MULTIDIMENSIONAL ANALYSIS OF SUCCESSIVE CATEGORIES (RATING) DATA BY DUAL SCALING

by

Maurice Jacob Odondi

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Education
University of Toronto

© Copyright by Maurice Jacob Odondi 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-28031-4
ABSTRACT

The collection of data by rating questionnaires is so commonplace in the behavioural sciences that the problem of quantifying the stimuli (items) rated has a long and varied history. Traditional Thurstonian scaling, based on the method of successive categories, provides a model for psychological judgement that estimates stimulus scale values and category boundaries. However, methods based on this model assume that all the raters are interchangeable and that all stimuli and categories are ordered on a unidimensional scale. In this thesis, the problem of finding multidimensional solutions for both stimuli and raters is addressed by a data analytic technique called dual scaling. A major purpose of the study has been to propose a way to deal with the difficulties caused by category boundary scale values that violate the assumed unidimensional ordering.

To illustrate the problem, consider rating students' performance using the categories “Poor,” “Mediocre,” “Average,” “Good,” “Excellent.” Such ratings exemplify “successive categories data,” the kind of data considered in this study. The category boundaries \( (\tau_k, k=1,2,3,4) \) are introduced as response variables such that \( \tau_1 \) is between “Poor” and “Mediocre,” \( \tau_2 \) between “Mediocre” and “Average,” \( \tau_3 \) between “Average” and “Good,” \( \tau_4 \) between “Good” and “Excellent,” and \( \tau_1 < \tau_2 < \tau_3 < \tau_4 \). If different raters (respondents) do not completely agree on how the students rank in their overall performance, then summarizing the ratings in a single
dimension of performance, with the category boundaries ordered as expected, would not be possible.

A two-step procedure is proposed in which the subjects (raters) are clustered into more homogeneous groups and a separate dual scaling solution for each cluster is found. The proposed method (DSMASC) is illustrated with artificial and real data. The major conclusion is that DSMASC provides multidimensional solutions for the stimuli with the category boundaries ordered as required. Limitations of the study and directions for future are discussed.
ACKNOWLEDGEMENTS

I wish to express my appreciation to many people who lent me the benefit of their thoughts, knowledge and expertise in the preparation of this thesis. I am grateful to the chair of my thesis committee, Professor Shizuhiko Nishisato for his academic leadership and guidance on methodological issues pertinent to this thesis. Also, I benefitted immensely from his encouragement, support and confidence in me as overcame many obstacles and frustrations.

For their reviews and comments leading up to the completion of this thesis, I want to thank the members of my thesis committee: Professor Ross Traub, for constantly checking the grammar and the logic of the thesis. Professor Phillip Nagy, for giving incisive criticisms and helpful suggestions, and Professor Merlin Wahlstrom, for his contribution. Of course, any errors or omissions that may remain in the thesis are the result of my own oversight.

My thanks also go to the following: Professor Richard Wolfe, the External Reader at my Departmental Oral and examiner at the Senate Oral, Professor Michael Friendly of York University, the External Examiner, who gave an insightful appraisal of the thesis and interesting suggestions for future research, Professor C.D. Shearing, the chair of the Senate Oral examination, and Professor Bruce Schneider, Extra-departmental examiner at the Senate Oral. I gratefully acknowledge the assistance with data from Dr. Howard Mzumara, a former student colleague and a friend of many years at the Ontario Institute for Studies in Education (OISE).

I will forever be indebted to my wife Nelly, who gave me unfailing support during my studies at the University of Toronto. I also owe my son Benjamin better times — he shared little quality time with me as I wrote the thesis. I owe much gratitude to my mother Beldina, who first taught me the virtues of hard work and waited patiently for so long to share with me the joy of completing this work. My heart also goes to my brothers, Ochieng, Otieno, Onyango, Uhuru, Odhiambo, Omollo, the late Mboya, Tabu, my sister Amondi, and my brother-in-law, Ajulu, all in Kenya. I thank each of you for keeping in touch and, at every opportunity, wishing me well in my studies.

I acknowledge assistance from the following:

- The Government of Kenya (Ministry of Education) who nominated me for a Canadian scholarship competition.
- The Canadian Commonwealth Scholarship and Fellowship Plan who funded a major part of my studies at the University of Toronto.
- Kenyatta University who granted me a leave to undertake my graduate studies at the University of Toronto.

Finally, to all individuals who, in one way or another, gave me any kind of support during my academic sojourn in Canada, I say, thank you.
DEDICATION

This is for my beloved son, Benjamin . . .
# THESIS TITLE ....................................................... 1

# ABSTRACT ......................................................... ii

# ACKNOWLEDGEMENTS ............................................ iii

# DEDICATION ....................................................... iv

# TABLE OF CONTENTS ............................................. vi

# LIST OF TABLES .................................................. xii

# LIST OF FIGURES ................................................ xiii

# LIST OF APPENDICES .......................................... xiv

## CHAPTER 1 ......................................................... 1

1.0 Introduction .................................................. 1

1.1 Dual Scaling .................................................. 3

1.2 Brief history of dual scaling ................................. 5

1.3 Forms of categorical data ..................................... 7

1.4 How dual scaling provides multidimensional analysis .... 8
   1.4.1 Solution extraction procedure .......................... 8
   1.4.2 One-dimensional analysis ............................... 8
   1.4.3 Multidimensional analysis ............................. 9

1.5 The parameter location vector .............................. 9

1.6 The research problem ....................................... 11
1.7 Significance of the study ................................................................. 14
1.8 Outline of the thesis ................................................................. 16

CHAPTER 2 .................................................................................. 17
DUAL SCALING OF SUCCESSIVE CATEGORIES DATA .......... 17
2.0 Introduction ................................................................. 17

2.1 Successive categories (rating) data ........................................... 17
  2.1.1 Conceptual form of data ......................................................... 19
  2.1.2 Assignment of numbers to categories ...................................... 21

2.2 Dual Scaling of successive categories (SC) data ....................... 21
  2.2.1 Definition of terms .......................................................... 21
  2.2.2 Steps in dual scaling of successive categories data .................. 22

2.3 Raw data transformations .......................................................... 23
  2.3.1 The rank-order procedure .................................................. 24
    2.3.1.1 Dominance matrix, $E$ .............................................. 25
  2.3.2 The design matrix procedure ............................................. 26
  2.3.3 Notation .................................................................. 29
  2.3.4 Derivation of the 'optimal' solution ...................................... 29
    2.3.4.1 Least-squares optimization to the optimal solution .......... 32
    2.3.4.2 Solution extraction procedure .................................. 33
    2.3.4.3 Deriving subject scores ........................................... 34

2.4 Summary ........................................................................ 34

CHAPTER 3 ................................................................................. 35
RELATED LITERATURE ................................................................. 35

3.1 The order constraint on variables ........................................... 35
  3.1.0 Introduction ................................................................. 35
3.1.1 The case of two category boundaries ........................................ 36
3.1.2 The case of four categories ............................................... 40

3.2 Points of view proposal ............................................................... 42
  3.2.1 Non-overlapping clustering (NOC) ....................................... 44
  3.2.2 Overlapping clustering (OC) ............................................... 44
  3.2.3 Stepwise non-overlapping clustering (SNOC) ......................... 45
  3.2.4 Discussion of OC, NOC and SNOC ..................................... 46
    3.2.4.1 The OC approach .................................................. 46
    3.2.4.2 The NOC approach ............................................... 46
    3.2.4.3 The SNOC approach ............................................. 46
    3.2.4.4 Summary of comments on OC, NOC, SNOC ...................... 47

3.3 The purpose of this investigation ............................................. 47

3.4 Summary ....................................................................................... 50

CHAPTER 4 .......................................................................................... 51
CLUSTER ANALYSIS METHODS .......................................................... 51

4.1 Introduction ................................................................................... 51

4.2 Definition of cluster analysis ..................................................... 52

4.3 Overview of clustering techniques ............................................. 53
  4.3.1 Optimization techniques ..................................................... 54
  4.3.2 Hierarchical techniques ...................................................... 54
  4.3.3 Clumping techniques ........................................................ 54

4.4 The input matrix for cluster analysis ........................................... 55

4.5 Choosing a clustering method ..................................................... 57
  4.5.0 Introduction ......................................................................... 57
  4.5.1 Guiding criteria for a cluster analysis method ....................... 58

4.6 Description of the QUICKCLUSTER computer program ............... 61
6.5.1 Introduction. .................................................. 115
6.5.2 Description of real data sets. ................................ 116
6.5.3 Analysis of Student Data ..................................... 117

CHAPTER 7 .............................................................. 129

SUMMARY, DISCUSSION AND CONCLUSIONS ..................... 129

7.0 Introduction ....................................................... 129

7.1 Summary of the study ............................................. 130

7.2 Discussion of findings ............................................ 131
   7.2.1 Order Constraint on Category Boundaries. ............... 131
   7.2.2 Orthogonality of solutions ................................. 135
   7.2.3 Total Variance to be Accounted For ..................... 136

7.3 Potential areas of application ................................. 137

7.4 Significance and Limitations of the study ..................... 142

7.5 Concluding remark .............................................. 144

References ........................................................... 146
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1-1.</td>
<td>20</td>
</tr>
<tr>
<td>A conceptual response-pattern form for successive categories (rating) data</td>
<td>20</td>
</tr>
<tr>
<td>Table 1-2.</td>
<td>20</td>
</tr>
<tr>
<td>A subjects-by-stimuli table of successive categories (rating) data</td>
<td>20</td>
</tr>
<tr>
<td>Table 3-1.</td>
<td>38</td>
</tr>
<tr>
<td>Six subjects, three stimuli, three categories</td>
<td>38</td>
</tr>
<tr>
<td>Table 3-2.</td>
<td>41</td>
</tr>
<tr>
<td>Inequality relations involving three category boundaries</td>
<td>41</td>
</tr>
<tr>
<td>Table 6-1.</td>
<td>89</td>
</tr>
<tr>
<td>Inter-point (parameter to subject) distances for the graph in Figure 4</td>
<td>89</td>
</tr>
<tr>
<td>Table 6-2.</td>
<td>91</td>
</tr>
<tr>
<td>Subjects-by-rank table for distances in Table 6-1</td>
<td>91</td>
</tr>
<tr>
<td>Table 6-3.</td>
<td>92</td>
</tr>
<tr>
<td>A 10-by-5 SC data generated from the graph in Figure 5</td>
<td>92</td>
</tr>
<tr>
<td>Table 6-41.</td>
<td>95</td>
</tr>
<tr>
<td>Dual scaling output for artificial data in Table 6-43.</td>
<td>95</td>
</tr>
<tr>
<td>Table 6-42.</td>
<td>97</td>
</tr>
<tr>
<td>A two-cluster representation of the dominance matrix in Table 6-41</td>
<td>97</td>
</tr>
<tr>
<td>Table 6-43.</td>
<td>98</td>
</tr>
<tr>
<td>Dual scaling results for the cluster, CL1.</td>
<td>98</td>
</tr>
<tr>
<td>Table 6-44.</td>
<td>99</td>
</tr>
<tr>
<td>Dual scaling results for the cluster, CL2.</td>
<td>99</td>
</tr>
<tr>
<td>Table 6-45.</td>
<td>100</td>
</tr>
<tr>
<td>Solutions (parameter location vectors) for CL1 and CL2.</td>
<td>100</td>
</tr>
<tr>
<td>Table 6-46.</td>
<td>102</td>
</tr>
<tr>
<td>Dual scaling solutions from the first residual matrix (RM)</td>
<td>102</td>
</tr>
<tr>
<td>Table 6-47.</td>
<td>103</td>
</tr>
<tr>
<td>Summary of solutions (parameter location vectors) for data in Table 6-3</td>
<td>103</td>
</tr>
</tbody>
</table>
Table 6-51. SDSS dual scaling results for a hypothetical three-category data set ................................................. 110

Table 6-52. DSMASC results for the dominance matrix in (B) of Table 6-51. ......................................................... 113

Table 6-61. First Solution for STUDENT1 data set. ...................... 121

Table 6-62. Solution from RM1 for STUDENT1 data set. .............. 121

Table 6-63. Solutions for clusters CL1, CL2, CL3 for RM1 .............. 122

Table 6-64. Two admissible solutions for STUDENT1 data set .......... 122

Table 6-7. Two Solutions for STUDENT2 Data Set. ..................... 125
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>Values of stimuli and category boundaries on a common scale . .</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>Midpoints and Regions Generated by Two Stimulus Points.</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>Boundary Mediators Generated by Three Response Variables in Two Dimensions.</td>
</tr>
<tr>
<td>Figure 4.</td>
<td>A two-dimensional plot of arbitrary points for subjects, stimuli and category boundaries.</td>
</tr>
<tr>
<td>Figure 5.</td>
<td>A two-dimensional plot of Solution 1 versus Solution 2</td>
</tr>
<tr>
<td>Figure 6.</td>
<td>Two-dimensional Plot of Solution I vs Solution II for STUDENT1 Data Set.</td>
</tr>
<tr>
<td>Figure 7.</td>
<td>Two-dimensional Plot of Solution I vs Solution II for STUDENT2 Data Set.</td>
</tr>
</tbody>
</table>
LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX 1A.</td>
<td>Description of real data examples</td>
</tr>
<tr>
<td>APPENDIX 1B.</td>
<td>Two 72 by 15 subjects-by-questions (stimuli) data matrices for students' responses.</td>
</tr>
<tr>
<td>APPENDIX 2.</td>
<td>DSOFE.BAS program for analyzing a dominance matrix</td>
</tr>
<tr>
<td>APPENDIX 3.</td>
<td>ORTHOG.BAS, a program for Gram-Schmidt orthogonalization.</td>
</tr>
<tr>
<td>APPENDIX 4.</td>
<td>DESIGN.BAS, a program that converts a dominance matrix into rank-order matrix.</td>
</tr>
<tr>
<td>APPENDIX 5.</td>
<td>RESIDUAL.BAS, a program that calculates ( \eta^2 ), ( y )-score and residual matrix.</td>
</tr>
</tbody>
</table>
CHAPTER 1

1.0 Introduction

Multidimensional analysis is a notion advanced in multivariate analysis to characterize and explain data structure. The task of multidimensional analysis is to search for some form of regularity in data. This regularity, when it is present, permits a description of the stimulus items (and sometimes of the individuals). Such description would be simpler than an exhaustive account of the response of every individual to every item, yet would tell much of what one may want to know about the data. One technique for multidimensional analysis of categorical data is dual scaling (Nishisato, 1980; see also Tenenhaus and Young, 1985). This data analytic technique is essentially a generalization of principal component analysis (PCA), a well-known technique for exploratory analysis of a set of quantitative variables.

Dual scaling provides multidimensional analysis for incidence data (for example, contingency tables, response-pattern tables) and two forms of dominance data (paired comparison data and rank-order data). However, certain problems arise when multidimensional analysis is attempted for a third form of dominance data
called successive categories (rating) data. This form of data is generated with a type of questionnaire item that requires a respondent to rate a set of statements on a set of successively ordered categories such as Poor, Fair, Good, Excellent. These category labels can be viewed as defining the ordering of response levels of some underlying variable if it makes intuitive sense to order them.

