, the place at which its sign changes is also a place at which it crosses zero, i.e., a root. However, the program may fail to converge in the allotted number of iterations, and hence return an erroneous value. Therefore as a final step, the program checks the obtained value of  $C$  to see if it indeed reduces Equation 12 to a number close to zero. If the obtained value is not so confirmed as a root, the error vector is discarded and a new error vector is generated. This appears to happen on approximately 5% of the matrices, not enough to significantly lower the efficiency of the program.

Different quartic equations, solved in the same manner, arise for Wishart error added to the vectors for scalar product matrices, and for asymmetric error. The quartic equations and their derivations are shown in Appendix 4.

By rescaling in this manner, the amount of error variation can be adjusted independently of the type of error.

Monotonic distortion. The purpose of monotonic distortion is to obscure the original scale of the matrix entries, while preserving their rank order. Presumably this reflects the case of ordinal level data, in which only rank order information can be assumed to be accurate. In areas such as multidimensional scaling, the ability of an analysis technique to

recover the original scale values after monotonic distortion is considered central to its usefulness (e.g., Kruskal, 1964a; Shepard, 1962).

Introducing monotonic distortion is a relatively simple process in the simulation program, the main task being to identify a way of quantifying and characterizing the distortion. It was felt at the outset that power functions of the type

$$y = x^k$$

would be quite useful as they permit both accelerating ( $k > 1$ ) and decelerating ( $k < 1$ ) distortions, whose magnitude is specified by the exponent. Moreover, the shape of distortion is both well-known and relevant to psychologists, due to its nearly universal appearance in psychophysics (e.g., Stevens, 1957). The chief problem is that in the simulation program matrix entries can take on negative values. Even, whole-numbered exponents  $k$  would then transform these into positive values, and thus fail to preserve order. Even, fractional values of  $k$ , common in psychophysics, would transform negative matrix entries into imaginary numbers, which would be exceedingly hard to interpret substantively. To solve these problems, the expedient was developed of stripping the sign before applying the transform, and then reattaching the sign afterwards

$$x'_{ij} = \text{sign}(x_{ij}) * [\text{abs}(x_{ij})]^k$$

This produces monotonic distortions similar to those shown in Figures 1a (for  $k < 1$ ) and 1b (for  $k > 1$ ).

Recovery Indices. Young (1970) operationalized metric recovery as the squared product moment correlation coefficient between the true and recovered matrix entries.



$$\text{Metric Recovery}_{(\text{Young, 1970})} = \frac{[\sum_{ij} (x_{ij} - x_{\text{mean}})(\hat{x}_{ij} - \hat{x}_{\text{mean}})]^2}{[\sum_{ij} (x_{ij} - x_{\text{mean}})^2][\sum_{ij} (\hat{x}_{ij} - \hat{x}_{\text{mean}})^2]} .$$

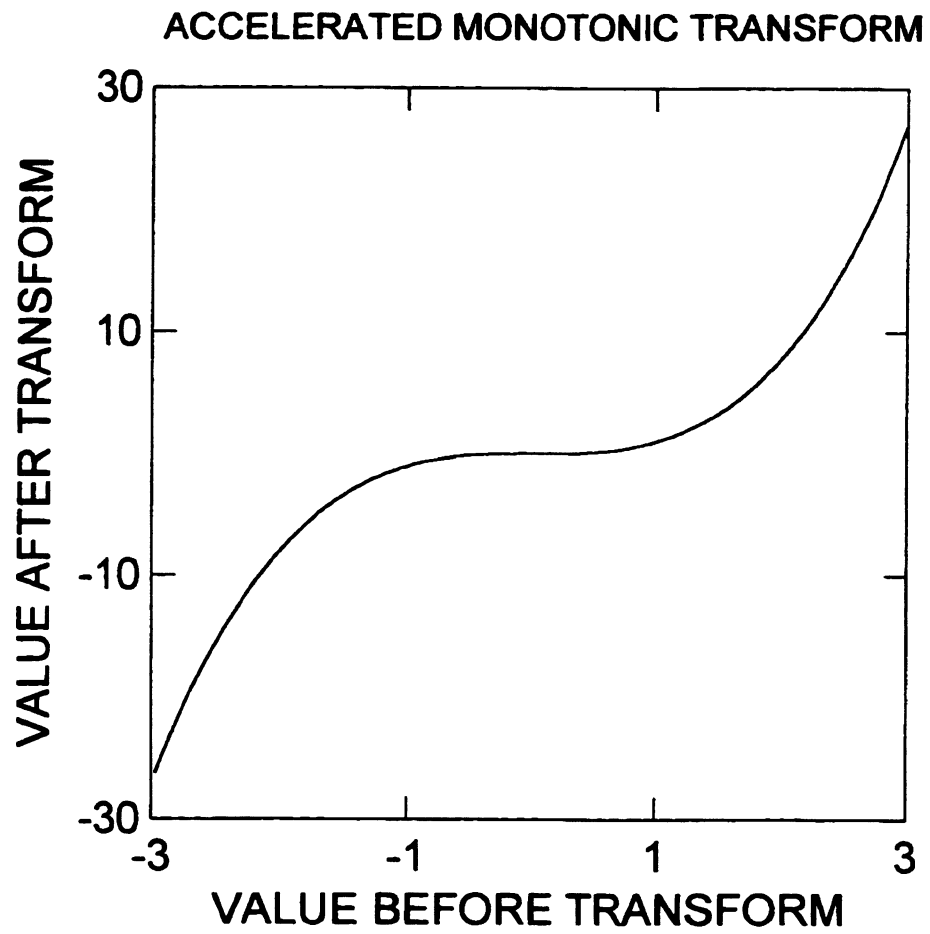
In the simulation program this framework is adopted, with two modifications. First, metric recovery is redefined as the squared ratio-level product moment correlation

$$\text{Metric Recovery}_{\text{Simulation Program}} = \frac{[\sum_{ij} (x_{ij} * \hat{x}_{ij})]^2}{[(\sum_{ij} x_{ij}^2)(\sum_{ij} \hat{x}_{ij}^2)]} .$$

as this is more appropriate to the (absolute level) canonical analysis of asymmetry. Second, a nonmetric recovery index was defined analogously, as the squared Spearman rank order correlation coefficient between the true and reconstructed matrix entries.

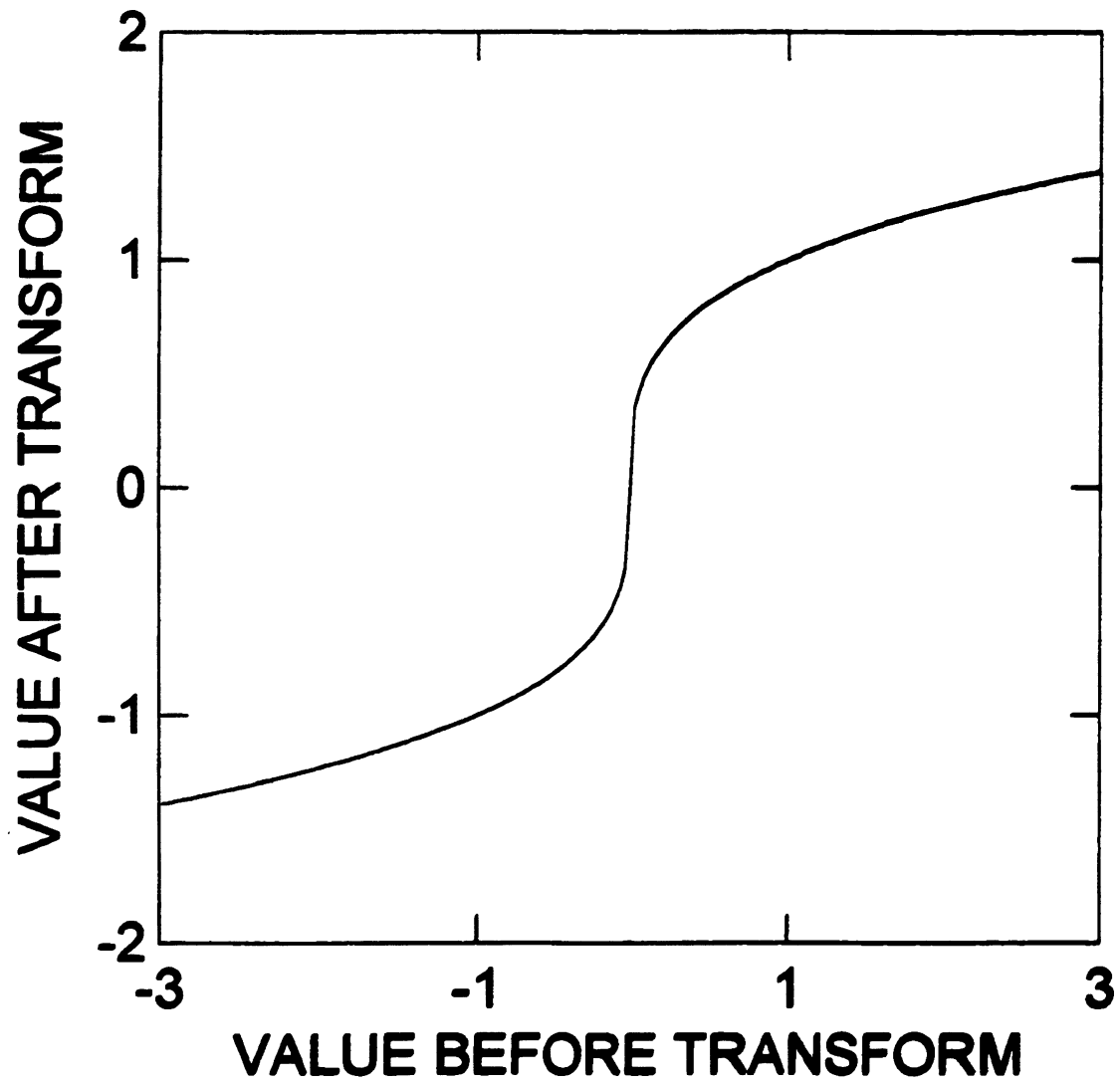
Eigenvalue Dropoff. One purpose to which simulation studies may be directed is determining how well an analysis technique distinguishes true from error variance. A condition that may bear on this is the extent to which the eigenvalues of the true factors display a different pattern than those of the error factors. To permit a systematic study of this the simulation program provides for two different patterns of true factors: in the fast drop-off condition each successive eigenvalue is lower than the preceding eigenvalue by a factor of 1.8. In the slow drop-off condition each eigenvalue is lower than the preceding eigenvalue by a factor of 1.18. Thus, in both drop-off conditions the eigenvalues show an exponential decay, which is characteristic of the principal components analysis of random data (e.g., Craddock & Flood, 1969). However, in the fast drop-off condition the true eigenvalues are separated from each other to a greater extent than is commonly seen in random matrices that are larger than 10 by 10 (see tables in Lautenschlager, 1989).

FIGURE 1A



# FIGURE 1B

## DECELERATED MONOTONIC TRANSFORM



**Random Numbers.** On each trial the true and error matrix components are created from random numbers. In the simulation program these are derived from the Wichmann and Hill (1982) random number generator, which gives numbers between 0 and 1 drawn from a uniform distribution. When random normal variates are needed, the output of the random number generator is filtered through Beasley and Springer's (1977) subroutine for the inverse normal cumulative density function. The Wichmann and Hill algorithm was chosen because it was published in a peer-reviewed journal (Applied Statistics), and was the subject of a number of subsequent commentary articles (McLeod, 1985; Wichmann & Hill, 1984; Zeisel, 1986).

The simulation program was thus designed to be quite broad in the range of conditions that it could be used to study. Of the numerous possible simulations, several were selected, as the ones most useful for guiding the application of CAA and NSKMDS to real world data.

## Chapter 8

### Simulation Studies: Quality Control

Initial efforts went to verifying that the simulation program itself was functioning effectively.

Accuracy of algorithms. First, analyses were conducted on matrices to which no error or monotonic distortion had been introduced. This was designed to test that the algorithms being used, and the subroutines by which they were implemented, had acceptable accuracy. I had written all of the source code for the analyses, including the Jacobi rotations used in the eigenvalue decomposition, and hence the accuracy could not be taken for granted.

One hundred error-free trials were conducted for each of the two analysis types. For CAA the mean metric recovery index value was 1.00000000, and the standard deviation was 0.00000000. These same values were obtained for NSKMDS. Thus, the results were perfect on these trials to eight decimal places, which is near the theoretical limit for single precision arithmetic.

Regulation of error variance. Second, the program was checked for its consistency in introducing the user-specified amount of error variance for each of the four types of error. One hundred trials were run for each of the four types of error at each of four error levels (sum of squared residuals set equal to 0.5, 1, 5, or 10 times the true symmetric sum of squares). On these trials, the mean absolute deviation, of the actual error level from the user-specified level, was less than  $10^{-5}$  times the user-specified level. Thus, if on a given trial the

error sum of squares should equal 227, say, the error sum of squares would actually equal, on average,  $227 \pm .00227$ .

Random number generator. Third, the numbers produced by the Wichmann and Hill (1982) subroutine were checked for apparent randomness, and to verify that they were being selected from a rectangular distribution between 0 and 1. A series of 1000 numbers was produced iteratively by the subroutine. The mean and variance of these numbers are shown in Table 2, where it may be seen that they differ only negligibly from the corresponding population parameters for a uniform 0,1 distribution ( $\mu = .50$ ,  $\sigma^2 = .0833$ ). Then the numbers were checked for periodicity by computing autocorrelations of lag 1 through 15. These, as shown in Table 2, indicate that the numbers selected by the subroutine do not depend in any consistent, obvious way on any of the 15 preceding numbers.

Confounding variance. Despite these precautions, a major source of variance in metric recovery, throughout the initial simulations, was the time of day at which the simulations were run. Analyses conducted after midnight yielded high metric recoveries; in early afternoon the recoveries were poor.

Table 2

## Diagnostic Statistics On The Random Number Generator

<b>Statistic</b>	<b>Value</b>
Mean	0.494
Variance	0.085
Autocorrelation Lag 1	0.01523
Autocorrelation Lag 2	0.01002
Autocorrelation Lag 3	0.04614
Autocorrelation Lag 4	0.03141
Autocorrelation Lag 5	-.01891
Autocorrelation Lag 6	-.00942
Autocorrelation Lag 7	-.02056
Autocorrelation Lag 8	0.03423
Autocorrelation Lag 9	-.00431
Autocorrelation Lag 10	0.02528
Autocorrelation Lag 11	-.03432
Autocorrelation Lag 12	0.03453
Autocorrelation Lag 13	0.00305
Autocorrelation Lag 14	0.01443
Autocorrelation Lag 15	0.00230

Note. Based on a series of 600 consecutive products of the random number generator, using clock seeds as described in the text.

Time of day enters explicitly into the program in only one place: in the seed variables passed to the random number generator:

```

DO 1 J=1, 10

    CALL GETTIM(IHR, IMIN, ISEC, I100TH)

    IX = IHR*10000 + IMIN*100 + I100TH
    IY = ISEC*10000 + IMIN*100 + I100TH
    IZ = IMIN*10000 + ISEC*100 + I100TH

    IF (J .NE. 1) THEN

        IX = IX*J

        IY = IY*AINT(RANDM*100)

        IZ = IZ*AINT(RANDM*100)

    END IF

    CALL RNDGEN(IX,IY,IZ,RANDM)

1 CONTINUE

```

where IHR, IMIN, ISEC, and I100TH are the hour, minute, second and hundredth of a second read from the computer's clock, IX, IY, and IZ are the seed variables used for the random number generator on a given iteration, and RANDM is the generator's output.

Statements of this type were used in all of the calls to the random number generator.

To assess the magnitude of the time-of-day effect, and to verify that it entered into the simulations through the value of IHR, that is, the first two digits in the first step of construction of the IX variable, "IHR" was replaced throughout the program, in statements such as (13), by another variable that was artificially set to 0, 3, 6, 9, 12, 15, 18, or 21.

Canonical analysis of asymmetry was conducted. The level of normal error was set to 6,



i.e., 6 times the level of true variation in the matrix, and the level of skew-symmetric error was set to 2. Monotonic distortion was set to 3, that is,

$$x'_{ij} = \text{sign}(x_{ij}) * [\text{abs}(x_{ij})]^3.$$

These levels were chosen to prevent "ceiling" and "floor" effects in the eight "time of day" conditions. That is, the error and monotonicity values were set so as to prevent two or more "time of day" conditions from becoming indistinguishable due to maximally high or maximally low metric recovery. The metric recovery of skew-symmetries was used as the dependent variable. Fifty trials were run for each of the 8 "hours". The results, shown in Figure 2, appeared to replicate the time of day effect. In the corresponding one-way ANOVA, the effect of time of day is significant,  $F(7,152) = 34.752$ ,  $p < .0005$  (see Table 3). In the sample, 61.5% of the variance is explained by time of day (correlation ratio =  $r^2 = .615$ ; Pearson, 1905). The proportion of explained variation estimated for hour in the population of trials is .59 ( $\omega^2 = .590$ ).

Table 3

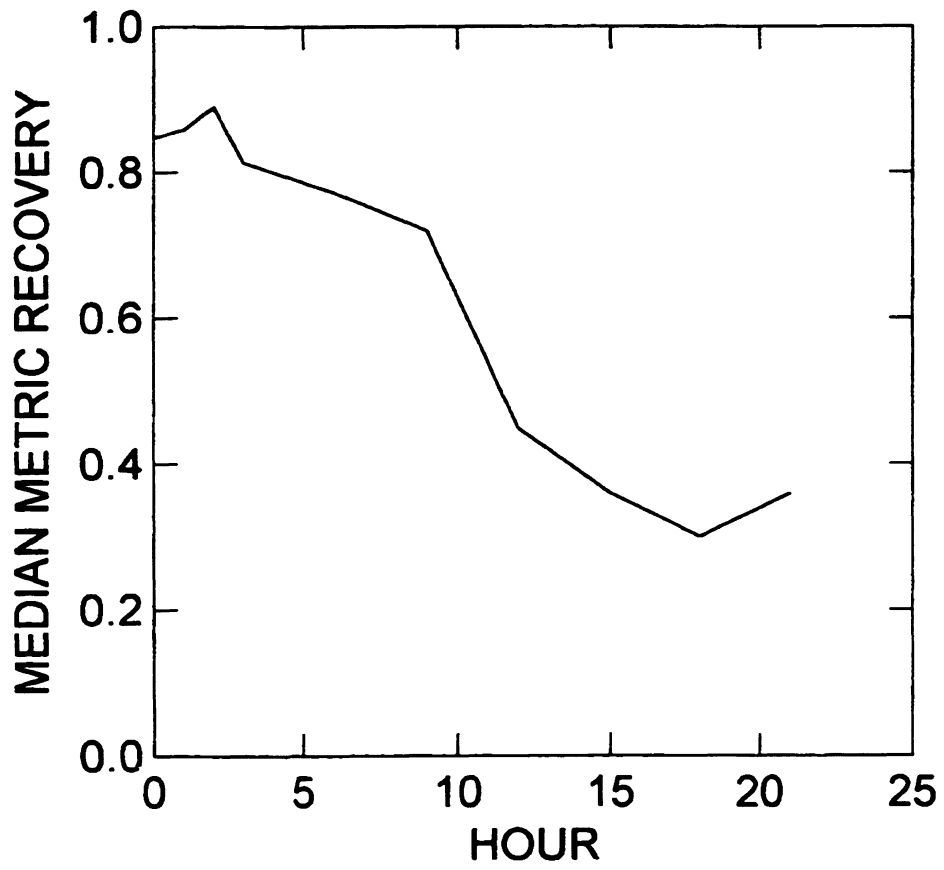
Time of Day Effect

Analysis of Variance

Source	SS	df	MS	F	p	$\omega^2$
Hour	6.304	7	.901	34.752	.000	.590
Error	3.939	152	.026			

### FIGURE 2

#### TIME OF DAY EFFECT



One solution, randomizing the IHR variable across trials, or replacing IHR with, say, ISEC in the random number generator calls, was immediately felt to be inadequate. "Time of day" pushes metric recovery through nearly its full range, and certainly far to either side of the threshold for acceptable values, say .80. Hence, randomizing time of day would add considerable error variance to the simulations. Moreover, because we do not know whether metric recovery is being suppressed in some conditions, or inflated in others, or whether both effects are occurring, we cannot be sure that randomization would give an unbiased estimate of true metric recovery. Thus, an attempt was made to determine the nature of the time of day effect in metric recovery, and the mechanism by which it was occurring.

The fact that the program automatically rescales the error vectors to produce the desired amount of perturbation should confer considerable resistance against problems with the random numbers. If, at some hours, the random numbers were small, the resulting error matrices would be small, but would be automatically scaled up. Similarly, the magnitude of skew-symmetric error would be suppressed if the random numbers were relatively constant, or if the first skew-symmetric error vector correlated with the second (high lag-10 autocorrelation, for 10-element vectors). However, the simulation program should automatically rescale the vectors upwards to eliminate these effects. Moreover, when 600 calls were made to the random number generator at each of the eight "hours", the resulting series show no departure from the expected mean (.50), variance (.0833), or lag-10 autocorrelation (0) for numbers randomly drawn from a uniform distribution (see Table 4). In fact, autocorrelations from lag 1 to lag 15 were computed for each of the eight "hours".

No trends emerged, and no autocorrelations exceeded .09 in absolute value. To appearances, the random number generator was functioning exceedingly well.

Table 4  
Diagnostic Statistics For The Random Number Generator

Time of Day Effects

"HOUR"	Mean	Variance	Lag 10 Autocorr.
00	0.494	0.085	0.025
03	0.493	0.086	0.082
06	0.491	0.083	0.030
09	0.497	0.082	0.052
12	0.490	0.082	0.020
15	0.486	0.086	0.033
18	0.500	0.082	0.028
21	0.495	0.082	0.034

Another possibility was then considered. It may be that the error in one randomly perturbed matrix was correlated with the error in another of the 30 random matrices in a trial. If the correlation were positive, error would tend to compound across matrices, rather than cancel as it should, and metric recovery would be suppressed. If the correlation were negative, the error would cancel at a higher-than-expected rate, and metric recovery would be inflated. Effects of this type did not seem likely because (1) there is an approximately 1.8 second interval between the production of one error matrix and the next in a trial.

Correlations between the error matrices would thus require that the random number generator behave in an approximately cyclical manner, with a 1.8 second period. Because there is no apparent mechanism that would induce a resonance between the random number generator and the simulation program as a whole, synchrony would be coincidental. (2) The rate of the simulation program as a whole should be affected by its iterative root-finding steps for rescaling the error vectors. Because these roots will in general differ for each error matrix, and hence be found more or less rapidly, the 1.8 second cycle-time is only approximate: The program will create some matrices at a faster, and some at a slower, rate. This should help reduce any coincidence in rhythm between the random number generator and the program as a whole. Moreover, as may be seen in (13), above, the seeds used for the random number generator are changed every hundredth of a second. Nonetheless, between-matrix effects were studied empirically. The skew-symmetric error vectors in particular were examined, as skew-symmetric error had shown the strongest time of day effect. First, the "hour" was set to 0. A single trial was run of 30 error-perturbed matrices, and hence 30 pairs of skew-symmetric error vectors. The elements of the 60 vectors were then subjected to a 2 (first vs. second vector) by 10 (position-within-vector) ANOVA, with the 30 replications giving the within cells variance. A significant vector by position interaction would indicate that error was positively correlated across matrices, but that the elements in the first vector differed from corresponding elements in the second vector. A significant main effect for position, in the absence of a significant interaction, would indicate that error was positively correlated across matrices, and that there was a positive lag-10 autocorrelation, causing entries in the first vector to resemble the corresponding entries in the second vector.

Five runs of 30 matrices were analyzed at each of the 8 "hours". None of the 40 position by vector interactions were significant. However, the position main effect was significant for all ANOVAs with "hour" greater than or equal to 12. The median of the five F ratios for the position main effect at each "hour", and the median metric recovery, are shown in Table 5. Increases in the variation associated with the position main effect appear to closely parallel decrements in metric recovery. Hence there is evidence that by 12 o'clock at the latest, i.e., when the first seed to the random number generator is initially of the form "1 2 \_ \_ \_ \_", error variance is positively correlated across matrices, suppressing metric recovery.

The fact that the position by vector interaction is not significant suggests that entries in the first vectors tend to resemble the corresponding entries in the second vectors. This implication of a lag-10 autocorrelation does not necessarily contradict the absence of such a correlation noted earlier. Before, we looked for oscillations that were sustained (stationary) in the 600-member series. Here our focus is on a short-term autocorrelation.

We may note that at IHR=0, the median  $\bar{F}$  is below 1. This discrepancy is not statistically significant. That is, if we invert the value ( $1/.69 = 1.45$ ) and test it at  $F(580,9)$  degrees of freedom (Feldt, 1965), the result is not significant at even the  $\alpha=.25$  level. However, to investigate this further, five trials were conducted, as above, at IHR=1 and IHR=2. The median  $\bar{F}$  ratios are shown in Table 5. From the trend, there is a suggestion that, at IHR=0 and IHR=1, the main effect of position has less variance than would be expected by chance. This would occur if the entries in some error matrices were negatively correlated with corresponding entries in other error matrices in the same trial. This would

Table 5  
Time of Day Effect

	<b>MEDIAN</b>	<b>MEDIAN F</b>	<b>POSITION M.E.</b>
00	0.847	0.691	<u>ns</u>
01	0.859	0.847	<u>ns</u>
02	0.890	1.099	<u>ns</u>
03	0.813	1.777	.070
06	0.772	1.571	.121
09	0.719	1.549	.127
12	0.448	3.016	.002
15	0.359	5.128	.000
18	0.300	5.231	.000
21	0.359	5.199	.000

artificially inflate metric recovery. Hence, it was felt prudent to consistently use values of the form "Q 2 \_ \_ \_ \_" as the first step in creating the first seed variable to the random number generator, to dampen what appears to be a 1.8 second periodicity. This strategy was adopted throughout the simulation studies. Thirty-two simulations in total were conducted. It is to the results of these studies that we will now turn our attention.

## Chapter 9

### Simulation Studies: Nonmetric Analysis of Skew-Symmetry

**General settings.** The simulations were designed to mimic likely conditions confronting a researcher. In particular, the simulations were intended to help guide the analysis of the similarity data whose collection is described in Chapters 16 and 17.

Therefore, in the simulations that follow, a ten by ten matrix of "true" asymmetric similarity values is assumed. Each matrix is replicated 30 times and is subject, in each replication, to monotonic and/or error distortion in the amount specified for that trial. Hence the 30 replications correspond to simulated data for 30 subjects. The NSKMDS analysis is then applied to the average of these 30 distorted matrices. The results of the analysis are used to reconstruct the ten by ten matrix, and the reconstruction is compared, cell by cell, with the original, "true" matrix.

Because our interest is in the techniques for representing skew-symmetries, only the skew-symmetric part of the true matrix was entered into the comparison. Two indices of comparison were used. Metric recovery was operationalized as the squared, ratio-level product-moment correlation between the true and reconstructed entries for the cells in the upper triangular half of the skew-symmetric matrix. Nonmetric recovery was operationalized as the squared, Spearman rank-order correlation coefficient between the true and reconstructed entries for the upper triangular half. Note that the diagonal cells in a skew-symmetric matrix are always zero, by the definition of skew-symmetry, which would inflate the correlation



between the true and reconstructed entries. Therefore the diagonal cells were excluded from calculations of the recovery indices.

Monotonic distortion. As noted, the nonmetric analysis program gives essentially perfect metric and nonmetric recovery of ideal data: data that has been altered by neither random error nor monotonic transform. In addition, however, a nonmetric analysis should be able to recover ordinal level information from data that has been subjected to a monotonic transform, that is, a transform that does not degrade or remove information on rank order.

To investigate whether the NSKMDS algorithm is able pass this criterion, 210 trials were run, spanning 21 levels of positively accelerated monotonic distortion. The levels varied between 1.0 (no distortion) and 3.0 (high distortion) in increments of 0.1. Thus, the NSKMDS algorithm was tested at distortion levels of 1.0, 1.1, 1.2, ...2.8, 2.9, and 3.0. The effects of this distortion on metric and nonmetric recovery are shown in Figure 3, and in Table 6. There it may be noted that mean nonmetric recovery, that is, the mean of the squared Spearman rank order correlation coefficients between the true and recovered elements in the upper triangular half matrix, remains above 0.97 for distortion levels between 1.0 and 1.9. Over the full 1.0 to 3.0 range the nonmetric recovery falls below 0.95 only twice: at 2.0 ( $r_s=0.939$ ) and at 2.4 ( $r_s=0.890$ ).

The NSKMDS algorithm, as those used in traditional nonmetric MDS, is only designed to recover the ordinal level properties of the data. Nonmetric MDS derives much of its utility, however, from the empirical result that metric recovery is often quite high as well (e.g., Borg, & Lingoes, 1987; Young, 1970). Metric recovery by NSKMDS is shown in Table 6 and Figure 3. The mean metric recovery is greater than 0.90 for all of the tested

distortion levels between 1.0 and 2.1, and is greater than 0.85 for all levels between 1.0 and 2.5. Thereafter, metric recovery appears to decline fairly rapidly.

The 2.5 level appears to reflect a fairly high distortion. For example, in Stevens' (1960) list of 22 power-law exponents determined empirically for psychophysical data, only one (electric shock) has an exponent greater than 1.7. Thus, NSKMDS appears to provide good metric recovery at the moderate distortion levels typical of psychophysical data. As nonmetric recovery remains high throughout the range studied, it seems likely that the decline in metric recovery at high distortion levels reflects a divergence between optimal metric and nonmetric solutions, rather than a failure of the algorithm.

The linear regression of mean metric recovery on distortion level is significant,  $E(1,19) = 149.421$ ,  $p < .0005$ , as is the linear regression of mean nonmetric recovery on distortion level,  $E(1,19) = 6.442$ ,  $p < .020$ . The slope for metric recovery,  $-0.124$ , suggests a faster decline than the slope for nonmetric recovery,  $-0.020$ . In both cases the regressions must be qualified by the heteroscedasticity seen in Table 6. There is no variability in recovery when the distortion level is 1 (no distortion). As the mean recovery indices decline, the variability in recovery at that distortion level increases.  $F_{\max}$  is essentially infinite if the no-distortion condition is included. If this condition is omitted,  $F_{\max}$  is 21609.00,  $p < .05$ , for metric recovery, and 15129.00,  $p < .05$ , for nonmetric recovery.

These results were replicated in a second simulation covering 8 monotonicity levels (1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0) whose results are described later (see Chapter 11).

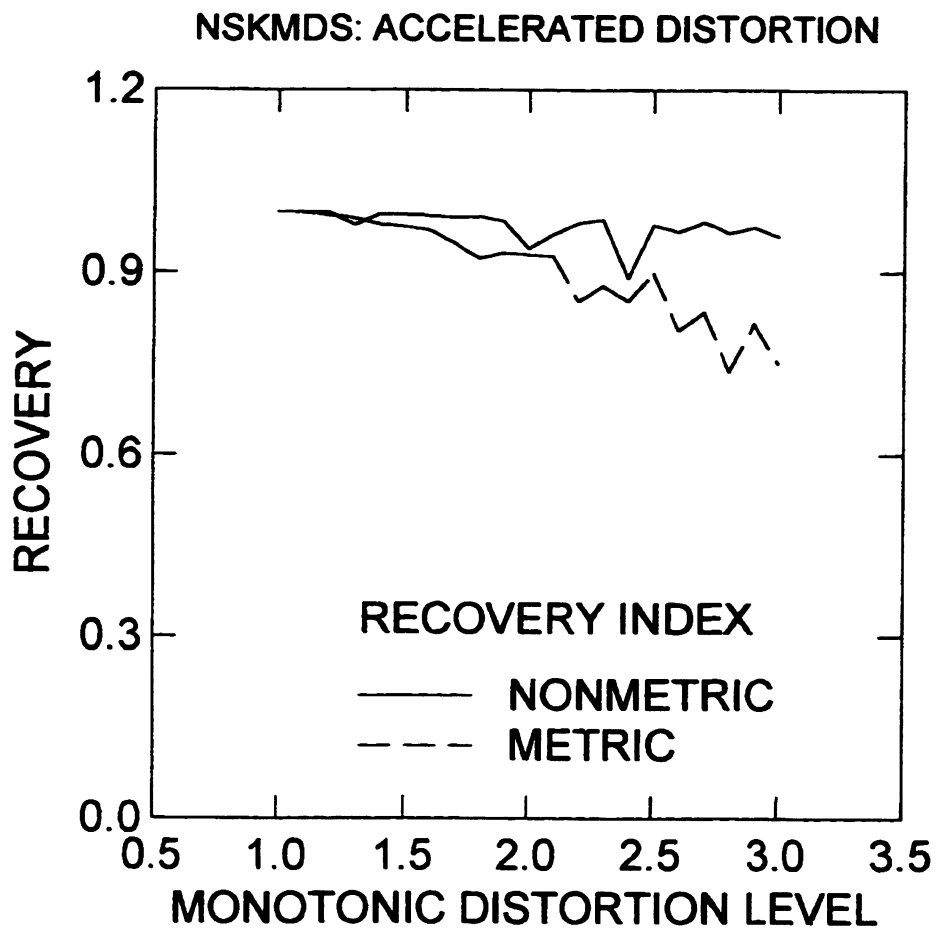
Table 6

## Nonmetric Analysis (NSKMDS)

## Recovery Under Positively Accelerated Monotonic Distortion

DISTORT. LEVEL	N	NONMETRIC RECOVERY		METRIC RECOVERY	
		MEAN	STD. DEV.	MEAN	STD. DEV.
1.0	9	1.000	0.000	1.000	0.000
1.1	7	0.999	0.000	0.998	0.001
1.2	12	0.998	0.001	0.994	0.002
1.3	12	0.978	0.048	0.989	0.003
1.4	7	0.996	0.002	0.979	0.009
1.5	11	0.995	0.002	0.975	0.012
1.6	12	0.993	0.005	0.969	0.022
1.7	11	0.990	0.004	0.947	0.023
1.8	10	0.992	0.003	0.923	0.030
1.9	12	0.984	0.018	0.932	0.031
2.0	7	0.939	0.123	0.929	0.060
2.1	7	0.963	0.073	0.926	0.062
2.2	11	0.981	0.013	0.852	0.071
2.3	12	0.987	0.009	0.878	0.054
2.4	13	0.890	0.239	0.852	0.064
2.5	10	0.978	0.014	0.899	0.071
2.6	19	0.967	0.029	0.804	0.114
2.7	12	0.983	0.014	0.835	0.066
2.8	9	0.965	0.022	0.736	0.147
2.9	8	0.975	0.020	0.817	0.115
3.0	9	0.959	0.019	0.749	0.094

FIGURE 3



Monotonicity exponents above 1.0 generate a positively accelerated distortion, expanding the range of values in the matrix. Entries whose absolute values are below 1.0 before the transformation will be brought closer to zero by the distortion. Entries whose magnitudes are greater than 1.0 will be increased, with the large entries increasing the most rapidly. In contrast, monotonicity levels below 1 compress the range of values, as shown in Figure 1b. Many of the power-law exponents in psychophysics, including those for the sone scale of loudness (0.3) and the bril scale for the brightness of a 5-degree target (0.3) and of a point source (0.5), compress the stimulus range in this manner (Stevens, 1957, 1960).

To test whether the NSKMDS algorithm is resistant to a decelerating monotonic transform, 20 trials were run at each of 8 monotonicity levels below 1.0: 0.67, 0.50, 0.40, 0.33, 0.29, 0.25, 0.22, and 0.20. These values are the reciprocals of 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0, respectively. Thus, the smaller coefficients produce higher levels of distortion. Results are shown in Table 7, and graphically in Figure 4. With increasing distortion level, the linear, downward trend is significant for both metric recovery,  $F(1, 6) = 6.29$ ,  $p < .05$ , and nonmetric recovery,  $F(1, 6) = 8.35$ ,  $p < .05$ . However, even at high distortion levels, metric recovery does not fall below .925, while the squared Spearman rank order correlation coefficient does not fall below 0.969. Thus, the algorithm appears able to correct for a wide range of monotonic distortions to recover metric and nonmetric information.

