

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

124
6059

MAPPING THE MIND WITH BROKEN THEODOLITES:
Contributions to Multidimensional Scaling methodology,
with special application to Triadic data,
and the Sorting and Hierarchical Sorting Methods.

A dissertation submitted in
partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in Human Development Studies
at Massey University

by

D. L. Bimler

Massey University

May 1995

Massey University Library

Thesis Copyright Form

Title of thesis: Mapping the Mind with Broken
Theodolites

Please delete section not applicable.

- (1) (a) I give permission for my thesis to be made available to readers in Massey University Library under conditions determined by the Librarian.
(b) I do not wish my thesis to be made available to readers without my written consent for months.
(2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.
(b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for months.
(3) (a) I agree that my thesis may be copied for Library use.
(b) I do not wish my thesis to be copied for Library use for months.

Signed [Signature]
Date 17.15/95

For Readers Only.

The copyright of this thesis belongs to the author. Readers must sign their name in the space below to show that they recognise this. They are asked to add their permanent address.

NAME and ADDRESS

DATE

ABSTRACT

This thesis focuses on the psychological applications of Multidimensional Scaling (MDS) theory and methodology. The results are investigated of treating certain kinds of dissimilarity data (triadic data, to begin with) as comparisons between dissimilarities. This is a familiar idea but many of its implications are unexplored.

First, when data are available from more than one subject, it becomes possible to apply models of individual variation, in non-metric form. The Weighted Euclidean (or INDSCAL) model is the one used most often in this thesis, but the more general IDIOSCAL model is used to investigate individual differences in the case of colour vision. The data sets need not be complete. This is important when the size of the stimulus set means that there are too many comparisons for a single subject to respond to them all.

Second, Maximum Likelihood Estimation (MLE) becomes a straightforward generalisation of the standard hill-descent algorithm for minimising Stress.

Third, data collected with the sorting and hierarchical sorting methods can also be regarded as dissimilarity comparisons. The convenience of the sorting method and the lesser demands it makes on subjects when the number of stimuli is large have led to its widespread use, but the best way of analysing such data is uncertain. A 'reconstructed dyad' analysis is described and shown to be better than the usual co-occurrence approach in a number of examples in which evidence about the true perceptual or conceptual space is available independently.

Finally, when the data are interpreted as dissimilarity comparisons, an interactive method of scaling large stimulus sets becomes possible, in which one selectively acquires incomplete data, concentrating on comparisons which are expected to contain most information about the configuration. This approach has been applied twice, with the stimuli being simple synthesised sounds in one example, and complex natural sounds (canine heartbeats) in the second, working well in both cases. The potential applications for training people to recognise sounds are briefly considered. Some possibilities for future research arising from this work are described.

ACKNOWLEDGEMENTS

Some of the ideas involved in this thesis were presented to the IVth International Workshop on Infant Cry Research, Munich, July 1992. I am grateful to the organisers of the Workshop, whose financial help made it possible for me to be there; also to Prof. Dr. Dietmar Todt, of the Institut für Verhaltensbiologie, Freie Universität Berlin, who provided hospitality and accommodation in Berlin, and to Peep Stalmeier, of the Nijmegen Institute for Cognition and Information, who made me equally welcome in Nijmegen.

The goals of my research were largely theoretical, but without the generosity of a number of people who made available data they had collected, for me to re-analyse, in addition to my own data, it is doubtful whether I could have made what progress I have. The people to whom I am indebted include Peep Stalmeier (colour triads), Willem van der Kloot (power-strategy F-sorts), Rich Faldowski and Mark Hollins (texture F-sorts), Stephen Morley (pain F-sorts), and the ESRC Data Archive at the University of Essex. As well as information, they provided interested encouragement, as did other researchers who could not help with the former.

In particular, the data collected by Emma Barraclough of Massey University – partly for her own research, on colour vision and market research, but partly on my request – have been invaluable. I can only hope that I have been as much of assistance to her as she has to me. The informants (judges, subjects) themselves deserve to be thanked many times over.

The loan of a computer by Dr Roger Jones, of the Research School Social Sciences, Australian National University, meant that a two-month exile in Canberra was not wasted.

I should thank my principal supervisor, John Kirkland, who has shown unfailing faith that I knew what I was doing, in the face of overwhelming evidence to the contrary. There are numerous other people I should thank, and I can only hope that they will understand that the omissions are from the pressure of time, and not from lack of appreciation of their help. I must also acknowledge the cogent and constructive criticism of journal reviewers.