Data collection by rating scales is so commonplace in the behavioural sciences that the problem of quantification of the stimuli (items) rated has a long and varied history. For instance, traditional Thurstonian techniques (see, e.g., Torgerson, 1958) based on the method of successive categories, provide a model for psychological judgement and estimates of stimulus scale values and category boundaries. However, these techniques assume that all raters are interchangeable and that all items and rating categories are ordered on a unidimensional scale. The problem addressed in this study concerns finding multidimensional solutions for both stimuli and raters. Although this problem had been addressed in previous research (Nishisato, 1986c), the purpose of the present study is to extend this approach to deal with the difficulties that arise when the estimates of category boundaries violate the assumed unidimensional ordering. Before explaining the nature of the problem in detail, a formal introduction of the dual scaling technique seems appropriate at this point.
1.1 Dual Scaling

Dual scaling is an analytical procedure for analyzing a two-way table of categorical data. The procedure is based on Guttman’s (1950) model, which was applied to a two-way table of categorical data consisting of subjects in the rows and stimuli in the columns. In this model, internal consistency was the primary criterion of success in deriving the category weights and subject scores. Saying nothing about weights, Guttman (1950) postulated that the most internally consistent sets of scores to assign people based on their responses to stimuli are those that satisfy the following condition:

All people who fall in one category of an item should have scores as similar as possible among themselves, and as different as possible from the scores of the people in other categories of the item; this should be true to the best possible extent for all people simultaneously (Guttman, 1950, p. 314).

Similarly, saying nothing about scores, he postulated that the most internally consistent values to assign the categories are those that satisfy the following condition:

All categories characterizing one person should have numerical values as similar as possible among themselves, and as different as possible from the value of categories that do not characterize this person; this should be true
to the best extent possible for all people simultaneously (Guttman, 1950. p. 314).

Although we may treat the derivation of scores and weights as separate processes, Guttman showed that the two solutions are equivalent: that is, a simple relationship exists between the optimum scores and the optimum weights. The most internally consistent score for an individual is proportional to the mean of the weights of the categories the individual has selected, and the most internally consistent weight of a category is proportional to the mean of the scores of the individuals who selected it. These two approaches lead to the same thing as maximizing the ratio of the between-subject (or between-option) sum of squares to the total sum of squares. This ratio is called the squared-correlation ratio\(^1\), which is a measure of internal consistency of the weights assigned to categories and scores assigned to subjects.

As earlier stated, dual scaling is essentially principal component analysis applied to categorical data arranged in a two-way table. Each row or column variable is ‘optimally quantified’ into one quantitative variable, followed by an ordinary principal component analysis carried out on the complete set of thus constructed quantitative variables. This involves the application of the Eckart-Young singular value decomposition to the “quantified” data matrix. The "duality" of the analysis

\(^1\)This statistic is sometimes simply called “correlation ratio.”
is based on its symmetry, that is, one can derive the weights for the columns
directly from the weights for the rows, and vice versa. The weights are optimal in
the sense that they represent the greatest possible between-columns discrimination
and between-rows discrimination.

In situations where subjects give responses to stimuli, dual scaling provides both
the scale analysis of the stimuli (see, e.g., Guttman, 1941) and individual
differences scaling (see, e.g., Torgerson, 1958). As such, dual scaling does not
only deal with the problems relating to the "measurement" of attributes, but also
with the equally important and much broader set of problems concerning the
systematic classification of the attributes.

1.2 Brief history of dual scaling

The first rigorous mathematical formulation of the dual scaling technique was
given by Hirschfeld (1935), later called simultaneous linear regression approach
(Lingoes, 1968). Another formulation given about the same time was the method
of reciprocal averages (Horst, 1933). Fisher (1940) presented additive scoring.
Maung (1941) presented the bivariate correlational, the canonical correlational and
the one-way-analysis-of-variance approaches to the dual scaling technique.
Guttman (1941) presented a thorough formulation of dual scaling as applied to multiple choice data. Guttman (1946) applied this approach to paired comparison and ranking data. The same dual scaling technique was called appropriate scoring by Fisher (1948).

A thorough mathematical exposition was given by Lancaster (1958) and Torgerson (1958). Further exposition was given by Lord's study on reliability (1958), a study on ordered categories by Bradley, Katti & Coons (1962), and McDonald's unified treatment of the weighting problem (1968). Later, Hill (1974) referred to the dual scaling analytical technique as "a neglected multivariate method." The dual scaling technique has been referred to by other names: Guttman weighting, canonical scoring, analyse factorielle des correspondances (Benzécri et al., 1973), correspondence analysis (Hill, 1974), principal component analysis of qualitative data (Torgerson, 1958), optimal scaling (Bock, 1960) and biplot (Gabriel, 1971).

Dual scaling is called Multiple Correspondence Analysis (MCA) in French literature where the technique has reached a high level of development and use (Tenenhaus and Young, 1985). Dual scaling research in English literature include works by Hayashi (1950, 1952, 1954), Nishisato (1972, 1973, 1976, 1978, 1979, 1980, 1994), Nishisato and Leong (1975), Healy and Goldstein (1976), de Leeuw

1.3 Forms of categorical data

For dual scaling, categorical data are divided into two groups, namely, *incidence data* and *dominance data* (see, e.g., Nishisato, 1993). *Incidence data* are absolute measurements given as presence, absence, or frequencies obtained by counting the number of objects (or incidences, events) of a certain class. Dual scaling has been applied to the following types of incidence data: *Contingency/frequency table data* (see, e.g., Hirschfeld, 1935; Fisher, 1940; Maung, 1941). *Multiple-choice data* (see, e.g., Guttman, 1941) and *Sorting data* (see, e.g., Takane, 1980).

*Dominance data* represent “dominance relations” between objects or between variables of interest. Dominance relations are order relations within a set of objects with respect to some property level. For example, consider the case in which some stimuli are rated on a set of ordered categories. Consider two stimulus items, A and B. Item A is considered “higher” than item B, with respect to some property, say, X if item A has been placed in a category “higher” than the category in which item B has been placed. Thus, item A “dominates” over item, B with
respect to the property, \( X \). As will be made clear later, dominance relations exist among stimulus items rated on a common set of successive categories.

1.4 How dual scaling provides multidimensional analysis

In dual scaling of dominance data, each categorical variable is quantified by means of one quantitative variable, and singular valued decomposition (Stewart, 1973) is done on the complete set of thus constructed quantitative variables. The column and the row eigenvectors and associated eigenvalues, when scaled, form a set of dual scaling solutions.

1.4.1 Solution extraction procedure. Although a detailed description is not given here, singular value decomposition (SVD) is the method of choice for many eigenanalysis problems. The SVD has applications to such problems as canonical correlation, generalized least-squares multivariate analysis of variance (MANOVA) and analysis of covariance. In dual scaling, solution sets are extracted one at a time via the power method (Ralston, 1962).

1.4.2 One-dimensional analysis: As standard dual scaling extracts solutions one at a time, all that is needed for a one-dimensional analysis is to apply the procedure
only once to a given two-way data table of categorical data. If the solution obtained happens to be associated with a small value of the squared correlation ratio, then there is a possibility for further solutions.

1.4.3 Multidimensional analysis: Theoretically, solutions can be successively extracted until one hundred percent of all variance is accounted for. These solutions form a multidimensional representation of the body of a data set (Nishisato, 1980). The solutions (or dimensions) are analogous to factors in principal factor analysis (FA) or components in principal component analysis (PCA).

1.5 The parameter location vector

A set of dual scaling solutions consists of a vector of (row) subject weights, a vector of (column) parameter weights (or a parameter location vector) and the associated squared correlation ratio. The derivation of scale values involves optimizing column ranks of both data values (stimuli, \( \mu_j \)) and model parameters (category boundaries, \( \tau_k \)). (See, for example, Nishisato, 1980a,b; Nishisato and Sheu, 1984; Nishisato, 1994).
The vector of parameter location values resembles Thurstonian scaling results (see, e.g., Torgerson, 1958) in that the *category boundaries* and the *stimuli* can be plotted on a common scale. Below is a graphic illustration of a solution (showing *three stimuli* in categories labelled 1, 2, 3 and *two category boundaries*). The solution depicted represents the response profile (3 2 1).

\[
\text{Category 1} \quad \quad \text{Category 2} \quad \quad \text{Category 3}
\]

\[\mu_3 \quad \tau_1 \quad \mu_2 \quad \tau_2 \quad \mu_1\]

If the categories labelled 1,2 and 3 are assigned increasing sequential numbers from left to right, then \(\tau_1 < \tau_2\). The elements of a dual scaling solution (a parameter location vector) are not necessarily integers. The position of a stimulus along the continuum is a weighted average of the responses endorsed by subjects. The weighting factor is subject consistency as measured by the squared correlation ratio (\(\eta^2\)).

A stimulus item that falls in the center of a broad response category is most probably assigned that response option by many highly-consistent subjects. Such a stimulus position suggests some certainty that the item best falls in that category for that group. When a stimulus item falls near a boundary between two narrow
response categories, it is possible that responses to the stimulus have been varied and inconsistent. Broad categories are those used often and by consistent subjects, whereas narrow categories are those used seldom or by inconsistent subjects.

1.6 The research problem

For theoretical reasons, researchers may impose the order constraint on response categories in evaluating a set of stimulus items. Often the intention is to define a clear ordering of the categories used by respondents in parallel with some underlying variable. Sometimes, it makes intuitive sense to treat response categories as ordered, especially when subsequent analysis shows this to give a more satisfactory representation of the data (see, e.g., Kiers, 1993).

For many reasons, some people might respond differently to some stimulus items while responding the same way to other stimulus items. One possible reason is that the number of categories given is more than the respondents can distinguish. Another possible reason is that the respondents use multiple dimensions in deciding the category to place a stimulus item. These possibilities suggest multidimensionality of successive categories data. The challenge for the data analyst or researcher is to isolate and identify dimensions of individual differences.
when multidimensionality is a possibility. With ordered categories, the analysis should preserve any order restrictions on response categories hence category boundaries.

This study investigates how dual scaling may provide multidimensional analysis for successive categories data with the order constraint on categories (therefore category boundaries). As Nishisato (1980a) pointed out, defining category boundaries as ordinal variables is crucial:

From a practical point of view, the use of order constraints on ordered categories can be considered as a way to extract more interpretable information, and possibly to increase the validity of the scores. From a technical point of view, it implies that the space becomes 'tighter' by confining the configuration of the variables within the permissible region (p.164).

As Kiers (1993) observed, the order constraint can be imposed only if it makes intuitive sense. Sometimes, treating categories as successively ordered may not be appropriate. Consider the case of using the categories, \{Young, Middle-aged, Old\} in evaluating the phrase, "proneness to accidents." Suppose the order constraint is imposed so that the category boundaries, $\tau_1$ and $\tau_2$ (where is $\tau_1 < \tau_2$) are introduced to define the continuum, $\text{Young} < (\tau_1) < \text{Middle-aged} < (\tau_2) < \text{Old}$. 
However, our common sense knowledge is that the Young and the Old are more “prone to accidents” than the Middle-aged. Therefore, a solution that violates the order constraint on category boundaries as, for example, Old, Young < Middle-aged makes intuitive sense because middle-aged people are generally stronger and, therefore, less likely to be prone to accidents than younger or older people.

If the order constraint on categories is warranted, the constraint is built in the ‘quantification’ matrix (the dominance matrix). However, it is not certain that a solution (a parameter location vector) resulting from singular value decomposition of the quantification matrix will satisfy the order constraint on category boundaries. Although it is anticipated that the first solution would satisfy the order constraint on category boundaries, (see, e.g., Nishisato, 1980), this has not been investigated fully. Interpreting a solution that violates the order constraint on category boundaries would be difficult. When the imposition of order constraint on categories is warranted, violation of this constraint makes a solution inadequate as a meaningful representation of the information in the raw data.

This study is concerned with how multiple dual scaling solutions that satisfy the order constraint on category boundaries can be extracted. In a dominance matrix E, the first m columns representing category boundaries (τ₁, τ₂, . . . τₘ) as response
variables are always ordered by design. Yet, as previously stated, solutions derived from $\mathbf{E}$ may not always maintain the order constraint on category boundaries. Apparently, no previous study has formally investigated this problem. Furthermore, no previous study has shown how one might extract any required solution, at all, when a dominance matrix fails to yield a required solution initially. The present study investigates this problem to outline a possible way to resolve it. The purpose of the investigation is to develop a procedure for obtaining several solutions (parameter location vectors) in which the order constraint on category boundary values is maintained.

1.7 Significance of the study

This study is significant to the extent it that contributes to the methodological literature in general and to dual scaling in particular, with an illustration of a multidimensional method for successive categories data. Dual scaling is historically conceived of as essentially exploratory, structure seeking method (see e.g., Greenacre, 1984). This technique can be used when the structure existing in data is unknown or only vaguely appreciated. In dual scaling analysis, no a priori segmentation of data as to variables or dimensions is done. Instead, the body of the data is taken as a whole and the underlying segmental structure is allowed to reveal
itself in the analysis. To this end, a multidimensional analysis elicits more information about data structure with an advantage over traditional item analysis.

In traditional item analysis, several statements are gathered and then administered to a population sample. The investigator always defines the dimension(s) sampled by ordered categories. However, it is not always known if the investigator defines the dimension adequately or specifies the dimensions that the respondent considers as relevant. Furthermore, it is not always known if all respondents decide to use only a single dimension (Messick, 1956). Many people consider two dimensions. Some no doubt consider more. Thus, analysing data as if the responses were given on a single, investigator-defined dimension may not only obscure relevant information but may distort it. As Napior (1972) pointed out, analytical procedures that are arbitrarily unidimensional preempt an early subjective decision about the particular variables or underlying continua to be used to represent the larger body of data. This occurs even in cases where a multidimensional approach to the data would be more fruitful.

Other approaches to analysing successive categories data ignore that fact that, for ordered categories, information on the size of categories is not available. A procedure that does recognize this fact and exploits it, is dual scaling. In dual
scaling, scores on category boundaries are considered unknown, but ordered. In the multidimensional dual scaling procedure to be proposed in this study, 'optimal' values of category boundaries estimated in each dimension satisfy the order restrictions imposed by the data.

1.8 Outline of the thesis

The thesis will begin with discussion of the standard dual scaling procedure for a multidimensional analysis, and in particular, for rating data in Chapter 2. Chapter 3 is devoted to a review of some literature on multidimensional analysis of categorical data. Chapter 4 presents an overview of cluster analysis methods to select one method to incorporate in this study. In Chapter 5, this study outlines the proposed multidimensional method for successive categories data, given the acronym, DSMASC. Computer programs written for certain computations in DSMASC are also given in Chapter 5. Chapter 6 presents results from analysis of numerical examples. Chapter 7 provides a summary of the thesis, a discussion of findings and conclusions.
CHAPTER 2
DUAL SCALING OF SUCCESSIVE CATEGORIES DATA

2.0 Introduction

This chapter contains a description of the standard procedure for analysing successive categories data by dual scaling. First, the nature of successive categories data is discussed. Then, the standard procedure in which dominance data are first transformed before being scaled is presented. Finally, the problem faced in the multidimensional analysis of successive categories data is highlighted and the statement of the purpose of this study is given.