Table 7

## Nonmetric Analysis (NSKMDS)

## Recovery Under Decelerating Monotonic Distortion

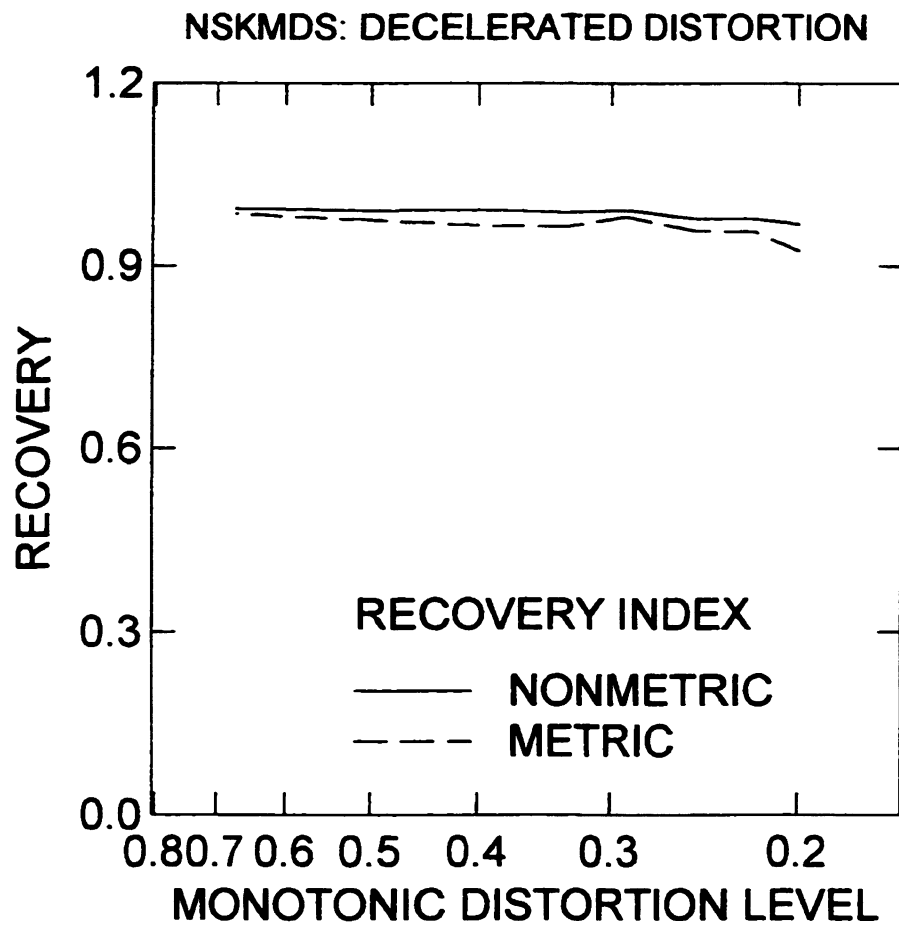
Distortion Level	Nonmetric Recovery		Metric Recovery	
	CAA	NSKMDS	CAA	NSKMDS
.67	.990	.995	.977	.986
.50	.976	.991	.946	.976
.40	.971	.993	.895	.967
.33	.956	.988	.877	.965
.29	.935	.992	.873	.981
.25	.929	.978	.840	.958
.22	.906	.978	.850	.957
.20	.898	.969	.836	.925

Note. Each entry is the mean recovery across 40 trials.  
of analysis of variance for these data.

See Tables 20 and 21 for results

Normal error. Kruskal and Shepard (1974) noted in regard to nonmetric factor analysis that its resistance to monotonic distortion seemed outweighed by its greater susceptibility to error. In nonmetric algorithms, small random distortions in the data can in theory lead to a substantial deterioration in the solution due to problems with local minima and degeneracy. Therefore, simulations were conducted to check the robustness of the NSKMDS algorithm when random normal error is added to the matrix of true scores.

FIGURE 4



One hundred and ten simulation trials were run, spanning eleven settings of normally distributed error. The error levels were sampled in increments of 0.5, between 0.0 (no error) and 5.0 (high error). Thus, the levels 0.0, 0.5, 1.0, ...4.0, 4.5, and 5.0 were included. In the simulations "level" indicates the sum of squared residuals in the error perturbed matrix, as a proportion of the sum of squares of the true score matrix. Thus, an error level of 0.5 indicates that the sum of squared residuals is half as large as the true score sum of squares. An error level of 1.0 indicates that the true and residual sums of squares are equal, and a level of 5.0 indicates that the residual sum of squares is five times as great as the true score sum of squares. Presumably error levels greater than 2.0 or 3.0 would be unusually high for psychological data.

Results are shown in Table 8, and graphically in Figure 5. Nonmetric recovery declined from 1.000 when no error was present, to 0.555 at an error level of 5.0. In the same conditions metric recovery decreased from 1.000 to 0.411. The linear trend in the nonmetric decline is significant,  $F_{lin} = 34.90$ ,  $p < .0005$ , as is the linear trend in the metric decline,  $F_{lin} = 105.76$ ,  $p < .0005$ . For both metric and nonmetric recovery,  $F_{max}$  is essentially infinite if the no-error condition is included. If the no-error condition is excluded, nonmetric recovery shows heteroscedasticity: as error level increases so does the trial-to-trial variability in recovery,  $F_{max} = 34.99$ ,  $p < .05$ . For metric recovery,  $F_{max} = 7.55$ , *ns*.

Thus, the NSKMDS algorithm is clearly vulnerable to the presence of normally-distributed error. However, most of the error levels tested here seem extreme for actual data. At an error level of 5.0, for example, the residual sum of squares is five times greater than the true score sum of squares, analogous to a test reliability of 0.17. If we restrict



Table 8

## Nonmetric Analysis (NSKMDS)

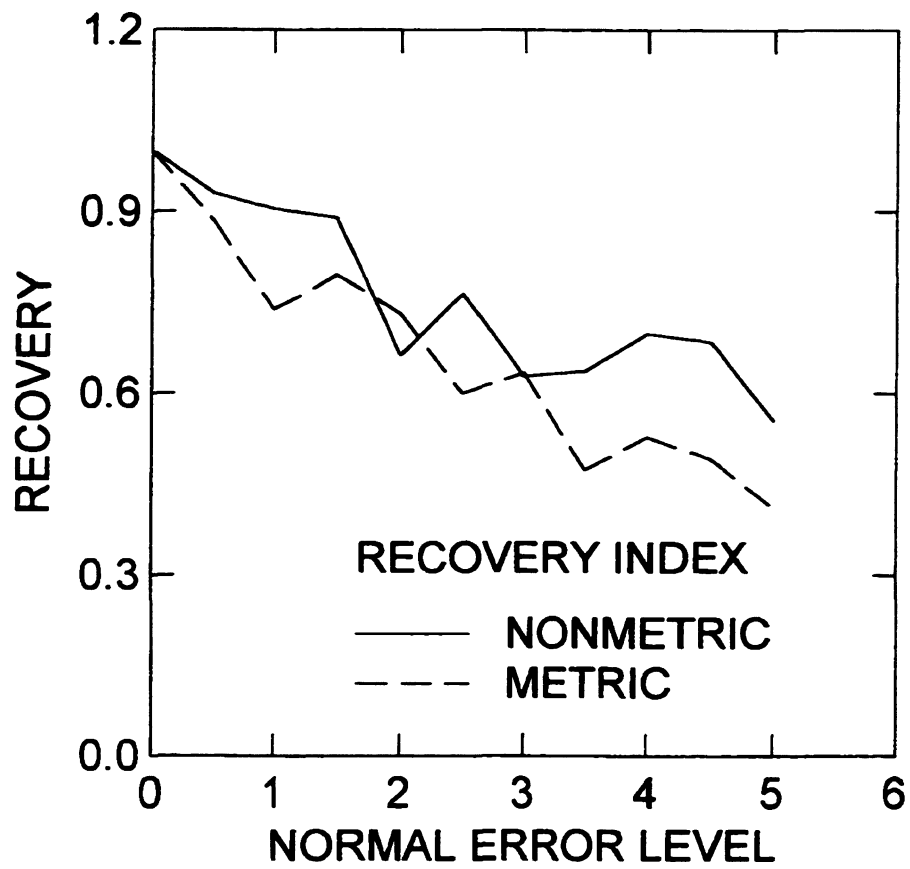
## Recovery as a Function of Normal Error

ERROR LEVEL	N	NONMETRIC RECOVERY		METRIC RECOVERY	
		MEAN	STD. DEV.	MEAN	STD. DEV.
0.0	10	1.000	0.000	1.000	0.000
0.5	9	0.931	0.050	0.885	0.115
1.0	11	0.906	0.070	0.739	0.220
1.5	6	0.891	0.047	0.796	0.262
2.0	9	0.663	0.278	0.732	0.232
2.5	12	0.765	0.122	0.601	0.282
3.0	11	0.628	0.233	0.634	0.266
3.5	13	0.637	0.212	0.474	0.316
4.0	10	0.699	0.147	0.528	0.246
4.5	8	0.685	0.128	0.492	0.246
5.0	11	0.555	0.218	0.411	0.279

consideration to error levels between 0 and 1.5, comparable to reliabilities between 0.40 and 1.00, the procedure seems to fare reasonably well. Nonmetric recovery declines only to 0.891, and metric recovery to 0.739.

FIGURE 5

NSKMDS: NORMAL ERROR LEVEL



Asymmetric Error. In addition to error corruption of the individual cells in a matrix, the vectors used in constructing the asymmetries may become contaminated by error. That is, the vectors underlying the asymmetries may reflect how the stimuli are represented cognitively. These internal representations, i.e., the vector elements, may themselves be susceptible to random fluctuations (e.g., Ramsay, 1969).

Error added to individual cells can be divided into symmetric and skew-symmetric components. In the ten by ten matrices with which we have been working, 55 of the 100 degrees of freedom are associated with the symmetries (Bishop, Fienberg, & Holland, 1975), suggesting that, on average, 55% of the random normal error will have no effect on a skew-symmetric technique such as NSKMDS. If the vectors underlying the asymmetries are corrupted however, all of the effect would pertain to the asymmetries specifically. Moreover, because the original vectors are being corrupted, it may be harder for an algorithm such as NSKMDS to recover the true configuration. Thus, asymmetric error poses a different and possibly harder challenge than cell-by-cell normal error.

To test the performance of NSKMDS under conditions of asymmetric error, 110 trials were conducted, covering error levels between 0 and 5.0, in increments of 0.5. Thus the levels 0, 0.5, 1.0, ... 4.0, 4.5, and 5.0 were included.

In the results, shown in Table 9 and in Figure 6, nonmetric recovery remains above .950 for error levels 0 to 1.5, declining rapidly thereafter, to as low as 0.580 at an error level of 4.0. A similar pattern is seen for metric recovery, with values decreasing from 1.000 at an error level of 0, to 0.898 at an error level of 1.5, and then sharply thereafter, reaching a low of 0.415 at an error level of 4.5. The linear trend component of the decline is significant

for metric recovery,  $F(1, 9) = 39.65$ ,  $p < .0005$ , and for nonmetric recovery,  $F(1, 9) = 18.47$ ,  $p < .005$ ). However, even after excluding the no-error condition, there was significant heteroscedasticity for both metric ( $F_{\max} = 200.02$ ,  $p < .05$ ) and nonmetric ( $F_{\max} = 497.63$ ,  $p < .05$ ) recovery.

As was true for normally distributed error, the nonmetric algorithm appears to offer excellent recovery of metric and nonmetric information at error levels below 1.5. If we assume that these are error levels typical for real-world data, the algorithm appears to function quite well.

Interaction. Thus, the NSKMDS algorithm appears relatively unaffected by reasonable levels of normal and asymmetric error, and monotonic distortion. The recovery levels seem high enough, however, to raise questions about how representative the simulations are of actual data. One possibility is that the two types of error, and monotonic distortion, interact in ways that are deleterious to the algorithm's functioning.

To test this, normal and asymmetric error, and monotonic distortion, were varied simultaneously in a 3-way, completely crossed, factorial design. Normal error level was set to either 2.0 or 4.0, monotonic distortion was set to either 1.5 or 3.0, and asymmetric error level was set to 2.0 or 4.0. The relatively high levels of error were adopted to prevent "ceiling" effects in recovery that could obscure true differences between cells in the design.

Table 9

## Nonmetric Analysis (NSKMDS)

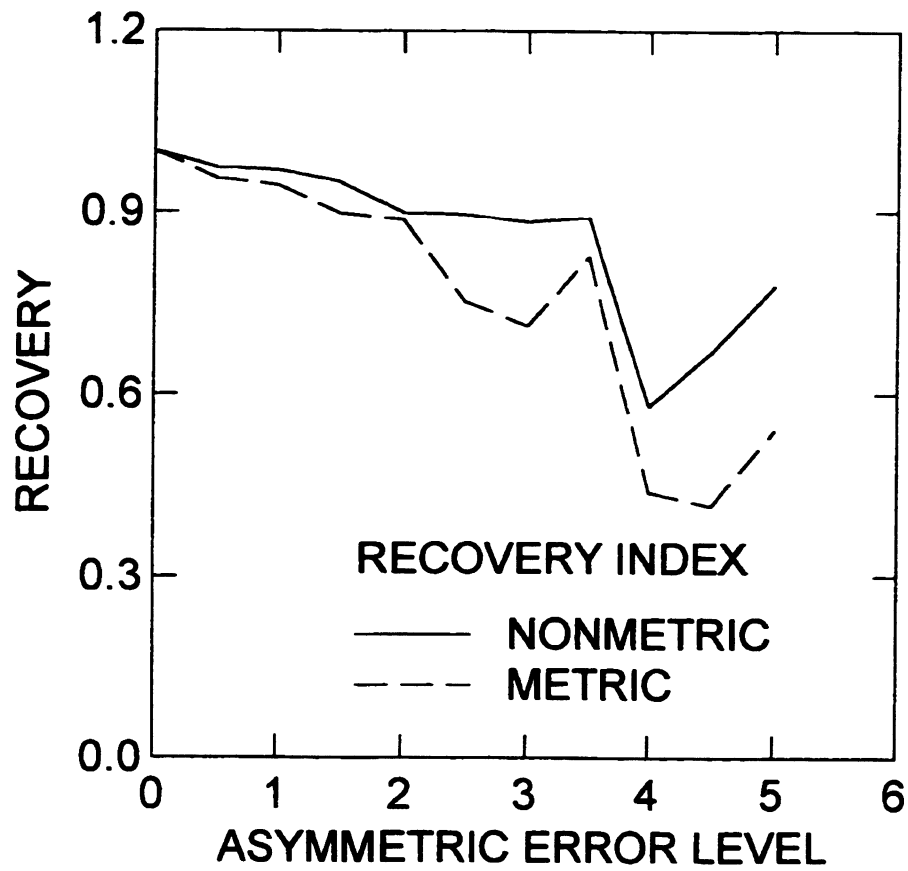
## Recovery As A Function of Asymmetric Error

ERROR LEVEL	N	NONMETRIC RECOVERY		METRIC RECOVERY	
		MEAN	STD. DEV.	MEAN	STD. DEV.
0.0	12	1.000	0.001	1.000	0.000
0.5	10	0.973	0.048	0.955	0.021
1.0	9	0.969	0.013	0.944	0.031
1.5	10	0.950	0.028	0.898	0.090
2.0	9	0.899	0.066	0.887	0.076
2.5	10	0.895	0.063	0.752	0.233
3.0	11	0.883	0.121	0.712	0.263
3.5	11	0.890	0.068	0.825	0.113
4.0	11	0.580	0.290	0.437	0.262
4.5	8	0.669	0.253	0.415	0.297
5.0	9	0.778	0.103	0.540	0.249

Each of the eight cells was represented by thirty trials. Initially, 240 simulation trials were run, with each trial randomly assigned at its outset to one of the eight combinations of error and monotonicity level. However, due to the anticipated heterogeneity of variance a balanced design seemed essential. Therefore, after 240 simulation trials had been run, trials were deleted at random from cells that had been over-represented, and additional trials were conducted to supplement under-represented cells.

FIGURE 6

NSKMDS: ASYMMETRIC ERROR



A two by two by two factorial analysis of variance was conducted. Results are shown in Table 10, treating normal error, asymmetric error, and monotonic distortion as fixed effects. Despite the leniency of this assumption and possibly some inflation in the Type I error rate for nonmetric recovery due to heterogeneity of variance (Keppel, 1982, p. 87;  $F_{\max} = 3.28, p < .05$ ) none of the interactions is significant for either metric or nonmetric recovery. As seen in Tables 10 and 11, the primary influence on recovery is the main effect for asymmetric error, and there is a smaller but significant main effect for monotonic distortion level. The main effect for normal error is inconsistent, attaining significance for nonmetric but not for metric recovery.

For metric recovery, none of the main effects is significant under a random effects model. None of the two-way interactions is significant when tested against the three-way interaction, as is appropriate for a random effects model. The significance test for the three way interaction is unaffected by the change to a random effects model, and remains nonsignificant.

Similarly, for nonmetric recovery no effects are significant under a random effects model. The absence of significant effects in the random effects models seems attributable to a marked loss of power, due to the low computed denominator degrees of freedom (1; Keppel, 1982, p. 642) for the tests.

Table 10  
Nonmetric Analysis (NSKMDS)

Analysis of Variance

Normal Error Level By Asymmetric Error Level By Monotonic Distortion

Metric Recovery (Fixed Effects Model)

SOURCE	SS	df	MS	F	p	$\omega^2$
NORMAL ERROR LEVEL	.009	1	.009	.150	ns	-
MONOT. DISTOR.	.485	1	.485	7.78	.006	.021
ASYMM. ERROR	5.150	1	5.150	82.66	.0005	.251
NORMAL BY MONOT.	.014	1	.014	.223	ns	-
NORMAL BY ASYMM.	.001	1	.001	.014	ns	-
MONOT. BY ASYMM.	.004	1	.004	.067	ns	-
NORMAL BY MONOT. BY ASYMM.	.119	1	.119	1.196	ns	-
WITHIN GROUPS	14.454	232	.062			



Table 10  
(Continued)

Nonmetric Analysis (NSKMDS)

Analysis of Variance

Normal Error Level By Asymmetric Error Level By Monotonic Distortion

Nonmetric Recovery (Fixed Effects Model)

SOURCE	SS	df	MS	F	p	$\omega^2$
NORMAL ERROR LEVEL	.277	1	.277	6.71	.010	.017
MONOT. DISTOR.	.601	1	.601	14.57	.0005	.040
ASYMM. ERROR	3.265	1	3.265	79.14	.0005	.230
NORMAL BY MONOT.	.079	1	.079	1.92	ns	-
NORMAL BY ASYMM.	.032	1	.032	.79	ns	-
MONOT. BY ASYMM.	.127	1	.127	3.071	ns	-
NORMAL BY MONOT. BY ASYMM.	.016	1	.016	.396	ns	-
WITHIN GROUPS	9.570	232	.041			

Table 11

## Nonmetric Analysis (NSKMDS)

Normal Error By Asymmetric Error By Monotonic Distortion

NML ERR	ASY ERR	MON DIST	NONMETRIC RECOVERY		METRIC RECOVERY	
			MEAN	STD DEV	MEAN	STD DEV
2	2	1.5	.782	.189	.636	.292
2	2	3	.781	.173	.495	.292
2	4	1.5	.635	.170	.311	.205
2	4	3	.509	.200	.242	.187
4	2	1.5	.791	.142	.568	.262
4	2	3	.684	.210	.546	.274
4	4	1.5	.564	.255	.324	.264
4	4	3	.398	.257	.196	.194

## Chapter 10

## Simulation Studies: Canonical Analysis of Asymmetry

Although canonical analysis of asymmetry is not a new technique, it seems to have rarely if ever been used. Therefore there is little experience with its performance, and simulations similar to those conducted with NSKMDS appear useful. The general settings for these studies are the same as those for the simulations involving the nonmetric technique.

Monotonic distortion. Two hundred and twenty trials were run, spanning monotonic distortion levels between 1.0 (no distortion) and 3.0, in increments of 0.1. Thus, the levels 1.0, 1.1, 1.2, ... 2.8, 2.9, and 3.0 were included in the study. Only positively accelerated distortion was used. Normal and asymmetric error levels were set to zero.

The results are shown in Table 12 and in Figure 7. Nonmetric recovery is almost perfect, with a mean recovery index above .96 for all levels of monotonic distortion. Metric recovery shows a steady decline, from 1.000 in the no-distortion condition, to .732 at a distortion level of 3.0. For metric recovery the linear trend is significant,  $F(1, 19) = 472.99$ ,  $p < .0005$ , with a slope of  $-.13$ . The linear trend for nonmetric recovery is significant as well,  $F(1, 19) = 13.09$ ,  $p < .005$ , although the slope is only  $-.01$ . There is heteroscedasticity for both the metric recovery index,  $F_{\max} = 5041.00$ ,  $p < .05$ , and the nonmetric index,  $F_{\max} = 9801$ ,  $p < .05$ .

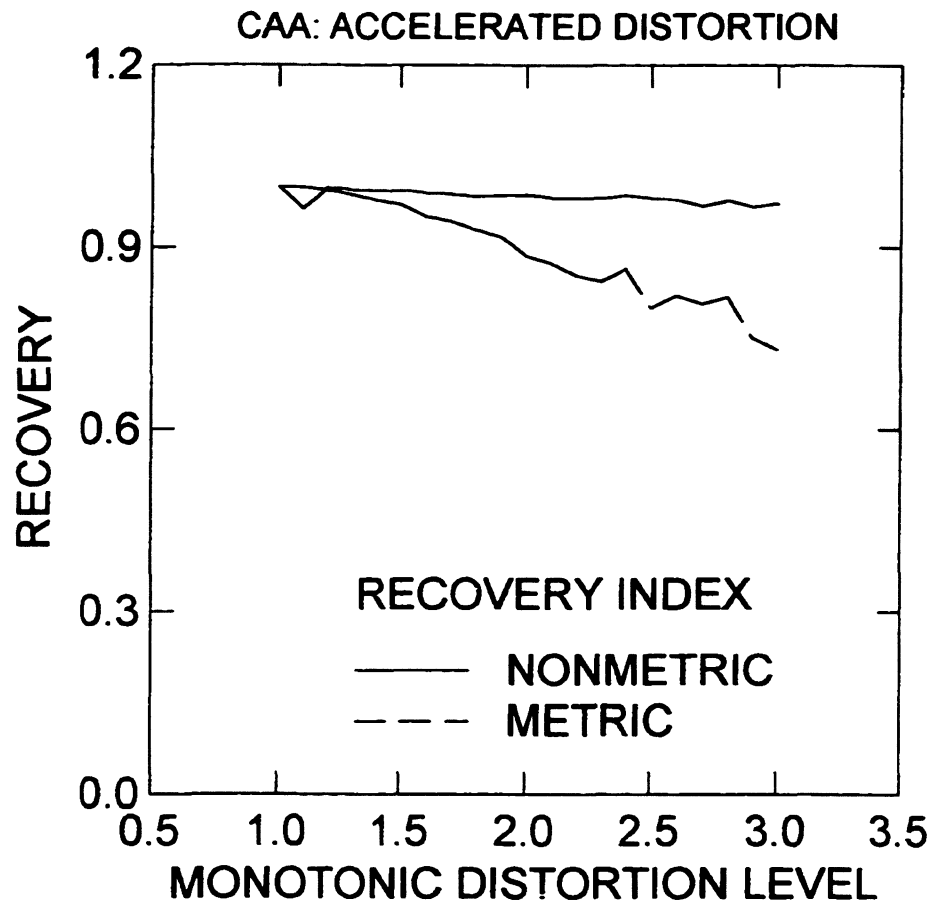
Table 12

## Canonical Analysis of Asymmetry (CAA)

## Recovery Under Positively Accelerated Monotonic Distortion

DISTORT. LEVEL	N	NONMETRIC RECOVERY		METRIC RECOVERY	
		MEAN	STD. DEV.	MEAN	STD. DEV.
1.0	6	1.000	0.000	1.000	0.000
1.1	8	0.964	0.099	0.999	0.000
1.2	6	0.998	0.001	0.994	0.001
1.3	11	0.995	0.002	0.987	0.004
1.4	14	0.993	0.004	0.978	0.010
1.5	11	0.994	0.002	0.971	0.009
1.6	13	0.990	0.005	0.951	0.011
1.7	12	0.988	0.007	0.943	0.018
1.8	8	0.984	0.011	0.929	0.034
1.9	8	0.986	0.004	0.917	0.027
2.0	13	0.987	0.005	0.886	0.021
2.1	10	0.982	0.009	0.874	0.035
2.2	5	0.981	0.008	0.854	0.031
2.3	17	0.981	0.009	0.844	0.050
2.4	4	0.986	0.008	0.865	0.040
2.5	11	0.981	0.010	0.800	0.070
2.6	12	0.979	0.012	0.821	0.054
2.7	14	0.968	0.014	0.807	0.049
2.8	9	0.978	0.012	0.819	0.044
2.9	14	0.967	0.017	0.751	0.067
3.0	14	0.973	0.009	0.732	0.071

FIGURE 7

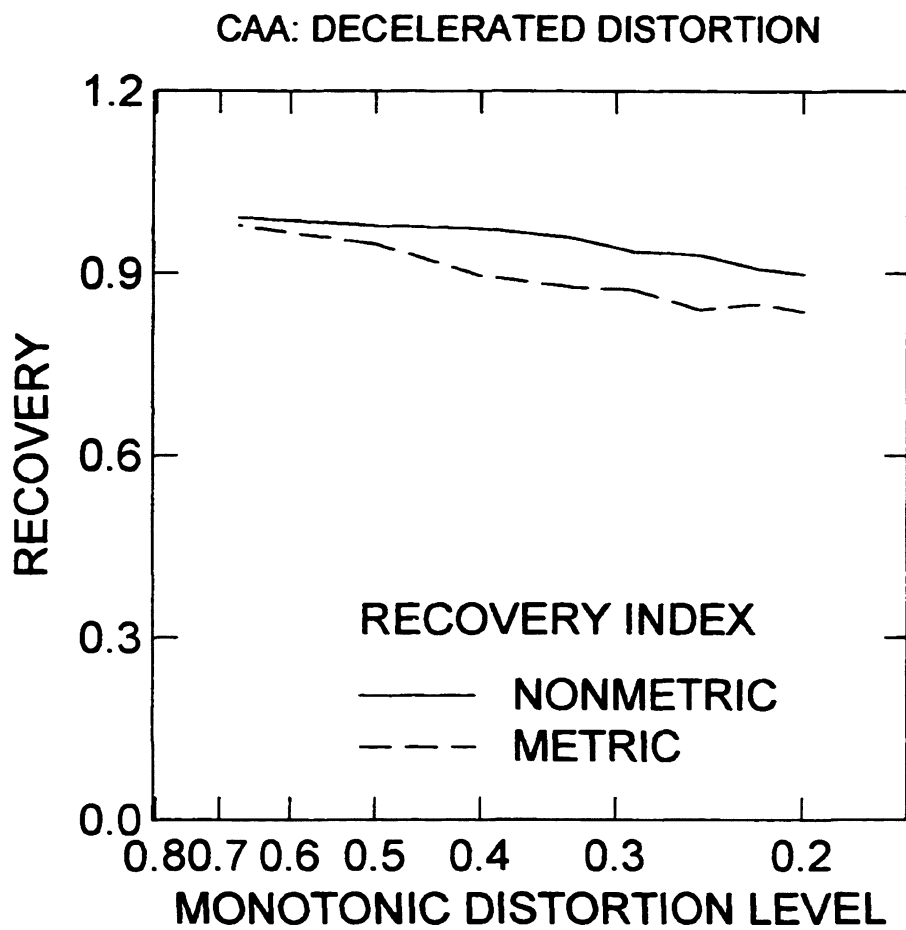


At first glance it seems counterintuitive that a metric technique would have relatively more difficulty with metric than with nonmetric recovery. However, in positively accelerated distortion, large matrix entries are made disproportionately even larger. A metric technique will attempt to reproduce these large entries. However, in doing so it will depart from the magnitude, although not necessarily the rank order, of many of the entries in the original, non-distorted matrix.

To study the effects of negatively accelerated distortion, monotonicity coefficients of .67, .50, .40, .33, .29, .25, .22, and .20 were used. These values are the reciprocals of representative levels of positively accelerated distortion, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0. Decreasing coefficients correspond to increasing levels of distortion. Forty trials were run at each distortion level. Normal and asymmetric error were set to zero.

The results are shown in Table 7, and in Figure 8. Metric recovery declines gradually from .977 at the .67 distortion level, to .836 at the .20 distortion level. For negatively accelerated distortion, nonmetric recovery also declines, albeit gradually, from .990 at the .67 distortion level, to .898 at the .20 distortion level. The linear downward trend involving the means is significant: For metric recovery,  $F(1,6) = 186.04$ ,  $p < .0005$ , with a slope of .31. For nonmetric recovery,  $F(1,6) = 30.57$ ,  $p < .001$ , with a slope of .19. The heteroscedasticity is significant, even after excluding conditions with no variability:  $F_{\max} = 129.39$ ,  $p < .05$  for metric recovery, and  $F_{\max} = 676.00$ ,  $p < .05$  for nonmetric recovery. As the trial-to-trial variability in recovery increases, the mean recovery level goes down.

FIGURE 8



Normal error. An advantage of standard factor analysis over nonmetric factoring, is the relative resistance of the metric technique to disruption by error. As recovery in NSKMDS declines steadily with error, it seemed reasonable to study canonical analysis of asymmetry under the same conditions. One hundred and ten trials, spanning 11 levels of normally-distributed error, were run. The normal error levels ranged from 0 (no error) to 5.0, in increments of 0.5. Asymmetric error and monotonic distortion were not included in these trials.

Results are shown in Table 13, and in Figure 9. Mean metric recovery declines gradually, from 1.000 at an error level of 0, to .826 and .873 at error levels of 4.5 and 5.0. The linear trend component is significant,  $F(1, 9) = 98.57$ ,  $p < .0005$ , and has a slope of  $-.03$ . Nonmetric recovery is more variable, although the downward linear trend is significant,  $F(1,9) = 10.88$ ,  $p < .01$ , with a slope of  $-.03$ . The conditions differ in their variability to a significant extent:  $F_{\max} = 112.89$ ,  $p < .05$  for metric recovery, and  $F_{\max} = 506.25$ ,  $p < .05$  for nonmetric recovery.

Nonmetric recovery is consistently lower than metric recovery. This contrasts with NSKMDS, for which nonmetric recovery generally appeared higher than metric recovery in the normal error conditions. However, inspection of Tables 8 and 13 suggests that nonmetric recovery may nonetheless be higher in an absolute sense, for the canonical analysis of asymmetry. The two techniques will be contrasted explicitly in the next chapter.



Table 13

## Canonical Analysis of Asymmetry (CAA)

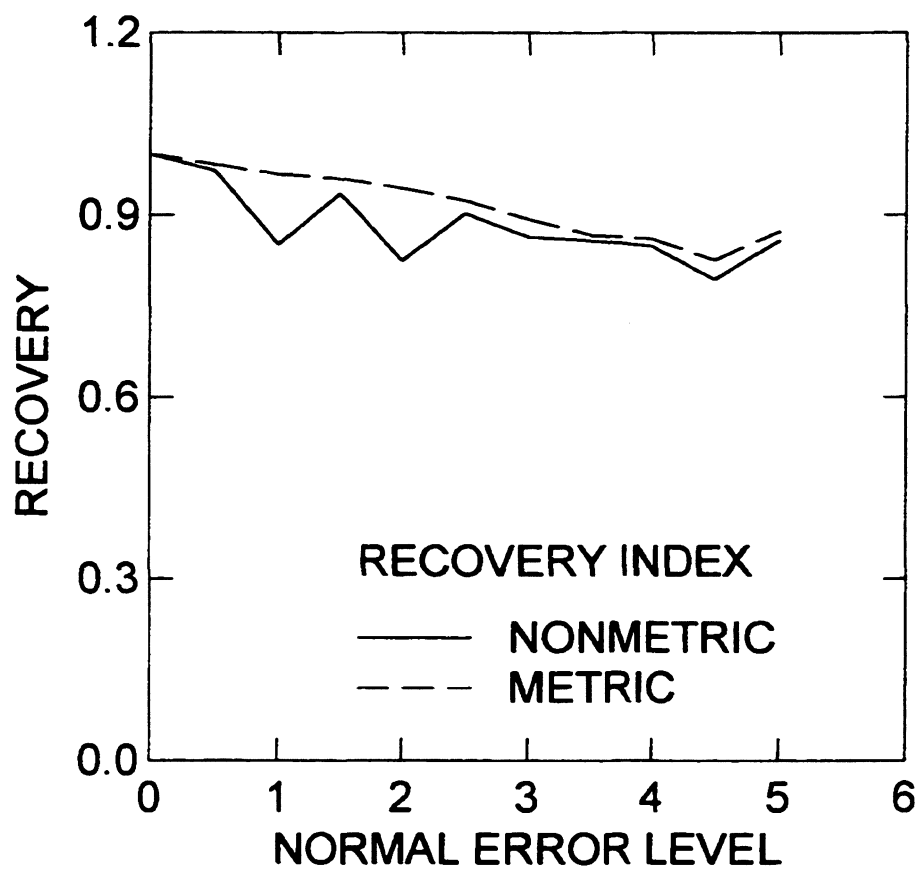
## Recovery As A Function Of Normal Error

ERROR LEVEL	N	NONMETRIC RECOVERY		METRIC RECOVERY	
		MEAN	STD. DEV.	MEAN	STD. DEV.
0.0	15	1.000	0.000	1.000	0.000
0.5	8	0.973	0.010	0.983	0.008
1.0	14	0.852	0.225	0.967	0.019
1.5	6	0.936	0.042	0.960	0.021
2.0	10	0.825	0.074	0.944	0.022
2.5	3	0.903	0.123	0.924	0.038
3.0	10	0.864	0.081	0.893	0.036
3.5	14	0.858	0.081	0.867	0.075
4.0	8	0.849	0.089	0.861	0.085
4.5	11	0.794	0.225	0.826	0.065
5.0	11	0.858	0.067	0.873	0.059

Asymmetric error. One hundred and ten trials were run, spanning eleven asymmetric error levels. Levels from 0 to 5, in increments of 0.5, were included in the study. Thus, metric and nonmetric recovery were assessed at asymmetric error levels of 0, 0.5, 1.0 ... 4.5, 5.0. Error level is the sum of squared differences between the true and error-perturbed matrices, as a

FIGURE 9

CAA: NORMAL ERROR LEVEL



proportion of the true sum of squares. Thus, an error level of 3 indicates that the residual sum of squares is 3 times greater than the sum of squares of the true matrix entries.

Results are shown in Table 14, and in Figure 10. Mean metric recovery declines gradually from 1.000 in the no-error condition, to .804 at an error level of 5.0. The linear trend component is significant,  $F(1, 9) = 50.41$ ,  $p < .0005$ , and has a slope of  $-.03$ .

Nonmetric recovery shows a similar pattern,

Table 14

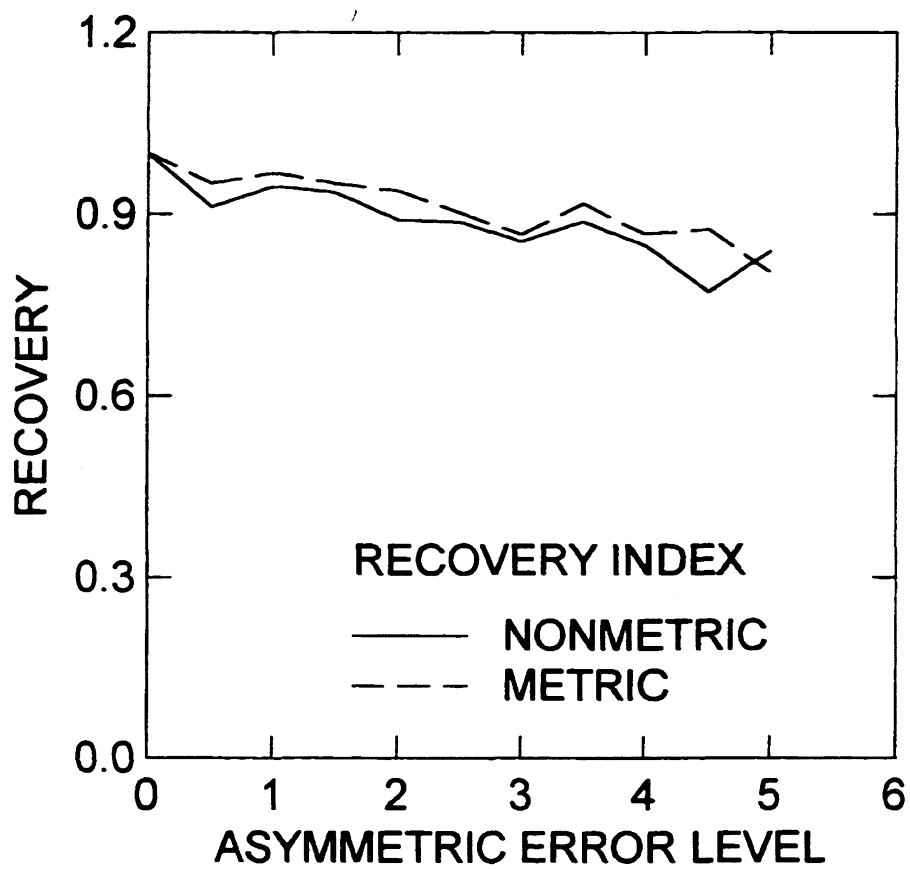
Canonical Analysis of Asymmetry (CAA)

Recovery As A Function Of Asymmetric Error

ERROR LEVEL	N	NONMETRIC RECOVERY		METRIC RECOVERY	
		MEAN	STD. DEV.	MEAN	STD. DEV.
0.0	9	1.000	0.000	1.000	0.000
0.5	9	0.912	0.064	0.951	0.033
1.0	11	0.946	0.060	0.968	0.014
1.5	11	0.936	0.037	0.951	0.030
2.0	16	0.891	0.096	0.939	0.028
2.5	6	0.887	0.041	0.902	0.044
3.0	16	0.855	0.101	0.867	0.089
3.5	10	0.888	0.029	0.918	0.026
4.0	8	0.849	0.058	0.868	0.050
4.5	6	0.772	0.124	0.875	0.074
5.0	8	0.839	0.090	0.804	0.105

FIGURE 10

CAA: ASYMMETRIC ERROR



declining from a mean of 1.000 at an error level of 0, to a mean of .772 at an error level of 4.5. Here, too, the linear trend component is significant,  $F(1, 9) = 31.91$ ,  $p < .0005$ , with a slope of  $-.03$ . The variance differs significantly among conditions, with  $F_{\max} = 56.25$ ,  $p < .05$ , for metric recovery, and  $F_{\max} = 18.28$ ,  $p < .05$ , for nonmetric recovery.