To Penny, who laughs at my jokes.

TABLE OF CONTENTS

Chapter 1	INTRODUCTION.....	1
Chapter 2	MULTIDIMENSIONAL SCALING	9
	Some Alternative Definitions	16
Chapter 3	TRIADS.....	22
	Variations on the Triadic Theme	33
	Element Numbers	37
	Triadic Experiments with Sounds (1)	38
	Triadic Experiments with Sounds (2)	42
	Intervals – Background	45
	Experiments with Sounds (3)	51
	Experiments with Sounds (4)	52
Chapter 4	GENERALISING THE MODEL.....	57
	Introduction	57
	Divers Weights	57
	Sparse Data sets	64
	The IDIOSCAL model	69
	The Points-of-View model	72
	A Residual-Forces Subject Scaling	77
	Divers Measures	79
	Implementation	82
Chapter 5	MAXIMUM LIKELIHOOD ESTIMATION	85
Chapter 6	FORMS OF RANKING DATA.....	97
	Ranked Dyads	97
	Conditional Ranking	102
	Preferences	113
	The Vector Model	116
	Curiosa	121
Chapter 7	SORTING AND HIERARCHICAL SORTING.....	125
	Sorting	125
	Kinship	144

	Facial Expressions – Lightfoot Series	146
	I-FEEL Faces	151
	Woodblocks	153
	F-Sorting Data	156
	Occupations	159
	Power Strategies	164
	Kinship	164
	Alternative Models for sorting data: Scales	166
	Alternative Models: Pick Any/ <i>N</i>	168
	Alternative models: Stochastic	169
	Alternative models: Low-resolution Proximities	170
	Variant vote-count distance estimates	174
	Individual differences	177
	Conclusions	180
Chapter 8	TRIADS REVISITED – ASPECTS OF DATA COLLECTION.....	182
	Balanced Incomplete Designs	182
	Unbalanced Designs	185
	Interactive Incomplete Designs	187
	PC-Triskele	194
	Tie-Handling in Triadic Data	195
	The Gestalt-fusion paradigm: a special case	199
	Mac-Triskele	201
	Canine Heartbeats	203
	Cross-modal supplementary data	207
	Implications for Training	209
Chapter 9	CONCLUSION.....	211
Appendix	WOOD-BLOCKS.....	217
	BIBLIOGRAPHY	

PARTIAL LIST OF SYMBOLS

N	number of elements (items, stimuli)
i, j, k, l	usually element indices, $1 \leq i, j, k, l \leq N$
(E_i, E_j) or (i, j)	dyad consisting of the i -th and j -th elements
N_d	number of possible dyads. $N_d = N(N-1) / 2$
$\langle E_i, E_j, E_k \rangle$ or $\langle i, j, k \rangle$	triad consisting of i -th, j -th and k -th elements
N_t	number of triads in a data set. For a complete set, $N_t = N(N-1)(N-2) / 6$
P	number of dimensions
p	usually a dimension index, $1 \leq p \leq P$
X	a reconstructed configuration of N points (a single point in $(N P)$ -dimensional configuration space)
x_i, x_{ip}	position of the i -th element in X
X	configuration considered as a $(N\text{-by-}P)$ matrix
Δ	matrix of dissimilarities
δ_{ij}	element of Δ . $\delta_{ij} > \delta_{kl}$ is equivalent to writing $(i, j) \gg (k, l)$.
D	matrix of reconstructed distances
d_{ij}	element of D
\mathfrak{D}	matrix of disparities ('pseudo-distances')
δ_{ij}	element of \mathfrak{D}
$\varepsilon_{ij,kl}$	distance comparison coefficient. $\varepsilon_{ij,kl} = 1$ if $\delta_{ij} > \delta_{kl}$, 0 otherwise.
M	the number of subjects
m	usually a subject index, $1 \leq m \leq M$
D_m, Δ_m	dissimilarities, reconstructed members for subject m (elements are $d_{m,ij}, \delta_{m,ij}$)
$\varepsilon_{m,ij,kl}$	distance comparison coefficient for subject m (comparing $\delta_{m,ij}$ and $\delta_{m,kl}$)
W	$(M\text{-by-}P)$ matrix of dimensional weights (salience)
w_{mp}	element of W
C_m	sorting data co-occurrence matrix for subject m
$c_{m,ij}$	element of C_m
E	matrix of averaged co-occurrences
e_{ij}	element of E . $e_{ij} = 1/M \sum c_{m,ij}$
f_{ij}	corrective force exerted between the i -th and j -th elements