2.1 Successive categories (rating) data

Successive categories (rating) data are obtained when individuals rate stimuli using some kind of a rating scheme in which the response categories are “successive.” In education, especially in the context of teacher evaluation studies, rating has been defined as “an estimate made according to some systematized procedure, of the degree to which an individual person or thing possesses any given characteristics.
and may be expressed qualitatively or quantitatively” (Price, 1979). A rating scale is considered to be a device used in evaluating products, attitude, or other characteristics of instructors or learners. In describing various forms of rating scales, Guilford (1954) has stated:

The types of rating scales are all alike in that they call for the assignment of objects by inspection, either along an unbroken continuum or in ordered categories along the continuum. They are alike in that the end result is the attachment of numbers to those assignments. They differ in the operations of placement of objects, in the kind and number of aids or cues, and the fineness of discrimination demanded by the rater (p. 263).

A rating questionnaire generally contains items evaluated by descriptive phrases or descriptive words such as Poor, Mediocre, Average, Good, Excellent or letters such as A, B, C, and so on. For example, in grading essays, the teacher may feel confident that the quality of the essay graded A is better than any graded B. Similarly, the essay graded B may be considered better than C, and so on. Yet the teacher may be unwilling to make such quantitative judgements as "An A paper is four times better than a C paper."
Descriptive words have been the most commonly used terms for assigning ratings in surveys of early teacher rating scales (Price, 1979). Such ratings are widely used today in social science research and other educational research. Their popularity can be attributed to the ease with which they can be constructed. From a theoretical standpoint, rating procedures facilitate the study of social stimuli as these have no obvious physical dimension along which they can be compared directly (Stevens, 1951).

2.1.1 Conceptual form of data

Table 1-1 is a conceptual illustration of a response-pattern form of data from a categorical rating study. This three-way table shows the $N$ subjects in the rows and $n$ stimuli in the columns. Nested in each stimulus are $(m+1)$ successive categories: this makes for $n(m+1)$ columns in total. An entry of 1 in any column in Table 1-1 shows that a subject has assigned a stimulus to the category represented by the column, an entry of 0 shows non-assignment.

Usually, the subject simply gives the "score" or number associated with the category chosen rather than the category itself. It is customary to record such
Table 1-1. A conceptual response-pattern form for successive categories (rating) data

<table>
<thead>
<tr>
<th>Stimuli (items)</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>j</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categories</td>
<td>1</td>
<td>2</td>
<td>...</td>
<td>m+1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1-2. A subjects-by-stimuli table of successive categories (rating) data.

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>j</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>...</td>
<td>m+1</td>
<td>...</td>
<td>m+1</td>
</tr>
<tr>
<td>2</td>
<td>m+1</td>
<td>1</td>
<td>...</td>
<td>2</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2</td>
<td>2</td>
<td>...</td>
<td>1</td>
<td>...</td>
<td>2</td>
</tr>
</tbody>
</table>
information in condensed form as in Table 1-2. Notice that each stimulus in Table 1-1 has \((m+1)\) columns that represent response categories labelled \(1, 2, \ldots, (m+1)\). In Table 1-2, each stimulus is represented by a single column, and the \((ij)\)th cell entry, \(k\), represents the number given to the category chosen by Subject \(I\) in respect of Stimulus \(j\). We will call this table a subjects-by-stimuli matrix, and a row vector of values a subject profile.

### 2.1.2 Assignment of numbers to categories

Sometimes higher numbers, treated as "scores," are considered better (more of something), and at other times, lower numbers, treated as rankings, are considered better (first, called '1', is better than second, called '2'). In this study, higher numbers were considered to represent 'more of something' or 'better' than lower numbers.

### 2.2 Dual Scaling of successive categories (SC) data

#### 2.2.1 Definition of terms

The terms used in the formulation of dual scaling are as follows:

1. Subjects refers to individuals (judges, observers) who are instructed to render categorical judgements on a set of stimuli into successive categories.
2. *Stimuli* refers to a set of objects (typically questionnaire items).

3. *Weights* refers to numerical values assigned by dual scaling to a set of stimuli.

4. *Scores* refers to numerical values assigned by dual scaling to subjects.

5. *Parameters* refers to stimuli and category boundaries estimated by dual scaling.

6. “*Successive*” refers to a set of ordered response categories used by all subjects in rating a set of stimuli.

7. “*Solution*” will refer to a vector of scale values for stimuli and category boundaries or “parameter location vector.”

2.2.2 Steps in dual scaling of successive categories data

To analyze successive categories data by dual scaling, the first step is to convert a *subjects-by-stimuli* table into a dominance matrix, which is a *subjects-by-stimuli plus category boundaries* matrix. The next step is to set a unit of scaling and analyze the resulting ‘matrix for eigenequation’ or ‘matrix for decomposition’ by a numerical procedure that is essentially principal component analysis. The results consist of optimal weights for the rows (subjects), a parameter location vector for the column variables, and the associated squared correlation ratio. The entries in the subjects-by-stimuli matrix are sequential numbers that are given to categories to represent their ordinal relations.
2.3 Raw data transformations

A dominance (subjects-by-parameters) matrix can be obtained from a subjects-by-stimuli matrix via two different procedures that lead to the same result. These are the *rank-order procedure* and the *design-matrix procedure* (Nishisato and Sheu, 1984). The two procedures are different to the extent that they require different computer programming statements.

To illustrate these two procedures, consider a hypothetical case in the study of attitudes toward education where students acting as the subjects evaluate a range of services. Suppose three school services itemized as Q1, Q2 and Q3 are evaluated using the categories *Poor*, *Good*, *Excellent*. Let us assume that these categories define a continuum and that they are assigned the numbers 1, 2, 3, respectively. If the item, Q1 is evaluated as *Excellent*, then it receives the number, 3; If the item, Q2 is evaluated as *Good*, then it receives the number, 2; If the item, Q3 is evaluated as as *Poor*, then it receives the number, 1. A subject providing these responses is placed in a row of a subjects-by-stimuli matrix with the response pattern vector as (3 2 1).
2.3.1 The rank-order procedure

With two category boundaries, \( \tau_1 \) (between categories Poor and Good), and \( \tau_2 \) (between Good and Excellent), the pattern \( (3 2 1) \) can be represented in terms of items and category boundaries that are ordered as \( Q_3, (\tau_1), Q_2, (\tau_2), Q_1 \). These ordered variables can be assigned the sequential numbers \( 1, 2, 3, 4, 5 \), respectively, to reflect the ordering. It is convenient to read the assigned numbers in the order \( \tau_1, \tau_2, Q_1, Q_2, Q_3 \), which corresponds to the response pattern \( (2 4 5 3 1) \).

Items placed in the same category share numbers. For example, when all three items, \( Q_1, Q_2 \) and \( Q_3 \) are placed in the Good category, the response pattern vector is \( (2 2 2) \). With category boundaries included, the new response variables are ordered as \( (\tau_1), Q_1, Q_2, Q_3, (\tau_2) \). These receive the sequential numbers \( (1 3 3 3 5) \) because the arithmetic average of 2, 3, and 4, the numbers that would have been assigned to \( Q_1, Q_2 \) and \( Q_3 \) is 3. Thus, for the two hypothetical subjects’ responses, the first two lines of a table of rank orders will be as follows:

<table>
<thead>
<tr>
<th>Subject</th>
<th>( (\tau_1) )</th>
<th>( (\tau_2) )</th>
<th>( Q_1 )</th>
<th>( Q_2 )</th>
<th>( Q_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
2.3.1.1 Dominance matrix, $E$: A dominance number is the number of times a
variable is placed or rated above the other variables minus the number of times it
is placed or rated below the other variables by a subject. In general, if $N$ subjects
evaluate $n$ items using $(m+1)$ categories, then the dominance number for Subject $i$
is defined as

$$e_i = 2K_i - (n + m + 1),$$

where $K_i$ is the rank number given by Subject $i$ to item $j$ in rank-order table. In the
foregoing numerical example. $n = 3$ and $m+1 = 3$. By using the above formula
with various values of $K_i$, the elements. $e_i$ of the dominance matrix are calculated.
Thus, vectors of rank numbers. $(2 \ 4 \ 5 \ 3 \ 1)$ for Subject 1 and $(1 \ 5 \ 3 \ 3 \ 3)$ for
Subject 2. correspond to the following rows in a dominance matrix:

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\tau_1$</th>
<th>$\tau_2$</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Subject 2</td>
<td></td>
<td></td>
<td>-4</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Total number of comparisons: Each dominance number is the outcome of $(n + m -
1)$ comparisons. that is, each subject compares an item or a category boundary with
the remaining $(n + m - 1)$ items and category boundaries. It follows that the total
number of comparisons (responses) for each subject (i.e., each row total) is the
product of the number of columns (parameters) and of comparisons (responses)
which is $(n + m)(n + m - 1)$. Likewise, the total number of comparisons in each
column (the column total) is \( N(n + m - 1) \). Consequently, in the dominance table, we have \( N(n + m)(n + m - 1) \) responses in total. The specified \( N \times (n + m) \) dominance table, \( E \) (with category boundaries in the first \( m \) columns, objects in the last \( n \) columns), is the appropriate input data for dual scaling.

### 2.3.2 The design matrix procedure

In this procedure, we first define a matrix, \( F \), of pair-wise contrasts for pairs of parameters to be estimated in dual scaling. One group of parameters is the set \{ \( \mu_1, \mu_2, \ldots, \mu_n \) \} representing items and the other group is the set \{ \( \tau_1, \tau_2, \ldots, \tau_m \) \} representing the category boundaries. In the case of three items and three categories, the between-group contrasts are \( (\tau_1 - \mu_1), (\tau_2 - \mu_1), (\tau_1 - \mu_2), (\tau_1 - \mu_3) \), and \( (\tau_2 - \mu_3) \), and the within-group contrasts are \( (\tau_1 - \tau_2), (\mu_1 - \mu_2), (\mu_1 - \mu_3), \) and \( (\mu_2 - \mu_3) \). Thus, the columns of \( F \) consist of the contrasts, \( (\tau_1 - \mu_1), (\tau_2 - \mu_1), (\tau_1 - \mu_2), (\tau_1 - \mu_3), (\tau_2 - \mu_3), (\tau_1 - \tau_2), (\mu_1 - \mu_2), (\mu_1 - \mu_3), \) and \( (\mu_2 - \mu_3) \). Each contrast is given a value of 1 or -1. That is, an element of matrix \( F \) is defined as, \( f_{ijk} = 1 \) if subject \( i \) places item \( j \) into the \( k \)-th or any preceding category, and \( f_{ijk} = -1 \), if subject \( i \) places item \( j \) into the category \( (k+1) \) or any higher category (see, e.g., Nishisato, 1980a; Bock & Jones, 1968). Having so defined an element of matrix \( F \), it is easy to see that the row vector \( (3 \ 2 \ 1) \) from a subjects-by-stimuli matrix translates into the row vector.
(-1 -1 -1 1 1 1 -1 1 1) of matrix $F$. Now, to obtain the dominance matrix $E$, matrix $F$ is postmultiplied by a design matrix, say $A$.

**Construction of design matrix, $A$:** A standard design matrix $A$ consists of -1's, 0's and 1's (see Bock & Jones, 1968). In our example, the set of parameters $\{\tau_1, \tau_2, \mu_1, \mu_2, \mu_3\}$, arranged in this order, as before, for convenience make up the columns of matrix $A$. Let us now briefly explain how columns of $F$ are related to columns of $A$. Consider the contrast $(\tau_1 - \mu_1)$, which is a column of $F$. The coefficients of $\tau_1$ and $\mu_1$ are 1 and -1, respectively. Therefore, the row of matrix $A$ that captures this contrast $(\tau_1 - \mu_1)$ is $(1 \ 0 \ -1 \ 0 \ 0)$. Similarly, every column of matrix $F$ is represented by a row of matrix $A$. A column of matrix $A$, which represents a given parameter, is designed to take the sum of the elements of rows of matrix $F$ in which the column parameter is involved. For example, the column of matrix $A$ that represents $\tau_1$ captures contrasts with a coefficient of 1; the contrasts are $1 \times (\tau_1 - \mu_1), 0 \times (\tau_2 - \mu_1), 1 \times (\tau_1 - \mu_2), 0 \times (\tau_2 - \mu_2), 1 \times (\tau_1 - \mu_3), 0 \times (\tau_2 - \mu_3), 0 \times (\mu_1 - \mu_2), 0 \times (\mu_1 - \mu_3),$ and $0 \times (\mu_2 - \mu_3)$ and these give the column of matrix $A$ in question as $(1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 0)$. Consequently, postmultiplying matrix $F$ by matrix $A$ amounts to taking the *inner product* of a row vector of matrix $F$ and a column vector of matrix $A$. 
In our example, the vector pattern (3 2 1) from the subjects-by-stimuli data matrix translates into the pattern, \((-1 -1 1 1 -1 1 1 1)\), which is the corresponding row of matrix \(F\). The inner product of \((-1 -1 1 1 -1 1 1 1)\) with \((1 0 1 0 1 0 1 0 0 0)\) gives -2. This is the number of times the parameter \(\tau_1\) has been judged or rated as ‘above’ the other parameters (represented by a column of matrix \(A\)) minus the number of times the same parameter has been judged or rated as ‘below’ the other parameters. The number in question refers to the subject whose responses are represented by the vector of number, (3 2 1). Therefore, postmultiplying matrix \(F\) with matrix \(A\) gives the dominance matrix, \(E\), that is, \(FA = E\).

A row of dominance numbers in the dominance matrix \(E\) provides a clear interpretation of the relations between the parameters to be weighted by dual scaling. For example, the vector \((-2 2 4 0 -4)\) represents the inequality relation, \(\mu_3 < \tau_1 < \mu_2 < \tau_2 < \mu_1\). Thus, the dominance matrix, which describes the relations between the \((n)\ items\) and the \((m) category\ boundaries,\) is the appropriate input matrix for dual scaling. In the next section, a brief discussion of the mathematical process involved in deriving dual scaling solution is provided.
2.3.3 Notation

The following notation is necessary for the exposition of the dual scaling framework:

\[ N = \text{number of subjects} \]
\[ n = \text{number of items (stimuli)} \]
\[ m = \text{number of category boundaries} \quad (m+1 = \text{number of categories}) \]
\[ f_i = N(n+m)(n+m-1) \cdot \text{total number of parameters involved in all pair-wise comparisons that lead to the dominance matrix} \]
\[ E = \text{dominance matrix} \]
\[ \theta' = ([n+m] \times 1): \text{the transpose of the optimal weight vector for columns of } E \]
\[ y = (N \times 1): \text{the optimal score vector for rows of } E \]
\[ D_c = N(n+m-1)I_{m-m} : \text{diagonal matrix for column totals of } E \]
\[ D_r = (n+m)(n+m-1)I_n : \text{diagonal matrix for row totals of } E \]
\[ \eta^2 = \text{the squared correlation ratio} \]
\[ SS_b = \text{the sum of squares between groups} \]
\[ SS_t = \text{the total sum of squares} \]
\[ SS_w = \text{the sum of squares within groups} \]

2.3.4 Derivation of the ‘optimal’ solution

There are four different approaches to dual scaling. For each of these approaches, the object is to maximize a specific criterion. These are:
(1) the principal component analysis approach in which the weights assigned the columns or rows maximize the variance of the composite scores;

(2) the analysis-of-variance approaches where the assigned weights maximize the between-row (column) sum of squares in relation to the total sum of squares (the squared correlation ratio);

(3) the bivariate correlation approach in which the product-moment correlation between responses weighted by row weights and those by column weights is maximized; and

(4) the simultaneous linear regression approach in which row and column weights are determined to make the regression of rows on columns and the regression of columns on rows simultaneously linear.