Hence, CAA, as NSKMDS, appears to be affected by asymmetric error. Comparing Tables 9 and 14 it appears that NSKMDS is affected more strongly by asymmetric error than is CAA. The two analysis techniques are compared directly in the next chapter.

Interaction. The canonical analysis of asymmetry appears to maintain acceptably high rates of recovery in the presence of even relatively large amounts of normal and asymmetric error, and monotonic distortion. Actual data, however, are likely to contain all three types of perturbation, and hence it is important to assess for interaction effects.

Two hundred and forty trials were run corresponding to the eight cells of a completely crossed two (normal error level) by two (asymmetric error level) by two (monotonic distortion level) factorial ANOVA. The normal error levels 2 and 4, the asymmetric error levels 2 and 4, and the monotonic distortion levels 1.5 and 3 were represented in the design. These levels were chosen to minimize "ceiling" and "floor" effects; that is, to prevent the results for two or more cells in the design from converging due to maximal or minimal recovery. For each of the 240 trials the combination of error and monotonic distortion conditions was assigned at random. However, heterogeneity of variance was expected, and therefore after the 240 trials had been run, entries were deleted at random from cells with more than 30 trials and additional simulations were run to supplement cells with less than 30 trials, to ensure a balanced design.

Results are shown in Tables 15 and 16. For metric recovery, under a fixed effects model only the main effects for monotonic distortion and asymmetric error are significant. The effect sizes are  $\omega^2 = .036$  for monotonic distortion and  $\omega^2 = .298$  for asymmetric error. None of the two-way interactions are significant, nor is the three-way interaction or the normal error level main effect. The results under a random effects model are strongly influenced by the virtual absence of a sum of squares for the three-way interaction. This ensures that each of the two way interactions is significant. Quasi-F ratios for the three main effects are not significant. As expected, there was marked heterogeneity of variance,  $E_{\max} = 992.25$ ,  $p < .05$ .

The fact that the mean square for the three way interaction is much lower than the pooled within cells mean square suggests that the three way interaction is not appropriate to use as an error term in this analysis. The fact that none of the interactions are significant when tested against the pooled within cells mean square suggests that the three types of perturbation are independent in their effects.

For nonmetric recovery, only the main effects of asymmetric error and monotonic distortion are significant under a fixed effects model. Effect sizes are  $\omega^2 = .024$  for monotonic distortion and  $\omega^2 = .256$  for asymmetric error. None of the interactions, nor the normal error main effect, are significant under a fixed effects analysis. Under a random effects model, the monotonic distortion and asymmetric error main effects remain significant. In addition the normal error by asymmetric error interaction achieves significance when tested against the mean square for the three-way interaction. Heterogeneity of variance was present, as indicated by Hartley's (1950) test,  $E_{\max} = 529.00$ ,  $p < .05$ .

The same concerns for the appropriateness of using the three-way interaction as an error term would seem to apply in the analysis of nonmetric recovery as pertained in the analysis of metric recovery.

Table 15

## Canonical Analysis of Asymmetry (CAA)

## Normal Error By Asymmetric Error By Monotonic Distortion

NML. ERROR	ASY. ERROR	MON. DIST.	NONMETRIC RECOVERY		METRIC RECOVERY	
			MEAN	STD. DEV.	MEAN	STD. DEV.
2	2	1.5	.899	.003	.923	.002
2	2	3	.827	.014	.844	.003
2	4	1.5	.630	.065	.658	.063
2	4	3	.507	.069	.504	.063
4	2	1.5	.845	.029	.858	.032
4	2	3	.789	.041	.827	.012
4	4	1.5	.651	.055	.678	.062
4	4	3	.584	.064	.572	.061

Table 16  
 Canonical Analysis of Asymmetry (CAA)  
 Analysis of Variance

Normal Error Level By Asymmetric Error Level By Monotonic Distortion

Metric Recovery (Fixed Effects Model)

SOURCE	SS	df	MS	F	p	$\omega^2$
NORMAL ERROR LEVEL	.000	1	.000	.003	ns	-
MONOT. DISTOR.	.516	1	.516	13.82	.0005	.036
ASYMM. ERROR	4.046	1	4.046	108.35	.0005	.298
NORMAL BY MONOT.	.034	1	.034	.920	ns	-
NORMAL BY ASYMM.	.107	1	.107	2.86	ns	-
MONOT. BY ASYMM.	.084	1	.084	2.25	ns	-
NORMAL BY MONOT. BY ASYMM.	.000	1	.000	.000	ns	-
WITHIN GROUPS	8.663	232	.037			



Table 16  
(Continued)

Canonical Analysis of Asymmetry (CAA)

Analysis of Variance

Normal Error Level By Asymmetric Error Level By Monotonic Distortion

Nonmetric Recovery (Fixed Effects Model)

SOURCE	SS	df	MS	F	p	$\omega^2$
NORMAL ERROR LEVEL	.000	1	.000	.003	ns	-
MONOT. DISTOR.	.382	1	.382	8.98	.003	.024
ASYMM. ERROR	3.655	1	3.655	85.99	.0005	.256
NORMAL BY MONOT.	.020	1	.020	.468	ns	-
NORMAL BY ASYMM.	.137	1	.137	3.22	ns	-
MONOT. BY ASYMM.	.015	1	.015	.355	ns	-
NORMAL BY MONOT. BY ASYMM.	.006	1	.006	.139	ns	-
WITHIN GROUPS	9.862	232	.043			

## Chapter 11

### Simulation Studies: Comparing Metric and Nonmetric Techniques

Considered by itself, then, the nonmetric analysis of skew-symmetry appears to function quite well in the presence of reasonable levels of normal and asymmetric error, and monotonic distortion. There is little evidence from the simulations to suggest that these three factors interact to lower the performance of the algorithm. Moreover, metric and nonmetric recovery levels have been high enough that it seems reasonable use the algorithm in the analysis of actual data. However, the performance of its metric counterpart, the canonical analysis of asymmetry, thus far appears at least as high. Therefore, several simulations were conducted to compare the two techniques.

Monotonic distortion. To study the effects of accelerating monotonic distortion (exponents  $> 1.0$ ), 20 trials were run for each of the two analysis methods, at each of eight distortion levels, in a completely crossed factorial design. The eight distortion levels were 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0. The variance components of interest were the main effect for analysis type, and the analysis type by distortion level interaction.

Results are shown in Tables 17 and 18, and in Figures 11a and 11b. For metric recovery, neither effect of interest is significant whether distortion level is treated as a fixed or random effect. For nonmetric recovery, the main effect for analysis type is significant under both a fixed effect and a random effect model. As may be seen in Tables 17 and 18, and in Figure 11a, there is a small but consistent tendency for the metric technique to give higher nonmetric recoveries than does the nonmetric technique.

Table 17

## Analysis Method By Positively Accelerated Monotonic Distortion

## Analysis of Variance For Metric Recovery

Fixed Effects Model

SOURCE	SS	df	MS	F	p	$\omega^2$
DISTORTION LEVEL	7.616	7	1.088	131.61	.0005	.739
METHOD	.006	1	.006	.707	ns	-
DISTORTION LEVEL BY METHOD	.082	7	.012	1.42	ns	-
WITHIN GROUPS	2.513	304	.008			

Method=Fixed Effect, Distortion Level=Random Effect

SOURCE	SS	df	MS	F	p
DISTORTION LEVEL	7.616	7	1.088	131.61	.0005
METHOD	.006	1	.006	.500	ns
DISTORTION LEVEL BY METHOD	.082	7	.012	1.42	ns
WITHIN GROUPS	2.513	304	.008		

Table 18

Analysis Method By Positively Accelerated Monotonic Distortion  
Analysis of Variance For Nonmetric Recovery

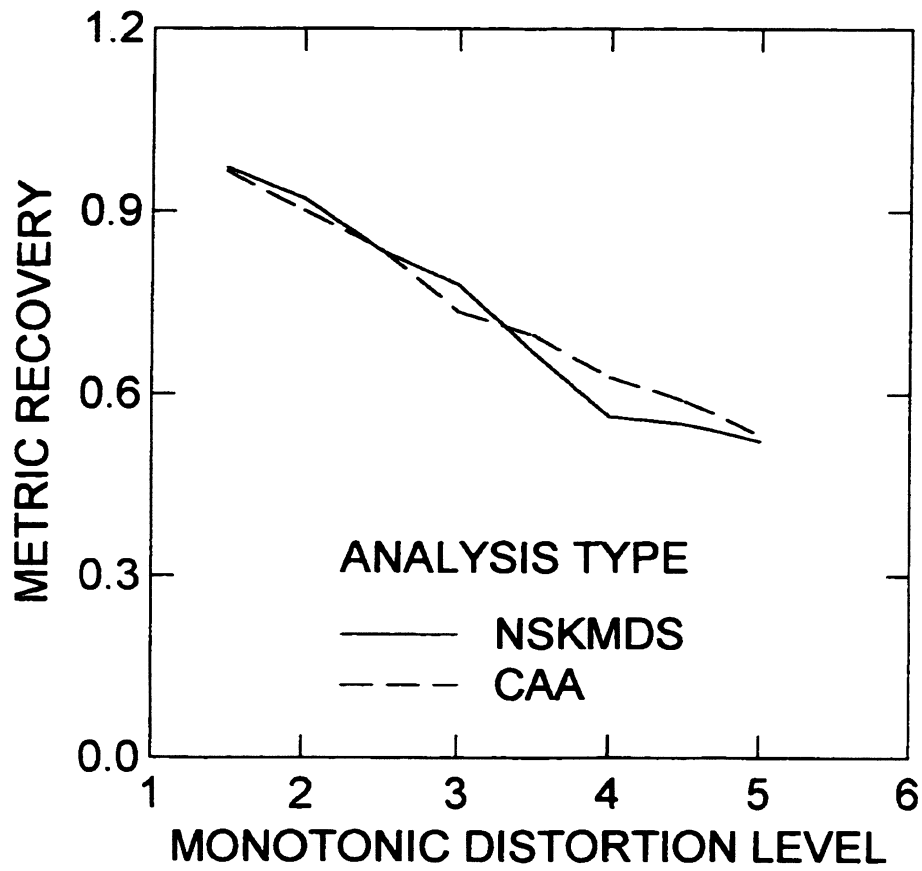
Fixed Effects Model

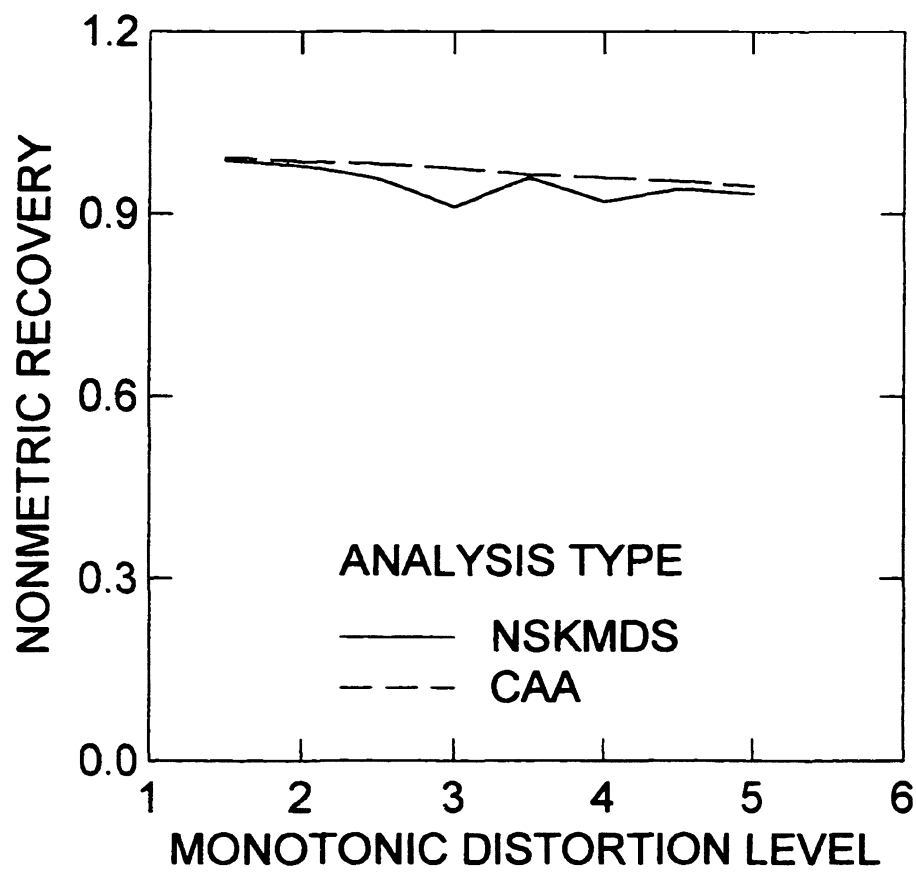
SOURCE	SS	df	MS	F	p	$\omega^2$
DISTORTION LEVEL	.114	7	.016	3.43	.005	.049
METHOD	.038	1	.038	7.93	.005	.020
DISTORTION LEVEL BY METHOD	.031	7	.004	.934	ns	-
WITHIN GROUPS	1.437	304	.005			

Method=Fixed Effect, Distortion Level=Random Effect

SOURCE	SS	df	MS	F	p
DISTORTION LEVEL	.114	7	.016	3.43	.005
METHOD	.038	1	.038	9.50	.05
DISTORTION LEVEL BY METHOD	.031	7	.004	.93	ns
WITHIN GROUPS	1.437	304	.005		

FIGURE 11A

METRIC RECOVERY V.  
ACCELERATED DISTORTION

**FIGURE 11B****NONMETRIC RECOVERY V.  
ACCELERATED DISTORTION**

This result, however, should be qualified in two ways. First,  $\omega^2$  for the main effect of analysis type is quite small, at .02. Second, both techniques appear to be providing nonmetric recovery at acceptable levels, greater than .91 for NSKMDS and greater than .94 for the canonical analysis of asymmetry.

Decelerating monotonic distortions were studied in a completely crossed two-way factorial design. Eight distortion levels, 0.67, 0.50, 0.40, 0.33, 0.29, 0.25, 0.22, and 0.20, were crossed with metric vs. nonmetric analysis type. The eight distortion levels included in the design were the reciprocals of the eight levels of accelerating distortion used in the previous simulation. Twenty trials were run for each of the sixteen cells in the design. The magnitudes of metric and nonmetric recovery were reported above, in Table 7.

The ANOVA for metric recovery is shown in Table 19. As may be seen, the main effect for analysis method, the main effect for distortion level, and the interaction are significant, both when distortion level is treated as random, and when it is treated as fixed.  $\omega^2$  for the main effect of analysis method is 0.332.  $\omega^2$  for the interaction is .064. As shown in Figure 12, metric recovery is consistently higher with the nonmetric algorithm. The effect appears to be particularly strong at the higher levels of distortion, which may account for the interaction. Not surprisingly, given the size of the main effect, the simple main effects are significant at all levels of distortion.

Table 19

## Analysis Method vs. Decelerated Monotonic Distortion Level

## Analysis of Variance For Metric Recovery

Fixed Effects Model

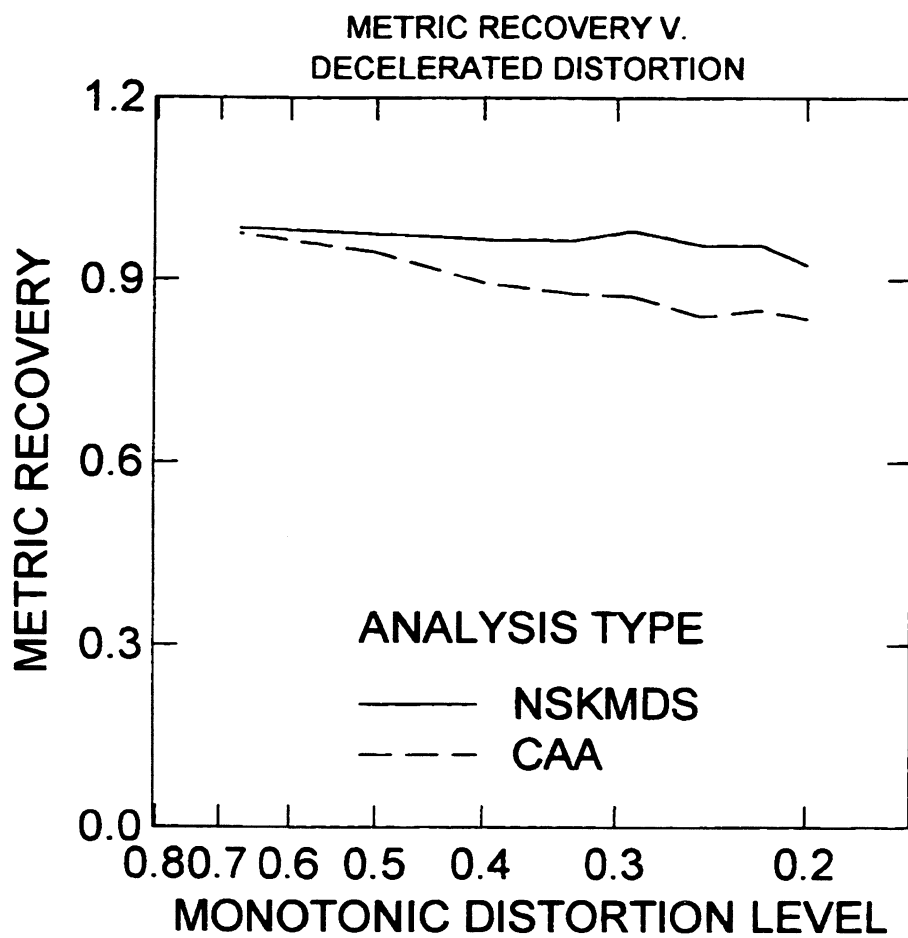
Source	SS	df	MS	F	p	$\omega^2$
DISTORTION LEVEL	.305	7	.044	24.04	.0005	.201
METHOD	.483	1	.483	266.40	.0005	.332
DISTORTION LEVEL BY METHOD	.107	7	.015	8.44	.0005	.064
WITHIN GROUPS	.551	304	.002			

Method=Fixed Effect, Distortion Level=Random Effect

Source	SS	df	MS	F	p
DISTORTION LEVEL	.305	7	.044	24.04	.0005
METHOD	.483	1	.483	32.20	.001
DISTORTION LEVEL BY METHOD	.107	7	.015	8.44	.0005
WITHIN GROUPS	.551	304	.002		



FIGURE 12



For nonmetric recovery, results are shown in Table 20 and in Figure 13.

Table 20

Analysis Method vs. Decelerated Monotonic Distortion Level

Analysis of Variance For Nonmetric Recovery

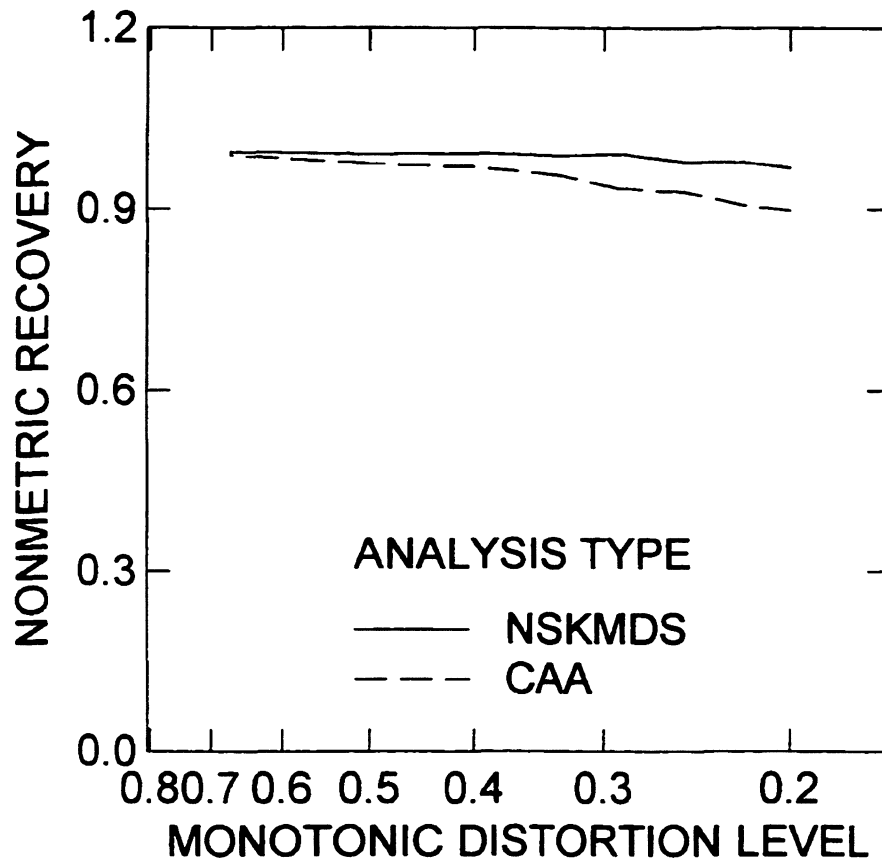
Fixed Effects Model

Source	SS	df	MS	F	p	$\omega^2$
DISTORTION LEVEL	.123	7	.018	26.74	.0005	.233
METHOD	.129	1	.129	196.54	.0005	.257
DISTORTION LEVEL BY METHOD	.046	7	.007	9.98	.0005	.078
WITHIN GROUPS	.199	304	.001			

Method=Fixed Effect, Distortion Level=Random Effect

Source	SS	df	MS	F	p
DISTORTION LEVEL	.123	7	.018	26.74	.0005
METHOD	.129	1	.129	18.43	.01
DISTORTION LEVEL BY METHOD	.046	7	.007	9.98	.0005
WITHIN GROUPS	.199	304	.001		

FIGURE 13

NONMETRIC RECOVERY V.  
DECELERATED DISTORTION

The two main effects and the interaction are significant, when distortion level is treated as random, and when it is treated as fixed.  $\omega^2$  for the main effect of analysis type is .257.  $\omega^2$  for the interaction is .078. The nonmetric algorithm gives higher nonmetric recovery levels. This effect appears to be strongest at the higher distortion levels, which seems to underlie the interaction. The simple main effects are significant at all monotonic distortion levels, as would be expected given the magnitude of the main effect.

Hence there is a suggestion that the nonmetric technique is superior for both metric and nonmetric recovery under a decelerating monotonic distortion. Moreover, the effect sizes associated with this appear reasonably high. Nonetheless, both techniques are giving a good reproduction of the true matrix entries. In all conditions, metric recovery is greater than .920 for NSKMDS, and greater than 0.830 for the canonical analysis of asymmetry. Nonmetric recovery is greater than .960 for NSKMDS and greater than .890 for the canonical analysis of asymmetry.

Normal error. Although the nonmetric analysis of skew-symmetry appeared to handle moderate levels of normal error reasonably well, susceptibility to error is generally a weakness of nonmetric algorithms. Therefore, NSKMDS was tested against the canonical analysis of asymmetry in the presence of normal error.

Four error levels, 1.0, 2.0, 3.0, and 4.0, were combined with analysis method (NSKMDS, the canonical analysis of asymmetry) in a completely crossed factorial design. Twenty trials were run for each of the eight cells.

Table 21

Analysis Method vs. Normal Error Level  
 Analysis of Variance For Metric Recovery

Fixed Effects Model

Source	SS	df	MS	F	p	$\omega^2$
ERROR LEVEL	.464	3	.155	5.60	.001	.054
METHOD	2.142	1	2.142	77.47	.000	.301
ERROR LEVEL BY METHOD	.186	3	.062	2.25	ns	
WITHIN GROUPS	4.203	152	.028			

Method=Fixed Effect, Error Level=Random Effect

Source	SS	df	MS	F	p
ERROR LEVEL	.464	3	.155	5.60	.001
METHOD	2.142	1	2.142	35.55	.01
ERROR LEVEL BY METHOD	.186	3	.062	2.25	ns
WITHIN GROUPS	4.203	152	.028		

For metric recovery, the main effects of error level and of analysis type are significant, regardless of whether error level is treated as a random or a fixed effect (Table 21). The interaction is not significant under either model. The main effects are shown in Figure 14a, where it may be noted that the canonical analysis of skew-symmetry provides

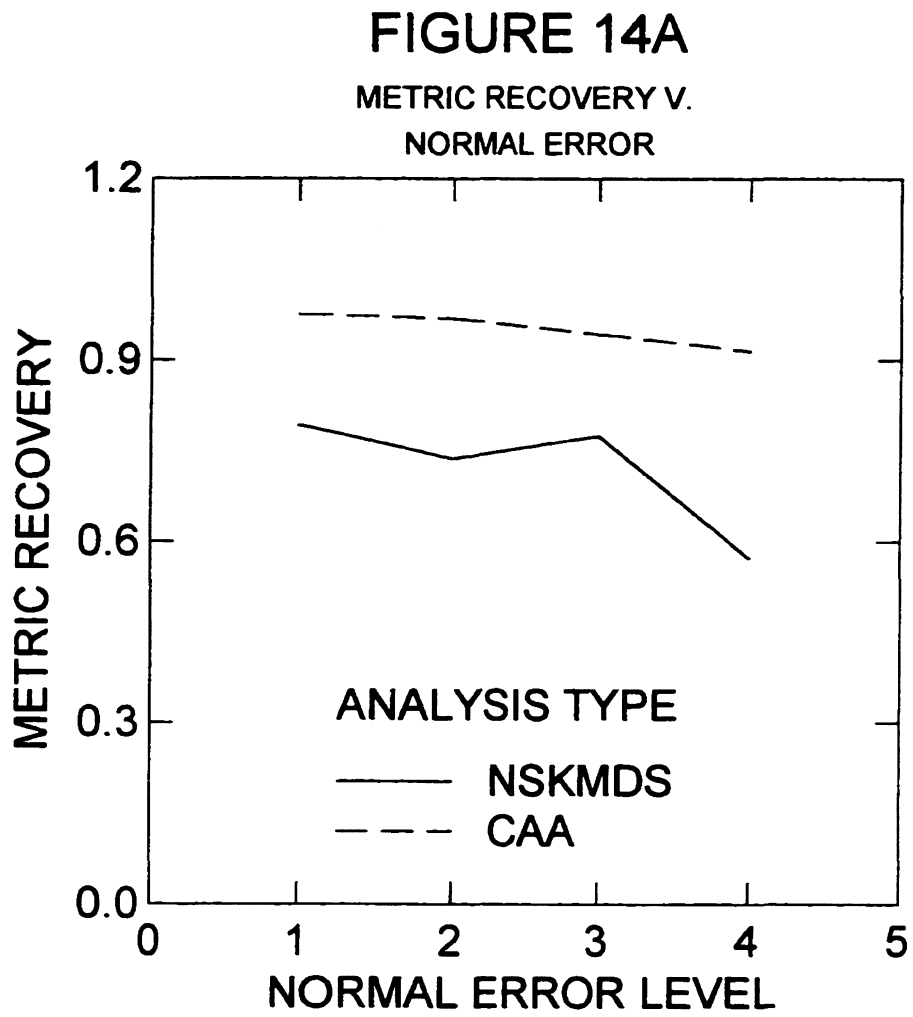


FIGURE 14B

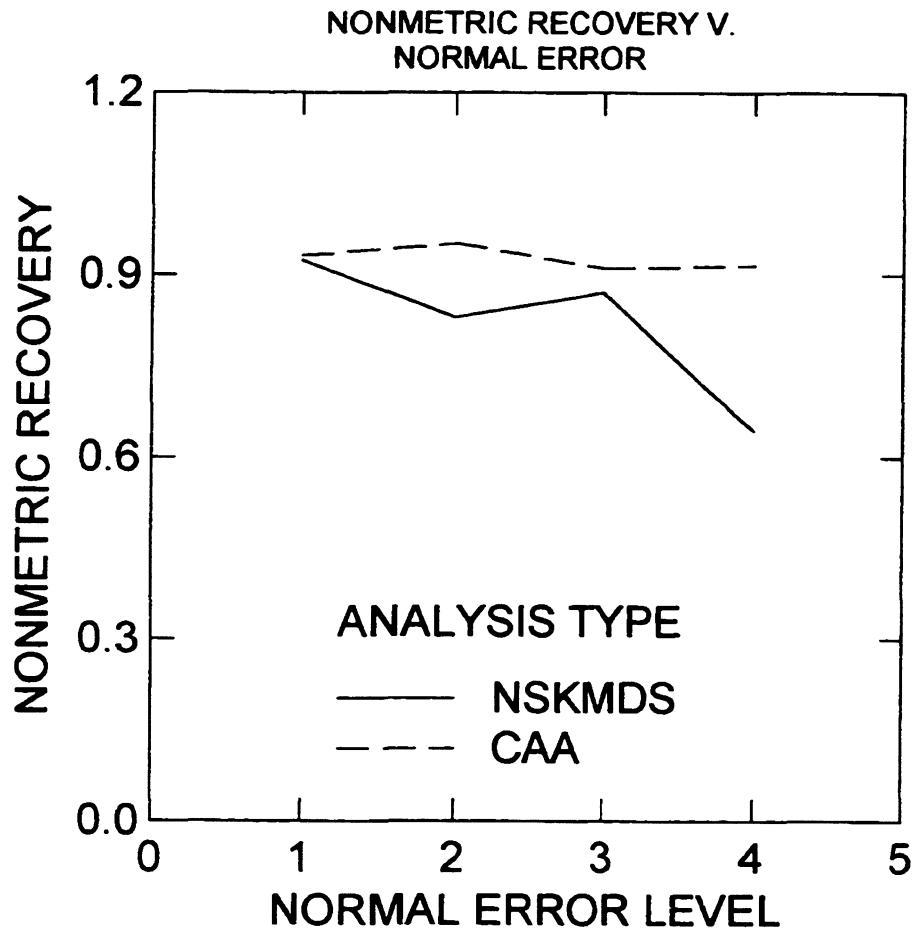


Table 22

Analysis Method vs. Normal Error Level  
Analysis of Variance For Nonmetric Recovery

Fixed Effects Model

Source	SS	df	MS	F	p	$\omega^2$
ERROR LEVEL	.495	3	.165	12.60	.0005	.134
METHOD	.488	1	.488	37.24	.0005	.140
ERROR LEVEL BY METHOD	.416	3	.139	10.57	.0005	.111
WITHIN GROUPS	1.993	152	.013			

Method=Fixed Effect, Error Level=Random Effect

Source	SS	df	MS	F	p
ERROR LEVEL	.495	3	.165	12.60	.0005
METHOD	.488	1	.488	3.51	ns
ERROR LEVEL BY METHOD	.416	3	.139	10.57	.0005
WITHIN GROUPS	1.993	152	.013		

superior metric recovery. The effect size is quite high, with  $\omega^2 = .301$  for the main effect of method.

For nonmetric recovery (Table 22) the main effect for error level and the interaction of error level with analysis type are significant, when error level is treated as a random effect



and when it is treated as a fixed effect. The main effect of analysis type is significant under the fixed effects but not the random effects model.  $\omega^2$  for the analysis type main effect is .140, and for the interaction  $\omega^2$  is .111. As suggested by Figure 14b, the canonical analysis of asymmetry provides higher nonmetric recovery than NSKMDS at error levels 2.0, and 4.0. The simple main effect is not significant at error levels of 1.0 and 3.0. It seems likely, however, that the absence of significance at the 3.0 level is a Type II error, and that the canonical analysis of asymmetry provides better recovery, at least at the higher error levels.

As in previous simulations, the differences between the two techniques appear greatest at higher error levels than would presumably be encountered in actual data. At the 1.0 error level the nonmetric recoveries do not differ significantly between the two techniques, and both metric and nonmetric recoveries appear acceptably high for NSKMDS. Nonetheless the magnitude of the main effect is large, and particularly when reasonably high levels of normal error are expected, CAA seems to be a better choice of analysis technique.

Asymmetric error. Because the effect of asymmetric error is potentially different from that of normal error, it was studied in a separate simulation.

Four levels of asymmetric error, 1.0, 2.0, 3.0, and 4.0, were combined with analysis method in a completely crossed factorial. Twenty trials were run for each of the eight cells.

The results for metric recovery are shown in Table 23. The main effect for error level and the analysis method by error level interaction are significant under a fixed effects model and when error level is regarded as a random effect. The main effect of method is

Table 23

## Analysis Method vs. Asymmetric Error Level

## Analysis of Variance For Metric Recovery

Fixed Effects Model

Source	SS	df	MS	F	p	$\omega^2$
ERROR LEVEL	1.523	3	.508	23.564	.000	.228
METHOD	.962	1	.962	44.656	.000	.147
ERROR LEVEL BY METHOD	.596	3	.199	9.228	.000	.083
WITHIN GROUPS	3.274	152	.022			

Method=Fixed Effect, Error Level=Random Effect

Source	SS	df	MS	F	p
ERROR LEVEL	1.523	3	.508	23.56	.0005
METHOD	.962	1	.962	4.83	ns
ERROR LEVEL BY METHOD	.596	3	.199	9.23	.0005
WITHIN GROUPS	3.274	152	.022		

significant only under a fixed effects model. Analysis of the simple main effects of analysis type indicates that metric recovery is significantly higher for the canonical analysis of asymmetries at error levels 2.0, 3.0, and 4.0 (see Figure 15).

For nonmetric recovery the results are shown in Table 24 and in Figure 16.

Table 24

Analysis Method vs. Asymmetric Error Level  
Analysis of Variance For Nonmetric Recovery

Fixed Effects Model

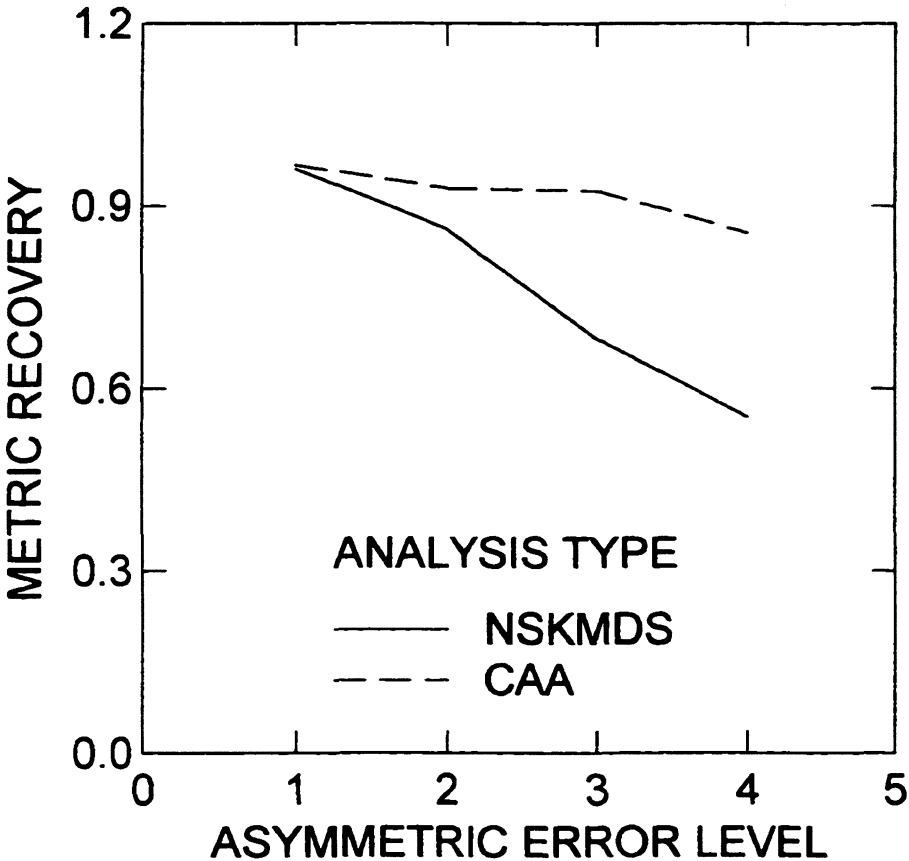
Source	SS	df	MS	F	p	$\omega^2$
ERROR LEVEL	.493	3	.164	12.07	.000	.161
METHOD	.040	1	.041	2.98	ns	-
ERROR LEVEL BY METHOD	.176	3	.059	4.30	.006	.048
WITHIN GROUPS	2.070	152	.014			

Method=Fixed Effect, Error Level=Random Effect

Source	SS	df	MS	F	p
ERROR LEVEL	.493	3	.164	12.071	.000
METHOD	.040	1	.041	.695	ns
ERROR LEVEL BY METHOD	.176	3	.059	4.304	.006
WITHIN GROUPS	2.070	152	.014		

The main effect for error level and the interaction of error level with analysis type, are significant whether error level is treated as fixed or random. The main effect of analysis type is not significant. Analysis of the simple main effects indicates significantly higher nonmetric recovery by the canonical analysis of asymmetry at the 3.0 error level. Differences at the

**FIGURE 15**  
METRIC RECOVERY V.  
ASYMMETRIC ERROR



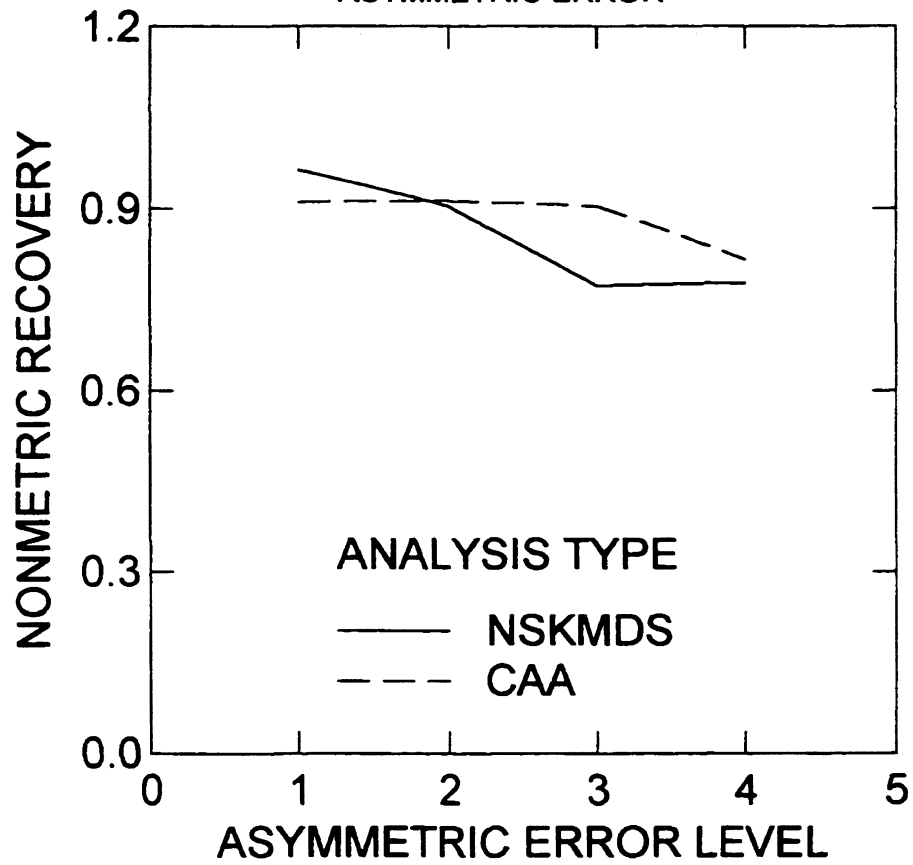
other levels were not significant. It seems likely that underlying this mix of results is a slightly better nonmetric recovery by the canonical analysis of asymmetry at the higher error levels, and that the absence of a significant difference at the 4.0 error level is a Type II error.

Realistic conditions. We have seen, then, that the canonical analysis of asymmetry provides better metric and nonmetric recoveries than NSKMDS, in the presence of high amounts of error in the input data. The two techniques do not seem to differ to a substantive extent in the presence of positively accelerated monotonic distortion, but NSKMDS provides better recovery with negatively accelerated distortions. Given this mix of findings it seemed desirable to compare the two techniques under combinations of error and monotonic distortion that might be expected to occur in real-world data.

Two simulations were run. In the first, normal error was set to 0.67, asymmetric error was set to 0.33, and monotonic distortion was set to 0.3. Thus, the error variation before monotonic distortion equaled the true sum of squares, analogous to a reliability of 0.5 at the level of the individual subject. The monotonic distortion level is the same as in the sone scale for loudness and the bril scale for brightness of a five degree target in sensory psychophysics (Stevens, 1960). One hundred trials were run for each analysis type.

Under these conditions, the relatively large sample size provides stable estimates of the expected recovery levels. For NSKMDS the mean metric recovery was 0.956, with a standard error of the mean of 0.004. The mean nonmetric recovery was 0.959, with a standard error of 0.003. For the canonical analysis of asymmetry the mean metric recovery is 0.963, with a standard error of 0.002, and the mean nonmetric recovery is 0.953, with a

FIGURE 16

NONMETRIC RECOVERY V.  
ASYMMETRIC ERROR

standard error of 0.010. Despite the stability of the estimates, the differences between the techniques are not statistically significant,  $t(198) = 1.76$ , *ns*, for metric recovery, and  $t(198) = -0.59$ , *ns*, for nonmetric recovery.

In the second simulation, normal error and asymmetric error were set to 0.67 and 0.33, respectively, as before. Monotonic distortion was set to 1.5, approximately equal to the power law exponent for the heaviness of lifted weights (Stevens, 1957). For NSKMDS the mean metric recovery was 0.965, with a standard error of 0.002, and the mean nonmetric recovery was 0.959, with a standard error of 0.009. For the canonical analysis of asymmetry the mean metric recovery was 0.975, with a standard error of 0.001, and the mean nonmetric recovery was 0.969, with a standard of 0.003. The difference in nonmetric recoveries is not significant,  $t(198) = 1.03$ , *ns*, while the difference in metric recoveries is significant,  $t(198) = 4.36$ ,  $p < .0005$ .

Thus, there are technical grounds for asserting that the canonical analysis of asymmetry provides better metric recovery under realistic error conditions, when the monotonic transform is positively accelerated. At this point, however, we are splitting hairs. Metric and nonmetric recovery are above 0.95 for each technique under each set of conditions. For all practical purposes they are providing solutions that are identical to the true configuration, and therefore to each other.

## Chapter 12

### Eigenvalue Ratio Test

For principal components analysis, several techniques are available for determining which components should be retained in the solution. Common practices include visually inspecting an eigenvalue plot for a scree-like appearance of the later factors (Cattell, 1966), retaining factors whose eigenvalue exceeds 1.0 (Kaiser, 1960, however see Cliff, 1988), testing the later eigenvalues for equality in the population (Bartlett, 1950), and comparing the eigenvalues in a principal components analysis with the eigenvalues that would be obtained from analysis of random normal variables of the same number and sample size ("parallel analysis"; Horn, 1965; Lautenschlager, 1989)<sup>3</sup>. It is unclear, however, whether these techniques can be transferred to the canonical analysis of asymmetry. In this section I describe a technique for testing the components of a skew-symmetric matrix, and present simulation studies of the technique. To my knowledge this technique is new to this dissertation, although as noted below, it builds on prior work by others.

Craddock and Flood (1969) noted that the last, presumably error, factors in a principal components analysis appear to show an exponential decline in eigenvalue with respect to factor number. Hence, they plotted the logarithms of the eigenvalues against the factor number and inspected the graph for a terminal straight line segment. A straight line having negative slope would indicate an exponential decay presumably more characteristic of

---

<sup>3</sup> See Borkum (1993) for a more extensive review and an alternative proposal.



error than of true factors.<sup>4</sup> Lautenschlager's (1989) extensive simulations of principal components with random data, which he presents in tabular form, do seem to demonstrate an exponential decay. Similar results have been reported by Farmer (1971) and by Mandel (1972). Thus, the ratios between successive error eigenvalues should be equal,

$$(\lambda_i/\lambda_j) = (\lambda_j/\lambda_k) ,$$

or

$$(\lambda_i/\lambda_j) / (\lambda_j/\lambda_k) = 1 \tag{14}$$

where  $j = i+1$  and  $k=j+1$ . If an exponential decay of eigenvalues is a general characteristic of random matrices, then it may be possible to use the left side of Equation 14 as a test statistic for determining the number of eigenvectors to retain. For skew-symmetric matrices, however, a modification is needed. Because the eigenvalues occur in pairs, the members of which are equal in magnitude, it would always be the case that  $\lambda_i = -\lambda_j$  when  $j = i + 1$ , and  $j$  is an even number. We can remedy this, however, by requiring that  $\lambda_i$ ,  $\lambda_j$ , and  $\lambda_k$  be from successive pairs – i.e., that  $j=i+2$ , and  $k=j+2$ .

Simulations were conducted to investigate whether this eigenvalue ratio approach shows promise for the canonical analysis of asymmetry.

First, 1000 trials were conducted in which the canonical analysis of asymmetry was used to analyze the skew-symmetries in random matrices. Normal error level was set to 1.5,

---

<sup>4</sup> Note, however, that the eigenvalues in Johnson et al.'s (1984) factor analysis of the MMPI also show an exponential decay, once the first factor is excluded. Jackson (1959) has argued that the MMPI is a one-factor instrument. If the succeeding factors are valid, however, then a log eigenvalue plot of the MMPI factors should presumably divide into two straight line segments, the first, "true" segment having a steeper slope than the second, "error" segment.

asymmetric error and true asymmetric variation were both set to 0, and monotonic distortion level was set to 1.0 (no distortion). The mean eigenvalues across the 1000 trials are shown in Table 25. The ratios between the mean

Table 25  
Eigenvalues of Skew-Symmetries for Random Normal Matrices

EIGENVALS	MEAN	RATIO TO PRIOR E. V.	STD. DEV.	SKEWNESS	KURTOSIS
1 and 2	1.059	—	0.109	-.163	0.361
3 and 4	0.746	0.7044	0.100	0.283	0.009
5 and 6	0.535	0.7172	0.077	0.091	-.351
7 and 8	0.356	0.6654	0.061	-.035	-.169
9 and 10	0.163	0.4579	0.070	0.118	-.501

Note. Based on a 1000 trial simulation.

eigenvalues of successive pairs are .7044, .7172, .6654, and .4579. With the exception of the last, the ratios appeared close enough to each other to justify further exploration.

In the next simulation the ratio of the first and third eigenvalues was divided by the ratio of the third and fifth eigenvalues,

$$(\lambda_1 / \lambda_3) / (\lambda_3 / \lambda_5) = \text{EVR1 (first eigenvalue ratio)}$$

and the distribution of this value across trials was tabulated. Normal error level was set to 0.67, asymmetric error level was set to 0.33, and monotonic distortion set to 1.5. These

conditions appeared reasonably representative of actual data. In one condition, two "true" vectors were used in constructing the asymmetries. In the second condition there was no true variance underlying the skew-symmetries. Asymmetric error was also set to zero, as there were no asymmetry vectors to perturb, and monotonicity was set to 1.0. To compensate for the loss of asymmetric error, normal error was increased proportionately, to 1.40. One hundred trials were run under each of the two sets of conditions.

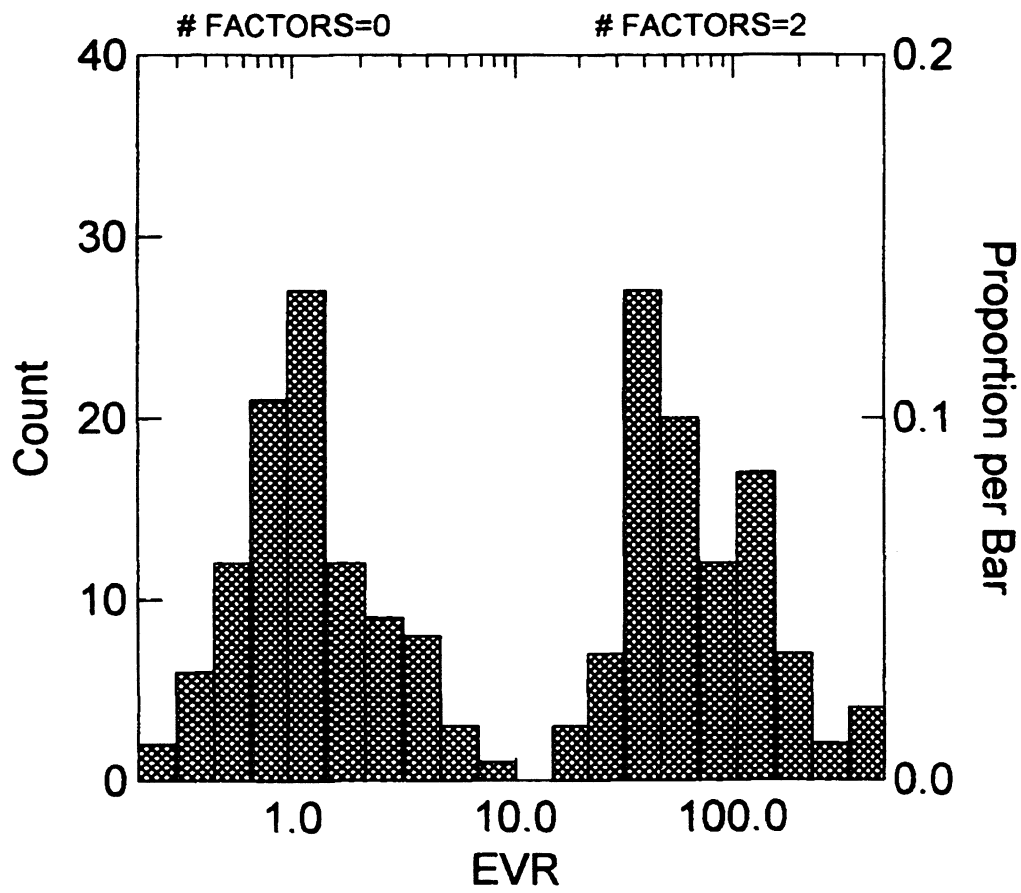
Under the true asymmetry condition, the eigenvalue ratio varied between 7.336 and 418.697, with a mean of 89.505. Under the pure error condition, the eigenvalue ratio varied between 0.255 and 6.380, with a mean of 1.479. As may be seen in Figure 17, there is no overlap between the distributions in the true and error conditions. Any cutoff value between 6.380 and 7.336 would, in this small sample, reduce the Type I and Type II error rates to zero. The mean under the pure error condition, 1.479, is significantly greater than 1.0,  $t(99) = 4.129$ ,  $p < .001$ .

In a second, exploratory test, a smaller effect size was mimicked by increasing the error level. In the "true asymmetries" condition, normal error was raised to 1.33 and asymmetric error was increased to 0.67. The level of monotonic distortion was kept at 1.5. In the "random" condition, the normal error level was increased proportionately to 3.0. One hundred trials were run in each condition.

When true vectors underlie the asymmetric part of the matrix, the eigenvalue ratio varies between 7.913 and 258.270, with a mean of 58.510. When the asymmetries consist only of error, the eigenvalue ratio varies between 0.147 and 4.341 with a mean of 1.288.

# FIGURE 17

## DISTRIBUTION OF EVR1



The mean of the error condition is significantly greater than one,  $t(99) = 3.064$ ,  $p < .01$ . With the increased error the means of the two distributions appear to have drawn closer. However, the standard deviations have also declined. As seen in Figure 18, there is no overlap, and any cutoff value between 4.341 and 7.913 would reduce the Type I and Type II error rates to zero in this limited number of trials.

As noted, the mean in both error conditions exceeded 1.0. This seems most likely to be an effect of using the arithmetic mean when the quantity varying is a ratio. That is, assume that  $\lambda_1$  and  $\lambda_3$  each have a mean of 1.0, but that fluctuations to 1.5 and 0.5 are equally likely. Then the ratios  $(1.5/.5) = 3.0$  and  $(.5/1.5) = .33$  are equally likely. However,  $(3.00+0.33)/2 = 1.67$ . The upward shift would not occur had we taken the geometric mean, as  $\text{SQRT}(3.0 \cdot .33) = \text{SQRT}(1) = 1$ . Moreover, the eigenvalues in the canonical analysis of asymmetry are the squares of those of the skew-symmetric matrix. The squaring would presumably increase the upward shift of the arithmetic mean. Therefore, in the remaining simulations, square roots of the eigenvalues were used.

To help gauge the usefulness of the test more precisely, and to extend it beyond the first pair of eigenvalues, 1000 trials were run on random matrices. Normal error was set to 1.50, asymmetric error and true variation were set to zero, and monotonic distortion was switched off. The following ratios were tabulated:

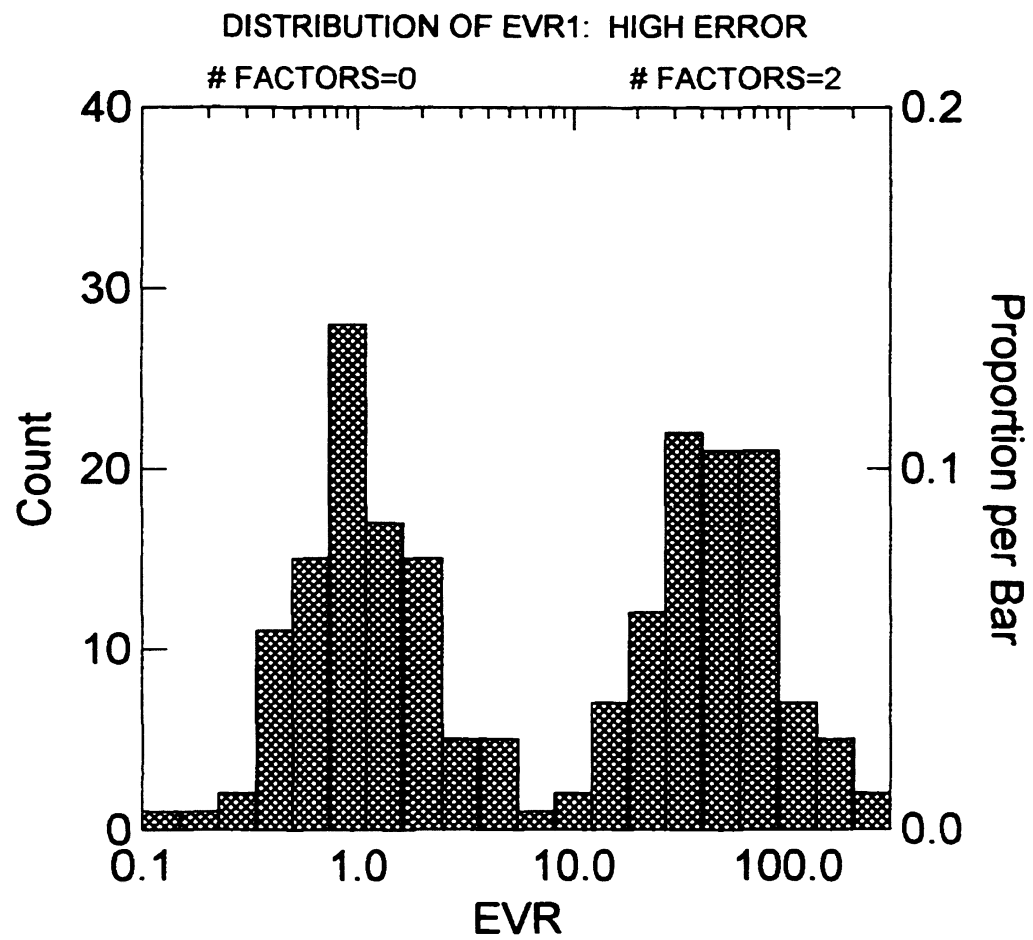
$$(\sqrt{\lambda_1} / \sqrt{\lambda_3}) / (\sqrt{\lambda_3} / \sqrt{\lambda_5}) \quad \text{first eigenvalue ratio (EVR1)}$$

$$(\sqrt{\lambda_3} / \sqrt{\lambda_5}) / (\sqrt{\lambda_5} / \sqrt{\lambda_7}) \quad \text{second eigenvalue ratio (EVR2)}$$

and

$$(\sqrt{\lambda_5} / \sqrt{\lambda_7}) / (\sqrt{\lambda_7} / \sqrt{\lambda_9}) \quad \text{third eigenvalue ratio (EVR3)}.$$

FIGURE 18



Aspects of the distributions of the three ratios are shown in Table 26. The means of the three distributions do not equal 1.0, nor do they equal one another, although the first two means,

Table 26

## Normative Values for the Eigenvalue Ratio Test

	MEAN	STD. DEV.	SKEWNESS	95th %ile	99th %ile	99.5 %ile
EVR1	1.160	0.564	1.623	2.211	2.957	3.475
EVR2	0.967	0.510	1.445	1.908	2.840	3.086
EVR3	0.629	0.598	2.490	1.722	3.148	3.661

Note. Based on a 1000 trial simulation with random normal matrices.

1.160 and 0.967, are close to 1.0. Despite the bell-shaped appearance of Figures 17 and 18, none of the three distributions here, after standardizing to zero mean and unit standard deviation, conform to a normal distribution, as assessed by the Komolgorov-Smirnov test. Therefore the upper percentiles, corresponding to  $\alpha$  levels of .05, .01, and .005, were obtained empirically, as described below.

Four additional simulations were then conducted. The simulations were identical to the 1000-trial condition just described, except that small amounts of true asymmetric variation were added. In the first simulation, true asymmetric variation was set to 0.03. On average, only 45% of the normal error variance contributes to the skew-symmetric part of a ten by ten matrix, the rest going towards the symmetric part and the diagonal. Therefore the true asymmetry level of .03 corresponds to a correlation ratio for the skew-symmetries of

$$.03/ (.03 + 1.5 * .45) = .05.$$

In the second simulation, the true asymmetry level was .12, corresponding to a correlation ratio for the skew-symmetries of .15. In the third simulation, the true asymmetry level was .55, giving a correlation ratio for the asymmetries of .45. In the fourth simulation the true asymmetry level was 1.01, corresponding to a correlation ratio for the asymmetries of .60. If we had defined the correlation ratio as the true variation in the matrix, divided by the true plus symmetric, asymmetric, and diagonal error variation, the correlation ratios for the four simulations would have been computed as .023, .074, .269, and .403, approximately one half the sizes noted above.

Each of the four simulations consisted of 100 trials. EVR1, EVR2, and EVR3 were computed on each trial and the distributions tabulated. In particular, the 1000 trial null simulation noted above was used to obtain the cutoff value for  $\alpha = .05$ , for each of the three ratios. Then, in each of the four non-null simulations described here, the number of trials were counted in which each of the ratios were above the cutoff value. For EVR1, corresponding to the two true asymmetric vectors, this permitted some assessment of the power of the test in the presence of different effect sizes in the population. For EVR2 and EVR3, the distributions provide a check on the presumed  $\alpha$  level.

Results are shown in Table 27. For EVR2 and EVR3, the  $\alpha$  levels appear to be approximately .05, with some variation in either direction. For EVR1, it appears that the test's power depends strongly on the effect size. At a .05 correlation ratio for the asymmetries (.023 overall correlation ratio),  $\beta$  was .90, corresponding to a power level of .10. For a correlation ratio of .15 for the skew-symmetries,  $\beta$  was .67, and power was .33.



For a correlation ratio of .45 for the skew-symmetries,  $\beta$  was .29, and power was .71, and for a skew-symmetry correlation ratio of .60,  $\beta$  was .01, and power was .99.

Thus, the EVR test appears generally able to distinguish when one pair of dimensions underlies a skew-symmetric matrix. Five additional simulations were then conducted to assess its capacity to detect that two pairs of dimensions should be extracted. These simulations provide some indication of the power or sensitivity of EVR2, and a check on the  $\alpha$  level for EVR3. Additionally, in these simulations we would want EVR1 to reject a 2-dimensional solution, even though the two dimensions that it assesses actually are present in the true, 4-dimensional solution. Thus, the simulations provide a check on the  $\alpha$  level for EVR1 under conditions different from those in the previous simulation.

One hundred trials were conducted for each of five correlation ratios for the second pair of skew-symmetric dimensions. Correlation ratios used in the study were .05, .15, .30, .45, and .60. The normal error level was set to 1.50, and monotonic distortion and asymmetric error were switched off. A fast drop-off was selected for the eigenvalues of the true asymmetric dimensions. In the fast drop-off, each successive pair of eigenvalues is lower by a factor of 1.80 than the preceding pair. Because the true eigenvalues thus show the kind of exponential decay thought to characterize error, the simulations presumably provide a fairly stringent test.

Table 27

## Empirically Obtained Type I and Type II Error Rates

## For the Eigenvalue Ratio Test

TRUE NUMBER OF DIMENSIONS	CORRELATION RATIO <sup>1</sup>	EVR1	EVR2	EVR3
2	.05	$\beta = .90$	$\alpha = .11$	$\alpha = .03$
2	.15	$\beta = .67$	$\alpha = .05$	$\alpha = .08$
2	.45	$\beta = .29$	$\alpha = .08$	$\alpha = .06$
2	.60	$\beta = .01$	$\alpha = .02$	$\alpha < .01$
4	.05	$\alpha = .03$	$\beta = .53$	$\alpha = .04$
4	.15	$\alpha = .02$	$\beta = .16$	$\alpha = .03$
4	.30	$\alpha = .02$	$\beta = .04$	$\alpha = .03$
4	.45	$\alpha < .01$	$\beta = .04$	$\alpha = .06$
4	.60	$\alpha = .03$	$\beta = .02$	$\alpha = .03$
6	.05	$\alpha < .01$	$\alpha = .02$	$\beta = .67$
6	.15	$\alpha < .01$	$\alpha = .04$	$\beta = .56$
6	.30	$\alpha = .01$	$\alpha = .07$	$\beta = .41$
6	.45	$\alpha = .02$	$\alpha = .06$	$\beta = .30$

<sup>1</sup> Note. Correlation ratio is the ratio of true to true plus error variation. "True" variation here refers to the true asymmetric variation of the last pair of dimensions, for example the third pair of dimensions in a 6-dimensional structure.

For each correlation ratio, the number of instances in which the test statistic, EVR1, EVR2, or EVR3, was above the  $\alpha=.05$  level in the 1000-trial normative simulation, was counted. The results are shown in Table 27. The obtained  $\alpha$  levels for EVR1 and EVR3 are close to the anticipated value of .05. For EVR1, the actual  $\alpha$  varied between .03 and  $<.01$ , suggesting that EVR1 may be slightly conservative under these conditions. The Type II error rate for EVR2 varied with the true correlation ratio. Power (1- $\beta$ ) was .47 at a correlation ratio of .05, .84 for a correlation ratio of .15, .96 for a correlation ratio of .30, .96 for a correlation ratio of .45, and .98 for a correlation ratio of .60.

The eigenvalue ratio test was then studied for its ability to detect that 3 pairs of dimensions should be extracted. This provides a check on the  $\alpha$  levels for EVR1 and EVR2, and on the  $\beta$  level for EVR3. One hundred trials were run for each of four correlation ratios for the third pair of dimensions underlying the skew-symmetries: .05, .15, .30, and .45. Normal error level was set to 1.50, monotonic distortion and asymmetric error were switched off, and a fast drop-off was selected for the true eigenvalue pairs.

As in the preceding simulations, the number of times that the test statistic, EVR1, EVR2, or EVR3, was above the  $\alpha=.05$  cutoff value in the 1000 trial norming study, was counted. This is shown in Table 27. The obtained  $\alpha$  level for EVR2 varies between .02 and .07, and appears generally to confirm the expected .05  $\alpha$  level. The obtained  $\alpha$  level for EVR1 varies between .02 and  $<.01$ , and suggests that EVR1 may be somewhat conservative under these conditions. For EVR3, the Type II error rate declines with increasing true effect size. For a correlation ratio of .05, power (1- $\beta$ ) was .33, for a correlation ratio of .15 power

was .44, for a correlation ratio of .30 power was .59, and for a correlation ratio of .45 power was .70.

Overall, then, the test appears to be reasonably effective at indicating the correct number of dimension pairs underlying the skew-symmetries. That is, it appears able to detect changes in the slope of the log-eigenvalue plot corresponding to a change between true and error variance. EVR1 and EVR3 seem to have considerably less power than researchers would ordinarily be used to. EVR1, for example, shows approximately the same relationship to population effect size as a 4-groups, one-way ANOVA with 3 subjects per group. This "under-powering" is not necessarily a disadvantage, however. One of the main criticisms of the use of metric MDS with nonmetric data is an apparent tendency to extract too many factors (Helm, 1960; Mellinger, 1956). A conservative stopping rule in factor extraction should help curtail a tendency to over-extraction. Moreover, in the current simulations the EVR stopping rule seems to accord reasonably well with the expected level of metric recovery. Thus, when one dimension pair underlay the asymmetries, power at the two lowest correlation ratios was .10 and .33, while the mean metric recovery at the two lowest correlation ratios was .57 and .83. For the two highest correlation ratios, power was .71 and .99, and mean metric recovery was .961 and .984. Hence, although the test appears quite conservative, it may help screen out poorly defined components. An additional advantage of the test, if further studies support its use, is its ease of computation, and the fact that it recasts the visual inspection of a scree plot into somewhat more definite terms.

## Chapter 13

## Analyses With Special Matrix Structures

One-Dimensional Asymmetries

The canonical analysis of asymmetry and NSKMDS require two dimensions to represent skew-symmetry. However, we may have a substantive theory in which the asymmetries are one-dimensional. A one-dimensional solution can occur in CAA and NSKMDS if the objects' coordinates fall along a straight line in the two-dimensional space.

This will happen, for example, if all of the loadings on one dimension are constant:

$$x_{ij} = r_i s_j - r_j s_i$$

and if  $r_i = r_j = 1$  for all  $i, j$ , then

$$x_{ij} = s_j - s_i .$$

That is, the asymmetries are explained as the difference between scores on a single set of additive bias coefficients,  $s_i$ .

Table 28 shows an error-free, asymmetric matrix generated in this manner. Because the bias coefficients are not in any particular order the one-dimensional structure is not easily accessible to inspection. However, it is readily apparent in Figure 19, which is the CAA solution for the data. The NSKMDS solution is identical, except for a reflection of one of the dimensions.

Simplex

Guttman (1954) and Shepard (1978) note that a one-dimensional ordering among objects is often the primary feature of a proximity matrix. For example, test questions may

Table 28

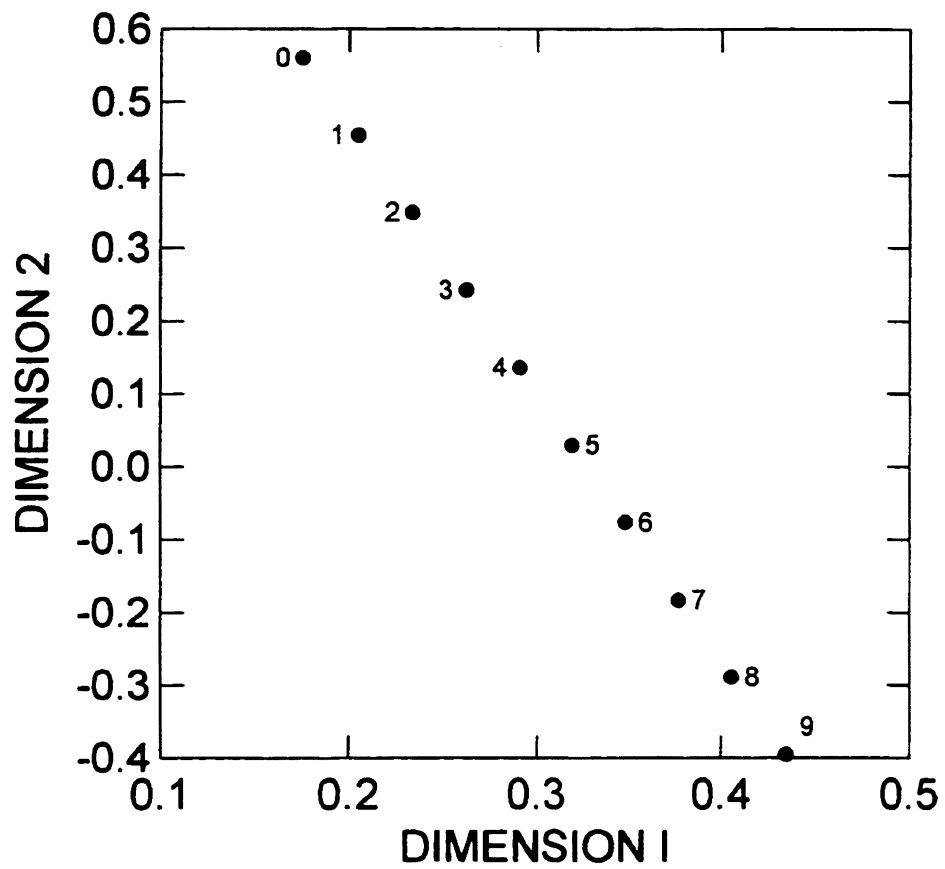
## One-Dimensional Skew-Symmetries

0	3	-1	4	1	-3	2	-4	-2	5
-3	0	-4	1	-2	-6	-1	-7	-5	2
1	4	0	5	2	-2	3	-3	-1	6
-4	-1	-5	0	-3	-7	-2	-8	-6	1
-1	2	-2	3	0	-4	1	-5	-3	4
3	6	2	7	4	0	5	-1	1	8
-2	1	-3	2	-1	-5	0	-6	-4	3
4	7	3	8	5	1	6	0	2	9
2	5	1	6	3	-1	4	-2	0	7
-5	-2	-6	-1	-4	-8	-3	-9	-7	0

differ in complexity, or depression inventory items in the severity of symptoms they reflect. In that case, the likelihood of a subject passing a question or endorsing an item will be closely related to the subject's likelihood of passing questions of similar difficulty level, or endorsing items of similar intensity. As the questions or items become more discrepant in their complexity or intensity, correlations between them would decline. A correlation matrix among the questions, arranged in order of item difficulty, would show a characteristic pattern: items next to each other in the order would have the highest correlations; as items became more removed the correlations between them would drop. Therefore, the highest entries would be near the main diagonal; the size of entries would decline with increasing distance from the diagonal.

## FIGURE 19

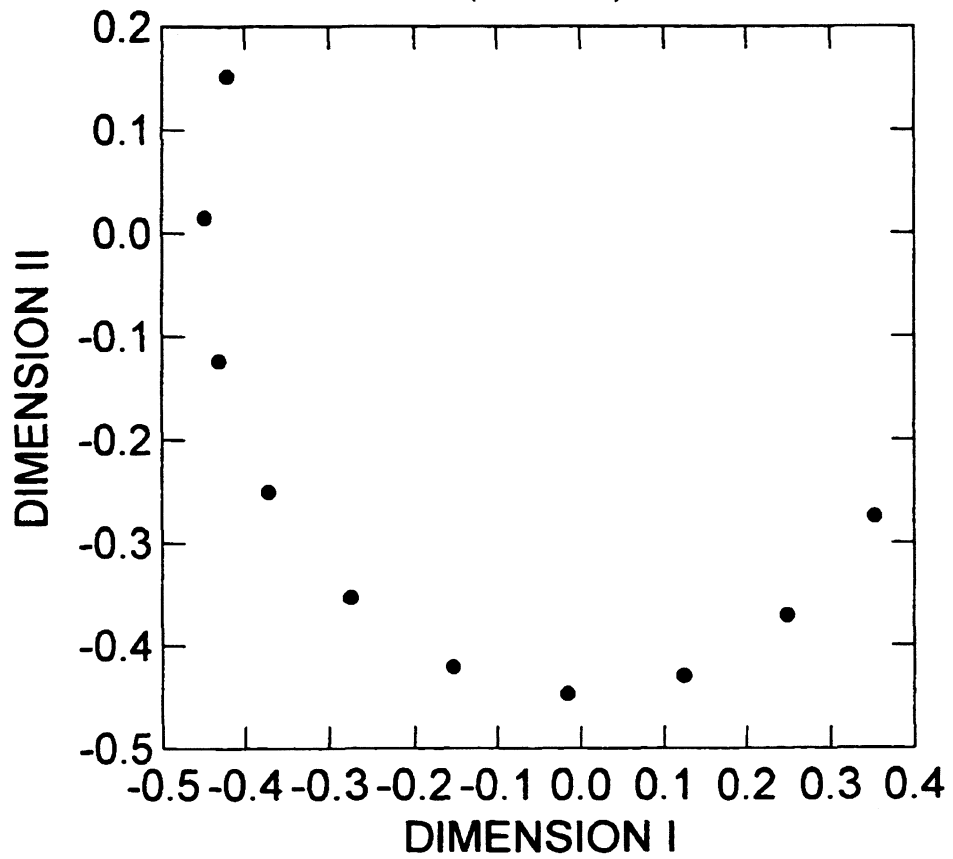
CAA SOLUTION FOR ONE-DIMENSIONAL DATA

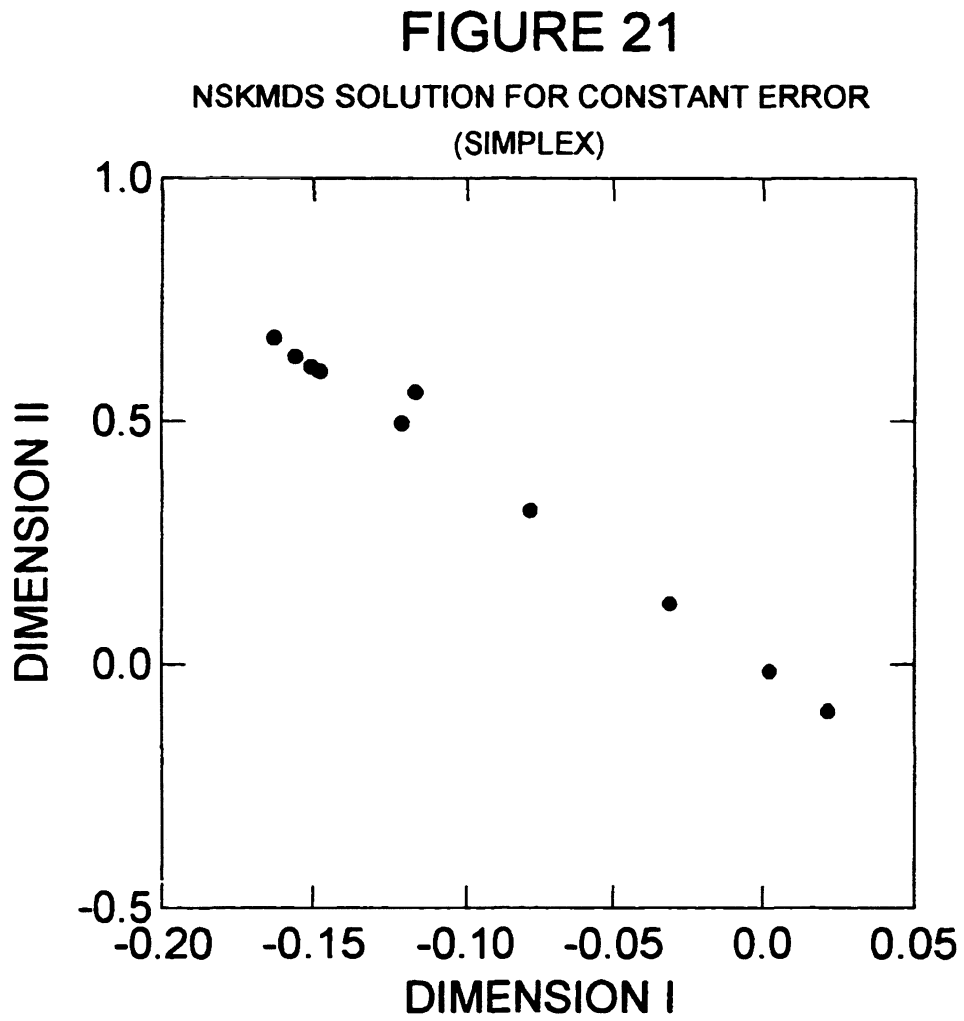


Guttman (1954) calls matrices of this type simplexes. Because of their frequent occurrence and parsimony he notes the importance of an ability to detect simplexes. As  $n$  items can be ordered in  $n!$  ways (for example, 3,628,800 orderings for 10 objects) it is often not possible to detect a simplex by visually inspecting or rearranging a matrix. Hence, Guttman (1966) uses the detection of simplexes as a key criterion for judging the adequacy of scaling techniques. Ideally, the technique should recover the one-dimensional ordering underlying the objects.