Since these approaches all lead to identical results, we consider only the analysis of variance (ANOVA) approach.

In a one way analysis of variance (ANOVA), the total sum of squares is orthogonally partitioned into the between and the within sum of squares, that is. \( SS_t = SS_b + SS_w \). This decomposition is uniquely determined if the data are continuous. However, if the data are categorical, as is the case with the dominance matrix \( E \), the relative magnitudes of the sums of squares vary, depending on weights.
assigned to the columns (parameters) or the rows (subjects). Therefore, the relative magnitudes of the sums of squares can be manipulated to find the optimal score vector, \( \mathbf{y} \), for the subjects and the optimal weight vector, \( \mathbf{\theta}' \), for the parameters. These vectors are those that maximize the between-column sum of squares, \( SS_b \), in relation to the total sum of squares, \( SS_t \). Mathematically, this is the same as maximizing the ratio \( SS_b/SS_t \), the squared correlation ratio, which is given the symbol \( \eta^2 \). That is,

\[
\eta^2 = \frac{SS_b}{SS_t}.
\]

The expressions for \( SS_b \) and \( SS_t \) that involve the dominance matrix, \( E \) and the optimal score vector, \( \mathbf{y} \) are:

\[
SS_b = \mathbf{y}' \mathbf{E} \mathbf{D}_c^{-1} \mathbf{E}' \mathbf{y} - \frac{\mathbf{y}' \mathbf{E} \mathbf{11}' \mathbf{E}' \mathbf{y}}{2nmN},
\]

\[
SS_t = \mathbf{y}' \mathbf{D}_c \mathbf{y}.
\]

In terms of the optimal weight vector, \( \mathbf{\theta} \), \( SS_b \) and \( SS_t \) are expressed as

\[
SS_t = \mathbf{\theta}' \mathbf{D}_c \mathbf{\theta}.
\]
2.3.4.1 Least-squares optimization to the optimal solution

The maximization scheme uses the well known Lagrangian (Apostol. 1957) multiplier, \( \lambda \), to define the objective function as

\[
L(\theta, \lambda) = SS_b - \lambda(SS_r - f_r) = \theta' E D_r^{-1} E \theta - \lambda(\theta' D_c \theta - f_r)
\]

and the system of partial derivatives of \( L \) with respect to \( \theta \) and \( \lambda \) set to zero:

\[
\frac{\partial L}{\partial \lambda} = \theta' D_c \theta - f_r = 0.
\]

\[
\frac{\partial L}{\partial \theta} = 2E' D_r^{-1} E \theta - 2\lambda D_c \theta = 0,
\]

After some algebraic manipulation of (2.7) and (2.8), it can be shown that \( \lambda \), the unknown multiplier is equal to the squared correlation ratio, the maximization criterion in (2.9).
Rearranging equation (2.9) gives (2.10), which is called the generalized eigenequation.

\[(2.9) \quad \lambda = \frac{\theta'ED_r^{-1}E\theta}{\theta'D_c\theta} = \eta^2,\]

\[(2.10) \quad (E'D_r^{-1}E - \eta^2D_c)\theta = 0\]

The standard form of this equation is obtained by simply pre-multiplying both sides of the equation by $D_c^{-1}$ to yield

\[(2.11) \quad (D_c^{-1}E'D_r^{-1}E - \eta^2I)\theta = 0.\]

### 2.3.4.2 Solution extraction procedure

In dual scaling analysis, a solution (an optimal weight vector) is derived from the eigenvector associated with the largest eigenvalue. Multiple dimensions are found sequentially via the power method (Ralston, 1965). A description of how the power method is employed in dual scaling can be found in Nishisato (1980).
2.3.4.3 Deriving subject scores

Once the optimal weight vector, $\theta$ is determined, subject scores are calculated using the "dual relations" formula,

$$y = (1/\eta) D_r^{-1} E \theta.$$

where $\eta$ is the square-root of the squared-correlation-ratio, $y$ is the vector of optimal subject scores.

2.4 Summary

In this chapter, the standard dual scaling for rating data (SDSS) has been presented. This procedure analyses a subject-by-stimuli data matrix to produce several dual scaling solutions. The elements of a solution (parameter location) vector represent category boundaries and stimuli. The question about maintenance of the order constraint on category boundaries in a solution will be discussed in the next chapter.
CHAPTER 3
RELATED LITERATURE

3.1 The order constraint on variables

3.1.0 Introduction: In the analysis of successive categories data by dual scaling, a subset of 'newly' defined response variables is put under the order constraint. These response variables are category boundaries. They are introduced when successive categories data are converted into a dominance matrix. As a result of dual scaling, the category boundary values are least-squares estimates obtained under the order constraint. This constraint is imposed by design on the first $m$ columns (where $m+1$ is the number of successive categories) of a dominance matrix within each subject in a row. The dominance matrix consists of subjects in the rows with category boundaries and stimuli in the columns.

So far we have seen that the order constraint is manifest in the dominance matrix. Also, we have seen that interpretability of dual scaling solutions derived from a dominance matrix depends on whether the estimated category boundary values satisfy the order constraint. The issue that has not been addressed, however, is how to handle cases that involve weak order relations between category boundaries. To
get the complete picture of potential outcomes, we consider two cases for
illustration: (1) The case of two category boundaries and, (2) The case of three
category boundaries.

3.1.1 The case of two category boundaries

In the case of two category boundaries, \( \tau_1 \) and \( \tau_2 \), there are three possible dual
scaling outcomes:

Case I: \( \tau_1 > \tau_2 \)

Case II: \( \tau_1 < \tau_2 \)

Case III: \( \tau_1 = \tau_2 \)

A solution represented by Case II above is admissible. However, a solution that
represented by the relation in Case I is not admissible. A reflection of the solution
represented by Case I results in a solution represented by the relation in Case II.
Therefore, with two category boundaries for which a strong order relation exists,
(as in Case I), an inadmissible solution can be transformed into an admissible
solution by reflection.

Solutions represented by Case I (after reflection) and Case II are linear scales that
reproduce categories 1, 2 and 3. However, a solution represented by the relation in
Case III does not reproduce Category 2. The implication of this for interpretation of a solution represented by the relation in Case III might be best understood with a numerical example.

In Table 3-1 is a dual scaling analysis of a hypothetical data set of ratings of three stimuli, $X_1$, $X_2$, $X_3$ ($n=3$) by six subjects ($N=6$) on three categories ($m+1=2$). As can be seen in (C) of Table 3-1, scale values for the two category boundaries and $X_2$ coincide ($\tau_1 = \tau_2 = \mu_2 = 0.0000$). Also, Category 2, which is bounded by $\tau_1$ and $\tau_2$, is diminished. The parameters $\tau_1$, $\tau_2$, and $\mu_2$ apparently contribute nothing to the variance in the dominance matrix. E. Inspection of these columns, in (B) of Table 3-1, reveals that the entries under each column are constants.

A commonsense way to evaluate the appropriateness of dual scaling solutions is to compare the response profile of the subject with the largest score to that profile the solution (a parameter location vector) represents. In Table 3-1, the first solution accounts for a large proportion of total variance ($8\% = 80$) as can be seen in (D) of Table 3-1. The parameter location vector depicts the following pattern:

Stimulus $X_1$ ($\mu_1 = -1.5811$) is located in Category 1.

Stimulus $X_2$ ($\mu_3 = 1.5811$) is located in Category 3.

Stimulus $X_2$ ($\mu_2 = -0.0000$) coincides with $\tau_1$, and $\tau_2$. 
Table 3-1. *Six subjects, three stimuli, three categories.*

(A) Data matrix

<table>
<thead>
<tr>
<th>SUBJECTS</th>
<th>STIMULI</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁, X₂, X₃</td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>3 2 1</td>
</tr>
<tr>
<td>2.</td>
<td>3 2 1</td>
</tr>
<tr>
<td>3.</td>
<td>3 2 1</td>
</tr>
<tr>
<td>4.</td>
<td>1 2 3</td>
</tr>
<tr>
<td>5.</td>
<td>1 2 3</td>
</tr>
<tr>
<td>6.</td>
<td>1 2 3</td>
</tr>
</tbody>
</table>

(B) Dominance matrix (E)

<table>
<thead>
<tr>
<th></th>
<th>-2</th>
<th>2</th>
<th>4</th>
<th>0</th>
<th>-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td></td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>-4</td>
</tr>
</tbody>
</table>

(C) Dual Scaling Solutions

<table>
<thead>
<tr>
<th>SUBJECTS</th>
<th>CATEGORY BOUNDARIES</th>
<th>STIMULUS VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>τ₁ -0.0000</td>
<td>X₁ -1.5811</td>
</tr>
<tr>
<td>2.</td>
<td>τ₂ 0.0000</td>
<td>X₂ -0.0000</td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td>X₃ 1.5811</td>
</tr>
</tbody>
</table>

(D) Statistics

Squared Correlation Ratio = 0.40000
Delta (Total Variance Accounted For): Partial = 80.00
Cumulative = 80.00
The stimulus and category boundary values can be plotted on a common scale as in Figure 1. Categories 1, 2 and 3 have been nominally assigned the numbers 1, 2 and 3, respectively. Clearly, in Figure 1, the profile represented by the solution (parameter location values) does not appear to reproduce any of the subjects' responses in (A) of Table 3-1 yet each subject has either a score of 1 or -1.

\[ \begin{align*}
\mu_1: & \text{Stimulus } X_1 \\
\tau_1: & \text{Category boundary 1} \\
\mu_2: & \text{Stimulus } X_2 \\
\tau_2: & \text{Category boundary 2} \\
\mu_3: & \text{Stimulus } X_3
\end{align*} \]

**Figure 1.** Values of *stimuli* and *category boundaries* on a common scale.

Incidentally, the solution depicted in Figure 1, captures the contributions of the group of subjects with the pattern \((1 \ast 3)\), where \(\ast\) represents 2. Although this solution does not violate the order constraint on category boundaries, interpreting it is difficult. The scale values for \(\tau_1, \tau_2\) and \(\mu_2\) are indistinguishable. Yet we see in the raw data that the stimulus item 2 was placed in Category 2. Clearly, tied *category boundaries* make interpretation difficult. Nicely separated category
boundaries should be easier to deal with. Let us now examine the case of four category boundaries.

3.1.2 The case of four categories

Let us discuss the case of four categories involving three category boundaries, \( \tau_1 \), \( \tau_2 \), and \( \tau_3 \). In a dominance matrix \( E \), the category boundaries \( \tau_1 \), \( \tau_2 \), and \( \tau_3 \) are ordered as \( \tau_1 < \tau_2 < \tau_3 \) within each subject. However, the parameter location vector derived by dual scaling may not always satisfy this order constraint on category boundaries (Nishisato, 1980). In theory, the order constraint on category boundaries can be maintained or violated in many ways. Consider the examples in Table 3-2. Cases 1, 4, 5 and 6, (with some reflected) would satisfy the admissibility criterion while Cases 2 and 3 would not. Cases 4 and 5 would each require inspection of the response patterns (raw data) of the subject receiving the largest absolute weight. This way, one may confirm whether Category 2 for Case 4 or Category 1 for Case 5 is an appropriate representation of reality in the raw data since the categories in question are diminished.
Table 3-2. Inequality relations involving three category boundaries

<table>
<thead>
<tr>
<th>Case</th>
<th>Relation</th>
<th>Reflection</th>
<th>Status after reflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\tau_1 &lt; \tau_2 &lt; \tau_3$</td>
<td>Not necessary</td>
<td>Admissible</td>
</tr>
<tr>
<td>2</td>
<td>$\tau_1 &gt; \tau_2 &lt; \tau_3$</td>
<td>$\tau_3 &gt; \tau_2 &lt; \tau_1$</td>
<td>Not admissible</td>
</tr>
<tr>
<td>3</td>
<td>$\tau_1 &lt; \tau_2 &gt; \tau_3$</td>
<td>$\tau_3 &lt; \tau_2 &gt; \tau_1$</td>
<td>Not Admissible</td>
</tr>
<tr>
<td>4</td>
<td>$\tau_1 &gt; \tau_2 = \tau_3$</td>
<td>$\tau_3 = \tau_2 &gt; \tau_1$</td>
<td>Admissible*</td>
</tr>
<tr>
<td>5</td>
<td>$\tau_1 = \tau_2 &gt; \tau_3$</td>
<td>$\tau_3 &gt; \tau_2 = \tau_1$</td>
<td>Admissible*</td>
</tr>
<tr>
<td>6</td>
<td>$\tau_1 &gt; \tau_2 &gt; \tau_3$</td>
<td>$\tau_1 &lt; \tau_2 &lt; \tau_3$</td>
<td>Admissible*</td>
</tr>
</tbody>
</table>

*These cases require inspection of subject scores to interpret.

As we saw in the case of three-category data, solutions that satisfy only a weak order constraint as in Cases 4 and 5, are difficult to interpret. Nevertheless, the appropriateness of the solutions can be checked by an inspection of the raw data. That is, the response pattern of a subject (category labels chosen for each stimulus rated) receiving the largest weight can be used to check the validity of the derived parameter location vector.

The results of the three-category example discussed previously (see Table 3-1) confirm the suspicion that failure of a solution to satisfy strong order constraints on category boundaries does not necessarily imply that an investigator's decision
to impose order is unwarranted. The near violation (tied category boundaries) of the order constraint by a solution might simply suggest multidimensionality of the data. If so, how can the quality of dual scaling solutions for successive categories data be improved? The present study addresses this question in addition to attempting to solve the problem of solutions that do not satisfy the order constraint on category boundaries. In the next section, literature related to the multidimensional analysis of successive categories is reviewed.

3.2 Points of view proposal

It is fair to say that the literature on dual scaling of successive categories is somewhat limited. Only one source in the literature seems to have discussed the problem of multidimensional analysis of successive categories by dual scaling. The paper by Nishisato (1986c) proposed point-of-view analysis (PVA) as a method for multidimensional analysis.

Acknowledging that category boundaries are always ordered as \( \tau_1 \leq \tau_2 \leq \ldots \leq \tau_m \) within each subject in a dominance matrix, \( E \), that is, \( e_{i1} \leq e_{i2} \leq \ldots \leq e_{im} \) for subject \( i, \ i = 1, 2, \ldots, N \), Nishisato (1986c) proposed a method for multidimensional analysis of successive categories data. The first step involved clustering subjects
(rows) into homogeneous subgroups. The second step involved submitting these
subgroups (clusters) to dual scaling.

The first solution from a dominance matrix consisting of any subset of rows
(subjects) was expected to satisfy the order constraint on category boundaries.
Each sub-matrix (cluster of subjects) was expected to tap a distinct dimension
while yielding an admissible solution. Three clustering procedures were suggested.
Each clustering procedure would operate on a matrix whose columns are subject
score vectors. These would be obtained as follows.

First, a dominance matrix, \( E \), is subject to singular value decomposition such that,
\[
Y^TE \Theta = c \Lambda
\]
where,
\[
Y = [Nxt] \text{ matrix of } t \text{ vectors of weights for the subjects.}
\]
\[
\Theta = [(m+n)xt] \text{ matrix of } t \text{ vectors of parameter location estimates.}
\]
\[
\Lambda = \text{ diagonal matrix of eigenvalues arranged in descending order.}
\]
\[
c = \text{ a scaling constant [usually } N(n+m)(n+m-1)].
\]

Second, \( Y \) is postmultiplied by \( \Lambda \) to give the matrix, \( YA \) whose columns comprise
of vectors of the subjects scores.

Three different clustering procedures to operate on \( YA \) to put subjects into clusters
were suggested: (1) *non overlapping clustering*, which we will call NOC; (2) *overlapping clustering*, which we will call OC; and (3) *stepwise non overlapping clustering*, which we will call SNOC.

### 3.2.1 Non-overlapping clustering (NOC)

To carry out this procedure, first obtain several solutions, say, $h$ ($h > 0$). Then, classify each subject (i.e., a row of $Y\Lambda$) into group $h$, where the absolute value of $y_{in}/h$ is the largest and $h = 1, 2, \ldots, p$. In other words, a subject is classified into one of $p$ groups where the subject's absolute score is the largest.

### 3.2.2 Overlapping clustering (OC)

Given the $N \times t$ matrix $Y\Lambda$, choose $s$ largest absolute values, with $s$ bounded as $2t \leq s \leq n/t$. where $t$ is the rank of $E$ and $n$, the number of stimuli. With respect to each column of $Y\Lambda$, the subjects so selected make up the first cluster of subject profiles. From the remaining subjects, the next cluster is formed in the same way. This cluster (say, $h=1$), which is a sub-matrix of $E$, is then submitted to dual scaling. Then, the optimal weight vector, $\theta_h$, associated with sub-matrix $E_h$ is checked to ensure that the order constraint on the category boundaries is satisfied.
If necessary, the weight vector $y_h$ for subjects in the cluster can be calculated using the ‘dual relations’ formula.

### 3.2.3 Stepwise non-overlapping clustering (SNOC)

This procedure was suggested as an alternative to the partitioning processes in 3.2.1 and 3.2.2 above. Here, the dominance matrix, $E$ is first submitted to dual scaling to obtain the first solution. Taking a column of $YA$ at a time, subjects with scores exceeding a criterion value (could be a number less than but close to 1.00 such as .95) are selected. This cluster of subjects is then analyzed by dual scaling. The subjects’ scores from this analysis are then compared with the criterion score. The aim is to isolate subjects whose scores are equal to or greater than the criterion score. A cluster of subjects finally obtained is then partitioned from the dominance matrix, $E$ and designated as the first cluster. The same steps are repeated with remaining partition of $E$ to give a second, third, ..., $p^{th}$ cluster of subjects. This procedure is continued until one is satisfied that the scores of remaining subjects do not define a cluster. The clusters of subjects ‘dominance’ profiles are ultimately submitted to dual scaling.

---

\(^2\)Note that unidimensional matrix $E$ is expected to yield weights of 1 or -1 for all the subjects.
3.2.4 Discussion of OC, NOC and SNOC

3.2.4.1 The OC approach. As for the overlapping clustering proposal, a cluster is defined by the number of absolute values, \( s \), to assign to a cluster, with \( s \) bounded as \( 2t \leq s \leq n/t \). If \( t \) is the rank of the matrix \( \mathbf{Y}\Lambda \), then the smallest possible value of \( t \) is 1 and the smallest possible number of items, \( n = 2 \). When \( t = 2 \), the number of items \( n \geq 8 \). Apparently, with eight stimulus items, not even a cluster of one subject can be defined. This condition seems to restrict the number of subjects in a cluster to one.

3.2.4.2 The NOC approach. In this approach, subjects with the largest scores from each column of \( \mathbf{Y}\Lambda \) are assigned to the same cluster. Technically, the number and size of clusters defined depends on how many columns of \( \mathbf{Y}\Lambda \) are included in the analysis. Thus, the definition of a cluster by this approach is also problematic.

3.2.4.3 The SNOC approach. In this approach (see Nishisato, 1980b: 1986c), the cutoff score for including subjects in a cluster is decided arbitrarily by the data analyst. This means that a data analyst employing a different cutoff score would come up with a different cluster, therefore, a different dual scaling solution (a parameter location vector).
3.2.4.4 Summary of comments on OC, NOC, SNOC. Considering how these approaches define clusters, different data analysts could define different clusters resulting in different solutions. To extend Nishisato's (1986c) PVA proposal, the present study will focus on a different way of defining clusters. This will include the idea of operating directly on the dominance matrix, \( E \) instead of subject scores derived without order constraints on category boundaries.

3.3 The purpose of this investigation

As we have noted in previous sections (see also Nishisato, 1986c: 1994), the dominance numbers for category boundaries always satisfy the order constraint within each subject in the dominance matrix. However, it has been suggested that a case of data with only two subjects might not yield a first solution that satisfies the order constraint but that empirically most cases of enough subjects would satisfy the order constraint (Nishisato, 1993, personal communication).

In the points of view analysis proposal (Nishisato, 1986), the first solution from a dominance matrix consisting of any subset of rows (subjects) of the entire dominance matrix is expected to satisfy the order constraint on category boundaries, \( \tau_k \). Each cluster of subjects would tap a distinct dimension, while
satisfying the order constraint on category boundaries. The present investigation essentially follows Nishisato’s (1986c) approach. However, this investigation differs from Nishisato’s approach in one major way. Instead of clustering subjects by operating on inter-subject distances derived from subject scores, the present study proposes using a traditional cluster analysis method to cluster subjects by operating on the dominance matrix itself.

The idea is to explore the structure of a dominance matrix that gives either an inadmissible solution or a solution that accounts for a small percentage of the total variance by clustering its rows. A resultant cluster would have the property that the weights of subjects tend to be close to 1 or -1. More specifically, we propose the following:

(1) The initial dominance matrix (IDM) that yields an inadmissible solution, or a solution that explains a small percentage of the total variance, is partitioned, initially, into two clusters (sub-matrices of the initial dominance matrix). If neither of the two clusters gives an admissible solution, IDM is partitioned into three clusters. If no admissible solution is obtained, then the search for an admissible solution may continue in the same way until the initial dominance matrix has been divided into m
clusters \((m = \text{number of category boundaries})\).

(2) When the IDM yields an *admissible solution* set, \([\eta_i^2, x_i, y_i]\), the effects of this solution set are then removed from IDM, leaving a *residual dominance matrix* \((RDM)\). Upon being subjected to dual scaling, RDM might yield the solution set, \([\eta_2^2, x_2, y_2]\) that satisfies the order constraint. If not, then RDM is subjected to *cluster analysis* as was done with IDM in (1) above.

(3) The search for admissible solutions stops when an adequate amount of the total variance has been explained by the solutions identified.

In the present study, the merits of using traditional cluster analysis on a dominance matrix as a standard procedure for multidimensional analysis of successive categories data will be discussed. For convenience, the analytical procedure to be developed will be called, "Dual Scaling Multidimensional Analysis of Successive Categories Data," also abbreviated as DSMASC. As before, the term, *admissible solution* will refer to a solution (a parameter location vector) that satisfies the order constraint on category boundaries. This will be a precondition for the interpretability of a solution.
3.4 Summary

This chapter contains a discussion of the potential problem concerning category boundary scale values not being maintained in certain solutions. Specifically, when order is imposed on response categories, a solution is *admissible* if it satisfies the order constraint. Category boundary values, along with stimuli, are so determined as to optimize Guttman's (1941) measure of internal consistency, the *squared correlation ratio*. In the case of *two* category boundaries, an *inadmissible solution* can be transformed into an admissible solution by reflection. However, transformation by reflection is not sufficient in cases with *more than two* category boundaries. This matter is further discussed in the next chapter.
CHAPTER 4
CLUSTER ANALYSIS METHODS

4.1 Introduction

The aim of this study is to develop a procedure for multidimensional analysis of successive categories data by dual scaling. It is known that the analysis of the so-called dominance matrix does not guarantee multiple solutions wherein the category boundaries are ordered by design. Therefore, the problem to be addressed concerns how to carry out a decomposition of a dominance matrix \( E \) in a constrained multidimensional space. The reader may recall that the first \( m \) columns of \( E \) (representing category boundaries \( \tau_1, \tau_2, \ldots, \tau_m \) as response variables) are ordered by design. The present study attempts to develop a method by which many admissible solutions may be derived from a given dominance matrix, \( E \). The procedure to be developed hopefully will provide multidimensional analysis for successive categories data.

The present study differs from Nishisato's (1986) PVA proposal in several ways: First, the PVA proposal was to cluster inter-subject distances calculated from dual
scaling results whereas the present study seeks to operate on a dominance matrix.

Second, the PVA's clustering procedures provided no objective criterion for defining clusters whereas the present study seeks to apply a traditional cluster analysis method with an established criterion for defining clusters. This chapter contains an overview of cluster analysis techniques and is concluded by the selection of one cluster analysis method to be used in the study.

4.2 Definition of cluster analysis

Cluster analysis models attempt to place the objects under study into groups in such a way as to best represent measures of similarity among them. The abstract clustering task is to distinguish clusters, for example, \( h_1, h_2, \ldots, h_n \) from a set of objects, \( H \) (where \( h_i \)'s are subsets of \( H \)). In the process of clustering, intra-cluster object similarity of each \( h_i \) is maximized, while the inter-object similarity over all \( h_i \)'s is minimized.

An object is described as a vector of variable values, where quantitative, nominal (categorical), and binary-valued variables may be allowed (see, e.g., Everitt, 1980). The similarity between two objects is the value of a numeric function applied to the descriptions of the two objects. The data analyst is typically
responsible for computing the pair-wise similarity of all objects in a data set. The similarity matrix is then used by the program to group objects that are most similar and to distinguish objects that are least similar.

Intra-cluster and inter-cluster similarity are computed by a function of the pair-wise similarities of the objects in each cluster. Given two subjects (objects), A and B, whose features are described by $A'$ and $B'$, respectively, a typical similarity measure between A and B has the form

$$\text{Similarity} (A, B) = f(A', B').$$

Such a similarity measure has been termed context-free, since the similarity between A and B is independent of A's and B's relationships to other subjects being clustered (Fisher & Langley, 1986).

### 4.3 Overview of clustering techniques

In a typical cluster analysis problem, one uses an amalgamation rule with the elements of a dissimilarity matrix to assign the subjects (in the rows of a dominance matrix) to a particular "cluster" or "class." Different clustering methods are distinguished by different amalgamation rules. Many classes of cluster analysis techniques have been described in three general headings: optimization techniques,
hierarchical techniques and clumping techniques. (See, e.g., Jain & Dubes, 1988.)

4.3.1 Optimization techniques. These techniques attempt to form an optimal K-partition over an object set (i.e., divide the object set into K mutually exclusive clusters) where K is supplied by the user. An extensive search is made for an optimal K-partition. Optimization techniques in traditional cluster analysis methods find clusters in the multidimensional space, and produce exclusive clusters wherein each object is assigned to only one cluster.

4.3.2 Hierarchical techniques. These techniques form binary classification trees, termed dendrograms, over object sets. Leaves of the tree represent individual objects and internal nodes represent object clusters. These techniques can be further divided into agglomerative and divisive techniques, which construct the dendrogram from bottom-up and top-down, respectively. Hierarchical techniques involve considerably more computations than optimization techniques.

4.3.3 Clumping techniques. These techniques return clusters in which constituent clusters may overlap. The possibility of overlap follows from independently considering some number of clusters as possible hosts for an object. A problem with some clumping techniques is that several renditions of the same object set
may be obtained. Dual scaling can be viewed as a "clumping" technique. As dual scaling projects the rows or columns of a table of categorical data as points in orthogonal sub-spaces, the objects or variables under study can contribute to more than one dimension.

4.4 The input matrix for cluster analysis

Generally, cluster analysis models attempt to place the objects under study into groups in such a way as to best represent measures of similarity among them. Most cluster analysis methods based on numerical taxonomy (Everitt, 1980) operate on what is generally called proximity data. This is a general term for any type of data that can be interpreted as similarities or distances among pairs of objects. Proximity measures include: Correlations between pairs of variables across subjects; direct ratings of similarity; average probabilities of confusion between pairs of objects; frequency of substitution of one object for another in certain tasks. Ratings of dissimilarity, travel or retrieval times between two "objects," are examples of types of data that can be interpreted as distance or dissimilarity data. However, unlike these types of data, the dominance numbers in a dominance matrix convey only ordinal information about the distances between the parameters (that is, category boundaries plus stimuli). As such, the rows of the
two-way dominance matrix can be converted to a symmetric matrix of "dissimilarities."

In the present study, we will employ the most frequently used similarity or dissimilarity index, a distance-function coefficient (Overall & Klett, 1972). A distance-function coefficient may be derived from the well-known Minkowski metric (see, e.g., Jain & Dubes, 1988) as follows. Consider subjects $j$ and $k$, with profile vectors of dominance numbers (real integers), $y_{ji}$ and $y_{ki}$, respectively, on a dimension $i$. Then, the Minkowski metric,

$$d_{jk}^{(p)} = \left( \sum_{i=1}^{n} |y_{ji} - y_{ki}|^p \right)^{1/p}$$

is a measure of the distance between a subject $j$ and a subject $k$ on a dimension $i$. Euclidean distance is defined when $p=2$. (However, $p$ can be any positive real number for other variants of the distance measure.) A distance-function similarity or proximity index can then be defined such that the distance between two subjects (profiles) is the sum of squared differences in values for each item (the column entity). For example, the index of dissimilarity between a subject $j$ and a subject $k$ measured over $n$ stimuli in a dominance matrix can be calculated as follows:

$$d_{jk} = \sum_{i=1}^{n} (y_{ji} - y_{ki})^2$$
Euclidean distance as a measure of profile similarity or dissimilarity is popular with cluster analysis studies carried out on constructed data (e.g., Blashfield, 1976; Gross, 1972; Kuiper & Fisher, 1975).

A useful property of Euclidean distance is that a similarity (or dissimilarity) value may reflect an infinite combination of differences in profile shape, elevation (e.g., arithmetic mean), and scatter (e.g., variance). So, two profiles with the same elevation may appear dissimilar due to differences in shape or scatter, whereas two profiles with identical shape and scatter may appear dissimilar due to a difference in elevation. Because it is so sensitive to different aspects of the profiles, the Euclidean distance is the most appropriate choice for creating a dissimilarity matrix from a two-mode (subjects-by-parameters) dominance matrix.