Simplexes can occur among skew-symmetries. For example, if the asymmetry in a matrix is due to constant error, such as a constant effect of spatial position or time order, the sum of cross products matrix derived from the skew-symmetries will have a simplex structure. Hence the canonical analysis of asymmetry will involve the representation of a simplex. Figure 20 shows the results of a CAA of a matrix in which all asymmetries are due to constant error. In the skew-symmetric part of this idealized matrix, all entries above the main diagonal are "2", and all entries below the main diagonal are "-2". (The results would be the same if we chose a different constant instead of "2"). As may be noted, the object points fall along a semicircular horseshoe curve. The results of an NSKMDS analysis of the same matrix are shown in Figure 21. The nonmetric analysis properly displays the one dimensional ordering of the object points along a straight line. In this example, then, NSKMDS provided a clearer representation of the simplex caused by constant error.



**FIGURE 20****CAA SOLUTION FOR CONSTANT ERROR  
(SIMPLEX)**



### Multiplicative Bias Coefficients

One of the simplest ways of inducing asymmetry in a matrix is to multiply each column or row by a different coefficient. Although neither the canonical analysis of skew-symmetry nor its nonmetric generalization is designed to indicate the presence of multiplicative bias coefficients (which can be done simply by attempting to symmetrize a matrix by dividing entries by the row or column sums) it would be preferable if the techniques were able to recover bias coefficients when they are present. To explore how CAA and NSKMDS would handle a bias-perturbed matrix, a brief trial was conducted using Kruskal and Wish's (1977) symmetric matrix of ratings of the perceived similarities among twelve countries. In one condition, each column in Kruskal and Wish's matrix was multiplied by a different, arbitrary number. In a second condition, each row was multiplied by a different number and in a third condition, each row and each column were multiplied by arbitrary coefficients.

For CAA the eigenvalue ratio test suggests that one pair of factors be extracted for the column bias condition. After varimax rotation the first factor has a Pearson product moment correlation of .885 with the column bias coefficients. This is actually a bit lower than the correlations with the unrotated factors (.948 with factor 1, .908 with factor 2) but presumably this would not be known to a researcher in actual practice. For NSKMDS a test statistic is not yet available, and the algorithm is currently programmed for only a two-dimensional solution. However, inspection of the sums of squares of the coordinates on the two dimensions (3.021 and .226) suggests that only the first should be retained as contributing significantly to the solution. It correlates -.988 with the vector of bias coefficients.

For the row bias condition, in CAA, the values of EVR1, EVR2, and EVR3 suggest that only the first pair of factors should be retained. After varimax rotation, the first factor correlates .986 with the row bias coefficients. The correlations of the bias coefficients with the two unrotated factors are .942 and -.930, respectively. In NSKMDS the sum of squared coordinates on the two dimensions (1.46 and .56) would probably suggest a one-dimensional solution. Although rotation is not available, this does not seem to be a hindrance, as the first dimension correlates .993 with the row bias coefficients.

When both row and column bias coefficients are used, the analysis techniques should recover the ratio of the two coefficients for each object (country in this case) in the matrix. For CAA the eigenvalue ratio test suggests that only the first pair of eigenvectors be retained. Their correlations with the ratio of the row and bias coefficients are .570 and -.910. However, in practice a researcher would not know to select the second eigenvector. The first vector after varimax rotation correlates .861 with the bias coefficient ratio. For NSKMDS, the sum of squared coordinates on the two dimensions (.314, 1.578) suggests that only the second dimension be interpreted. In fact this dimension correlates .944 with the bias coefficient ratio.

Hence, in this brief demonstration there was some advantage to the nonmetric technique, in that it appears to recover multiplicative bias coefficients a bit more accurately than CAA, and without requiring rotation.

## Chapter 14

### Published Data Sets

To check the usefulness of the canonical analysis of asymmetry and nonmetric, skew-symmetric MDS with real-world data, several previously published data sets were analyzed. The NSKMDS solutions were constrained to one and two dimensions because the technique has not yet been programmed for greater than two dimensions. As will be seen, however, this does not impede its usefulness for the data sets to follow. Moreover, for each of these data sets the eigenvalue ratio test suggests a two dimensional CAA solution, which should be an upper bound for the dimensionality of the corresponding nonmetric solution.

#### Correlations Among Cognitive Tests

Ham and Parsons (1997) administered six cognitive tests to 131 individuals who met the diagnostic criteria for alcohol dependence, at the end of a 2-3 week detoxification program. The test battery was also administered to 83 control subjects matched for handedness, education, and age. None of the subjects in either condition was judged to have a medical, neurological, or psychiatric disorder other than alcohol dependence that could affect neuropsychological functioning, and none were taking psychoactive medications. The alcohol-dependent subjects had lower scores on the cognitive tests. Ham and Parsons' primary interest, however, was on comparing the covariance structures for the alcohol-dependent and control subjects. Using chi-square testing in LISREL (Joreskog & Sorbom, 1989, cited in Ham & Parsons, 1997) they determined that the two covariance matrices differed to a statistically significant degree. LISREL extracted 3 factors from the 6 tests. These factors

were labelled Verbal (Shipley Vocabulary test and Wechsler Adult Intelligence Scale-Revised, Information subtest), Problem Solving (Shipley Abstraction test and Booklet Categories Test), and Visual-Spatial (Trails B and Wechsler Adult Intelligence Scale-Revised, Block Design subtest). It appeared from the LISREL results that the Verbal and Visual-Spatial factors were more highly correlated in the alcohol-dependent subjects, compared with the control subjects. In addition, the LISREL results suggested that the Verbal and Problem-Solving factors were less highly correlated in the alcohol-dependent subjects than in the control subjects. Ham and Parsons interpreted the increase in Verbal-Visual Spatial correlation as indicating that the alcohol-dependent subjects were using verbal skills to compensate for losses in spatial ability. They explained the decrease in the Verbal-Problem Solving correlation as a sign that the alcohol-dependent subjects were less able to access verbal skills for the purpose of problem-solving.

Ham and Parsons' data can also be examined via CAA and NSKMDS. First, we would construct a matrix in which the upper triangular half contains the test intercorrelations for the alcohol-dependent subjects, and the lower triangular half gives the intercorrelations for the control subjects (Table 29). We would then extract and analyze the asymmetries in this hybrid matrix. The asymmetries indicate the change in correlations as one shifts from the control to the alcohol-dependent subjects. The asymmetries can then be displayed visually using CAA or NSKMDS. In contrast to LISREL, CAA and NSKMDS provide a visual display of the changes in the correlations themselves, rather than the changes in factors derived from the correlations.

Figure 22 shows the two-dimensional CAA solution for the hybrid correlation matrix. The two visual-spatial tests, Block Design and Trails B, are at the far left of the figure. The verbal tests, Information and Shipley Vocabulary, are at the far right. At the center, close to the origin, are the problem solving tests, Booklet Categories and Shipley Abstraction. By picturing the triangle formed by two object points and the origin, we can identify the magnitude of change in a correlation as one shifts from the alcohol-dependent subjects to the controls. Comparison with the hybrid matrix shows that separations of less than 180 degrees in the figure, in a clockwise direction, correspond to larger correlations for the alcohol-dependent subjects compared with the controls.

The largest triangular areas are between visual-spatial and verbal tests: Trails and Information, Trials and Shipley Vocabulary, Block Design and Information, and Block Design and Shipley Vocabulary. These are the correlations showing the greatest increases in the alcohol-dependent subjects. This aspect of the solution matches Ham and Parsons' LISREL results and is readily verified by inspection of the test intercorrelation matrix. Trails B also appears to dominate Categories and Shipley Abstraction. This reflects the larger Trails B-Categories and Trails B-Shipley Abstraction correlations in the alcohol-dependent subjects. There appears to be some tendency for the correlation between Information and Shipley Vocabulary to be higher in the alcohol-dependent subjects.

Table 29  
Intercorrelations Among Cognitive Tests  
For Alcohol-Dependent and Control Subjects

	<b>Shipley Vocabulary</b>	<b>WAIS-R Information</b>	<b>Shipley Abstraction</b>	<b>Booklet Categories</b>	<b>Trails B</b>	<b>WAIS-R Block Design</b>
<b>Shipley Vocabulary</b>	--	.692	.543	.176	.284	.323
<b>WAIS-R Information</b>	.491	--	.440	.250	.247	.391
<b>Shipley Abstraction</b>	.498	.419	--	.449	.476	.487
<b>Booklet Categories</b>	.156	.188	.364	--	.294	.292
<b>Trails B</b>	.019	.011	.287	.089	--	.393
<b>WAIS-R Block Design</b>	.239	.144	.536	.332	.277	--

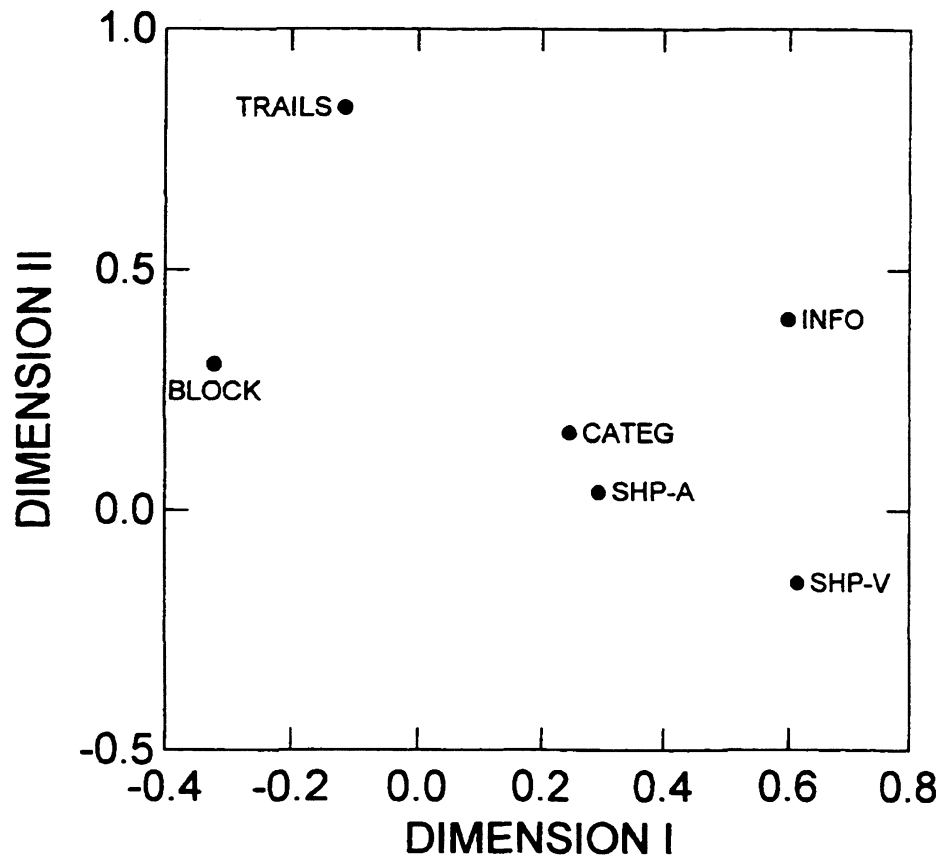
Note. Correlations for alcohol-dependent subjects are above the main diagonal, correlations for control subjects are below the main diagonal.

Source. Adapted from Ham & Parsons, 1997



**FIGURE 22**

CAA SOLUTION FOR COGNITIVE TEST DATA

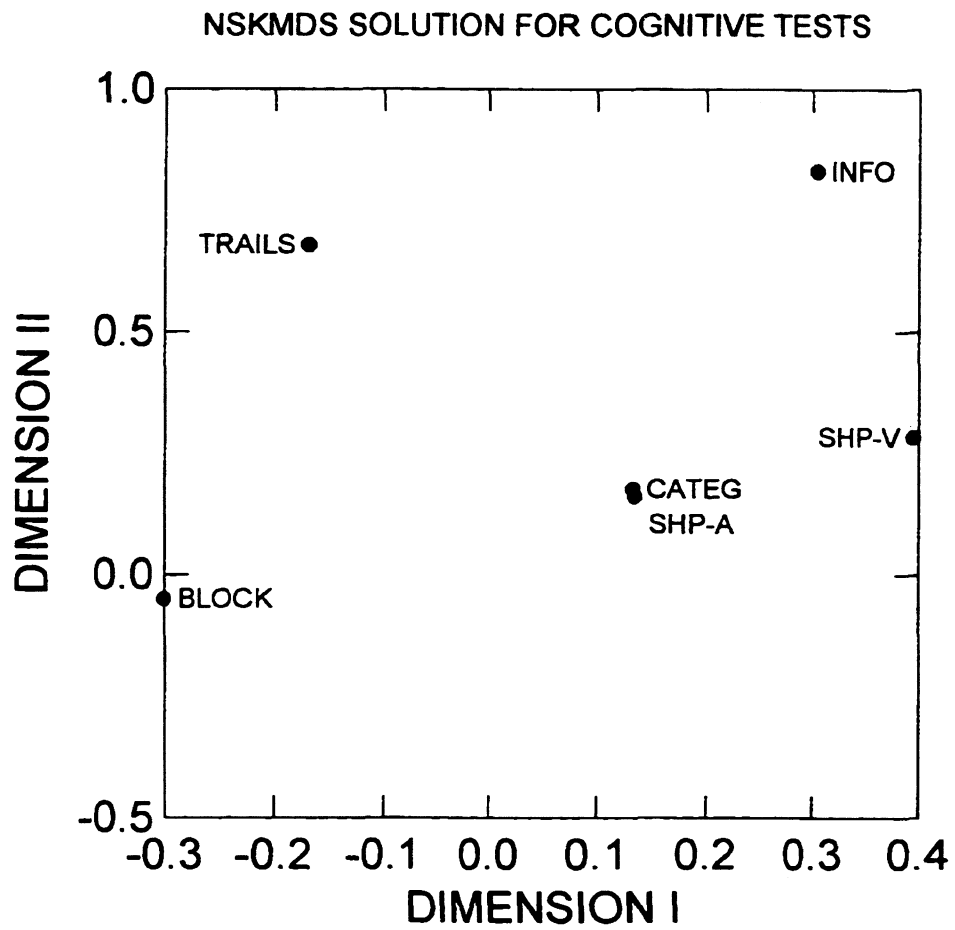


NOTE: DATA FROM HAM AND PARSONS (1997)

Note, too, that in Figure 22, the problem solving tests, Categories and Shipley Abstraction, form only small triangular areas with the verbal tests, Information and Shipley Vocabulary. From the figure there is little indication that these correlations change as one transitions from the alcohol-dependent to the control subjects. Indeed, inspection of the correlation matrix indicates that the largest change in correlations among these four tests is a .062 increase in shifting to the alcohol-dependent subjects. In this respect Ham and Parsons' LISREL solution and the CAA representation differ. In the LISREL solution the problem solving tests showed a lower correlation with the verbal tests in the alcohol-dependent sample. However, inspection of the correlation matrix shows that all four of the intercorrelations between verbal and problem solving tests are higher in the alcohol-dependent subjects.

Figure 23 shows the NSKMDS display of the same hybrid correlation matrix. The results are nearly identical to those of CAA. Two differences, however, may be noted. First, the points separated by the smallest distance in the CAA solution, Categories and Shipley Abstraction, have essentially merged into a single point. That is, discrimination of the two closest points was lost. Second, in the NSKMDS solution four of the object points fall along a straight line, raising the possibility that the changes in correlations among these tests in going from the control to the alcohol-dependent subjects corresponds to constant error. The four tests showing this property are all of the tests in the battery except for Trails B and Information. In fact, examination of the correlation matrix shows that of the 7 correlations showing a change in magnitude of .10 or more, all 7 involve Trails B, Information, or both. The remaining correlations show approximately equal, and very small, changes as one shifts from the control to the alcohol-dependent subjects.

FIGURE 23



NOTE: DATA FROM HAM AND PARSONS (1997)

### International Trade

Data on the balance of trade between nations in 1993 was extracted from the Direction of Trade Statistics Yearbook (International Monetary Fund, 1994). As the purpose of the analysis was a demonstration of the techniques rather than a study in economics, only ten nations were drawn from those listed in the yearbook. An attempt was made to include representative nations of major economic groups (European Organization for Economic Cooperation and Development, the Organization of Petroleum Exporting Countries, North American Free Trade Agreement) and each continent. Nations were specifically excluded if their reports of imports and exports did not match those of their trading partners, or if their volume of trade was so low that the country would be likely to drop out of the analysis. The countries selected as meeting these criteria were: Brazil, China, France, Japan, Nigeria, Russia, Saudi Arabia, South Korea, the United Kingdom, and the United States. The trade data are shown in Table 30.

In CAA, eigenvalue ratio test values for the first three dimension pairs were 24.91, 0.56, and 0.24, suggesting that two dimensions (one pair) should be extracted. These two dimensions are shown in Table 31, and in Figure 24. All countries have loadings of approximately zero on both dimensions, except for the United States, Japan, and China. In Figure 24 it may be seen that most of the trade imbalances in the matrix as a whole represent a favorable balance of Japan and China in relation to the United States. Although petroleum exporting countries (Saudi Arabia, Nigeria, Venezuela, and Russia), other countries having large populations (Brazil, Russia), and other industrialized nations (France, the United

Table 30

## International Trade In 1993

	<b>BRAZIL</b>	<b>CHINA</b>	<b>FRANCE</b>	<b>JAPAN</b>	<b>NIGERIA</b>	<b>RUSSIA</b>	<b>SAUDI ARABIA</b>	<b>KOREA</b>	<b>UNITED KINGDOM</b>	<b>USA</b>
<b>BRAZIL</b>	--	78	79	231	24	22	42	54	114	803
<b>CHINA</b>	15	--	368	2065	12	296	58	286	199	3373
<b>FRANCE</b>	69	160	--	515	56	162	139	120	1853	1569
<b>JAPAN</b>	152	1735	549	--	49	151	411	1919	1210	11042
<b>NIGERIA</b>	14	0.1	75	1	--	0	0	5	17	561
<b>RUSSIA</b>	10	453	237	188	0	--	92	89	112	185
<b>SAUDI ARABIA</b>	147	12	271	893	10	0	--	388	192	843
<b>SOUTH KOREA</b>	32	536	123	1174	13	66	134	--	162	1778
<b>UNITED KINGDOM</b>	53	111	1615	497	95	91	274	120	--	2239
<b>UNITED STATES</b>	606	877	1327	4795	89	297	667	1478	2638	--

Note. In tens of millions of U.S. dollars. Rows = exports; columns = imports.  
Data from International Monetary Fund (1994)

Table 31

## International Trade Data: Skew-Symmetric Dimensions

Nation	Canonical Analysis of Asymmetry		Nonmetric Skew-Symmetric MDS	
	Dimen. I (Imports)	Dimen. II (Exports)	Dimen. I (Exports)	Dimen. II (Imports)
Brazil	-.0136	.0302	.0000	.0000
China	-.0470	.3584	.2759	-.0129
France	.0178	.0411	.0000	.0000
Japan	.0124	.9262	.7936	.0455
Nigeria	.0059	.0668	.0272	.0013
Russia	-.0135	-.0164	.0000	.0000
Saudi Arabia	-.0666	.0288	.0033	-.0256
South Korea	.0885	.0434	.0216	.0267
United Kingdom	.1021	-.0597	-.0304	.0243
United States	.9870	.0087	-.0958	.8340

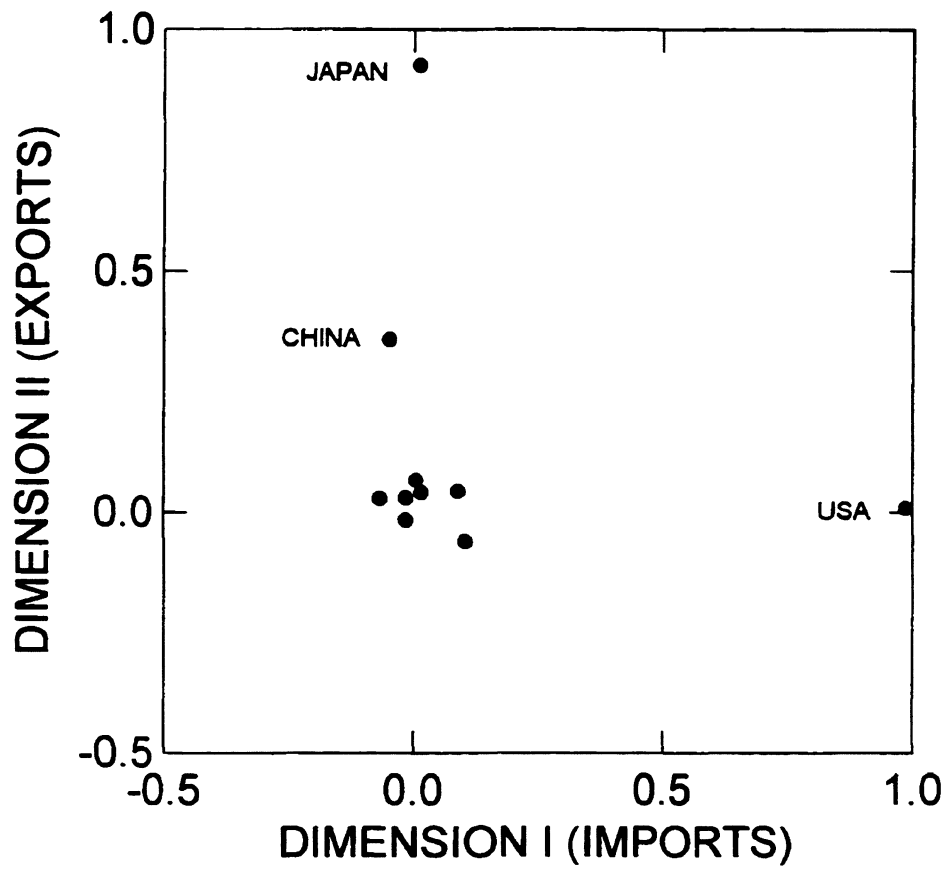
Kingdom) are included in the matrix, the imbalances are essentially restricted to the United States, Japan, and China.

The nature of the two dimensions helps to clarify this. The first dimension has a Pearson product moment correlation of .909 (.853 after varimax rotation) with the amount of imports by a country. The second dimension correlates .911 (.892 after varimax rotation) with the amount of exports. Hence, only those nations having a large absolute volume of trade participate significantly in the solution. The interpretation of the two dimensions as "sending" and "receiving" appears quite straightforward in this case. Note, too, that if each country's trade were perfectly balanced, the object points would fall along a straight, 45 degree line in Figure 24. All triangles between two object points and the origin would have zero area, indicating an absence of asymmetry.

The results of an NSKMDS analysis of the trade data are shown in Table 31. The sum of squared loadings on the first two dimensions are .695 and .722, suggesting that both contribute to the solution. The first NSKMDS dimension has a Pearson product moment correlation of .928 with total exports. The second NSKMDS dimension has a Pearson product moment correlation of .914 with total imports. Thus, the NSKMDS solution is almost identical to that of CAA. However, the NSKMDS dimensions appear to have slightly higher correlations with imports and exports. As may be seen in Table 31, there appears to be some degeneracy in the NSKMDS solution, with three countries (Brazil, France, and Russia) having loadings of almost exactly zero on the two dimensions. This does not impair goodness of fit, however, because the actual trade volume of these countries is far below that of the others.

FIGURE 24

CAA SOLUTION FOR INTERNATIONAL TRADE



NOTE: DATA FROM INTERNATIONAL MONETARY  
FUND (1994)



### Journal Citations

The study of cross-citations between journals often involves asymmetry. The number of times that articles in Psychometrika, for example, cite those in Psychological Bulletin will not necessarily equal the number of citations flowing in the reverse direction. Everett and Pecotich (1991) note that the bias coefficient approach of Bishop, Fienberg, and Holland (1975) is the preferred approach for this because citations are frequency data, for which log-linear analysis provides significance tests. However, we will apply CAA and NSKMDS to citation data, because these two techniques provide graphical displays that can complement the purely numerical approach of Bishop, et al. The graphical representations may be particularly helpful if the asymmetries require a greater than one-dimensional set of bias coefficients. An early cross-citation matrix was collected by M. Levine and published in Coombs (1964). It shows the citations among 10 psychology journals in 1960, and is reproduced in Table 32. The rows correspond to the journals in which the citations are found; the columns represent the journals whose articles are cited.

Coombs regards the matrix as an example of "row conditional data". Because journals differ in the number of articles they contain, and in the number of references per article, numbers in different rows are not on the same scale, and are not strictly comparable. In the unfolding analysis that he discusses, between-row comparisons are excluded from consideration. In the present analyses we will take the simpler approach of norming each row to a mean of 1. That is, we will divide each entry by the sum for its row, and then multiply this quotient by 10.

Table 32

## Journal Citation Data For 1960

	AJP	JASP	APP	JCPP	JCP	EDU	EXP	BUL	REV
AJP	122	4	1	23	4	2	135	17	39
JASP	23	303	9	11	49	4	55	50	48
APP	0	28	84	2	11	6	15	23	8
JCPP	36	10	4	304	0	0	98	21	65
JCP	6	93	11	1	186	6	7	30	10
EDU	6	12	11	1	7	34	24	16	7
EXP	65	15	3	33	3	3	337	40	59
BUL	47	108	16	81	130	14	193	52	31
REV	22	40	2	29	8	1	97	39	107
PKA	2	0	2	0	0	1	6	14	5

**Note.** AJP=American Journal of Psychology, JASP=Journal of Abnormal and Social Psychology, APP=Journal of Applied Psychology, JCPP=Journal of Comparative and Physiological Psychology, JCP=Journal of Consulting Psychology, EDU=Journal of Educational Psychology, EXP=Journal of Experimental Psychology, BUL=Psychological Bulletin, REV=Psychological Review, PKA=Psychometrika.

Numbers in the body of the table give the frequency with which articles in the column journal were cited by articles in the row journal.

**Source.** Coombs (1964)

A canonical analysis of asymmetry on the normed data yields eigenvalues of  $\lambda_1 = \lambda_2 = \pm.06172$ ,  $\lambda_3 = \lambda_4 = \pm.01429$ ,  $\lambda_5 = \lambda_6 = \pm.01210$ ,  $\lambda_7 = \lambda_8 = \pm.00076$ , and  $\lambda_9 = \lambda_{10} = \pm.00005$ . The eigenvalue ratio test gives values  $EVR1 = 3.658$ ,  $EVR2 = .0743$ ,  $EVR3 = 1.0047$ , suggesting a two-dimensional solution. This solution is shown in Table 33, and in Figure 25.

After varimax rotation the first factor correlates .917 with the column sum for the journal, that is, its total "export" of citations to the journals in the matrix. The second rotated factor correlates -.604 with the journal's ratio of imports to exports, that is, the ratio of the number of times the journal cites other journals, to the number of times it is cited by them. As may be seen in Figure 25, Psychological Bulletin and the Journal of Experimental Psychology have high loadings on the first factor because articles in them tend to be cited frequently. On the second factor, Journal of Experimental Psychology has a large positive loading, and Psychological Bulletin a large negative loading, because the former journal is primarily an exporter of citations, while the latter is primarily an importer.

As shown in Table 33, the NSKMDS solution is similar to the CAA solution, but contains marked degeneracy. The four journals with the highest loadings on the CAA dimensions show the same pattern on the NSKMDS dimensions. However, in NSKMDS the other six loadings on each dimension are tied at essentially zero. Use of locality parameters between -1 and -10 did not significantly deter NSKMDS from this solution.

Table 33

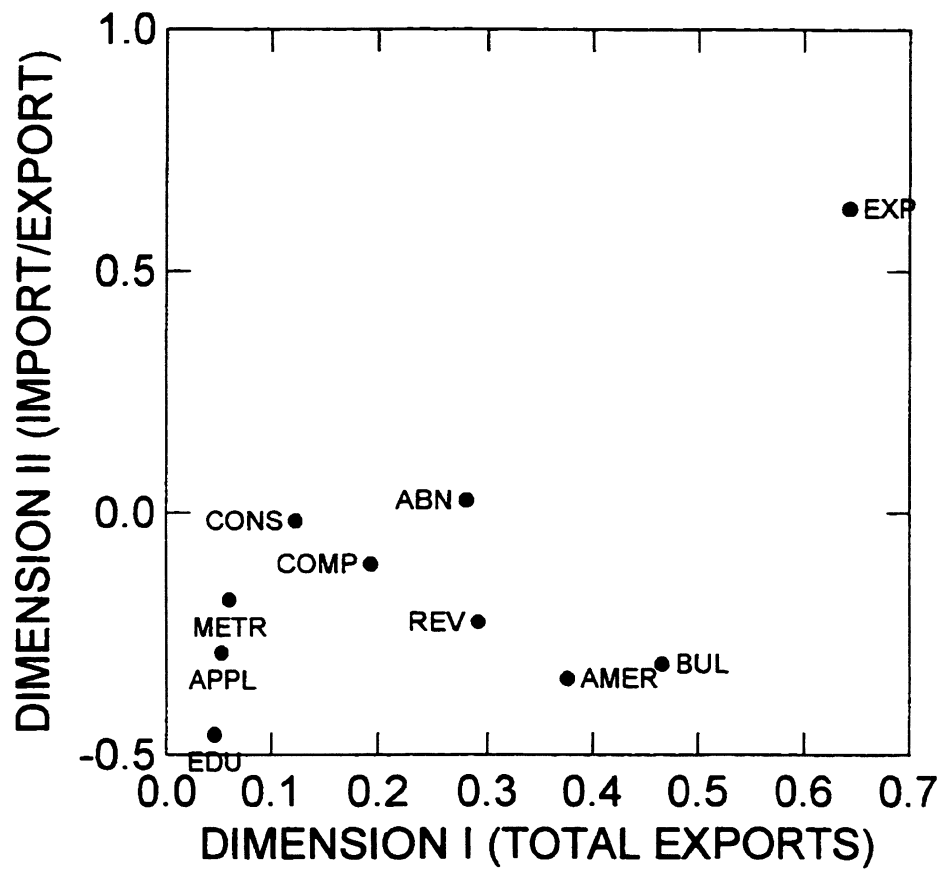
## Representation of Asymmetries in Journal Citation Data

	CAA		NSKMDS	
	Dimen. I <sup>1</sup>	Dimen. II	Dimen. I	Dimen. II
<b>American Journal of Psychology</b>	.3771	-.3426	.5247	-.2897
<b>J. of Abnormal and Social Psychology</b>	.2808	.0250	.0000	.0000
<b>J. of Applied Psychology</b>	.0533	-.2896	.0000	.0000
<b>J. of Comparative and Physiological</b>	.1940	-.1063	.0001	-.0001
<b>J. of Consulting Psychology</b>	.1234	-.0181	.0000	.0000
<b>J. of Educational Psychology</b>	.0463	-.4596	.3863	-.2133
<b>J. of Experimental Psychology</b>	.6438	.6280	.2241	.7301
<b>Psychological Bulletin</b>	.4667	-.3123	.5112	-.2822
<b>Psychological Review</b>	.2918	-.2259	.0002	-.0001
<b>Psychometrika</b>	.0607	-.1802	.0000	-.0001

<sup>1</sup> Note. CAA dimensions are shown after Varimax rotation.

## FIGURE 25

CAA SOLUTION FOR JOURNAL CITATIONS



NOTE: DATA FROM COOMBS (1964)

## Chapter 15

### Rationale for Experiments

#### Overview

In the two studies described here the plan was to

(1) use absolute magnitude estimation and rating scales to collect parametric similarity data on representative data sets;

(2) test whether the asymmetries are statistically significant;

(3) if the asymmetries are significant, determine which assumptions about the asymmetries are supported and, through this, state which special purpose models could justifiably be used;

(4) conduct exploratory analyses of the asymmetries using the canonical analysis of asymmetry (CAA) and its nonmetric generalization; and

(5) for parametric CAA, to determine the degree of relationship between the dimensions underlying the symmetric and skew-symmetric parts of the data.

Stimulus sets on which similarity ratings were collected were verbal exemplars of the natural categories vegetables and tools. Typicality ratings were collected separately to allow a test of Tversky's (1977) contrast model.

#### Experiment 1. Similarity.

The goal underlying Experiment 1 was to obtain similarity matrices that contain statistically significant asymmetries, that are measured on a parametric level, and that are

reliable enough to permit substantive interpretation of even minor variance components of the asymmetries.

(1) Significant asymmetries. A study by Tversky and Hutchinson (1986) is useful in guiding our search for asymmetries. Tversky and Hutchinson studied the appropriateness of multidimensional scaling models to proximity data, by computing two ordinal-level test statistics on the proximity matrices. Of particular interest to us is their reciprocity statistic,  $R_i$ .

Imagine that we are studying the similarity of four fruits. The row for "apple", for example, would contain the similarity values of three other fruits to an apple. Using this row we can find the fruit that is rated as most similar to an apple, say fruit "i". We can then use row "i" of the matrix to make the reverse determination: how similar to fruit "i" is an apple. This value is denoted  $R_i$ : the similarity to fruit "i" of the object to which "i" is most similar.

For example, consider Table 34. The values in each row have been reduced to rank orders.

Table 34

Hypothetical Similarity Matrix

To an	How similar is an			
	APPLE	PEAR	ORANGE	GRAPE
APPLE	—	1	3	2
PEAR	2	—	3	1
ORANGE	2	1	—	3
GRAPE	3	2	1	—

"Pear" is rated as most similar to "Apple", so  $R_{\text{apple}}$  is the similarity of an apple to a pear, in this case "2". "Grape" is rated as most similar to "Pear", so  $R_{\text{pear}}$  is the similarity rank of a pear to a grape, "2" here. "Pear" is ranked as most similar to "Orange" in the data, so  $R_{\text{orange}}$  is the similarity rank of "Orange" to "Pear", "3". "Orange" is most similar to "Grape" so  $R_{\text{grape}}$  is the similarity of grape to orange, "3" in this data set.

The reciprocity,  $\underline{R}$ , of this data set is

$$\begin{aligned}\underline{R} &= [1/(K+1)] \sum_i R_i \\ &= (1/5)[R_{\text{apple}} + R_{\text{pear}} + R_{\text{orange}} + R_{\text{grape}}] \\ &= (1/5)[2 + 2 + 3 + 3] = 2.\end{aligned}$$

The reciprocity statistic is a limited measure of the asymmetry of the matrix. It can be rewritten as

$$\underline{R} = (1/2) \sum_{i,j \text{ restricted}} (X_{ij} - X_{ji} + 1),$$

the difference (plus one) between values on one side of the main diagonal and the corresponding values on the other side. However, the sum only involves those  $X_{ij}$ ,  $X_{ji}$  pairs for which at least one of the two values equals 1. Thus,  $\underline{R}$  is based on the sum of  $K$  different  $R_i$  values. Note that a full accounting of the asymmetries would make use of all  $K^2 - K$  off-diagonal cells and, for parametric data, would not involve a prior conversion to ranks.

Nonetheless, the reciprocity statistic is a rough guide to the degree of asymmetry in a data set. Tversky and Hutchinson (1986) examined 100 previously studied proximity matrices and found significant asymmetries in 33 of them. Almost invariably, the asymmetric data sets consisted of (a) word associations; or (b) subjective similarity values for objects belonging to a semantic category and varying in their typicality. Examples of the latter would be a car and a



skateboard, both members of the category "vehicles" (loosely defined), but varying in their typicality, the degree to which they are good examples of the category. In the present study, subjective similarity will be used as a source of presumably asymmetric proximity data.