4.5 Choosing a clustering method

4.5.0 Introduction. In this study, a clustering technique is employed as a tool that makes possible the discovery and the grouping of subjects with similar response patterns. Apparently, many clustering methods exist that have been described in different ways in different sciences (Anderberg, 1973). However, rarely is there
any a priori theory or guidelines for selecting a clustering procedure to use for a specific social science problem.

4.5.1 Guiding criteria for a cluster analysis method. In this study, the kind of (1) classification wanted, (2) representation required, and (3) data to be analyzed served as the guiding criteria for choosing a clustering technique to use with DSMASC. Let us begin with the kind of classification wanted. As dual scaling is essentially a technique for exploratory analysis of categorical variables, whatever structure underlies the data is revealed by the analysis. As such, a clustering technique based on some kind of optimization to define clusters would be appealing.

As to the kind of representation needed, the question often asked in cluster analysis concerns the number of clusters necessary to represent the underlying structure in the data adequately. In this study, the purpose of clustering was not to identify and characterize clusters of subjects but to sort these into homogeneous groups. Consequently, the question of how many clusters necessary to represent the underlying structure was not critical. Instead, it was the property of being able to define exclusive clusters, systematically, that influenced the decision to use a
traditional cluster analysis method. At least the cluster analysis computer program to be sought would be one in which the user specifies the number of clusters.

As to the kind of data to be analyzed, the dominance matrix E was to be clustered by the rows. In this matrix, dominance numbers of the category boundaries satisfy the order constraint within subjects. A row of E represents the ranking of the parameters (stimuli and category boundaries) by a subject. The columns of E are partially ranked as to category boundaries, and therefore, may not be clustered. The inter-subject (inter-row) distances calculated from a dominance matrix provide a suitable dissimilarity matrix to submit to cluster analysis.

Traditional hierarchical clustering methods employing various types of agglomerative or divisive rules (see Hartigan, 1975, for a review) such as single-linkage, average-linkage, complete-linkage, centroid-linkage or median-linkage are available in most statistical software packages. However, these different hierarchical procedures, when applied to the same dissimilarity matrix, typically lead to different tree structures, as has been well documented in classification literature (Jain & Dubes, 1988). Second, rarely do any a priori theory or guidelines exist for choosing a hierarchical clustering procedure to apply to a specific social science research problem.
Ward's (1963) method purportedly finds good (separated) partitions (see, e.g., Blashfield, 1976); however, it does not define exclusive clusters that we seek to define on the rows of a dominance matrix. For these reasons, traditional hierarchical clustering procedures were not considered for use in this study.

Clumping clustering techniques generally provide 'inclusive' clusters in which an object is a member of two clusters simultaneously. Our decision to define exclusive clusters technically excludes clumping techniques from consideration. This leaves us with Partitional clustering techniques to consider.

Partitional clustering methods are used more frequently in applications where single partitions are important. They are especially appropriate for the efficient representation and compression of large data bases. The partitional clustering problem has been formally stated as follows (Jain & Dubes, 1988, p. 90). Starting with \(n\) patterns (subject profiles) in a \(d\)-dimensional space, we partition the patterns (subject profiles) into \(K\) groups, or clusters, such that the patterns in a cluster are more similar to each other than patterns in different clusters. The value of \(K\) may or may not be specified. Once the data are partitioned, the resulting partitions are then modified into "better" partitions.
The way "better" is defined leads to various methods. A criterion must be selected, and from among all possible partitions containing \( K \) clusters, the partition that optimizes the criterion is accepted. The most commonly used partitional clustering strategy is based on the squared-error criterion. A general objective is to obtain that partition that, for some clusters, minimizes the squared error. This is conceived from analysis of variance (ANOVA) methodology, wherein partitioning leads to clusters for which the sum-of-squares-within (SS\(_w\)) is minimum and the sum-of-squares-between (SS\(_b\)) is maximum.

A computer program written for a partitional clustering technique may not be commonly found in the marketplace. However, such a program is more likely to be found as part of a statistical package. One such computer program called QUICKCLUSTER is available in the SPSS/PC+ statistical package (SPSS inc., 1988, p.841). Due to ease of its accessibility, QUICKCLUSTER was chosen for use in this study. In the next section, a brief description of the QUICKCLUSTER computer program is given.

4.6 Description of the QUICKCLUSTER computer program

This program does cluster analysis in three steps. In the first step, initial cluster centers (seed points) are selected. A center is an estimate of the average value of
each clustering variable for the cases in a cluster. Thus, a center includes one value for each variable. The user can supply the initial centers. Alternatively, the user can request QUICKCLUSTER to select \( k \) cases of well-separated, non missing values as initial cluster centers.

In the second step, the values of initial cluster centers are updated to derive the classification cluster centers. When a case (a subject profile) is assigned, the center is updated to a mean for the cases that are thus far in the cluster. As the cases are processed, the centers move to concentrations of observations (subject profiles). In the third step, each case is reassigned to the nearest of the updated cluster centers, yielding the final clusters. The QUICKCLUSTER procedure uses squared Euclidean distance, which equally weights all clustering variables. The QUICKCLUSTER computer program was chosen due to (1) the investigator's familiarity with this computer program, and (2) the fact that this program was readily available. QUICKCLUSTER can handle both interval and ratio data.
CHAPTER 5

THE PROPOSED DSMASC PROCEDURE

5.0 Introduction

Using a combination of cluster analysis and a scaling procedure to explore data structure is not a completely new idea. Napier (1972) presented a strategy for multidimensional analysis of metric data in which a data matrix was first clustered by a clustering technique and then the resulting clusters subjected to multidimensional scaling (MDS). Napier (1972) suggested that analyzing the interrelations among objects in the clusters might provide a critical clues about data structure. In fact, he viewed the application of MDS to clustered data as a valuable strategy to ‘fine-tune’ the analysis.

In the points of view analysis (PVA) proposed by Tucker and Messick (1963) and the individual differences scaling problem addressed by Meulman and Verboon (1993), cluster analysis has an apparent role to play:

“... it is the purpose of the present paper to show there is still room for the PVA model, explicitly when one is interested in finding subsets of individuals (clusters
of sources) that have the same frame of reference. Thus, PVA is truly different from doing separate scalings because it is not known a priori which sources belong to the same point of view; therefore, we will perform a clustering task that assigns the sources to different points of view.” (p.9)

The problem of scaling subjects' responses is particularly challenging. It would be rare if all subjects in a large group of people "perceived" stimuli in the same way. It would also be rare if all subjects “perceived” stimuli using no dimensions in common. Characterization of individual responses takes two extreme approaches (Green and Carmone, 1970):

(1) One extreme would be to assume homogeneity of responses across people and then, by simple aggregation, find one configuration that purportedly typifies the group.

(2) The other extreme, would be to develop each individual's configuration separately, not only an unwieldy task but also an approach that militates against any form of generalization (p.61).
The approach in (1) defines clusters with many members whereas approach (2) defines clusters with a single member. To derive the benefits of both approaches, Green and Carmone (1970) suggested a middle course approach:

"...to search for 'points of view' in which each point of view may be shared by a number of subjects, but, in turn, is presumed to be different from other points of view. That is, if homogeneity is found by analysis - not ordained by assumption - then aggregation over subjects can be done in a more reasonable manner"(p.61).

Nishisato (1986) proposed finding homogeneity by three clustering procedures operating on inter-subject distances calculated from subject scores obtained by dual scaling without the order constraint imposed on category boundaries. Following Nishisato's approach to the problem, the present study seeks to find homogeneity by a traditional cluster analysis method. Each homogeneous group of subjects is then analyzed by dual scaling to find the 'optimal' configuration to describe the members of the group. A distinct group identified in this way could be related to some identifiable characteristic or feature, (for example, socioeconomic or demographic) of the subjects. The motivation for employing a cluster analysis method, therefore, is partly due to the desire to discover relations
that might exist between a feature and a group of subjects. The assumption here is that each group of subjects is determined by some feature or a set of features that the subjects in the group share in their responses to stimuli.

5.1 The proposed method

As we have seen in Chapter 2, a \textit{subjects-by-stimuli} data matrix is transformed into a \textit{dominance} matrix, which is the appropriate input matrix for dual scaling. The rows of the dominance matrix sum to zero (since these elements are calculated within subjects). Let us call this matrix the initial \textit{dominance matrix} and outline the DSMASC methodology as follows.

An idea is to first obtain a set of dual scaling solutions from the initial \textit{dominance matrix}. This set of solutions is used to obtain an approximation to the initial \textit{dominance matrix}; we call an approximation matrix the \textit{constructed matrix}. The difference between the initial \textit{dominance matrix} and the \textit{constructed matrix} is called the \textit{residual matrix}. This matrix can be analysed by dual scaling just like the initial dominance matrix. When a residual matrix yields an \textit{admissible solution} and the effects of this solution are removed, then another residual matrix results. The
next set of dual scaling solutions may be extracted from this residual matrix. Let us define the following notation:

\( \mathbf{B} = \) Subjects-by-stimuli data matrix

\( N = \) number of subjects (rows) in \( \mathbf{B} \)

\( n = \) number of stimuli (columns) in \( \mathbf{B} \)

\( m = \) number of category boundaries for successive category scale used to generate entries in \( \mathbf{B} \)

\( f_i = N(n+m)(n+m-1) \), the total number of comparisons for \( \mathbf{E} \)

\( \mathbf{E} = \) Dominance matrix derived from \( \mathbf{B} \)

\( \mathbf{E}_{\text{resd}} = \) Residual dominance matrix

\( \mathbf{E}_{\text{appx}} = \) Constructed dominance matrix

\( \mathbf{x} = \) Optimal solution (vector of parameter weights)

\( \mathbf{y} = \) Optimal subject scores (vector of subject weights)

\( D_r = \) Diagonal matrix of "total row comparisons" for \( \mathbf{E} \)

\( D_c = \) Diagonal matrix of "total column comparisons" for \( \mathbf{E} \)

\( \mathbf{C}_1 = \) Matrix for decomposition derived from \( \mathbf{E} \).

\( \eta^2 = \) Squared correlation ratio - the ratio of the between-row sum of squares to the total sum of squares generated by \( \mathbf{x} \) from \( \mathbf{E} \).

\( \rho = \) Positive square-root of \( \eta^2 \)

\( \delta\% = \) Percentage of variance in \( \mathbf{E} \) explained by solution \( \mathbf{x} \)
$tr(C_1) = \text{Total variance (trace of } C_1\text{), which is the sum of diagonal elements of } C_1$

Note that $tr(C_1) = \sum_{l=1.2, \ldots}^{n-1, n+m-1} \eta^2_l$. which is the "Total variance to be accounted for (TVAF)" in $C_1$ (the matrix for decomposition or the matrix for eigenequation), has an upper bound (see, e.g., Nishisato, 1993: 1994) of

$$\text{(5.8) } tr(C_1) = \frac{n+m+1}{3(n+m-1)}.$$

Now, the structural equation for the dominance matrix $E$ (see, e.g., Nishisato, 1994), can be expressed as

$$\text{(5.9) } E = \frac{1}{f_t} D_x (\sum_{k=1}^{\ell} \rho_{k} y_{k} x_{k}^{\ell}) D_e$$

where $t$ is the rank of $E$ minus one or $t=\tau(C_1)$. This equation, also called the formula for reconstitution (Bénzecri et al., 1973), provides the means to approximate the input data (here, the dominance matrix $E$) by a subset of solutions. However, these solutions do not always satisfy the order constraint on category boundaries, especially when more than two category boundaries are involved (see the numerical example in Chapter 2). Therefore, to extract multiple
solutions with correctly ordered category boundaries, a new approach based on some modification of the SDSS is proposed.

It has been pointed out that subsets of the rows of the dominance matrix \( E \) may be associated with the principal axes necessary to summarize the information in a given data set statistically (see, e.g., Nishisato, 1986). Similarly, the rows of a residual matrix, \( E_{(resd)} \) with a substantial information left to be explained might be associated with principal axes yet to be discovered. This suggests the possibility that a cluster of rows of a dominance matrix or a residual matrix could represent an interesting structure to be captured by dual scaling. A cluster analysis method is used to search for clusters (or subsets) of rows of the residual matrix that would yield an admissible solution. A solution so obtained would be one that accounts for the largest possible variance in the residual matrix, \( E_{(resd)} \).

Although dual scaling has been previously applied to successive categories, the idea of approximating the structure of a data matrix with solutions derived from its sub-matrices is new. The salient steps in the DSMASC can be summarized as follows:
Using SDSS, extract as many solutions as possible until a given solution vector $x_k$ violates the order requirement on category boundary values.

Construct a matrix approximation, $E_{\text{app}(k)}$ (APM) to the input matrix (IDM) or any of the subsequent residual matrices from the solution set produced by the input matrix.

Obtain a residual matrix, $E_{\text{res}(k)}$, (that is, the data matrix from which the effects of the first $k$ admissible solutions have been removed) using the formula, $E_{\text{res}(k)} = E_k - E_{\text{app}(k)}$, where $k=1,2, \ldots t$, and $t = \text{tr}(C_1)$ or rank of $C_1$ (the matrix for the eigenequation in section 2).

If the input matrix $E_k$ or the residual matrix $E_{\text{res}(k)}$ produces a solution that violates the order constraint on category boundaries, do any of the following:

(i) If category boundaries are entirely reversed, reflect the solution vector to obtain the admissible solution;
(ii) If the admissible solution cannot be obtained by reflection, cluster the rows of the dominance matrix that has produced an inadmissible solution. The clusters obtained (c) are sub-matrices of $E_{rest(k)}$. Each of these sub-matrices is then analysed by dual scaling to produce $c$ sets of solutions.

(5) Suppose that of the $c$ solutions derived from $c$ clusters of rows, $l$ are admissible. Then, we take each of the $l$ admissible solutions, in turn, and orthogonalize it with respect to any preceding admissible solutions. Now, of the orthogonalized $l$ solutions, we select those that are admissible.

(6) Now we calculate the percentage of information explained by each of the $l$ solutions that satisfy the order constraint. For example, the contribution of the solution, $x_{k+1}$, toward accounting for the information in the input matrix, $E_{k-1}$ can be calculated as follows.

First the proportion of the total variance to be accounted for (TVAF) in the input matrix $E_{k-1}$ by solution $x_{k+1}$ is calculated as the squared-correlation-ratio, $\eta_{k+1}^2$, using the formula
Of the $l$ solutions, the one that accounts for the largest possible variance (squared-correlation-ratio) in the residual matrix is designated as the next admissible solution. Then, a subject score vector, $y_{k+1}$, corresponding to the solution, $x_{k+1}$, is calculated using the "dual relations" formula (see, e.g., Nishisato, 1980b),

$$y_{k+1} = \frac{1}{\rho_{k-1}} D_r^{-1} E_0 x_{k-1}$$

where $\rho_{k+1}$ is the square-root of the squared-correlation-ratio. Now, the structure of the residual matrix $E_{\text{resd}(k)}$ (RM) represented by the solution set given by $[\rho_{k+1}, y_{k+1}, x_{k+1}]$ is in the constructed matrix $E_{\text{app}(k+1)}$. The residual matrix, $E_{\text{resd}(k+2)}$ (RM) is obtained as

$$E_{\text{resd}(k-2)} = E_{\text{resd}(k)} - E_{\text{app}(k+1)} = E_{\text{resd}(k)} - \frac{D_r (\rho_{k+1} y_{k+1} x_{k+1}')} {f_t} D_c,$$
where \( f_i = N(n+m)(n+m-1) \), the total number of comparisons that lead to the elements of the approximation matrix, \( E_{app(k+1)} \).