Parametric data. There are at least two well-validated techniques for collecting parametric-level data. Anderson (1974) found that rating scales give interval-level data provided that subjects are familiar with the scale and the stimuli, and provided end effects are avoided. End effects may be understood as follows: Picture that we study objects representing a range of values on a particular attribute. The ratings for objects near the center of the range will tend to be distributed symmetrically and approximately normally over several adjacent categories. However, the ratings for extreme stimuli will "pile up" in the outside categories. The distribution of ratings for extreme stimuli will be skewed towards the center of the scale. The ends of the rating scale truncate the distribution of ratings, lowering discrimination among extreme stimuli and producing a bias towards the center of the scale (Thurstone, 1929). Anderson (1974) found that end effects can be removed from the data by including more extreme stimuli than those we intend to analyze. The unwanted "pile up" of extreme stimuli is thus sequestered by "dummy objects" that have no substantive role in the study.

A different approach to obtaining parametric data is provided by absolute magnitude estimation. Stevens (1956) proposed that people associate with numbers a subjective sense of "numerosity", the subjective magnitude that the numbers convey. For example, people might regard the number ".01" as inherently small and "1,000,000" as inherently large. Absolute magnitude estimation is a cross-modality match between this sense of numerosity and the

subjective magnitude produced by a stimulus (Hellman & Zwillocki, 1961). People are asked to "assign a number to every [stimulus] so that the subjective magnitude of the number is equal to the subjective magnitude of the [stimulus]" (Zwiloeki, 1983, p.463).

Absolute magnitude estimation appears to produce ratio-level data while avoiding many of the context effects that arise in ordinary magnitude estimation (Collins, & Gescheider, 1989; Gescheider, & Hughson, 1991). In the present study absolute magnitude estimation of similarity will be used in addition to rating scales, to help verify that the results are independent of specific experimental task.

The AME and rating scale tasks are straightforward and the stimuli familiar. This should help increase the reliability of the ratings. Dimensional analyses will be conducted on the mean ratings averaged across subjects. Therefore increasing the number of subjects will enhance reliability. This runs into a problem of diminishing returns, of course, as the standard error of the mean declines only in proportion to the square root of the sample size. Fifty subjects seems a reasonable balance between economy and reliability.

### Experiment 2. Typicality

Rosch (1975a) found that the psychological distance from a poor example of a category (e.g., a line at an 85 degree inclination) to a prototype (e.g., a vertical line) was shorter than the reverse distance. Similarly, Tversky found that similarity ratings were higher when comparing a low salience to a high salience stimulus. Tversky described salience as dependent on the "intensity, frequency, familiarity, good form, and informational content" (Tversky, 1977, p. 332). Certainly one source of asymmetries in similarity data is thus likely to be stimulus typicality or salience. Tversky and Rosch tested their theories against the null

hypothesis of no asymmetry in the context of salience or typicality difference. Thus, their studies were not designed to detect the presence of a multidimensional structure underlying the asymmetries. However, pilot studies to the present investigation (Borkum, 1991a, 1991b) suggest that a multidimensional structure is quite likely. This raises three possibilities about the relationship between asymmetry and typicality. Typicality may turn out to be one factor among many that brings about asymmetries: typicality would be a strictly unidimensional construct that correlated with one of the asymmetry dimensions. Alternatively, typicality may itself turn out to be a multidimensional construct. The various dimensions, aspects, or types of typicality may correlate with those of the asymmetries. The most complex possibility would be that typicality is multidimensional, but that only one aspect of typicality produces a directional effect in similarity judgments.

To distinguish between these it is important to determine the number and nature of dimensions underlying typicality judgments.

## Chapter 16

### Experiment 1: Similarity

#### Method

##### Stimuli

Rosch (1975b) prepared a relatively exhaustive list of semantic categories and their exemplars. Using the word frequency norms of Kučera and Francis (1967) she found 17 categories of concrete objects that were represented by five or more items. Seven of these categories had specific characteristics or ambiguities that made them difficult to use as stimuli. For the remaining ten categories, she obtained fifty to sixty exemplars each, varying widely in degree of association to the category, from the Battig and Montague (1969) category production norms. Rosch obtained typicality ratings for each of these exemplars, and presented them in rank order of typicality.

These semantic categories, with items differing widely in typicality, should provide the stimuli necessary for a study of asymmetries. For pairwise similarity judgments, ten stimuli from a category seems sufficient. Beyond this number and we are subject to diminishing returns. For example, we can determine the asymmetric relationships among ten stimuli with  $10^2 - 10 = 90$  judgments, but to study an additional two stimuli would increase the number of judgments to  $12^2 - 12 = 132$ .

Our first preference might be to select ten stimuli at equal intervals along the typicality continuum: this would provide a representative set of the instances of the category. Alternatively, we could sacrifice representativeness and select only high and low typicality

items, so as to maximize the number of likely asymmetric pairs. However, I suspect that neither approach would be optimal. Typicality is likely to exert an influence in the experiment only if the category is kept firmly in mind. This seems best accomplished by increasing the number of high typicality exemplars. Therefore, in categories with c. 50 exemplars I have chosen stimuli at approximately the following ranks: 1, 3, 5, 8, 12, 17, 23, 30, 38, and 47. For categories with c. 60 exemplars, stimuli with ranks of approximately 1, 3, 6, 10, 15, 21, 28, 36, 45, and 55 have been drawn.

Two categories – carpenter's tools and vegetables – were used for the pairwise similarity judgments. The list of stimulus words for these two categories is given in Appendix 5.

For the rating scale study the  $10^2 - 10 = 90$  unique pairs of 10 stimuli in a category were formed. These were supplemented by the 10 pairs of identical stimuli (e.g., "sander - sander"), and by 10 pairs in which each of the ten stimuli in turn is matched with a term from a different category (e.g., "sander - magazine rack"). This brings the total number of stimulus pairs rated to  $110 * 2 \text{ categories} = 220$ . The identical pairs and extraneous pairs are the extreme stimuli to help eliminate end effects.

### Subjects

Eighty-seven undergraduates drawn from the PSY 100 subject pool (43 in the magnitude estimation and 44 in the rating scale condition) participated in the group study in exchange for one hour of extra credit towards their course grade. Standard informed consent procedures were used and the research conformed to American Psychological Association guidelines on the ethical treatment of human subjects.

### Procedure

Category ratings. Subjects were asked to complete two ratings booklets. One booklet contained the 110 pairs of carpenter's tools, and the other booklet contained the vegetables. Subjects completed the booklets in random order, and the pairs within booklets were randomized individually for each subject. The instructions were

On the pages that follow you will see 110 pairs of objects. Almost all of these objects are carpenter's tools. For each pair, please decide how similar the first tool is to the second tool. Then, please rate the similarity of the first tool to the second tool on this 1 to 9 scale:

1	2	3	4	5	6	7	8	9
not at all								very
similar								similar

Here are some sample pairs, to show you the type of objects involved. Please rate them.

1. How similar are PLIERS to WOOD? \_\_\_\_\_
2. How similar is a SAW to NAILS? \_\_\_\_\_
3. How similar is a HINGE to CEMENT? \_\_\_\_\_

[Ten examples were given, including a pair of identical objects, and an extraneous pair.]

[The 110 pairs were presented on the following pages of the subject's booklet, in a different random order for each subject. Each pair was embedded in the full sentence, "how similar is a...?", to help ensure that the judgments would be directional.]

Absolute magnitude estimations. The procedure for the two category booklets was the same as for the ratings, except that: (1) the booklet and the 10 practice pairs did not include identical, nor extraneous pairs. (2) Instead of the 9 point scale, subjects were instructed:

This is a matching study.

People often have an internal, intuitive feel for the size of numbers. For each number, they have a sense, or a feeling, of how large or small it is.

On the pages that follow you will see 110 pairs of objects. Almost all of these objects are carpenter's tools. For each pair, please decide how similar the first tool is to the second tool. Then, please choose a number that matches this similarity. The feeling of how large the number is should match the feeling of how similar the objects are. Please do not use a rating scale. Rather, please try to match your sense of the amount of similarity with a number that has the same size.

You can use whole numbers, fractions, decimals, whatever number best reflects the amount of similarity. Treat every pair individually and do not worry about the numbers you gave to pairs that came before.

Here are some sample pairs, to show you the type of objects involved. To each of them please give a number that shows the amount of similarity.

## Results

The ten extraneous pairs, in which stimuli were matched with objects from outside the category, were not included in the analyses.

Tool category. For each subject, an intraclass correlation coefficient (ICC) was computed from the ten repeated pairs. The ICC was selected to assess absolute agreement. That is, subjects who assigned ratings of "1", "2", and "3" to two pairs at the first presentation, and "7", "8", and "9" at the second presentation, would be considered

unreliable, even though the product moment correlation coefficient between the two sets of ratings is 1.00. Reliability was assessed for the individual presentation of the stimulus pairs, rather than for the average of the two presentations. To determine the significance of an ICC, the mean square for the stimulus pairs was tested against the pooled mean square within stimulus pairs in a one-way ANOVA (McGraw & Wong, 1996). Only subjects whose reliabilities differed significantly from zero at the  $\alpha=.05$  level were included in subsequent analyses. This corresponded to an ICC of approximately 0.63.

Of the 43 subjects in the magnitude estimation condition, 21 did not meet the criterion of reliability. Hence, the magnitude estimation data for 22 subjects was retained. Similarly, 24 of the 44 subjects giving rating scale data did not meet the reliability threshold and were excluded from further analysis.

In examining the data, it appeared that nearly all subjects in the magnitude estimation condition had in fact used a rating scale, that is, a scale with definite end-points. Of the 22 subjects whose data was retained, 10 used only whole numbers between 0 and 100, with the occasional addition of half-steps, 6 used only whole or half steps between 0 and 10 and 3 used only whole or half steps between 0 and 9.

In order to equate the differing scales used in the magnitude estimation data, each subject's judgments were first transformed into z-scores. The ratings and magnitude estimations were then averaged across subjects, and the scatterplot was examined, of ratings vs. magnitude estimations for the 100 stimulus pairs. The concave downward function, that generally characterizes the relationship of category judgments to magnitude estimations (Stevens & Galanter, 1957) did not emerge. Rather, the scatterplot was very nearly linear,



suggesting that magnitude estimation subjects had in fact produced rating scale data.

Therefore, data from the rating scale and magnitude estimation conditions were combined in all following analyses.

As noted, the data for a subject was retained if it showed a high internal consistency across the ten repeated pairs. The agreement or internal consistency across subjects was assessed as well. Conservative estimates were obtained by excluding the diagonal elements, that is, the comparisons of objects with themselves, from the calculations. Subjects invariably agreed in assigning the highest possible similarity values to these pairs, but it was not clear that this would properly reflect their agreement on the pairs in which the objects differed. The average similarity ratings across the 42 subjects had an absolute agreement ICC of .9192. This is approximately the anticipated ratio-level correlation the average ratings would have with the average ratings of a new sample of 42 subjects drawn randomly from the same population. At the level of the individual subject the absolute agreement ICC was .1122. Thus, an individual subject's ratings would be expected to have a ratio-level product moment correlation of .1122 with those of another subject drawn randomly from the same population. The significance of the reliabilities was checked by testing the mean square for stimulus pairs against the mean square of the stimulus pair by subject interaction. The reliabilities were significant,  $F(89,3649) = 12.37, p < .001$ .

Reliabilities were computed for the rating scale and magnitude estimation conditions separately. The average similarity judgments for the 20 subjects in the rating scale condition had an absolute agreement ICC of .9903. The absolute agreement ICC for individual subjects was .5323. The reliabilities were significant, with  $F(89,1691) = 103.44, p < .001$ . For the

magnitude estimations, the average similarity judgments had an absolute agreement ICC of .9360, with an ICC at the individual subject level of .1398. These reliabilities were significant,  $F(89,1869) = 15.62, p < .001$ .

Average similarity matrices were computed for the magnitude estimation condition, the rating condition, and the two conditions combined. Disregarding the diagonal elements, 86.36% of the average magnitude estimation sums of squares were symmetrical and 99.72% of the average ratings sums of squares were symmetrical. In the two conditions combined, 97.45% of the sums of squares of the average similarity judgments were symmetrical. The Hubert and Baker (1979) coefficient, that is, the Pearson product moment correlation of the rows with the columns, was .9243 for the average magnitude estimations, .9914 for the average ratings, and .9201 for the average similarity judgments across conditions, indicating a high degree of symmetry.

The asymmetries were then isolated for each subject in each condition. Overall, the asymmetries had an absolute agreement ICC of .0000 at the level of the individual subject, and -.0002 for the average of the 42 subjects. Of course these were not statistically significant. For the magnitude estimates,  $F(44,924) = 1.20, ns$ ; for the ratings,  $F(44,836) = .90, ns$ ; and for the two conditions combined,  $F(44,1804) = 1.00, ns$ .

A power analysis indicated that, given the number of subjects in the two conditions combined (42) and the asymmetry by subject interaction mean square, which was used to gauge error, a correlation ratio in the population of .048 would be detected by this experiment 90% of the time. Hence, if similarity judgments are asymmetric, the effect size associated with the asymmetries would be judged from this experiment to be below .048.

In the typicality judgments given by subjects in this study (see experiment 2) "wood" emerged as the most typical, and "paintbrush" as the least typical, of the tools. A paired t-test was conducted to determine whether, across the 42 subjects, the judged similarity of "wood" to "a paintbrush" differed from the similarity of "a paintbrush" to "wood". The similarity judgments were higher when the least typical exemplar was compared to the most typical exemplar, but the difference was not significant:  $t(41) = 1.27, p > .10$ .

The symmetric part was then isolated from the similarity judgments averaged across subjects. A principal components analysis was then conducted on this symmetric matrix. The scree plot does not show definite breaks, but seems most consistent with a one-factor solution. Across tools, the first component has a Pearson product-moment correlation of .800 ( $p < .01$ ) with the subjects' average typicality judgments.

Vegetable category. Similar analyses were conducted for the judged similarities among vegetables as for the similarities among tools. Absolute agreement ICCs were computed for each subject using the ten repeated pairs. The significance of the ICCs was assessed by testing the mean square across pairs against the pooled mean square within pairs. Significance at the  $\alpha = .05$  level generally corresponded to an absolute agreement ICC at the level of the individual judgment (not the average of the two judgments for each pair) of .63. The data of 18 of the 43 magnitude estimation subjects and 17 of the 44 subjects using rating scales, were reliable by this criterion, and retained for further analyses. Of the 18 subjects retained in the magnitude estimation condition, 3 used only whole or half steps between 0 and 9, 7 used only whole or half steps between 0 and 10, and 5 used only whole or half steps between 0 and 100. In a scatterplot for the 100 stimulus pairs, the average rating across

subjects shows a nearly linear relation to the average magnitude estimation. Thus, it seems likely that the subjects in the magnitude estimation condition were in fact making category judgments. For each subject, similarity judgments were transformed to z-scores. Data from the two conditions were then combined as noted for most of the analyses to follow.

An intraclass correlation coefficient was used to assess absolute agreement across subjects. A conservative test was arranged by omitting the diagonal elements (comparison of an object with itself) from the reliability and significance calculations. It did not seem certain that judgments of self-similarity, which showed essentially perfect agreement, would adequately reflect the reliability for pairs in which the stimuli differed. In the magnitude estimation condition, the average similarity judgments had an ICC of .9691, and the individual subject's judgments had an ICC of .2583. Testing the mean square for stimulus pairs against the stimulus pair by subject interaction, the reliabilities were significant:  $F(89,1602) = 32.33, p < .001$ . In the rating scale condition the average judgments had an ICC of .9894, and the individual subjects' judgments had an ICC of .1456. The intersubject agreements were significant, with  $F(89,1424) = 94.43, p < .001$ . When the rating scale and magnitude estimation conditions are combined, the average judgments have an absolute agreement ICC of .9416, with a absolute agreement ICC of .1519 at the level of the individual subject. These agreement coefficients are significant:  $F(89,3026) = 17.12, p < .001$ .

Average similarity judgments were computed across subjects for the magnitude estimation condition, the rating scale condition, and the two conditions combined. Disregarding the sums of squares on the diagonal (all of which would count toward the

symmetries if it were included), 93.82% of the average magnitude estimation sums of squares, 97.88% of the average rating scale sums of squares, and 98.47% of the average combined sums of squares, was symmetric.

The asymmetric part of the similarity judgments was then extracted for each subject in each condition. In the rating scale condition, the asymmetries had an absolute agreement ICC of .0008 at the individual subject level, and .0351 at the level of the average rating. For magnitude estimations the individual subject ICC was -.0013, while the average judgments had an absolute agreement ICC of -.2413<sup>5</sup>. For the two conditions combined, the absolute agreement at the individual subject level was -.0033, and for the average judgments it was -.1721. These are not statistically significant. For the rating scale condition,  $F(44,704) = 1.04$ , *ns*. For the magnitude estimation condition,  $F(44,748) = .81$ , *ns*. For the two conditions combined,  $F(44,1496) = .85$ , *ns*.

A power analysis indicated that, given the number of subjects whose data was retained (35), and the level of error variance, estimated as the asymmetry by subject interaction, a population correlation ratio of .057 would be detected by this experiment 90% of the time. Hence, if similarity judgments are asymmetric, the effect size of the asymmetries would be expected from this research to be below .057.

As discussed in Experiment 2, below, the subjects in this study judged "peas" to be the most typical, and "dandelion" the least typical, of the stimuli. These results were also

---

<sup>5</sup> A negative reliability estimate is obtained whenever the mean square for error exceeds the mean square for the effect. That is, "negative reliability" occurs when the F ratio is below 1. Unless error variance has been artificially inflated through a mistake in the experimental design, a negative ICC would presumably reflect sampling error from a population whose reliability is zero.

obtained by Rosch (1975b). In the present experiment, the average judged similarity of peas to a dandelion exceeds the judged similarity of a dandelion to peas, contrary to hypothesis. The difference is not statistically significant, however, as assessed by a paired t-test:  $t(34) = -.82$ ,  $p > .25$ .

A principal components analysis was then conducted of the squared, average symmetric part of the vegetable similarity judgments. A two-factor solution is suggested by examination of the scree plot. The first factor, after varimax rotation, has a Pearson product-moment correlation of  $.802$  ( $p < .01$ ) with the subjects' average typicality judgments.

## Chapter 17

### Experiment 2: Typicality

The current study was designed to investigate typicality for a multidimensional structure. This was accomplished with Q-factor analysis (Borkum, 1989). In this technique, typicality ratings of stimuli are collected from a large number of subjects. Principal components are extracted from the subject by subject correlation matrix. These components are the dimensions underlying the typicality ratings. Rotation to simple structure (e.g., varimax) may be conducted, and would correspond to the hypothesis that each typicality dimension is used solely and exclusively by a different cluster of subjects. With or without rotation, the dimension is labelled by comparing objects with high component scores to those with low component scores.

#### Stimuli

The stimuli were the ten exemplars for each of the two categories, as given in Appendix 5.

#### Subjects

Eighty seven undergraduates were recruited for the group study from the departmental subject pool in accordance with the procedures noted above. These were the same subjects who gave the similarity judgments described in Study 1.

#### Procedure

Typicality judgments were not requested until after the similarity judgments were complete. Half of the subjects were requested to use a 1-9 scale for their judgments, and half

were asked to use absolute magnitude estimation. Subjects used the same judgment modality, rating scale vs. magnitude estimation, for typicality judgments as they had for the similarity judgments. Subjects wrote their typicality judgments in a booklet. Instructions for the rating scale condition for tools were:

"On the next page you will see 12 object names. For each object please decide how typical it is of the category "Carpenter's Tools. That is, please decide how much it is a typical or good example of the category "Carpenter's Tools", vs. being an atypical or bad example of a Carpenter's Tool. Please rate this typicality on the following 1 to 9 scale:

1	2	3	4	5	6	7	8	9
Not at all typical of a carpenter's tool					Very typical of a carpenter's tool			

Instructions for subjects asked to use absolute magnitude estimation were:

This is a matching study.

People often have an internal, intuitive feel for the size of numbers. For each number, they have a sense, or a feeling, of how large or small it is.

On the next page you will see 12 object names. For each object please decide how typical it is of the category "Carpenter's Tools". That is, please decide how much it is a typical or good example of the category "Carpenter's Tools", vs. being an atypical or bad example of a Carpenter's Tool.

Then, please choose a number that matches this typicality. The feeling of how large the number is should match the feeling of how much the object is a typical carpenter's tool. Please do not use a rating scale. Rather, please try



to match your sense of the amount of typicality with a number that has the same size.

You can use whole numbers, fractions, decimals, whatever number best reflects the amount of typicality. Treat every object individually and do not worry about the numbers you gave to objects that came before."

On a single page, following the instructions, 12 words were printed, arranged vertically down the page. Ten of the words were the ten exemplars of the category. One word ("motorcycle" for the vegetables category and "duck" for the tools) was obviously extraneous and drawn from a different category. The remaining word was the category name itself. The extraneous word and category name were designed to be extreme stimuli to reduce end effects in the use of the rating scales. The 12 words were presented in a different random order for each subject.

### Results

Subjects whose similarity data had been excluded as possibly random were excluded from the typicality analyses. The 35 subjects remaining for judgments of typicality of vegetables showed good absolute agreement, with an intraclass correlation coefficient of .964 for the average ratings, and .728 for the individual subjects' ratings. These reliabilities were significantly different from zero,  $F(9,306) = 27.77, p < .001$ . A principal components analysis was conducted on the subject by subject correlation matrix. The eigenvalues are shown in Table 35. The first seven values of the eigenvalue ratio test are 4.68, 0.95, 0.80, 1.13, 1.35, 0.45, and 1.92, suggesting a one-factor solution. EVR values after the seventh cannot be calculated because eigenvalues 10-35 are approximately zero. This is largely artifactual: because there are only ten typicality ratings underlying the subject intercorrelations,

all eigenvalues after the tenth are necessarily zero. In these data, the tenth eigenvalue is also near zero, suggesting some further linear dependence across subjects. The object scores on the first principal component have a Pearson product-moment correlation of  $-.87$  with the ratings obtained by Rosch (1975b). The primary difference between the ratings obtained here, and those reported by Rosch, appears to be a greater typicality accorded to potatoes and to peppers in the current data (see Table 36). The direction of the scale in the present study differed from that used by Rosch, which accounts for the negative sign in the correlation.

Table 35

## Eigenvalues of Vegetable Typicality Judgments

Principal Component Number	Eigenvalue
1	21.749
2	3.665
3	2.893
4	2.168
5	1.292
6	0.868
7	0.786
8	0.323
9	0.255
10	0.000

The 42 subjects remaining for judgments of the tools category showed acceptable interjudge agreement, with an intraclass correlation coefficient of .928 for the average ratings and .565 for the individual judges. The reliabilities differ significantly from zero,  $F(9,369) = 13.98, p < .001$ . A principal components analysis of the typicality judgments gave the eigenvalues shown in Table 37. The first eight values of the eigenvalue ratio test are 1.73, 1.29, 0.68, 1.17, 1.00, 1.05, 0.89, and 0.00, suggesting a one-factor solution. Eigenvalues 10-35 are essentially zero, for the same reasons as noted for the vegetable typicality judgments. The tools' scores on the first principal component have a Pearson

Table 36  
Typicality Judgments of Vegetables

VEGETABLE	MEAN TYPICALITY	Rosch (1975b)
Pea	0.784	1.07
Green Beans	0.676	1.18
Broccoli	0.653	1.28
Brussels Sprouts	0.189	1.72
Beets	0.025	2.08
Eggplant	-.019	2.38
Potato	0.603	2.89
Peppers	0.278	3.21
Avocado	-.618	3.62
Dandelion	-2.577	5.20

product moment correlation of  $-.84$  with the ratings reported in Rosch (1975b). Subjects in the current sample rated nails and, in particular, wood, as more typical of carpenters' tools than did subjects in Rosch's study (see Table 38).

Table 37

Eigenvalues of Typicality Judgments of Tools

<b>PRINCIPAL COMPONENT NUMBER</b>	<b>EIGENVALUE</b>
1	15.528
2	6.070
3	4.105
4	3.573
5	2.108
6	1.449
7	0.992
8	0.716
9	0.458
10	0.000

Table 38  
Typicality Judgments of Tools

<b>TOOL</b>	<b>MEAN TYPICALITY</b>	<b>Rosch (1975)</b>
<b>Saw</b>	0.594	1.04
<b>Ruler</b>	0.568	1.48
<b>Nails</b>	0.901	1.67
<b>Sander</b>	0.410	1.79
<b>Chisel</b>	0.084	2.26
<b>Pliers</b>	0.346	2.56
<b>Wood</b>	0.942	2.77
<b>Hinge</b>	-.317	3.12
<b>Paintbrush</b>	-1.836	3.81
<b>Cement</b>	-1.690	4.91

## Chapter 18

## Discussion

Nonmetric, skew-symmetric MDS (NSKMDS)

In developing a nonmetric version of the canonical analysis of asymmetry, the explicit goal was to be able to accurately represent the rank order of the skew-symmetries. The NSKMDS algorithm appears quite satisfactory for this. In the simulation studies it was relatively unaffected by accelerating and decelerating monotonic transforms of the asymmetries. Nonmetric recovery, operationalized as the squared Spearman rank order correlation coefficient between the true and recovered asymmetries, was above .89 for all of the monotonic transforms sampled between  $x \rightarrow x^{0.2}$  and  $x \rightarrow x^{2.5}$ . Nonmetric recovery by NSKMDS declined for exponents above 2.5. However, these are fairly drastic transforms that seem unlikely to occur in actual data.

It is surprising at first that monotonic transforms would have any effect on a nonmetric technique. However, NSKMDS relies on the canonical analysis of asymmetry (CAA) for its starting configuration. As the correct metric and nonmetric solutions diverge at higher distortion levels, NSKMDS is required to travel further from the starting configuration to the final results. Thus there is presumably more opportunity for local minimum and degeneracy problems to arise.

NSKMDS relies only on ordinal-level information in the data. In this respect it is similar to the nonmetric multidimensional scaling of symmetric data. Standard nonmetric MDS is useful primarily because it generates solutions with desirable metric properties, and

with more parsimony than metric MDS solutions. This, indeed, was noted in early MDS studies of hue (e.g., Helm, Messick, & Tucker, 1961; Helm, 1960; Mellinger, 1956), and attributed to subjects' tendencies to underestimate the larger color differences. Metric techniques, that attempted to reproduce the judgments at an interval level, produced a surfeit of dimensions. The first two dimensions, the color wheel, were substantive, and the other dimensions reflected the subjects' decelerating monotonic distortion.

Similar results are found for the nonmetric MDS of asymmetry. The NSKMDS algorithm shows very high metric recovery, despite relying only on ordinal level information. For the sampled monotonic transforms between  $x \rightarrow x^{20}$  and  $x \rightarrow x^{2.5}$ , the squared ratio-level product moment correlation between the true and recovered asymmetries was consistently greater than .85. The canonical analysis of asymmetry tended to have a marginally better nonmetric recovery than NSKMDS for positively accelerated transforms, that is, transforms that exaggerate the differences between large and small asymmetries. A much stronger effect, however, was the higher metric and nonmetric recovery by NSKMDS, compared with CAA, for decelerating transforms.

This advantage of the nonmetric technique, NSKMDS, is qualified in several ways. First, although it is statistically significant, and larger the more decelerating the transform, it is probably not large enough to be of practical importance. Moreover, it is balanced out by a higher susceptibility of the nonmetric technique to error. Indeed, one of the most surprising results of the simulations was the robustness of the canonical analysis of asymmetry to even large amounts of normal or asymmetric error, and to monotonic transforms of the input data.

In shifting from large simulation runs to the analysis of selected matrix structures and published data sets, a different pattern of strengths and weaknesses emerged. NSKMDS appeared much better able to represent a simplex, and thus to show when some or all of the asymmetries were likely to reflect constant error, such as a time or spatial position error. There was a suggestion, although one that would need verification, that the nonmetric technique may better recover multiplicative bias coefficients. However, NSKMDS solutions appear quite prone to a particular type of degeneracy, in which the smallest asymmetries are set equal to zero. Thus, NSKMDS appears less able to distinguish among the smallest asymmetries than is CAA. For some analyses this tendency did not significantly impair the solutions, as the smallest asymmetries were small enough to disregard. In other analyses, however, it led to discomfoting lacunæ in the solutions.

The pattern of strengths and weaknesses of NSKMDS leads to the hypothesis that the technique tends to increase the larger asymmetries in the data, and shrink the smaller asymmetries. This would account for its ability to "straighten the horseshoe" of a simplex, representing it as a straight line. It would also explain why NSKMDS was able to outperform CAA for negatively accelerated transforms, but performed slightly more poorly than CAA for positively accelerated transforms. From this perspective, its strength is also its weakness, for the reduction of the smallest asymmetries to zero would simply be an extreme form of this tendency. The NSKMDS algorithm contains no explicit instructions to differentially increase the large asymmetries and decrease the small asymmetries, nor, equivalently, to increase the variance of the skew-symmetries. Hence, this is an unexpected property of the algorithm. Its most likely source seems to be at the monotone regression



stage. Small and nearly equal asymmetries would be more likely to be in the wrong rank order, and hence to be set equal to each other by the monotone regression. Particularly small asymmetries are susceptible not only to being in the wrong order, but to having the wrong sign, and would be set equal to their mean, a number close to zero. Hence, at each iteration the monotone regression module tends to divide the disparities into small, tied values, and large, untied values.

This hypothesis suggests directions for future work in developing NSKMDS. Increasing the size of large vs. small asymmetries is an excellent dimension-reduction technique. Indeed the first nonmetric MDS program, the analysis of proximities, explicitly maximized the variance among distances so as to produce parsimonious solutions (Shepard, 1962). NSKMDS contains no such explicit instructions, nor instructions to achieve greater parsimony than the pre-set dimensionality of the solution. However, NSKMDS could be modified to seek explicitly a more parsimonious solution. Indeed, at the simplest level, NSKMDS could be readily adapted to produce monotonic one-dimensional solutions, by replacing  $x_{III}$ ,  $x_{II}$  with 1 in the equations used to derive the objects' coordinates. The squared product-moment correlation between the asymmetries in the data, and those derived from a one-dimensional model, would indicate the appropriateness of the solution.

More generally, future work with NSKMDS should attempt to delineate the conditions under which a degenerate solution will form, and to develop modifications to prevent degeneracy. The locality parameter was not successful for this.

In its present form, the NSKMDS algorithm is limited to two-dimensional solutions. This limitation appears easily circumvented, however. Monotonic regression can be used to

generate values that are as close as possible to the skew-symmetries in the data, while strictly maintaining the rank order of the skew-symmetries in the two-dimensional solution. These predicted values would be subtracted from the skew-symmetric data to leave "monotonic residuals" that could themselves be analyzed into a two-dimensional solution. The two dimensions of the residualized data would be interpreted as the third and fourth dimensions of the skew-symmetric data. The residualization process could then be conducted again, and the twice-residualized data analyzed to give the fifth and sixth skew-symmetric dimensions, etc.

Pending these developments, the experience with the techniques reported in this dissertation can give some guidance about whether CAA or NSKMDS should be used in a given analysis. In a psychophysical study with a sensory modality whose power law exponent is known to be below 1, and whose error level is likely to be low, NSKMDS would probably be the preferred approach. NSKMDS could also be useful for examining whether some or all of the asymmetries in a matrix are likely to be due to constant error. All other investigations, including studies for which the type of distortion or amount of error are unknown, should use the canonical analysis of asymmetry due to its greater robustness.

### Simulation Program

To my knowledge, metric and nonmetric MDS have not been systematically compared in the way CAA and NSKMDS have been compared in this dissertation. Generally, nonmetric MDS is recommended due to its ability to recover metric configurations from the nonmetric properties of the data. However, we have seen that under some conditions (positively accelerated transforms plus high error) a metric technique, CAA,

provides better metric and nonmetric recovery than a nonmetric technique, NSKMDS. Nonmetric MDS received much of its impetus from analyses of negatively accelerated data, judged similarities among hues. Moreover, Shepard diagrams, relating disparities to distances, nearly always show the distances to be a positive, exponential function of the data. Thus there is reason to believe that nonmetric multidimensional scaling, like NSKMDS, is deriving much of its advantage from a tendency to differentially expand large differences, and that its advantage disappears, or is reversed, for positively accelerated transforms. The current studies of skew-symmetric MDS, then, raise the hypothesis that the advantage of nonmetric MDS does not pertain to ordinal level data in general, but is specific to data in which the differences between large and small dissimilarities have been minimized by a decelerated transform, such as a power law exponent below 1 in psychophysics. The simulation program developed for this dissertation, with its capacity to generate symmetric matrices using either a component or a distance model, and to apply normal, Wishart, chi-square, or asymmetric error whose magnitudes can be adjusted separately from the type of error, would seem well-suited to a systematic comparison of metric and nonmetric MDS.

Surprisingly, virtually no studies of metric recovery by the various factor analysis techniques have been conducted (Seber, 1984). The simulation program developed for this dissertation should, with the addition of subroutines for factor analysis, suffice for this purpose.

### Eigenvalue Ratio Test

In studying the multidimensional structure of asymmetries it was useful to have a guideline for the number of dimensions to extract. The eigenvalue ratio test was developed here for this purpose. Although its cutoff values were determined empirically rather than through a theoretical analysis of its sampling distribution, the derived  $\alpha$  levels appear stable across simulations. The power associated with the test was relatively low, increasing with effect size to roughly the same degree as would a one-way, four-groups ANOVA with 3 subjects per condition. This conservativeness may be an advantage, however, in that we would ordinarily seek factors that are not simply statistically significant, but that have enough true score variance to be well-defined.

Although the eigenvalue ratio test was derived empirically for skew-symmetric data, its development was suggested by the exponential decay in successive eigenvalues seen in standard principal components (e.g., Craddock, & Flood, 1969; Lautenschlager, 1989; Mandel, 1972). Moreover, the general purpose, to quantify changes in the slope of a log-eigenvalue plot, should apply to standard principal components analysis. Hence, simulation studies of the type reported in this dissertation may be useful for extending the eigenvalue ratio test for use with standard principal components analysis.

Recent work in psychology has focused on "parallel analysis", that is, retaining in a principal components analysis those factors whose eigenvalues exceed the eigenvalues obtained from an analysis of random data. Although tables of eigenvalues for random data have been published, to my knowledge there has been no systematic investigation of the sensitivity or specificity of parallel analysis, nor of whether factors that are judged to be statistically

significant also tend to be well-identified, that is, show good metric recovery. The simulations used to study the eigenvalue ratio test in this dissertation, could be extended to parallel analysis as well.

### Similarity Judgments

Experiment 1 was an attempt to collect similarity judgments containing systematic asymmetries, by selecting objects that belonged to the same category, but differed in typicality. Higher similarity judgments have been reported by Rosch (1975a) and by Tversky (1977) when a lower-typicality exemplar is compared to a higher-typicality exemplar. The goal was to obtain asymmetric similarity judgments that could be examined for a multidimensional structure using CAA and NSKMSD. However, no asymmetries were found.

Because asymmetric similarity judgments have been reported previously by Rosch (1975a) and by Tversky (1977), including in a large-scale analysis of previously published similarity matrices (Tversky & Hutchinson, 1986), the absence of asymmetries in the current data might be a Type II error. The studies conducted for this dissertation were able to detect asymmetries with 90% power ( $\beta = .10$ ) down to a correlation ratio of .047 for the vegetable category, and .057 for the tools. This is relatively sensitive, but smaller effects do occur in psychological research (Cohen, 1977).

Several aspects of the results, however, give the absence of asymmetries some significance. First, the subjects whose data were retained for analysis showed high within-subject reliability in their similarity judgments. Moreover, the average judgments of these

subjects showed a high between-subjects reliability. For the vegetable category, the average similarity judgments had a reliability of .942, and the average typicality judgments had a reliability of .964. For the tools, the reliability of the average similarity judgments was .919, and the reliability of the average typicality judgments was .928. Thus, the subjects were able to give reasonably precise estimates of similarity and typicality. The asymmetric portion of these estimates, however, shows no indication of containing true-score variance. The reliability of the average asymmetries is estimated at  $-.003$  for the vegetable category, and  $.000$  for the tools.

Second, subjects appear to have been influenced by typicality in the symmetric portion of their similarity judgments. In a principal components analysis of the symmetries, the one-factor solution for the tool similarity judgments correlated  $.800$  with the average typicality judgment. For the vegetable category, a two-factor solution was indicated. After varimax rotation, the first factor correlated  $.802$  with the average typicality judgment. Thus, typicality appears to have been the primary influence on the symmetric part of the similarity judgments, but, contrary to hypothesis, it generated no asymmetries.

Third, an "extreme groups" analysis, involving the directional similarity judgments between the most and least typical objects, did not support the hypothesis. The difference between the two directions of comparison was not significant for either the vegetable or the tool category. For the vegetables, even the direction of difference was contrary to the hypothesis: higher similarity judgments were obtained when the most typical exemplar was compared to the least typical.

The most obvious weakness in the current study is the high proportion of subjects who appear to have responded randomly. Much of the random responding seems attributable to the study's conditions: subjects were run as a group and were free to leave as soon as they had recorded their judgments. The study was run on a Friday, spring afternoon, nearly the last day of the semester. I suspect that subjects did not have a clear sense of the importance of their similarity judgments, and that the study lost a competition with the weather.

Random responding was readily identified by checking the consistency of each subject's ratings across ten repeated pairs, and the unreliable subjects were excluded. However, the attrition entails a loss of power that may be quite unfortunate, given the presumably small effect size of asymmetries. Of course attrition can also produce an unrepresentative sample. However, because the subject sign-up process itself yields unrepresentative samples, and because no individual differences have been proposed for the asymmetries, this does not seem likely to be a problem. Certainly I cannot identify a reason to expect that the lost subjects, had they given reliable judgments, would have generated asymmetries.

Some of the attrition may be due to the group nature of the study. Were the study hypotheses to be followed up, it would be preferable to run subjects individually. Computer-presentation of the stimuli might also lend a greater credibility of the research to the subjects. The number of trials (240 similarity judgments and 20 typicality judgments) may have been fatiguing, and a reduction to, say 8 stimuli for each category, for a total of 160 similarity and 16 typicality judgments, may give better results.

If a follow-up study were to be conducted, the judgments should probably be collected as ratings. Although Hellman and Zwislocki (1961) have argued that absolute magnitude estimations are less likely to contain response set variance, this seems more than outweighed by the subjects' unfamiliarity with the task. Indeed, many of the magnitude estimation subjects whose judgments were reliable seem to have disregarded the instructions altogether and used a rating scale. For the detection of asymmetries, it would be crucial to reduce error variance as far as possible. This does not seem well-served by attempting to train subjects in an unfamiliar response mode.

However, although improvements can be made, I would tend to recommend against following up the hypotheses by collecting more similarity ratings. As noted, there was no trace of asymmetry in the current, reasonably sensitive experiment. Moreover, CAA and NSKMDS "take the asymmetries at face value", that is, they attempt to represent all of the asymmetries in the matrix. Therefore if the asymmetries contained a significant amount of error, the CAA and NSKMDS solutions, and particularly multidimensional solutions, could be quite misleading. Tests of a priori hypotheses are safer in this regard, in that they focus on a single source of variance, and include a check on the likely replicability of the finding. Further research on asymmetric similarity could perhaps best be conducted instead by examining the data sets identified by Tversky and Hutchinson (1986) as containing asymmetries.

### Typicality

Previous research had suggested that ratings of the complexity and meaningfulness of geometric figures could be analyzed into multiple dimensions using Q-factor analysis



(Borkum, 1989). Hence, it seemed plausible that typicality would also be multidimensional, and that these dimensions might account for multidimensional asymmetries. The typicality judgments collected for this dissertation, however, appear to have a strongly one-dimensional structure. Because there is no compelling theoretical reason, to my knowledge, to expect typicality to be multidimensional, this research direction may not be useful for further studies.

#### Random number generator

The initial problems seen with the random number generator are worth considering in more detail. The generator was published in peer-reviewed journal (Wichmann, & Hill, 1982). It was the subject of some later commentary (McLeod, 1985; Wichmann, & Hill, 1984; Zeisel, 1986), in which no difficulties of the type encountered here were reported. In this dissertation, a series of numbers produced by the generator showed no significant autocorrelations to suggest a periodicity. Nonetheless, in the context of the simulation program, the generator was found to be producing numbers that correlated across trials. The trials were separated by a slightly variable interval of approximately 1.8 seconds. The periodicity appears to have been controlled by the first two digits of one of the three six-digit seeds.

In this dissertation the concern was simply to locate the periodicity well enough to eliminate it from the simulations. However, the Wichmann and Hill algorithm is presumably in widespread use, as it has been adopted as the random number generator in SYSTAT (Wilkinson, Hill, Welna, & Birkenbeuel, 1992). Moreover, in Monte Carlo studies of

statistics, the randomness of the generated numbers is often taken for granted. Hence, the cause of the difficulties may be important beyond the current work.

The Wichmann and Hill algorithm is a variant of the multiplicative congruential generator developed by Lehmer in 1949 (Knuth, 1981). Generators of this type produce a sequence of numbers. Each element in the sequence is the modulus, or remainder, left after the preceding element is multiplied by one constant, "a", and then divided by another constant, "m":

$$X_{i+1} = \text{mod} [(a \cdot X_i) / m] .$$

The output of the generator on one iteration is its input on the next. The sequence necessarily repeats itself, with a period less than or equal to m. However within the period, for carefully chosen values of a and m, the successive numbers can be quite random (Knuth, 1981).

Moreover, Wichmann and Hill's method contains a refinement: The algorithm is based on three simultaneously running, iterative congruential generators. The random number returned on each trial is derived from the sum of the three generator outputs. Wichmann and Hill note that the algorithm's results pass a number of statistical tests for randomness.

The difficulties encountered in the simulation program almost certainly pertain not to the generator itself, but to how it was used. Many of the seeds used here were outside the proper values for the generator. Moreover, the generator is designed to run iteratively, starting with a single, arbitrary set of seeds: The output from successive iterations should then be random with respect to the earlier iterations. In the simulation program, however, the generator was called repeatedly, with new seeds on each call. Hence, instead of using the generator's output across iterations, we were using its output across different seeds.

Multiplicative congruential generators have not been tested for this type of operation.

Wichmann and Hill note that the seeds, once set "should not be changed other than by calls of the algorithm" (p. 189). Certainly, they did not intend for each random number to be generated by a separate call.

Since the studies reported in this dissertation were conducted, the simulation program has been modified. The seeds to the Wichmann and Hill generator are set to proper values using a data statement at the beginning of the program, and are not reset while the program is running. A second random number generator has been added, from Knuth (1981). The second generator is of a subtractive rather than a multiplicative congruential type. Thus it can provide some check on simulation results obtained with the multiplicative generator, and vice versa.

Press, Teukolsky, Vetterling, and Flannery (1992) caution that random number generators should be tested for the application in which they are being used. Hence, the investigations reported in Chapter 8, to check for and correct violations of randomness, will need to be retained in the modified simulation program.

## REFERENCES

- Anderson, N. H. (1970). Functional measurement and psychophysical judgment. Psychological Review, *77*, 153-170.
- Anderson, N. H. (1974). Algebraic models in perception. In E. C. Carterette & M. P. Friedman(Eds.), Handbook of perception. Vol. II: Psychophysical judgment and measurement. NY: Academic Press.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. British Journal of Psychology. Statistics Section, *3*, 77-85.
- Basilevsky, A. (1983). Applied matrix algebra in the statistical sciences. NY: North-Holland.
- Battig, W. R., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monograph, *80*, (3, Pt. 2).
- Beasley, J.D., & Springer, S.G. (1977). Algorithm AS 111. The percentage points of the normal distribution. Applied Statistics, *26*, 118-121.
- Birkhoff, G., & Mac Lane, S. (1953). A survey of modern algebra (rev. ed.). NY: The Macmillan Company.
- Bishop, Y. M., Fienberg, S. E., & Holland, P. W. (1975). Discrete multivariate analysis: Theory and practice. Cambridge, MA: The MIT Press.
- Borg, I., & Lingoes, J. (1987). Multidimensional similarity structure analysis. NY: Springer-Verlag.
- Borkum, J. (1989, March). Multidimensional unfolding and the recovery of multiple dimensions from a single rating scale. Paper presented at the Eastern Psychological Association, Boston, MA.
- Borkum, J. (1991a, June). Asymmetries in a proximity matrix of phoneme identification. Paper presented at the Maine Psychological Association, Lewiston, Maine.
- Borkum, J. (1991b, August). Analyzing the asymmetries in proximity matrices. Poster presented at the American Psychological Association, Division of Evaluation, Statistics, and Measurement, San Francisco, California.
- Borkum, J. (1992, June). Analyzing the asymmetries in similarity ratings: Two proposed inferential techniques. Paper presented at the Maine Psychological Association, Bangor, Maine.

- Borkum, J. (1993, August). Determining the number of principal components: A dual-criterion approach. Poster presented at the American Psychological Association, Division of Evaluation, Statistics, and Measurement, Toronto, Canada.
- Borkum, J. (1995, August). Nonmetric analysis of skew-symmetry using alternating least squares. Poster presented at the American Psychological Association, Division of Evaluation, Statistics, and Measurement, New York City.
- Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.
- Chino, N. (1978). A graphical technique for representing the asymmetric relationship between N objects. Behaviormetrika, 5, 23-40.
- Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. Psychological Bulletin, 103, 276-279.
- Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev. ed.) NY: Academic Press.
- Collins, A. A., & Gescheider, G. A. (1989). The measurement of loudness in individual children and adults by absolute magnitude estimation and cross-modality matching. Journal of the Acoustical Society of America, 85, 2012-2021.
- Constantine, A. G., & Gower, J. C. (1978). Graphical representations of asymmetric matrices. Applied Statistics, 27, 297-304.
- Coombs, C. H. (1964). A theory of data. NY: Wiley.
- Craddock, J. M., & Flood, C. R. (1969). Eigenvectors for representing the 500 mb. geopotential surface over the Northern Hemisphere. Quarterly Journal of the Royal Meteorological Society, 95, 576-593.
- Craig, C. C. (1936). On the frequency function of  $xy$ . Annals of mathematical statistics, 7, 1-15.
- Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. Psychometrika, 1, 211-218.
- Everett, J. E., & Pecotich, A. (1991). A combined loglinear/MDS model for mapping journals by citation analysis. Journal of the American Society for Information Science, 42, 405-413.

- Farmer, S. A. (1971). An investigation into the results of principal component analysis of data derived from random numbers. Statistician, 20, 63-72.
- Feldt, L. S. (1965). The approximate sampling distribution of Kuder-Richardson Reliability Coefficient Twenty. Psychometrika, 30, 357-370.
- Fréchet, M. (1906). Sur quelques points du calcul fonctionnel. Rendiconti del Circolo Matematico di Palermo, 22, 1-74.
- Gescheider, G. A., & Hughson, B. A. (1991). Stimulus context and absolute magnitude estimation: A study of individual differences. Perception and Psychophysics, 50, 45-57.
- Gower, J. C. (1977). The analysis of asymmetry and orthogonality. In J. Barra et al. (Eds.), Recent developments in statistics. Amsterdam: North-Holland.
- Guttman, L. (1954). A new approach to factor analysis: The radex. In: P. F. Lazarsfeld (Ed.), Mathematical thinking in the social sciences. Glencoe, IL: Free Press.
- Guttman, L. (1966). Order analysis of correlation matrices. In R. B. Cattell (Ed.), Handbook of multivariate experimental psychology. Chicago: Rand McNally and Co.
- Ham, H.P., & Parsons, O.A. (1997). Organization of psychological functions in alcoholics and nonalcoholics: A test of the compensatory hypothesis. Journal of Studies on Alcohol, 58, 67-74.
- Harris, W. P. (1957). A revised law of comparative judgment. Psychometrika, 22, 189-198.
- Harshman, R. A. (1978, August). Models for analysis of asymmetrical relationships among N objects or stimuli. Paper presented at the First Joint Meeting of the Psychometric Society and The Society for Mathematical Psychology, McMaster University, Hamilton, Ontario.
- Harshman, R. A. (1981). DEDICOM multidimensional analysis of skew-symmetric data. Part I: Theory (Unpublished technical memorandum). Murray Hill, NJ: Bell Laboratories.
- Hartley, H. O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. Biometrika, 37, 308-312.
- Hellman, R. P., & Zwislocki, J. J. (1961). Some factors affecting the estimation of loudness. Journal of the Acoustical Society of America, 33, 687-694.

- Helm, C. E. (1960). A successive intervals analysis of color differences: A technical report. (Office of Naval Research Contract Nonr 1858-(15), Project Designation NR 150-088). Princeton, NJ: Educational Testing Service.
- Helm, C. E., Messick, S. J., & Tucker, L. R. (1961). Psychological models for relating discrimination and magnitude estimation scales. Psychological Review, *68*, 167-177.
- Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, *30*, 178-185.
- Horst, P. (1965). Factor analysis of data matrices. NY: Holt, Rinehart, & Winston.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. I. Journal of Educational Psychology, *24*, 417-441.
- Hubert, L.J., & Baker, F.B. (1979). Evaluating the symmetry of a proximity matrix. Quality and Quantity, *13*, 77-84.
- Indow, T. (1988). Multidimensional studies of Munsell Color Solid. Psychological Review, *95*, 456-470.
- International Monetary Fund. (1994). Direction of trade statistics yearbook. Washington DC: Author.
- Jackson, D. (1924). The trigonometry of correlation. American Mathematical Monthly, 275-280.
- Johnson, J. H., Butcher, J. N., Null, C., & Johnson, K. N. (1984). Replicated item level factor analysis of the full MMPI. Journal of Personality and Social Psychology, *47*, 105-114.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. Econometrica, *47*, 263-291.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, *20*, 141-151.
- Keppel, G. (1982). Design and analysis: A researcher's handbook (2nd. ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Knuth, D. E. (1981). The art of computer programming. Vol. 2: Seminumerical algorithms (2nd ed.). Reading, MA: Addison-Wesley.

- 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.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. Cognitive Psychology, 11, 346-374.
- Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 29, 1-27.
- Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. Psychometrika, 29, 115-129.
- Kruskal, J. B., & Shepard, R. N. (1974). A nonmetric variety of linear factor analysis. Psychometrika, 39, 123-157.
- Kruskal, J. B., & Wish, M. Multidimensional scaling. Beverly Hills, CA: Sage Publications.
- Kučera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.
- Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. Multivariate Behavioral Research, 24, 365-395.
- Leon, S. J. (1990). Linear algebra with applications (3rd ed.). NY: MacMillan Publishing Co.
- Lingoes, J. C. (1972). A general survey of the Guttman-Lingoes nonmetric program series. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), Multidimensional scaling: Theory and applications in the behavioral sciences. NY: Seminar Press.
- Lorr, M. (1983). Cluster analysis for social scientists. San Francisco, CA: Jossey-Bass.
- Luce, R. D. (1989). Autobiography. In: G. Lindzey (Ed.) The history of psychology in autobiography, Vol. 8. Palo Alto, CA: Stanford University Press.
- Mandel, J. (1972). Principal components, analysis of variance and data structure. Statistica Neerlandica, 26, 119-129.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1, 30-46.



- McLeod, A.I. (1985). A remark on Algorithm AS 183. An efficient and portable pseudo-random number generator. Applied Statistics, 34, 198-200.
- Meeker, W.Q. Jr., Cornwell, L.W., & Aroian, L.A. (1981). The product of two normally distributed random variables. In: W.J. Kennedy, R.E. Odeh, and J.M. Davenport (Eds.) Selected tables in mathematical statistics (Vol. 7). Providence, RI: American Mathematical Society.
- Mellinger, J. J. (1956). An investigation of psychological color space. Unpublished doctoral dissertation, University of Chicago.
- Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America, 27, 338-352.
- Mosteller, F. (1951). Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. Psychometrika, 16, 3-9.
- Münsterberg, H. (1894). Studies from the Harvard Psychological Laboratory. Psychological Review, 1, 34-60.
- Needham, J. G. (1934). The time-error in comparison judgments. Psychological Bulletin, 31, 229-243.
- Nosofsky, R. M. (1991). Stimulus bias, asymmetric similarity, and classification. Cognitive Psychology, 23, 94-140.
- Pearson, K. (1905). Mathematical contributions to the theory of evolution. XIV. On the general theory of skew correlation and non-linear regression. Draper's Company Research Memoirs, Biometric Series II.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). Numerical recipes in FORTRAN: The art of scientific computing(2nd ed.). NY: Cambridge University Press.
- Ramsay, J. O. (1969). Some statistical considerations in multidimensional scaling. Psychometrika, 34, 167-182.
- Rosch, E. (1975a) Cognitive reference points. Cognitive Psychology, 7, 532-547.
- Rosch, E. (1975b). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104, 192-233.

- Saffir, M. A. (1937). A comparative study of scales constructed by three psychophysical methods. Psychometrika, 2, 179-198.
- Saito, T. (1986). Multidimensional scaling to explore complex aspects in dissimilarity judgment. Behaviormetrika, 20, 35-62.
- Seber, G.A.F. (1984). Multivariate observations. NY: John Wiley and Sons.
- Shepard, R. N. (1958). Stimulus and response generalization: tests of a model relating generalization to distance in psychological space. Journal of Experimental Psychology, 55, 509-523.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. Psychometrika, 27, 125-140.
- Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. Psychometrika, 39, 373-421.
- Shepard, R. N. (1978). The circumplex and related topological manifolds in the study of perception. In S. Shye (Ed.), Theory construction and data analysis in the behavioral sciences (pp. 29-80). San Francisco: Jossey-Bass.
- Shepard, R. N. (1980). Multidimensional scaling, tree fitting, and clustering. Science, 210, 390-398.
- Singh, S., Woods, D., & Becker, G. (1972). Perceptual structure of 22 prevocalic English consonants. Journal of the Acoustical Society of America, 52, 1698-1713.
- Stevens, S. S. (1946). On the theory of scales of measurement. Science, 103, 677-680.
- Stevens, S. S. (1956). The direct estimation of sensory magnitude – loudness. American Journal of Psychology, 69, 1-25.
- Stevens, S. S. (1957). On the psychophysical law. Psychological Review, 64, 153-181.
- Stevens, S. S. (1960). Ratio scales, partition scales, and confusion scales. In: H. Gulliksen and S. Messick (Eds.), Psychological scaling: Theory and applications. NY: Wiley.
- Stevens, S. S., & Galanter, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 54, 377-411.
- Stewart, D.K., & Love, W.A. (1968). A general canonical correlation index. Psychological Bulletin, 70, 160-163.

- Takane, Y., Young, F.W., & de Leeuw, J. (1977). Nonmetric individual differences in multidimensional scaling: An alternating least square method with optimal scaling features. Psychometrika, 42, 7-67.
- Thurstone, L. L. (1927). A law of comparative judgment. Psychological Review, 34, 273-286.
- Thurstone, L. L. (1929). Fechner's Law and the method of equal-appearing intervals. Journal of Experimental Psychology, 12, 214-224.
- Tobler, W. R. (1976-77). Spatial interaction patterns. Journal of Environmental Systems, 6, 271-301.
- Tobler, W. R. (1979). Estimation of attractivities from interactions. Environment and Planning A, 11, 121-127.
- Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. Psychometrika, 17, 401-420.
- Torgerson, W. S. (1958). Theory and methods of scaling. NY: Wiley.
- Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.
- Tversky, A., & Hutchinson, W. (1986). Nearest neighbor analysis of psychological spaces. Psychological Review, 93, 3-223.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 321-327.
- Wichmann, B.A., & Hill, I.D. (1982). An efficient and portable pseudo-random number generator. Applied Statistics, 31, 188-190. (Algorithm AS 183)
- Wichmann, B.A., & Hill, I.D. (1984). Correction to Algorithm AS 183. An efficient and portable pseudo-random number generator. Applied Statistics, 33, 123.
- Wickelgren, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. Journal of the Acoustical Society of America, 39, 388-398.
- Wilkinson, L, Hill, M. A., Welna, J. P., & Birkenbeuel, G. K. (1992) Systat. Statistics. Chicago: SPSS Inc.
- Wilson, W. A. (1931). On quasi-metric spaces. American Journal of Mathematics, 53, 675-684.

- Wish, M. (1967). A model for the perception of Morse code like signals. Human Factors, 9, 529-539.
- Young, F. W. (1970). Nonmetric multidimensional scaling: Recovery of metric information. Psychometrika, 35, 455-473.
- Young, F. W. (1975, August). An asymmetric Euclidian model for multi-process asymmetric data. Paper presented at the U.S. - Japan Seminar on the theory, methods and applications of multidimensional scaling and related techniques, UCSD, La Jolla, CA.
- Young, F. W. (1984). Scaling. Annual Review of Psychology, 35, 55-81.
- Zeisel, H. (1986). A remark on Algorithm 183. An efficient and portable pseudo-random number generator. Applied Statistics, 35, 89.
- Zielman, B., & Heiser, W. J. (1993). Analysis of asymmetry by a slide-vector. Psychometrika, 58, 101-114.
- Zwislocki, J. J. (1983). Group and individual relations between sensation magnitude and their numerical estimates. Perception and Psychophysics, 33, 460-468.

## Appendix I

## Orthogonality of Symmetric and Skew-Symmetric Parts

The symmetric and skew-symmetric parts of an asymmetric matrix are orthogonal.

This can be seen by computing their sum of cross-products:

$$\begin{aligned}
& \sum_{i,j} (\bar{x}_{ij \text{ symm}} - \bar{x}_{..})(\bar{x}_{ij \text{ asym}} - \bar{x}_{..}) \\
&= \sum_{i,j} (\bar{x}_{ij \text{ symm}} \bar{x}_{ij \text{ asym}} - \bar{x}_{ij \text{ symm}} \bar{x}_{..} - \bar{x}_{ij \text{ asym}} \bar{x}_{..} + \bar{x}_{..}^2) \\
&= \sum_{i,j} [.5(\bar{x}_{ij} + \bar{x}_{ji}) * .5(\bar{x}_{ij} - \bar{x}_{ji}) - .5(\bar{x}_{ij} + \bar{x}_{ji})\bar{x}_{..} \\
&\quad - .5(\bar{x}_{ij} + \bar{x}_{ji})\bar{x}_{..} + \bar{x}_{..}^2] \\
&= \sum_{i,j} [.25(\bar{x}_{ij}^2 - \bar{x}_{ji}^2) - .5(\bar{x}_{ij}\bar{x}_{..} + \bar{x}_{ji}\bar{x}_{..}) \\
&\quad - .5(\bar{x}_{ij}\bar{x}_{..} - \bar{x}_{ji}\bar{x}_{..}) + \bar{x}_{..}^2] \\
&= \sum_{i,j} [.25(\bar{x}_{ij}^2 - \bar{x}_{ji}^2) - \bar{x}_{ij}\bar{x}_{..} + \bar{x}_{..}^2] \\
&= .25\sum_{i,j}\bar{x}_{ij}^2 - .25\sum_{i,j}\bar{x}_{ji}^2 - \bar{x}_{..}\sum_{i,j}\bar{x}_{ij} + k^2\bar{x}_{..}^2 \\
&= .25\sum_{i,j}\bar{x}_{ij}^2 - .25\sum_{i,j}\bar{x}_{ji}^2 - (\sum_{i,j}\bar{x}_{ij}/k^2)\sum_{i,j}\bar{x}_{ij} + k^2(\sum_{i,j}\bar{x}_{ij}/k^2)^2 \\
&= -(\sum_{i,j}\bar{x}_{ij})^2/k^2 + (\sum_{i,j}\bar{x}_{ij})^2/k^2 \\
&= 0 .
\end{aligned}$$

## Appendix II

## Canonical Analysis of Asymmetry: Computational Features

To understand Gower's approach, recall first that prototypical similarity data is symmetric ( $d_{ij} = d_{ji}$ ) and that prototypical dominance data is skew-symmetric ( $d_{ij} = -d_{ji}$ ). Generally in MDS a square data matrix is forced to symmetry by replacing  $d_{ij}$  and  $d_{ji}$  with their arithmetic mean:  $.5*(d_{ij} + d_{ji})$ . When performed over the entire matrix this is equivalent to averaging the matrix with its transpose:

$$S = .5*(D + D')$$

where  $D$  is the original, square, asymmetric data matrix, and  $S$  is the derived, symmetrized matrix. That  $S$  is symmetric is easily verified, for it meets the definitional requirement of being equal to its own transpose:

$$\begin{aligned} S' &= .5*(D + D')' \\ &= .5*(D'' + D') \\ &= .5*(D + D') \\ &= S \end{aligned}$$

$S$ , however, is not  $D$ , and information is lost when we symmetrize. This information is the difference between  $S$  and  $D$ :

$$\begin{aligned}
D - S &= D - .5*(D + D') \\
&= D - .5*D - .5*D' \\
&= .5*D - .5*D' \\
&= .5*(D - D') \\
&= A
\end{aligned}$$

A, the lost information, is skew-symmetric, for it meets the definitional requirement of being equal to the negative of its transpose:

$$\begin{aligned}
A' &= .5*(D - D')' \\
&= .5*(D' - D'') \\
&= .5*(D' - D) \\
&= -.5*(D - D') \\
&= -A
\end{aligned}$$

Thus, any matrix can be partitioned into a purely symmetric and a purely asymmetric (i.e., skew-symmetric) part (Basilevsky, 1983). The symmetric part is easily analyzed of course, with principal components analysis, factor analysis, multidimensional scaling, and the like. Hence, Gower focusses on the asymmetries.

A symmetric matrix can be analyzed into very simple, vector components, with optimal least squares properties. The symmetric matrix is analyzed into a set of simple matrices, each of which is the outer product of a vector with itself:

$$\begin{aligned}
S &= S_1 + S_2 + S_3 \dots + S_k \\
&= s_1s_1' + s_2s_2' + s_3s_3' \dots + s_ks_k'
\end{aligned}$$

The vectors  $\mathbf{s}_i$  are the eigenvectors of the matrix. In the outer product of a vector with itself, each row is proportional to all the other rows, and each column is proportional to all the other columns. Thus, each outer product matrix contains only 1 independent row and 1 independent column. The matrix is of rank one. Its information content can be displayed along a (one-dimensional) line, corresponding to the principal component vector. Hence, the symmetric matrix is analyzed into simple and easily described components.

Each outer product of a vector with itself is necessarily symmetric:

$$\begin{aligned}\mathbf{s}_i' &= (\mathbf{s}_i \mathbf{s}_i')' \\ &= \mathbf{s}_i'' \mathbf{s}_i' \\ &= \mathbf{s}_i \mathbf{s}_i' \\ &= \mathbf{s}_i\end{aligned}$$

This makes the outer product of a vector with itself unsuitable for representing skew symmetric data. We are forced to a more complex representation. The outer product of one vector with another is not necessarily symmetric:

$$\begin{aligned}(\mathbf{s}_i \mathbf{s}_j')' &= \mathbf{s}_j'' \mathbf{s}_i' \\ &= \mathbf{s}_j \mathbf{s}_i'\end{aligned}$$

which need not equal  $\mathbf{s}_i \mathbf{s}_j'$ . The outer product is still of rank one, however, because each row is proportional to the other rows, and each column is proportional to the other columns. If  $\mathbf{s}_i$  and  $\mathbf{s}_j$  have the same number of elements  $\mathbf{s}_i \mathbf{s}_j'$  will be square and its transpose can be subtracted from it:

$$\mathbf{s}_i \mathbf{s}_j' - \mathbf{s}_j \mathbf{s}_i' \tag{15}$$

The difference matrix is skew-symmetric, for it equals the negative of its transpose:



$$\begin{aligned}
& (s_j s_j' - s_i s_i')' \\
&= s_j'' s_i' - s_i'' s_j' \\
&= s_j s_i' - s_i s_j' \\
&= -(s_i s_j' - s_j s_i')
\end{aligned}$$

The difference matrix is rank two (must be represented as a plane rather than a vector), and is the simplest unit into which skew-symmetric matrices can be analyzed without losing the skew-symmetric property. Gower (1977) points out that the singular value decomposition of a skew-symmetric matrix is into difference matrices such as these. Hence, using the singular value decomposition to give a least squares approximation of the asymmetries, leads to rank two difference matrices. In matrix terms, the singular value decomposition of a skew-symmetric matrix, as for all matrices, is

$$S = U D V'$$

However, the vectors in  $V$  are the same as those in  $U$ , except for the order and signs within pairs of vectors. Therefore, a simpler form of the singular value decomposition is

$$= U D J U'$$

where  $J$  is a block diagonal skew-symmetric matrix.

Thus, while we analyze symmetric matrices into rank one components, or vectors, the fundamental unit of a skew-symmetric matrix is a rank two component, or plane.

A skew-symmetric component plane is not a metric space. This may be seen by considering how distance is defined in a vector space.

Distance is inherently bound up with the notion of vector length. The length of a vector is the distance from its end point to the origin. The distance between the endpoints of

2 vectors is the length of the vector joining the endpoints. Length, in turn, is defined as the inner product of a vector with itself.

The scalar product (sum of cross products) of two vectors is the most common example of an inner product. The scalar product of a vector with itself is the sum of squares of its elements.

$$\mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 + x_3^2 \dots$$

The square root of this is its Euclidean length (Pease, 1965, pp. 51-53).

Slightly more generally, we can modify the scalar product as  $\mathbf{x}'\mathbf{K}\mathbf{x}$ , where  $\mathbf{K}$  is a square matrix. When  $\mathbf{K} = \mathbf{I}$ , the identity matrix, the usual scalar product emerges as a special case. When  $\mathbf{K} \neq \mathbf{I}$ , the scalar product still equals the squared distance, but computed now using a different set of axes. When the columns of  $\mathbf{K}$  are orthogonal the new axes are orthogonal rotations of the original axes. When the columns are not orthogonal, an oblique rotation occurs. When the sum of squares of a column are greater (less) than one, the axis corresponding to the column has been expanded (contracted).

For  $\mathbf{x}'\mathbf{K}\mathbf{x}$  to meet the definitional requirements of an inner product  $\mathbf{K}$  must be symmetrical. When the vector spaces are derived from skew symmetric data, the asymmetries are defined as

$$\begin{aligned} & x_1y_2 - x_2y_1 \\ & = (y_1, y_2)' \mathbf{K} (x_1, x_2) \end{aligned}$$

where  $\mathbf{K}$  is an elementary skew-symmetric matrix:

$$\mathbf{K} = \begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}$$

Because  $\mathbf{K}$  is not symmetric, it cannot be used to define an inner product relation, and therefore does not indicate distance in the usual sense.

## Appendix III

## Geometric Interpretation of Skew-Symmetric Matrices

In the canonical analysis of asymmetry, and in its nonmetric extension, a skew-symmetric matrix is represented as the sum of elementary skew-symmetric matrices, each constructed from two dimensions:

$$\dot{x}_{ij} = a_{iI}a_{jII} - a_{jI}a_{iII} \quad (16)$$

where  $a_{iI}$  is the coordinate of object  $i$  on dimension I,  $a_{iII}$  is the coordinate of  $i$  on dimension II,  $a_{jI}$  is the coordinate of  $j$  on dimension I, and  $a_{jII}$  is the coordinate of  $j$  on II.

This elementary skew-symmetric relation between objects  $i$  and  $j$  can be represented geometrically in two-dimensional space as the signed area of the triangle whose vertices are object 1, object 2, and the origin. The area is signed. We label it as, say, positive if object  $i$  is less than 180 degrees from  $j$  in a clockwise sweep around the origin, and negative if object  $i$  is more than 180 degrees from  $j$  in a clockwise sweep.

There are two ways to see the basis for this interpretation. The first involves recognizing that the relation given by Equation 16 above is the determinant of a 2 by 2 matrix whose rows represent the objects and whose columns represent the two dimensions. One would then follow standard proofs (e.g., Birkhoff & Mac Lane, 1953) that the determinant is equal to the area of a parallelogram, two of whose sides are the lines drawn between the origin and the two object points. The area of a triangle formed by the two object points and the origin is simply one-half the area of the trapezoid.

Here I will simply illustrate the correspondence with reference to Figure 26, below.

In Figure 26, we are seeking to demonstrate that the area of the inner triangle equals one half the quantity given in Equation 16. We can obtain the area of the inner triangle by subtracting the areas of triangles 1, 2, and 3 from the area of the circumscribing rectangle:

$$\text{Area of rectangle} = a_I a_{II}$$

$$\text{Area of triangle 1} = \frac{1}{2} a_{II} a_{III}$$

$$\text{Area of triangle 2} = \frac{1}{2} a_I a_{III}$$

$$\begin{aligned} \text{Area of triangle 3} &= \frac{1}{2} (a_I - a_{II}) (a_{II} - a_{III}) \\ &= \frac{1}{2} (a_I a_{II} - a_I a_{III} - a_{II} a_{II} + a_{II} a_{III}) \end{aligned}$$

$$\text{Area of central triangle} = \text{Area of rectangle}$$

$$- \text{Area of triangle 1}$$

$$- \text{Area of triangle 2}$$

$$- \text{Area of triangle 3}$$

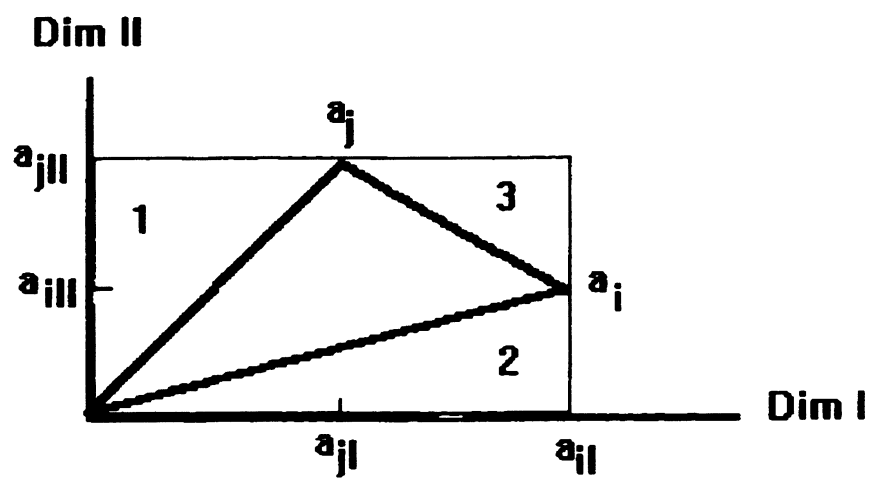
$$= a_I a_{II} - \frac{1}{2} (a_{II} a_{III} + a_I a_{III} + a_I a_{II} - a_I a_{III} - a_{II} a_{II} + a_{II} a_{III})$$

$$= a_I a_{II} - \frac{1}{2} (a_I a_{III} + a_{II} a_{II})$$

$$= \frac{1}{2} (a_I a_{II} - a_{II} a_{II})$$

as in Equation 16, above.

FIGURE 26



## Appendix IV

## Rescaling Factors For Wishart, Chi-Square, and Asymmetric Error

Wishart Error

Let  $t_i, t_j$  be elements in a vector used to create the true symmetric part of a matrix.

Under the factor model, cell  $x_{ij}$  in the true matrix is constructed as

$$x_{ij} = t_i * t_j,$$

i.e., as the cross product of the appropriate scores on the true vector.

Now let  $e_i, e_j$  be random normal deviates in an error vector that is added to the true vector to corrupt it. If we create an error perturbed matrix from the corrupted vector, using the factor model, cell  $c_{ij}$  will be given as

$$c_{ij} = (t_i + e_i) * (t_j + e_j),$$

that is, as the cross product of the appropriate scores on the corrupted vector.

The amount of error perturbation for cell  $x_{ij}$ , that is, the squared residual at  $i,j$  thus equals

$$\begin{aligned} \text{Residual}^2 &= [c_{ij} - x_{ij}]^2 \\ &= [(t_i + e_i) * (t_j + e_j) - t_i * t_j]^2 \end{aligned}$$

Across all cells, the total perturbation, or sum of squared residuals equals

$$SS_{\text{resid}} = \sum_i \sum_j [(t_i + e_i) * (t_j + e_j) - t_i * t_j]^2$$

We want to multiply all elements  $e$  in the error vector by a rescaling factor,  $K$ , so that  $SS_{\text{resid}}$  equals the desired amount of error perturbation, or DAP:

$$\text{DAP} = [(t_i + Ke_i) * (t_j + Ke_j) - t_i * t_j]^2$$

$$= \Sigma_i \Sigma_j [t_{ij} + K t_{ej} + K t_{ei} + K^2 e_{ej} - t_{ij}]^2$$

$t_{ij}$  and  $-t_{ij}$  cancel, leaving

$$= \Sigma_i \Sigma_j [K t_{ej} + K t_{ei} + K^2 e_{ej}]^2$$

After squaring we have

$$\begin{aligned} &= \Sigma_i \Sigma_j [K^2 t_{ej}^2 + 2K^2 t_{ej} t_{ei} + 2K^3 t_{ej} e_{ej} + K^2 t_{ei}^2 \\ &\quad + 2K^3 t_{ei} e_{ej} + K^4 e_{ej}^2] \\ &= K^2 (\Sigma_i t_{ij}^2) (\Sigma_j e_j^2) + 2K^2 (\Sigma_i t_{ei}) (\Sigma_j t_{ej}) + 2K^3 (\Sigma_i t_{ei}) (\Sigma_j e_j^2) \\ &\quad + K^2 (\Sigma_j t_{ij}^2) (\Sigma_i e_i^2) + 2K^3 (\Sigma_j t_{ej}) (\Sigma_i e_i^2) + K^4 (\Sigma_i e_i^2) (\Sigma_j e_j^2) \end{aligned}$$

Some condensing is possible because

$$\Sigma_i t_{ij}^2 = \Sigma_j t_{ij}^2$$

$$\Sigma_i e_i^2 = \Sigma_j e_j^2$$

and

$$\Sigma_i t_{ei} = \Sigma_j t_{ej}$$

This yields

$$DAP = K^4 (\Sigma_i e_i^2) (\Sigma_j e_j^2) + 4K^3 (\Sigma_i t_{ei}) (\Sigma_j e_j^2) + 2K^2 [(\Sigma_i t_{ei}) (\Sigma_j t_{ej}) + (\Sigma_i t_{ij}^2) (\Sigma_j e_j^2)]$$

or

$$\begin{aligned} 0 = & K^4 (\Sigma_i e_i^2) (\Sigma_j e_j^2) + 4K^3 (\Sigma_i t_{ei}) (\Sigma_j e_j^2) + 2K^2 [(\Sigma_i t_{ei}) (\Sigma_j t_{ej}) \\ & + (\Sigma_i t_{ij}^2) (\Sigma_j e_j^2)] - DAP \end{aligned} \tag{17}$$

in which  $K$  is the unknown rescaling coefficient. The rescaling coefficient is obtained by solving Equation 17 for a real root.