(7) As before, \( E_{res(k+2)} \) is analysed by dual scaling to obtain the solution vector, \( x_{k+2} \). This vector becomes the next solution if it is admissible. Otherwise, this residual matrix is submitted to cluster analysis as in (4) above.

Usually, the percentage of TVAF in the input matrix \( E \) by solution a solution set \([\rho_x, y_x, x_k]\) designated as \( \delta_k \) is given by

\[
\delta_k = \frac{\eta^2_k}{\text{tr}(C'_k)} \times 100.
\]

The entire DSMASC procedure can be summarized in the following algorithm.

**5.1.1 Summary of DSMASC procedures**

**Step 1:** Let \( j=0, c=1 \). Convert a data matrix \( B \) into a table of "choice frequencies" or "dominance matrix," \( E_j \).
Step 2: Submit $E_j$ to dual scaling to get a solution vector $x$, subject scores vector $y$, the associated squared-correlation-ratio $\eta^2$, and the percentage of information explained by this solution, $\delta\%$. If $x_j$ satisfies the order requirement on category boundary values, designate the vector as an admissible solution; if $x_j$ shows a reflected admissible solution, reflect the vector and designate it as an admissible solution. Designate this solution as $j^{th}$ solution and go to Step 3. Otherwise, let $c = c+1$ and go to Step 5.

Step 3: Calculate $E_j$(appr) (APM), the matrix whose structure is represented by $x$, $y$ and $\eta^2$ obtained in Step 1 and go to Step 4.

Step 4: Remove the structure represented by $x_j$, $y_j$ and $\eta^2$ obtained in Step 1 by subtracting $E_j$(appr) obtained in Step 2 from $E_j$ calculated in Step 1. Let $j=j+1$ and designate the residual matrix as $E_j$. (This is the matrix $E_{j-1}$(residual)). Execute Step 2.

Step 5: Submit $E_j$ to cluster analysis. Divide $E_j$ into $c$ clusters. Submit each of the $c$ clusters to dual scaling. Orthogonalize each of the $c$ cluster-derived admissible solutions with respect to any preceding
admissible solutions. Select the orthogonalized solutions that are admissible and calculate their respective $\eta^2$ and $y$ for $E_j$. Select the solution with the largest $\eta^2$ and designate it as the next solution. If the cumulative value of $\eta^2$, for the admissible solutions so far determined, is small, go to Step 2. Otherwise, stop.

5.2 Descriptions of Computer programs

Four computer programs were written in QBASIC programming language to carry out calculations in some stages of the DSMASC procedure. All listings for computer programs appear in the Appendix section. These programs were set to handle data sets with up to a maximum of eighty subjects, nine stimuli and five successive categories. However, with adequate computer memory available, these dimensions could be increased.

5.2.1 DESIGN.BAS. With this program, a subjects-by-stimuli data matrix is converted into a dominance matrix. The following information must be supplied:

- Name of computer file where the subjects-by-stimuli data matrix is stored
- Number of stimuli
- Number of successive categories
5.2.2 *DSOFE.BAS.* With this program, a dual scaling analysis is carried out on a dominance matrix resulting in (1) parameter weights, (2) subject scores and (3) the squared-correlation-ratio. The following must be supplied:

- Name of computer file where the dominance matrix is located
- Number of rows (subjects) of the dominance matrix
- Number of stimuli
- Number of successive categories
- Name of computer file where optimal weights are to be stored (*.x)
- Name of computer file where subject scores are to be stored (*.y)

The squared-correlation-ratio has to be recorded from the screen output.

5.2.3 *RESIDUAL.BAS.* With this program, the following are calculated: (1) the variance component, \( \eta^2 \) associated with an admissible solution \( x \), (2) the vectors of subject scores \( y \), and (3) the residual matrix (RM) left after the removal of the effects of a solution set \( [\eta, x, y] \) from an input dominance matrix. The following information must be supplied:

- Name of computer file where the input dominance matrix is located
- Number of rows (subjects) of the dominance matrix
- Number of stimuli
- Number of successive categories
- Name of computer file where optimal weights are located (*.x)
- Name of computer file where subject scores are located (*.y)

5.2.4 **ORTHOG.BAS.** With this program, cluster-derived dual scaling solutions are orthogonalized. Two computer files must be first created: (1) a file with cluster-derived solutions, and (2) a file with the appropriate matrix for eigenequation $C_1$. The program prompts the user to supply the following information:

- Name of the computer file containing the cluster-derived solutions
- Number of cluster-derived solutions
- Name of the computer file containing matrix $C_1$ for the data set under consideration
- Number of observations (subjects) in the data under consideration
- Number of successive categories
- Number of stimuli
- Name of the output file

The output from this program is an orthogonal series of solutions. The proportion of variance explained by each solution for the dominance matrix analysed is also given.
5.3 Summary of Chapter

In this chapter, the proposed method for multidimensional analysis of successive categories data with dual scaling has been outlined. In this method, a dominance matrix not yielding an admissible solution is to be subjected to cluster analysis. The resulting clusters are to be submitted to dual scaling. Cluster-derived solutions are subsequently be mutually orthogonalized, and a statistic showing their relative contributions to the total variance reported. This process forms the strategy for exploring structures in clusters that lead to the extraction of multiple admissible dual scaling solutions. Computer programs for carrying out calculations at various stages of the proposed method have also been described.
CHAPTER 6

ANALYSIS AND RESULTS

6.0 Introduction

This chapter contains three sections. In Section I, a method used to generate an artificial successive categories (SC) data set is described. This data set was used to illustrate the DSMASC method. In Section II, the results of multidimensional analysis are presented for (i) the artificial SC data set, (ii) a comparison of SDSS and DSMASC using two small numerical examples, and (iii) an empirical educational data set.

SECTION I

6.1 Construction of an Artificial SC Data Set

6.1.1 Introduction. Successive categories data can be considered as rank-order data of stimuli ($\mu_i$) and category boundaries ($\tau_k$). As outlined in section 2.3.1, a matrix of SC data is first transformed into a subjects-by-parameters rank-order
table before being submitted to dual scaling. The rank-order table yields a set of 

*projected weights* for the parameters (category boundaries plus stimuli),

represented by the vector, $\theta$; a set of *projected normed weights* for the subjects,

represented by the vector, $y$; and the associated squared-correlation-ratio, $\eta^2$.

whose positive square root is denoted by $\rho$. Such a solution is, in fact, a least-
squares type approximation of the correct ranking of parameters by the subjects.

In the total space (multidimensional), spanned by all possible dual scaling

solutions, the parameters are ranked according to their relative distances from a

subject's position in the space. This model of data structure corresponds to a

multidimensional generalization of Coomb's method of unfolding (see, e.g.,


---

**Coombs' Multidimensional Unfolding Model.** Coombs' model (Bennett & Hayes, 1960; Coombs, 1964; Coombs & Kao, 1960; Hayes & Bennet, 1961) assumes a

"joint space" of *stimulus points and person points*. In the unidimensional case, the

model supposes that both *subjects* and *stimuli* can be represented by points on a

line segment (the single attribute under investigation). Each subject ranks all of the

stimuli in order of their directions. A subject is deemed to pick up the continuum at

his own position as one might pick up a string, letting the ends swing together and

fuse. The analytic task is the unfolding of these rankings, to recover the original
ordering of subjects and stimuli on the continuum. The method used to accomplish this task requires no definition of "zero," "addition," or the "unit interval," and in fact employs no properties of the real line except the ordering of its points and the comparability of (some) intervals in size. Internal checks provide that even these assumptions will justify themselves in practice; that is, the unfolding method is a scaling criterion and a scaling method, and any given set of rankings may or may not unfold.

**Unfolding in one-dimensional space.** In the unidimensional case explored by Coombs, the whole one-dimensional space is segmented by the midpoint into regions, within each of which every subject will give the same ranking of the stimuli under study, although within any one region different subjects' ideals may lie at different distances from the stimuli.

Now consider what position in space a subject's ideal must occupy to present a particular ranking of two response variables, for instance, \( \tau_1 \) and \( \tau_2 \). About these variables, a subject may make one of three judgements: \( \tau_1 \) is above \( \tau_2 \), symbolized \( \tau_1 > \tau_2 \); the converse \( \tau_1 < \tau_2 \); and third, the subject is undecided between them, symbolized \( \tau_1 = \tau_2 \), which means that the subject's ideal is equidistant from \( \tau_1 \) and \( \tau_2 \). In one-space (that is, an infinite straight line), only the midpoint on a straight
line segment that joins the points, \( r_1 \) and \( r_2 \), is exactly equidistant from these points. All the points on the infinite straight line to one side of that midpoint will be closer to one stimulus, and all on the opposite side will be closer to the other stimulus. Only a subject with an ideal lying exactly on the midpoint would judge \( r_1 = r_2 \), that is, indecision between the two. (See Figure 2.)

![Diagram showing midpoints and regions generated by two stimulus points.](Figure 2. Midpoints and Regions Generated by Two Stimulus Points.)

*Unfolding in a two-dimensional space.* In a two-dimensional space (a plane), the locus of equidistance from two response variables is a line, the perpendicular bisector (mediator) of the line segment connecting those response variables. The subjects (individuals) on one side of this line, the side containing the response variable \( r_1 \), will judge \( r_1 > r_2 \); all subjects on the opposite side will judge \( r_1 < r_2 \); only those exactly on the line will judge \( r_1 = r_2 \). With three response variables \((r_1, r_2, r_3)\) in a two-dimensional space not all on the same line, there will be three perpendicular line bisectors, \( M(r_1r_2) \), \( M(r_1r_3) \), \( M(r_2r_3) \), which will meet at a point equidistant from all three response variables, \((r_1, r_2, r_3)\). (See Figure 3.)
6.1.2 Solving Coombs' problem by dual scaling. Dual scaling provides a practical method to solve Coombs' problem. The weight vector for the subjects, \( y \), can be derived from the weight vector for the parameters, \( \theta \), through the 'dual relations' formula, \( y = (1/\rho)D_r^{-1}E\theta \), where \( \rho \) is the square root of the squared correlation ratio; \( D_r \) is the diagonal matrix for the row totals of \( E \) (dominance matrix). (See section 2.3.4.3.) By the dual relations formula, the ideal position of a subject, say, \( y_i \), is a projection of the subject's position on the stimulus sub-space and vice-versa.

In Coombs's (1964) model, a rank order of a stimulus is generated from the proximity of the subject's position, \( y_i \), to the projection of a stimulus onto the subject sub-space. If we call this projection \( px \), then the proximity in question is the absolute difference, \(|px - y_i|\). Suppose the ideal position of Subject \( i \) (\( y_i \)) is closer to the position of Stimulus \( k \) (\( x_k \)) than the position of Stimulus \( r \) (\( x_r \)). Then, Subject \( i \) has judged Stimulus \( k \) to be above Stimulus \( r \).

Now, dual scaling determines \( \rho \), \( x_k \), \( x_r \), and \( y_i \) so that

\[ |px_k - y_i| < |px_r - y_i| \]

whenever the rank of Stimulus \( k \) is lower than the rank of Stimulus \( r \).
Figure 3. Boundary Mediators Generated by Three Response Variables in Two Dimensions.

The inequality $|px_k - y_i| < |px_r - y_i|$ holds true only with respect to the distance relations in the multidimensional space, not with respect to a single solution (Nishisato. 1994). Thus, all possible dual scaling solutions, however small the variance each solution may explain, are necessary to reproduce the stimuli and category boundary rankings correctly.
6.2 Construction of artificial data

A hypothetical successive categories data set with ten subjects (N=10), four categories (m+1=4) and five stimuli (n=5) was constructed. A graphical method was used to accomplish this task. For the sake of simplicity, a two-dimensional (total) space was considered. As in Coombs's model, the graphical method employed required no definition of "origin," or the "unit interval." In fact, the method uses no properties of the real number line; the exception is the ordering of its points and the comparability of intervals in size.

6.2.1 Plotting category boundaries and stimuli. Employing Coomb's model, an arbitrary set of stimuli and category boundaries was first plotted on the two-dimensional space in Figure 4. Because of the order constraint on them, the category boundaries, τ1, τ2 and τ3 were plotted first. Then, line segments joining adjacent category boundaries were drawn. This was followed by the drawing of perpendicular line bisectors, M(τ1,τ2) and M(τ2,τ3). Finally, the points representing the stimuli, X1, X2, X3, X4, X5 were distributed in the two-dimensional space. The positions of stimuli with respect to category boundaries were varied to create different structures.
6.2.2 **Plotting subjects.** To maintain the order constraint on the category boundaries, the subjects' points were plotted only in the area marked **PERMISSIBLE REGION** in Figure 4. This area is bounded by the perpendicular line bisectors labeled $M(r_1, r_2)$ and $M(r_2, r_3)$. The permissible region was defined as the area from which any subject point, $Y_i$, would always be closest $r_1$, followed by $r_2$, and then $r_3$, in this order. This condition would ensure that if $r_1$, $r_2$, and $r_3$ were assigned sequential numbers according to their respective measured distances from a subject point, $Y_i$, the order relation, $r_1 < r_2 < r_3$ would always be satisfied. For example, in Figure 4, Subject $Y_8$ is 3.00 arbitrary units from $r_1$, 4.25 arbitrary units from $r_2$ and 6.50 arbitrary units from $r_3$. If the three distances were assigned sequential numbers from 1 to 3, with the shortest distance receiving the lowest number, then the order relation, $r_1 < r_2 < r_3$ would be satisfied.

6.2.3 **Distances and ranks.** The graph in Figure 4 was used to generate distances between subject points (i.e., "an ideal point I" in Coomb's unfolding technique: Coombs, 1964) and the parameter points (stimuli and category boundaries). For convenience, the distances were measured in centimeters, otherwise treated as arbitrary units.
Figure 4. A two-dimensional plot of arbitrary points for subjects, stimuli and category boundaries.
To illustrate the process of generating rank-order data for category boundaries and stimuli, consider the subject point, \( Y_8 \) in Figure 4. The measured distances from this subject point to various category boundaries and stimuli were as follows: \( Y_8 \) to \( \tau_1 \) was 3.00, \( Y_8 \) to \( \tau_2 \) was 4.25, \( Y_8 \) to \( \tau_3 \) was 6.50. \( Y_8 \) to \( X_1 \) was 6.40, \( Y_8 \) to \( X_2 \) was 1.95. \( Y_8 \) to \( X_3 \) was 5.95. \( Y_8 \) to \( X_4 \) was 6.10. \( Y_8 \) to \( X_5 \) was 5.60. These distances were arranged into a sequential order. For example, if we denote the distance between a subject’s ‘ideal position’ \( Y \) to a stimulus point, \( X \) by \( D(Y, X) \), then the distances for Subject \( Y_8 \) (see Table 6-1) would be arranged in increasing order as follows: \( D(Y_8, X_2) < D(Y_8, \tau_1) < D(Y_8, \tau_2) < D(Y_8, X_5) < D(Y_8, X_3) < D(Y_8, X_4) < D(Y_8, X_1) < D(Y_8, \tau_3) \). This ordering corresponds to the sequential numbers (in parentheses in Table 6-1), (1), (2), (3), (4), (5), (6), (7), (8). The rest of Table 6-1 was constructed in a similar manner.