### Chi-Square Error

We could add random normal deviates to a "true" vector, and then use this now-corrupted "true" vector to generate the cells of a symmetric matrix under a squared Euclidean distance model:

$$x_{ij} = [(t_i + e_i) - (t_j + e_j)]^2$$

If the vector had not been corrupted by the addition of error, the squared Euclidean distances would have been

$$x_{ij} = [t_i - t_j]^2$$

and therefore the sum of squared residuals is given by

$$x_{ij} = \{[(t_i + e_i) - (t_j + e_j)]^2 - [t_i - t_j]^2\}^2.$$

To rescale the error vectors so as to obtain the desired amount of residual sum of squares

$$\begin{aligned} \text{DAP} &= \sum_i \sum_j \{[(t_i + Ke_i) - (t_j + Ke_j)]^2 - [t_i - t_j]^2\}^2 \\ &= \sum_i \sum_j \{[t_i + Ke_i - t_j - Ke_j]^2 - [t_i - t_j]^2\}^2 \\ &= \sum_i \sum_j \{[t_i^2 + 2Kt_i e_i - 2t_i t_j - 2Kt_i e_j + K^2 e_i^2 - 2Ke_i t_j \\ &\quad - 2K^2 e_i e_j + t_j^2 + 2Kt_j e_j + K^2 e_j^2 - t_i^2 + 2t_i t_j - t_j^2]\}^2 \\ &= \{2Kt_i e_i - 2Kt_i e_j + K^2 e_i^2 - 2Ke_i t_j - 2K^2 e_i e_j + 2Kt_j e_j \\ &\quad + K^2 e_j^2\}^2 \end{aligned}$$

$$\begin{aligned}
0 = & K^4(n\Sigma_i e_i^4 - 4\Sigma_i e_i^3 \Sigma_j e_j + 6\Sigma_i e_i^2 \Sigma_j e_j^2 - 4\Sigma_j e_j^3 \Sigma_i e_i + n\Sigma_j e_j^4) \\
& + K^3(4n\Sigma_i t_i e_i^3 - 12\Sigma_i t_i e_i^2 \Sigma_j e_j + 12\Sigma_i t_i e_i \Sigma_j e_j^2 - 4\Sigma_i t_i \Sigma_j e_j^3 - 4\Sigma_j t_j \Sigma_i e_i^3 \\
& + 12\Sigma_j t_j e_j \Sigma_i e_i^2 - 12\Sigma_i e_i \Sigma_j t_j e_j^2 + 4n\Sigma_j t_j e_j^3) \\
& + K^2(4n\Sigma_i t_i^2 e_i^2 - 8\Sigma_i t_i^2 e_i \Sigma_j e_j - 8\Sigma_i t_i e_i^2 \Sigma_j t_j + 16\Sigma_i t_i e_i \Sigma_j t_j e_j + 4\Sigma_i t_i^2 \Sigma_j e_j^2 \\
& - 8\Sigma_i t_i \Sigma_j t_j e_j^2 + 4\Sigma_i e_i^2 \Sigma_j t_j^2 - 8\Sigma_i e_i \Sigma_j t_j^2 e_j + 4n\Sigma_j t_j^2 e_j^2) - \text{DAP}
\end{aligned}$$

As in the case of Wishart error, this fourth degree polynomial in K can be solved for a real root, which would be the desired rescaling factor.

### Asymmetric Error

The entries in an elementary skew-symmetric matrix are constructed from two vectors:

$$x_{ij} = (t_i t_{jn} - t_j t_{in})$$

If we add an error vector to each of these true vectors, an error perturbed elementary skew-symmetric matrix can be formed:

$$x_{ij} = [(t_i + e_i)(t_{jn} + e_{jn}) - (t_j + e_j)(t_{in} + e_{in})]$$

The sum of squared differences, across cells, between the true and error perturbed matrices is given by

$$\begin{aligned}
\Sigma(\text{Resid})^2 &= \Sigma_i \Sigma_j [(t_i + e_i)(t_{jn} + e_{jn}) - (t_j + e_j)(t_{in} + e_{in}) - (t_i t_{jn} - t_j t_{in})]^2 \\
&= \Sigma_i \Sigma_j [t_i e_{jn} + t_{jn} e_i + e_i e_{jn} - t_j e_{in} - t_{in} e_j - e_{in} e_{jn}]^2
\end{aligned}$$

Because there are two error vectors,  $e_i$  and  $e_{in}$ , corresponding to the two true vectors, we will need two rescaling coefficients:

$$\text{DAP} = \Sigma_i \Sigma_j [t_i K_{in} e_{jn} + t_{jn} K_{ei} + K_i K_{in} e_i e_{jn} - K_i t_{in} e_j - K_{in} t_j e_{in} - K_i K_{in} e_{in} e_{jn}]^2$$

However, in the canonical analysis of skew-symmetry, the two true vectors defining an elementary skew-symmetric matrix have the same sum of squares. Therefore, we can simplify the present task by requiring that the two error vectors, after rescaling, have the same sums of squares. Thus we require that

$$\begin{aligned}
 \Sigma_i(K_{II}e_{III})^2 &= \Sigma_i(K_I e_{II})^2 \\
 K_{II}^2 \Sigma_i e_{III}^2 &= K_I^2 \Sigma_i e_{II}^2 \\
 K_{II}^2 &= K_I^2 * [\Sigma_i e_{II}^2 / \Sigma_i e_{III}^2] \\
 K_{II} &= K_I * [\text{SQRT}(\Sigma_i e_{II}^2 / \Sigma_i e_{III}^2)]
 \end{aligned} \tag{18}$$

We can then replace  $K_{II}$  with  $K_I R$ , where  $R$  equals  $\text{SQRT}(\Sigma_i e_{II}^2 / \Sigma_i e_{III}^2)$ :

$$\begin{aligned}
 \text{DAP} &= \Sigma_i \Sigma_j [t_{II} K_I R e_{III} + t_{III} K_I e_{II} + K_I^2 R e_{II} e_{III} - K_I t_{III} e_{II} \\
 &\quad - K_I R t_{II} e_{III} - K_I^2 R e_{III} e_{II}]^2 \\
 0 &= 2R^2 K_I^4 [\Sigma_i e_{II}^2 \Sigma_j e_{III}^2 - (n \Sigma_i e_{II} e_{III})^2] + 4K_I^3 [R^2 (\Sigma_i t_{II} e_{II}) (\Sigma_j e_{III}^2) \\
 &\quad - R^2 (\Sigma_i t_{II} e_{III}) (\Sigma_j e_{II} e_{III}) + R (\Sigma_j t_{III} e_{III}) (\Sigma_i e_{II}^2) - R (\Sigma_i e_{II} e_{III}) (\Sigma_j t_{II} e_{II}) \\
 &\quad + 2K_I^2 [R^2 (\Sigma_i t_{II}^2) (\Sigma_j e_{III}^2) + 2R (\Sigma_i t_{II} e_{II}) (\Sigma_j t_{III} e_{III}) - 2R (\Sigma_i t_{II} t_{III}) (\Sigma_j e_{II} e_{III}) \\
 &\quad - R^2 (\Sigma_i t_{II} e_{III}) (\Sigma_j t_{II} e_{III}) - (\Sigma_i t_{III} e_{II}) (\Sigma_j t_{II} e_{II}) + (\Sigma_i t_{III}^2) (\Sigma_j e_{II}^2)] - \text{DAP}
 \end{aligned} \tag{19}$$

All quantities in Equation 19 are known except for the rescaling coefficient for the first error vector,  $K_I$ , which can be determined with root-finding methods. Once  $K_I$  is established, we can use Equation 18 to find  $K_{II}$ . Source code in FORTRAN 77 for solving Equation 19 is given in Appendix 7.

## Appendix V

## Stimulus Words

CATEGORY	EXEMPLAR	TYPICALITY	
		RANK	MEAN
Carpenter's Tool	saw	1	1.04
	ruler	3	1.48
	nails	6	1.67
	sander	10	1.79
	chisel	16	2.26
	pliers	22	2.56
	wood	27.5	2.77
	hinge	34.5	3.12
	paintbrush	45	3.81
	cement	55	4.91
Vegetable	pea	1	1.07
	green beans	3	1.18
	broccoli	6	1.28
	brussels sprouts	10	1.72
	beets	15	2.08
	eggplant	21	2.38
	potato	29	2.89
	peppers	37	3.21
	avocado	44	3.62
	dandelion	54	5.20

Source: Rosch, E. (1975b).



```

DATA CT1/0/,INDX/0/,I/0/,J/0/,F/0/,H/0/,L/0/,M/0/,R/0/,A/0/,B/0/
DATA I2/0/,I3/0/,I4/0/,II/0/,JJ/0/,HH/0/,JB/0/,D/0/,E/0/,Q8/0/
DATA N/0/,IND2/0/
DATA ((DSQ(I,J),I=1,10),J=1,10)/100*0/
DATA ((SKDST(I,J),I=1,10),J=1,10)/100*0/
DATA ((EIGVEC(I,J),I=1,10),J=1,10)/100*0/
DATA ((DSKW(I,J),I=1,10),J=1,10)/100*0/
DATA (EIGVAL(I),I=1,10)/10*0/
DATA (ORD1(I),I=1,45)/45*0/
DATA (ORD2(I),I=1,45)/45*0/
DATA (ORD3(I),I=1,45)/45*0/
DATA (ORD4(I),I=1,45)/45*0/
DATA (ORD5(I),I=1,45)/45*0/
DATA (ORD6(I),I=1,45)/45*0/
DATA (ARDSP(I),I=1,100)/100*0/
DATA (ARDSKW(I),I=1,100)/100*0/
DATA (ARDSK1(I),I=1,100)/100*0/
DATA (ARDST(I),I=1,100)/100*0/
DATA (ARCOUN(I),I=1,100)/100*0/
DATA (ARINDX(I),I=1,100)/100*0/
DATA (ARIND2(I),I=1,100)/100*0/

```

C

```

NFACTR=2
LOCPAR = -LOCPAR
PS1B = 1E10
EVSUM2 = 1E10
SLOPE=0
INTERC=0
RSQ=0
SD=0
DO 801 I=1, 10
    SSDIM(I) = 0
801 CONTINUE
DO 1802 I=1, K
    DO 1803 J=1, K
        DSKW(I,J)=NDSKW(I,J)
1803 CONTINUE
1802 CONTINUE

```

C

KS=K\*K

C

```

C First be sure that an even number of factors is
C being extracted.

```

C

```

IF ((NFACTR - 2*INT(NFACTR/2)) .NE. 0 .AND. NFACTR .LT. K)
+ THEN
  NFACTR=NFACTR+1
ELSE IF ((NFACTR - 2*INT(NFACTR/2)) .NE. 0 .AND.
+ NFACTR .EQ. K) THEN
  NFACTR=NFACTR-1
END IF
C
DO 200 JB = 1, NFACTR/2
C
  S16 = 0
  DO 4190 I=1, K
    DO 4191 J=1, K
      S16 = S16 + DSKW(I,J)**2
4191   CONTINUE
4190   CONTINUE
  S16 = SQRT(S16)
  DO 4192 I=1, K
    DO 4193 J=1, K
      DSKW(I,J) = DSKW(I,J)/S16
4193   CONTINUE
4192   CONTINUE
  DO 111 I=1, K
    DO 112 J=1, K
      DSQ(I, J) = 0
112   CONTINUE
111   CONTINUE
C
C   The starting point of the analysis is the extraction
C   of the first two ratio-level factors (canonical analysis
C   of asymmetry) from the matrix being analyzed.
C
C   First square the skew-symmetric matrix:
C
C
C   CALL EIGSQR(K, DSKW, DSQ)
C
C   Then check for a null matrix and extract the principal
C   components
C
  SSTOTL = 0
  DO 1181 I=1, K
    DO 1182 J=1, K
      SSTOTL = SSTOTL + DSQ(I, J)**2

```

```

1182     CONTINUE
1181     CONTINUE
      IF (SSTOTL .LT. 1E-10) GOTO 200
C
      CALL JACOBI(K, DSQ, EIGVEC, EIGVAL)
      DO 887 J=1, NFACTR
          CONS = 1
          IF (EIGVAL(J) .LT. 0) THEN
              WRITE(UNIT=*,FMT=1251) EIGVAL(J)
              CONS = -1
              EIGVAL(J) = -EIGVAL(J)
          END IF
          EIGVAL(J) = SQRT(EIGVAL(J))
          EIGVAL(J) = SQRT(EIGVAL(J))
887     CONTINUE
1251  FORMAT(' EIGVAL < 0 = ',F9.4)
      DO 387 I=1, K
          DO 388 J=1, NFACTR
              EIGVEC(I, J) = EIGVEC(I, J)*EIGVAL(J)*CONS
388     CONTINUE
387     CONTINUE
C
C
C     Now use the first two metric dimensions to start the
C     nonmetric analysis.
C
C
      SUM7 = 0
      DO 1 A=1, K
          DO 2 B=1, K
              SKDST(A, B) = EIGVEC(A, 1)*EIGVEC(B, 2)
+           - EIGVEC(B, 1)*EIGVEC(A, 2)
              SUM7 = SUM7 + DSKW(A, B)**2
2         CONTINUE
1         CONTINUE
      S1=0
      DO 616 I=1, K
          DO 617 J=1, K
              S1 = S1+SKDST(I,J)*NDSKW(I,J)
617     CONTINUE
616     CONTINUE
      SUM7 = SQRT(SUM7)
      IF (S1 .LT. 0) THEN
C         DO 817! A=1, 2

```



```

        DO 619 I=1,K
            EIGVEC(I,1) = -EIGVEC(I,1)
C         B = 3-A
C         TMP=EIGVEC(I,A)
C         EIGVEC(I,A) = EIGVEC(I,B)
C         EIGVEC(I,B) = TMP
619      CONTINUE
C8171    CONTINUE
8811     DO 1019 A=1, K
            DO 1020 B=1, K
                SKDST(A, B) = EIGVEC(A, 1)*EIGVEC(B, 2)
+         - EIGVEC(B, 1)*EIGVEC(A, 2)
1020     CONTINUE
1019     CONTINUE
        END IF
C
C
C       To compute the disparities, we first need to put
C       the skew-symmetries data, and the distances, into
C       arrays, for later sorting.
C
C
        DO 113 L=1, K
            DO 114 M=1, K
                Q2 = (L-1)*K + M
                ARDSKW(Q2) = DSKW(L, M)
                IF (JB .EQ. 1) ARDSK1(Q2) = ARDSKW(Q2)
                ARDST(Q2) = SKDST(L, M)
                ARINDX(Q2) = Q2
114      CONTINUE
113      CONTINUE
        S16=0
        S17=0
        DO 4183 I=1, KS
            S16 = S16 + ARDSKW(I)**2
            S17 = S17 + ARDST(I)**2
4183     CONTINUE
        S16 = SQRT(S16)
        S17 = SQRT(S17)
        DO 4184 I=1, KS
            ARDSKW(I) = ARDSKW(I)/S16
            ARDST(I) = ARDST(I)/S17
4184     CONTINUE
        S16=0

```

```

S17=0
C
C
C   Then the skew-symmetric data, in array form, are
C   sorted. The distances are put in the same order as
C   the data, as is an index array. The index array is
C   used later to put the updated distances into the same
C   order as the original skew-symmetric data.
C   The sorting algorithm used here is a simple (and
C   rather inefficient) insertion sort, modified slightly
C   to improve its efficiency a bit.
C
C
DO 115 II=2, KS
  IF (ARDSKW(II) .GE. ARDSKW(II-1)) GOTO 115
  DO 17 JJ=1, II-1
    IF (ARDSKW(II) .GE. ARDSKW(JJ)) GOTO 17
    DSKW1 = ARDSKW(II)
    DST = ARDST(II)
    INDX = ARINDX(II)
    DO 18 HH=II-1, JJ, -1
      C
      ARDSKW(HH+1) = ARDSKW(HH)
      C
      ARDST(HH+1) = ARDST(HH)
      C
      ARINDX(HH+1) = ARINDX(HH)
      C
    18   CONTINUE
    ARDSKW(JJ) = DSKW1
    ARDST(JJ) = DST
    ARINDX(JJ) = INDX
  17   CONTINUE
115  CONTINUE
C
DO 5050 I=1, KS
  IND2 = ARINDX(I)
  ARIND2(IND2) = I
5050 CONTINUE
C
C
C   Now from the distances, which are in the same
C   order as the original skew-symmetric input

```

```

C      data, the disparities are computed, using Kruskals
C      block averaging algorithm.
C
C      Iteratively we (1) compute the disparities;
C      (2) compute the Pseudo-Stress 1 measure;
C      (3) change the coordinates; (4) update the
C      distances; (5) standardize the disparities.
C
C
      J8 = 0
31     IF (J8 .GE. 50) GOTO 80
C
      J8 = J8 + 1
C
C      CALL LINREG(KS,ARDST,ARDSKW,SLOPE,INTERC,RSQ)
C      SLOPE=0
      DO 8340 I=1, KS
        SLOPE=SLOPE + ARDST(I)*ARDSKW(I)
8340  CONTINUE
      IF (SLOPE .LT. 0) THEN
        WRITE(UNIT=*,FMT=8339)
        DO 8009 I=1, KS
          ARDST(I) = -ARDST(I)
8009  CONTINUE
        END IF
8339  FORMAT('  SLOPE < 0, BOUNCE ')
C
      DO 24 I2=1, KS
        ARDSP(I2) = ARDST(I2)
24    CONTINUE
C
C
      DO 22 I=2, KS
        DO 27 J=1, KS
          ARCOUN(J) = 0
27    CONTINUE
        CT1 = 1
        SUM1 = ARDSP(I)
        IF (ARDSP(I)+1E-12 .GE. ARDSP(I-1)) THEN
          ARDSP(I)=ARDSP(I)
          GOTO 22
        END IF
        DO 23 I3=I-1, 1, -1

```

```

                IF (ARDSP(I) .LT. ARDSP(I3)) THEN
                    CT1 = CT1 + 1
                    SUM1 = SUM1 + ARDSP(I3)
                    ARCOUN(I3) = 1
                    ARCOUN(1+I3) = 1
                END IF
23      CONTINUE
          AVGDST = SUM1/CT1
          DO 26 I4=1, I
            IF (ARCOUN(I4) .EQ. 1) ARDSP(I4) = AVGDST
26      CONTINUE
22      CONTINUE
C
C
C
C
C      Now that we have both distances and disparities
C      we can compute "Pseudo-Stress 1", which is a goodness
C      of fit measure for skew-symmetric data that is similar
C      to Kruskal's Stress formula 1 for symmetric data.
C
C
C
C
          PS1A = 0
          DO 127 I=1, KS
            PS1A = PS1A + (ARDSP(I) - ARDST(I))**2
127      CONTINUE
C      WRITE(UNIT=*,FMT=8117)PS1A
C8117  FORMAT('STRESS=',F9.4)
          IF (PS1A/ABS(S1) .GT. .5) THEN
            DO 8810 I=1, K
              EIGVEC(I,1) = -EIGVEC(I,1)
              EIGVEC(I,2) = -EIGVEC(I,2)
8810      CONTINUE
          WRITE(UNIT=*,FMT=8813)
8813  FORMAT(' BOUNCE X 2 ')
          GOTO 8811
          END IF
          IMPROV = PS1B - PS1A
C      IF (IMPROV .LT. -.10) GOTO 80
          PS1B = PS1A
C
C
C      Update the coordinates.
C

```

C

```

DO 130 A=1, K
  DO 32 B=1, 2
    C = 3 - B
    SUM2 = 0
    SUM3 = 0
    SUM4 = 0
    DO 34 D = 1, K
      IF (D .EQ. A) GOTO 34
      N = (A-1)*K + D
      M = ARIND2(N)
      SUM2 = SUM2 + ARDSP(M)*EIGVEC(D, C)*
+      EXP(LOCPAR*ABS(ARDSP(M)))
      SUM3 = SUM3 + EIGVEC(D, B)*EIGVEC(D, C)*
+      EXP(LOCPAR*ABS(ARDSP(M)))
      SUM4 = SUM4 + EIGVEC(D, C)*EIGVEC(D, C)*
+      EXP(LOCPAR*ABS(ARDSP(M)))
34    CONTINUE
      IF (B .EQ. 2) SUM2 = -SUM2
      IF (SLOPE .LT. 0) SUM2 = -SUM2
      EIGVEC(A, B) = (SUM3*EIGVEC(A, C)+SUM2)/SUM4
32    CONTINUE
130  CONTINUE

```

C

C After the coordinates for a point have been updated,  
C recompute the distances between it and the other  
C points.  
C

```

DO 4130 A=1,K
  DO 36 E=1, K
    M = (A-1)*K + E
    N = ARIND2(M)
    ARDST(N) = EIGVEC(A, 1)*EIGVEC(E, 2) -
+    EIGVEC(E, 1)*EIGVEC(A, 2)
36    CONTINUE
4130  CONTINUE

```

C

C

C

C

C

C

C

C

After the disparities have been recomputed,  
they are restandardized, to prevent a trivial  
solution in the "MDS", and so that we can  
disregard the denominator of the Pseudo Stress 1  
measure.

```

C
  SUM5 = 0
  SUM15 = 0
  DO 38 F=1, KS
    SUM5 = SUM5 + ARDST(F)*ARDST(F)
    SUM15 = SUM15 + ARDSP(F)*ARDSP(F)
38  CONTINUE
  SUM5 = SQRT(SUM5)
  SD = SQRT(SUM15)
  DO 42 H=1, KS
    ARDST(H) = ARDST(H)/SUM5
    ARDSP(H) = ARDSP(H)/SD
42  CONTINUE
C
C
  IF (IMPROV .GT. 0 .AND. IMPROV .LT. .00005) GOTO 80
  IF (PS1B .EQ. 0) GOTO 80
  GOTO 31
C
80  DO 2350 R = 1, KS
    T = INT((R-1)/K) + 1
    U = R - (T-1)*K
    M = ARIND2(R)
    DSKW(T, U) = ARDSKW(M)
    SKDST(T, U) = ARDST(M)
    IF (R .LT. 46) ORD3(R)=R
    IF (R .LT. 46) ORD5(R)=R
2350 CONTINUE
C
  JJ=0
  DO 2031 I=1, K
    DO 2032 J=I+1, K
      JJ = JJ+1
      ORD1(JJ)=DSKW(I,J)
      ORD2(JJ)=SKDST(I,J)
2032 CONTINUE
2031 CONTINUE
C
DO 5115 II=2, 45
  IF (ORD1(II) .GE. ORD1(II-1)) GOTO 5115
  DO 5117 JJ=1, II-1
    IF (ORD1(II) .GE. ORD1(JJ)) GOTO 5117
    DSKW1 = ORD1(II)
    INDX = ORD3(II)

```

```

DO 5118 HH=II-1, JJ, -1
C
ORD1(HH+1) = ORD1(HH)
C
ORD3(HH+1) = ORD3(HH)
C
5118 CONTINUE
ORD1(JJ) = DSKW1
ORD3(JJ) = INDX
5117 CONTINUE
5115 CONTINUE
AVGDST=0
CT1=0
C
DO 5119 I=1, 45
IND2 = ORD3(I)
ORD4(IND2) = I
5119 CONTINUE
C
CT1=0
AVGDST=0
DO 5142 I=1, 44
IF (ORD1(I) .NE. ORD1(I+1) .AND. CT1 .EQ. 0) GOTO 5142
IF (ORD1(I) .EQ. ORD1(I+1) .AND. CT1 .EQ. 0) THEN
AVGDST=AVGDST+2*I+1
CT1=CT1+2
ELSE IF (ORD1(I) .EQ. ORD1(I+1) .AND. CT1 .GT. 0) THEN
AVGDST=AVGDST+I+1
CT1=CT1+1
END IF
IF (ORD1(I) .NE. ORD1(I+1) .AND. CT1 .GT. 0) THEN
DO 5143 J=I-CT1+1,I
TMP = ORD3(J)
ORD4(TMP) = AVGDST/CT1
5143 CONTINUE
CT1 = 0
AVGDST = 0
END IF
5142 CONTINUE
C
DO 6115 II=2, 45
IF (ORD2(II) .GE. ORD2(II-1)) GOTO 6115
DO 6117 JJ=1, II-1
IF (ORD2(II) .GE. ORD2(JJ)) GOTO 6117

```

```

        DSKW1 = ORD2(I)
        INDX = ORD5(I)
        DO 6118 HH=I-1, JJ, -1
C
            ORD2(HH+1) = ORD2(HH)
C
            ORD5(HH+1) = ORD5(HH)
C
6118     CONTINUE
            ORD2(JJ) = DSKW1
            ORD5(JJ) = INDX
6117     CONTINUE
6115     CONTINUE
C
        DO 6119 I=1, 45
            IND2 = ORD5(I)
            ORD6(IND2) = I
6119     CONTINUE
C
        CT1=0
        AVGDST=0
        DO 6142 I=1, 44
            IF (ORD2(I) .NE. ORD2(I+1) .AND. CT1 .EQ. 0) GOTO 6142
            IF (ORD2(I) .EQ. ORD2(I+1) .AND. CT1 .EQ. 0) THEN
                AVGDST=AVGDST+2*I+1
                CT1=CT1+2
            ELSE IF (ORD2(I) .EQ. ORD2(I+1) .AND. CT1 .GT. 0) THEN
                AVGDST=AVGDST+I+1
                CT1=CT1+1
            END IF
            IF (ORD2(I) .NE. ORD2(I+1) .AND. CT1 .GT. 0) THEN
                DO 6143 J=I-CT1+1,I
                    TMP = ORD5(J)
                    ORD6(TMP) = AVGDST/CT1
6143     CONTINUE
                    CT1 = 0
                    AVGDST = 0
                END IF
6142     CONTINUE
C
        SUMDSQ=0
        DO 6121 I=1, 45
            SUMDSQ = SUMDSQ + (ORD4(I)-ORD6(I))**2
6121     CONTINUE

```



```

SQSPMN = (1 - (6*SUMDSQ/91080))**2
C
SUM8 = 0
SUM9 = 0
SUM10 = 0
SUM11 = 0
DO 1478 I=1, K
  DO 1481 J=1, K
    M = (I-1)*K + J
    SUM10 = SUM10 + ARDSP(M)
    SUM11 = SUM11 + SKDST(I, J)
1481  CONTINUE
1478  CONTINUE
SUM10 = SUM10/KS
SUM11 = SUM11/KS
DO 1493 I=1, K
  DO 1494 J=1, K
    M = (I-1)*K + J
    ARDSP(M) = ARDSP(M) - SUM10
    SKDST(I, J) = SKDST(I, J) - SUM11
    SUM8 = SUM8 + ARDSP(M)**2
    SUM9 = SUM9 + SKDST(I, J)**2
1494  CONTINUE
1493  CONTINUE
SUM8 = SQRT(SUM8)
SUM9 = SQRT(SUM9)
DO 2481 I=1, K
  DO 2478 J=1, K
    M = (I-1)*K + J
    IF (SUM8 .NE. 0) THEN
      ARDSP(M) = SUM7*ARDSP(M)/SUM8
    END IF
    IF (SUM9 .NE. 0) THEN
      SKDST(I, J) = SUM7*SKDST(I, J)/SUM9
    END IF
2478  CONTINUE
2481  CONTINUE
C
C   To extract the next two dimensions we first residualize
C   the data matrix. The disparities have just been restandarized
C   to the same units as the data, so we need only subtract the
C   disparities from the data.
C
C

```

```

DO 110 Q=1, K
  DO 411 I=1, K
    M = (Q-1)*K + I
    DSKW(Q, I) = DSKW(Q, I) - ARDSP(M)
411  CONTINUE
110  CONTINUE
C
DO 121 J=1, 2
  DO 122 I=1, K
    J2 = 2*JB + J - 2
    FNCOOR(I, J2) = EIGVEC(I, J)
    SSDIM(J2) = SSDIM(J2) + EIGVEC(I, J)*EIGVEC(I, J)
122  CONTINUE
121  CONTINUE
200  CONTINUE
C
DO 300 I=1, KS
  ARDST(I) = 0
300  CONTINUE
C
C   Predict the skew-symmetric data using all of the
C   dimensions that were extracted, and store the
C   predicted values in ARDST. This is a different
C   use of the array than above. We overwrite the
C   array and use it for a different purpose to reduce
C   memory overhead a bit.
C
DO 301 L=1, K
  DO 302 J=1, K
    Q8 = (L-1)*K + J
    DO 303 I=1, INT(NFACTR/2)
      M2 = I*2 - 1
      M3 = I*2
      ARDST(Q8) = ARDST(Q8) +
+      FNCOOR(L, M2)*FNCOOR(J, M3) -
+      FNCOOR(J, M2)*FNCOOR(L, M3)
303  CONTINUE
302  CONTINUE
301  CONTINUE
C
C
RETURN
END
C

```

## Appendix VII

Sample FORTRAN 77 Source Code  
For Bracketing and Bisection of Asymmetric Error Vectors

```

C
C
C   And then we create and orthogonalize the asymmetry error vectors.
C
C
C
1236 IF (SKWERR .EQ. 'Y' .AND. AELEVEL .NE. 0) THEN
      DO 1237 I=1, NDIMA
        DO 1238 J=1, TNVAR
          CALL GETTIM(IHR,IMIN,ISEC,I100TH)
          IX=ILR*10000+IMIN*100+I100TH
          C      IX=I100TH*10000+IMIN*100+I100TH
          IY=ISEC*10000+IMIN*100+I100TH
          IZ=IMIN*10000+ISEC*100*I100TH
          IF (I .NE. 1 .OR. J .NE. 1) THEN
            IX=IX*J
            IY=IY*AJNT(RANDM*100)
            IZ=IZ*AJNT(RANDM*100)
          END IF
          CALL RNDNML(IX, IY, IZ, RANDM2)
          ERRSKW(J,I)=RANDM2
1238   CONTINUE
1237 CONTINUE
C
C
C   As in previous cases, we use Gramm-Schmidt othogonalization to
C   make the error vectors orthonormal.
C
C
C
      DO 1380 I=2, NDIMS
        SUM1=0
        IF (I .EQ. 2) THEN
          DO 1282 C=1, TNVAR
            SUM1 = SUM1+ERRSKW(C, 1)**2
1282   CONTINUE
          DO 1283 C=1, TNVAR
            IF (SUM1 .NE. 0) THEN
              ERRSKW(C,1) = ERRSKW(C,1)/SQRT(SUM1)

```

```

                END IF
1283      CONTINUE
      END IF
C
      DO 1284 C=1, TNVAR
        INNVEC(C)=0
1284      CONTINUE
C
      DO 1285 L=I-1, 1, -1
        DO 1286 C=1, TNVAR
          INNVEC(L)=INNVEC(L)+ERRSKW(C,I)*ERRSKW(C,L)
1286      CONTINUE
1285      CONTINUE
      DO 1287 L=I-1, 1, -1
        DO 1288 C=1, TNVAR
          ERRSKW(C,I) = ERRSKW(C,I)-INNVEC(L)*ERRSKW(C,L)
1288      CONTINUE
1287      CONTINUE
      SUM1=0
      DO 1289 C=1, TNVAR
        SUM1=SUM1+ERRSKW(C,I)**2
1289      CONTINUE
      DO 1290 C=1, TNVAR
        IF (SUM1 .NE. 0) THEN
          ERRSKW(C,I) = ERRSKW(C,I)/SQRT(SUM1)
        END IF
1290      CONTINUE
1380      CONTINUE
      SUM1=0
      DO 2173 I=1, TNVAR
        SUM1=SUM1+ERRSKW(I,1)**2
2173      CONTINUE
      IF (SUM1 .NE. 0) THEN
        DO 2174 I=1, TNVAR
          ERRSKW(I,1)=ERRSKW(I,1)/SQRT(SUM1)
2174      CONTINUE
      END IF

```

C  
C

C

C And then we rescale the skew-symmetry error vectors so that  
C they introduce the desired amount of perturbation. This amount  
C is operationalized as the sum of squared differences between

C the entries in the skew-symmetric matrix that would be formed  
 C from the ASYMAT vectors before the error vectors are added to  
 C them, and the corresponding entries in the skew-symmetric matrix  
 C that would be formed after the error has been added to the ASYMAT  
 C vectors. As is generally true in this simulation program, the rescaling  
 C factor is given as a root of a quartic equation. (The nature of  
 C the quartic equation is different for each type of vector –  
 C skew-symmetric true scores, skew-symmetric error, Wishart error,  
 C and noncentral chi-square error.)  
 C

```

DO 1239 P=1, NDIMA/2
  SUMSK4 = 0
  SUMSK5 = 0
  SUMSK6 = 0
  SUMSK7 = 0
  SUMSK8 = 0
  SUMSK9 = 0
  SUMS10 = 0
  SUMS11 = 0
  SUMS12 = 0
  SUMS13 = 0
DO 1240 I=1, TNVAR
  SUMSK4 = SUMSK4 + ERRSKW(I,2*P)**2
  SUMSK5 = SUMSK5 + ERRSKW(I,2*P-1)**2
  SUMSK6 = SUMSK6 + ERRSKW(I,2*P)*ERRSKW(I,2*P-1)
  SUMSK7 = SUMSK7 + ASYMAT(I,2*P-1)*ERRSKW(I,2*P-1)
  SUMSK8 = SUMSK8 + ASYMAT(I,2*P)*ERRSKW(I,2*P)
  SUMSK9 = SUMSK9 + ASYMAT(I,2*P)*ERRSKW(I,2*P-1)
  SUMS10 = SUMS10 + ASYMAT(I,2*P-1)*ERRSKW(I,2*P)
  SUMS11 = SUMS11 + ASYMAT(I,2*P-1)*ASYMAT(I,2*P-1)
  SUMS12 = SUMS12 + ASYMAT(I,2*P)*ASYMAT(I,2*P)
  SUMS13 = SUMS13 + ASYMAT(I,2*P-1)*ASYMAT(I,2*P)
1240 CONTINUE
IF (SUMSK5 .NE. 0) THEN
  SKWRT3 = SUMSK4/SUMSK5
ELSE
  SKWRT3 = 0
END IF
COEF1=2*SKWRT3*(SUMSK4*SUMSK5 - SUMSK6*SUMSK6)
COEF2=4*(SQRT(SKWRT3)*SUMSK7*SUMSK4
+   +SKWRT3*SUMSK8*SUMSK5
+   -SKWRT3*SUMSK9*SUMSK6-SQRT(SKWRT3)*SUMS10*SUMSK6)
COEF3=2*SUMS11*SUMSK4+2*SKWRT3*SUMS12*SUMSK5
+   +4*SQRT(SKWRT3)*SUMSK7*SUMSK8-
```

```

+      *SQRT(SKWRT3)*SUMS13*SUMSK6
+      -2*SUMS10*SUMS10-2*SKWRT3*SUMSK9*SUMSK9
COEF4=AELEVEL*CONSK2(2*P)
X2 = 0
X1 = X2
FX1 = -1
FX2=COEF1*(X2**4) + COEF2*(X2**3) + COEF3*(X2**2) - COEF4
DO 1241 I=1, 300
  IF (ABS(FX2) .LT. 1E-3) THEN
    ROOT1=X2
    GOTO 1241
  ELSE IF (FX2*FX1 .LT. 0) THEN
    GOTO 1241
  ELSE
    X1 = X2
    X2 = X2 + .05
    FX2=COEF1*(X2**4)+COEF2*(X2**3)+COEF3*(X2**2)-COEF4
  END IF
1241  CONTINUE

```

C  
C  
C  
C  
C  
C

Now, for the asymmetry error, the root is bracketed between X1 and X2. We use the method of bisection to converge on it.

```

DO 1242 I=1, 200
  X3=(X1+X2)/2.
  FX1=COEF1*(X1**4)+COEF2*(X1**3)+COEF3*(X1**2)-COEF4
  FX3=COEF1*(X3**4)+COEF2*(X3**3)+COEF3*(X3**2)-COEF4
  IF (ABS(FX3).LT..001) THEN
    ROOT1=X3
  ELSE IF (FX3*FX1.LT.0)THEN
    X2=X3
  ELSE IF (FX3*FX1.GT.0) THEN
    X1=X3
  END IF
1242  CONTINUE

```

C

ROOT1 is probably the desired root, or rescaling factor, but it could be that the program failed to find the root, in which case it is still at its default value, 1. To check for this, compute the value of F(X) at the putative root. The value should be close to zero. If it is not, skip the trial and add 1 to the tally of number of skipped trials.

```
C
  FX1=COEF1*(ROOT1**4) + COEF2*(ROOT1**3) +
+ COEF3*(ROOT1**2) - COEF4
  IF (ABS(FX1).GT..001) THEN
    SKIPS=SKIPS+1
    WRITE (UNIT=*, FMT=9110) SKIPS, FX1, AELEVEL
    GOTO 1236
  END IF
9110  FORMAT('— AE ROOT PROBLEM; TIME = ',I6,' CONVERGE = ',F9.5,
+ ' AELEVEL = ',F9.5)
C
C  If the root is okay, go ahead and rescale the error vector.
C
  ROOT2 = SQRT(SKWRT3)*ROOT1
  DO 1243 I=1, TNVAR
    ERRSKW(I,2*P) = ERRSKW(I,2*P)*ROOT1
    ERRSKW(I,2*P-1) = ERRSKW(I,2*P-1)*ROOT2
1243  CONTINUE
1239  CONTINUE
C
  END IF
```