Consideration was given to a potential problem posed by the possibility of tied ranks. A subject point, \( Y_i \), placed the same distance from two adjacent category boundary points, \( \tau_k \) and \( \tau_{k+1} \), (that is, when \( D(\tau_k, Y_i) = D(\tau_{k+1}, Y_i) \)) would imply a diminished category. Also, a subject point, \( Y \), placed the same distance from a category boundary point, \( \tau_k \), and a stimulus point, \( X_j \) (that is, when \( D(\tau_k, Y_i) = D(X_j, Y_i) \)) would preempt the classification of \( X_j \) into a category, when \( X_j \) and \( \tau_k \)
Table 6-1. Inter-point (parameter to subject) distances for the graph in Figure 4.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Parameters</th>
<th>Category boundaries</th>
<th>Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_1$</td>
<td>$\tau_2$</td>
<td>$\tau_3$</td>
</tr>
<tr>
<td>$Y_1$</td>
<td>3.50</td>
<td>4.85</td>
<td>6.05</td>
</tr>
<tr>
<td></td>
<td>(3)*</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>2.70</td>
<td>3.20</td>
<td>4.35</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>3.75</td>
<td>4.20</td>
<td>5.05</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>3.90</td>
<td>4.10</td>
<td>6.05</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(6)</td>
</tr>
<tr>
<td>$Y_5$</td>
<td>3.55</td>
<td>4.25</td>
<td>6.35</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(7)</td>
</tr>
<tr>
<td>$Y_6$</td>
<td>2.96</td>
<td>3.45</td>
<td>5.60</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(6.5)</td>
</tr>
<tr>
<td>$Y_7$</td>
<td>1.95</td>
<td>3.45</td>
<td>5.60</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
<td>(6)</td>
</tr>
<tr>
<td>$Y_8$</td>
<td>3.00</td>
<td>4.25</td>
<td>6.50</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(8)</td>
</tr>
<tr>
<td>$Y_9$</td>
<td>2.60</td>
<td>2.70</td>
<td>4.90</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(3)</td>
<td>(6)</td>
</tr>
<tr>
<td>$Y_{10}$</td>
<td>2.30</td>
<td>2.95</td>
<td>3.85</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
</tbody>
</table>

*Number in a parenthesis is the rank assigned to the parameter next to it.
occupy the same position on a common scale. This anticipated problem was preempted by allowing no ties in rank between any two category boundaries or between a stimulus and a category boundary. A tied rank was eliminated by shifting a subject point within the permissible region until no two or more category boundary points were the same distance from a subject’s position. This also applied to a category boundary and a stimulus point the same distance away from a subject’s position.

A rearrangement of Table 6-1 gave Table 6-2. In this table, the subjects are in the rows and the numbers sequencing the category boundaries and the stimuli are in the columns. An example of what Table 6-2 shows is as follows. The sequence of the column variables for Subject Y8 is X2 (1), τ1, τ2, X5 (3), X3 (3), X4 (3), X1 (3). τ3, where the numbers in brackets indicates category labels. For example, the stimulus, X2 is in Category 1, whose upper bound is τ1. No stimulus is in Category 2, whose lower bound is τ1 and upper bound, τ2. Also, no stimulus is in Category 4, whose lower bound is τ3. Four stimuli, X5, X3, X4 and X1 are in Category 3, whose lower bound is τ2 and upper bound, τ3.

A further rearrangement of Table 6-2 gave Table 6-3, which shows the generated *subjects-by-stimuli* matrix of successive categories data.
Table 6-2. *Subjects-by-rank* table for distances in Table 6-1.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Assigned sequential numbers or ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$Y_1$</td>
<td>$X_1$ (1)*</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$X_2$ (1)</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$X_3$ (1)</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>$X_2$ (1)</td>
</tr>
<tr>
<td>$Y_5$</td>
<td>$X_2$ (1)</td>
</tr>
<tr>
<td>$Y_6$</td>
<td>$X_2$ (1)</td>
</tr>
<tr>
<td>$Y_7$</td>
<td>$X_4$ (1)</td>
</tr>
<tr>
<td>$Y_8$</td>
<td>$X_2$ (1)</td>
</tr>
<tr>
<td>$Y_9$</td>
<td>$X_2$ (1)</td>
</tr>
<tr>
<td>$Y_{10}$</td>
<td>$X_3$ (1)</td>
</tr>
</tbody>
</table>

*Number in a parenthesis is the label of the category in which a stimulus ($X_i$) has been placed.*
Table 6-3. A 10-by-5 SC data generated from the graph in Figure 5.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
<th>X₄</th>
<th>X₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y₁</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y₂</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y₃</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y₄</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Y₅</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Y₆</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Y₇</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y₈</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Y₉</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Y₁₀</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>
The columns of Table 6-3, from left to right, correspond to the stimuli, \( X_1, X_2, X_3, X_4, X_5 \), in this order. The numbers entered into Table 6-3 are labels of categories chosen by a subject for a stimulus. For example, the responses of Subject \( Y_8 \) corresponds to the pattern \( (3 \ 1 \ 3 \ 3 \ 3) \). This generated data matrix (Table 6-3) was investigated by the proposed method, DSMASC.

SECTION II

6.3 Results of analyses

*Introduction.* The DSMASC method was applied to the artificial *subjects-by-stimuli* data matrix in Table 6-3. For convenience, acronyms were invented as names for some computational matrices as follows: IDM—initial dominance matrix; APM—approximation matrix; RM—residual matrix. The cluster analysis component of DSMASC method was applied when, (1) the "parameter location" vector failed to satisfy the order constraint on category boundaries, and (2) the total variance for a data matrix had not been sufficiently accounted for.

Shown in Table 6-41 are the dual scaling results for the artificial 10-by-5 *subjects-by-stimuli* data matrix. In the section labeled \( \mathbf{E} \), it can be seen that the solution (a parameter location vector) obtained was not admissible since category boundary
values violated the order constraint \((\tau_1=0.2150 < \tau_2=0.2459 > \tau_3=-0.3702)\). Not even the weak order \(\tau_1 \leq \tau_2 \leq \tau_3\) was satisfied. This solution could not be transformed into an admissible solution by reflection. Thus, the data set (Table 6-3) that produced this solution represented a case in which the first solution did not satisfy the order constraint on category boundaries.

**Cluster analysis of Initial Dominance Matrix (IDM).** The subjects-by-stimuli data matrix was converted into the dominance matrix shown in section (B) of Table 6-41. This matrix was submitted to cluster analysis using the QUICKCLUSTER computer program from SPSSPC+ (SPSS inc., 1988, p.841).

At first, two clusters were obtained. The Quickcluster output that shows a two-cluster representation of the subjects' membership can be seen in Table 6-42. Subjects \(Y_1, Y_2, Y_3, Y_7\) and \(Y_{10}\) were assigned to the cluster labeled 1 and Subjects \(Y_4, Y_5, Y_6, Y_8\) and \(Y_9\) assigned to the cluster labelled 2. For easy identification when presenting the results of analysis, the sub-matrices of IDM corresponding to the cluster labeled 1 and the cluster labeled 2 were renamed CL1 and CL2, respectively.
Table 6-41. Dual scaling output for artificial data in Table 6-43.

(A) Raw data (subjects-by-stimuli) matrix

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Y5</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Y6</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Y7</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Y8</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Y9</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Y10</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

(B) Transformed (dominance) matrix

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>-3</td>
<td>-1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Y2</td>
<td>-3</td>
<td>-1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Y3</td>
<td>-3</td>
<td>-1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Y4</td>
<td>-5</td>
<td>-3</td>
<td>3</td>
<td>0</td>
<td>-7</td>
</tr>
<tr>
<td>Y5</td>
<td>-5</td>
<td>-3</td>
<td>5</td>
<td>1</td>
<td>-7</td>
</tr>
<tr>
<td>Y6</td>
<td>-5</td>
<td>-3</td>
<td>5</td>
<td>1</td>
<td>-7</td>
</tr>
<tr>
<td>Y7</td>
<td>-3</td>
<td>-1</td>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Y8</td>
<td>-5</td>
<td>-3</td>
<td>7</td>
<td>2</td>
<td>-7</td>
</tr>
<tr>
<td>Y9</td>
<td>-5</td>
<td>-3</td>
<td>3</td>
<td>0</td>
<td>-7</td>
</tr>
<tr>
<td>Y10</td>
<td>-3</td>
<td>-1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

(C) The total variance to be accounted for = 0.061224

(D) Accompanying statistics

SQUARED CORRELATION RATIO = 0.26080
MAXIMUM PRODUCT-MOMENT CORRELATION = 0.51068
PERCENTAGE HOMOGENITY = 26.08
DELTA (TOTAL VARIANCE ACCOUNTED FOR): PARTIAL = 64.22
CUMULATIVE = 64.22

(E) OPTIMAL WEIGHT VECTORS

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Category Boundaries</th>
<th>Stimulus values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>1.0842</td>
<td>τ_1: 0.2150</td>
</tr>
<tr>
<td>Y2</td>
<td>1.0842</td>
<td>τ_2: 0.2459</td>
</tr>
<tr>
<td>Y3</td>
<td>1.0842</td>
<td>τ_3: -0.3702</td>
</tr>
<tr>
<td>Y4</td>
<td>-1.1048</td>
<td>X_1: 0.6688</td>
</tr>
<tr>
<td>Y5</td>
<td>-0.9079</td>
<td>X_2: 1.5464</td>
</tr>
<tr>
<td>Y6</td>
<td>-0.9079</td>
<td>X_3: -1.6397</td>
</tr>
<tr>
<td>Y7</td>
<td>0.8888</td>
<td>X_4: 0.6688</td>
</tr>
<tr>
<td>Y8</td>
<td>-0.6470</td>
<td>X_5: -1.3349</td>
</tr>
<tr>
<td>Y9</td>
<td>-1.1048</td>
<td></td>
</tr>
<tr>
<td>Y10</td>
<td>1.0842</td>
<td></td>
</tr>
</tbody>
</table>
**Obtaining the first solution.** Both CL1 and CL2 yielded admissible solutions when submitted to dual scaling (see Tables 6-43 & 6-44). (The computer program DSOFE.BAS described in section 5.2.2 was used to carry out the relevant calculations.) These solutions were orthogonalized by a Gram-Schmidt orthogonalization process written in the program labeled ORTHOG.BAS in section 5.2.4. According to the output presented in Table 6-45, the percentage of variance associated with the solution yielded by CL1 was more ($\delta=97.79\%$) than that associated with the solution given by CL2 ($\delta=92.50\%$). Consequently, the CL1 solution was designated as the first solution and labeled **Solution 1**.

**Obtaining a second solution.** This process began by first removing the effects of **Solution 1** from IDM. The computer program RESIDUAL.BAS\(^3\), described in section 5.2.3, was used to calculate: The approximation matrix (APM) represented by **Solution 1**; the variance component, $\eta^2$, contributed by **Solution 1**; the residual matrix (RM) — obtained by subtracting from IDM the APM of **Solution 1**. As the results in section (B) of Table 6-46 show, the solution yielded by RM was admissible ($\tau_1 = -1.4438 < \tau_2 = -0.8054 < \tau_3 = 1.2604$).

---

\(^3\) This program applies equation 5.9 to a solution set, $[\theta, y, \rho]$ to calculate APM, which is a matrix approximation to IDM.
Table 6-42: A two-cluster representation of the dominance matrix in Table 6-41.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>CLUSTER LABEL</th>
<th>INTERCLUSTER DISTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y₁</td>
<td>1</td>
<td>.154</td>
</tr>
<tr>
<td>Y₂</td>
<td>1</td>
<td>.154</td>
</tr>
<tr>
<td>Y₃</td>
<td>1</td>
<td>.154</td>
</tr>
<tr>
<td>Y₄</td>
<td>2</td>
<td>.730</td>
</tr>
<tr>
<td>Y₅</td>
<td>2</td>
<td>4.964</td>
</tr>
<tr>
<td>Y₆</td>
<td>2</td>
<td>4.964</td>
</tr>
<tr>
<td>Y₇</td>
<td>1</td>
<td>4.536</td>
</tr>
<tr>
<td>Y₈</td>
<td>2</td>
<td>6.880</td>
</tr>
<tr>
<td>Y₉</td>
<td>2</td>
<td>.730</td>
</tr>
<tr>
<td>Y₁₀</td>
<td>1</td>
<td>.154</td>
</tr>
</tbody>
</table>
Table 6-43. Dual scaling results for the cluster, CL1.

| (A) Cluster 1 - a sub-matrix of the matrix in Part (B) of Table 6-4a |
|---|---|---|---|---|---|---|---|---|
|  | Y1 | Y2 | Y3 | Y4 | Y5 | Y6 | Y7 | Y8 |
| t1 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 |
| t2 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| t3 | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| X1 | 5  | 5  | 5  | 5  | 5  | 5  | 5  | 5  |
| X2 | 5  | 5  | 5  | 5  | 5  | 5  | 5  | 5  |
| X3 | -6 | -6 | -6 | -6 | -6 | -6 | -6 | -6 |
| X4 | 5  | 5  | 5  | 5  | 5  | 5  | 5  | 5  |
| X5 | -6 | -6 | -6 | -6 | -6 | -6 | -6 | -6 |

| (B) The dual scaling solutions for Cluster 1(OPTIMAL WEIGHT VECTORS) |
|---|---|
| \( \eta^2 = .3971772 \) : \( \delta \% = 97.79 \) |
| Subjects | Category boundaries |
| Y1 | 1.0047 | \( \tau_1 \) -0.6800 |
| Y2 | 1.0047 | \( \tau_2 \) -0.2267 |
| Y3 | 1.0047 | \( \tau_3 \) 0.3156 |
| Stimulus values | |
| Y7 | 0.9810 | X1 1.1778 |
| Y10 | 1.0047 | X2 0.9554 |
| | | X3 -1.3600 |
| | | X4 1.1778 |
| | | X5 -1.3600 |
Table 6-44. Dual scaling results for the cluster. CL2.

(A) CL2 matrix

<table>
<thead>
<tr>
<th></th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>x5</th>
<th>x6</th>
<th>x7</th>
<th>x8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y4</td>
<td>-5</td>
<td>-3</td>
<td>3</td>
<td>0</td>
<td>-7</td>
<td>6</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Y5</td>
<td>-5</td>
<td>-3</td>
<td>5</td>
<td>1</td>
<td>-7</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Y6</td>
<td>-5</td>
<td>-3</td>
<td>5</td>
<td>1</td>
<td>-7</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Y7</td>
<td>-5</td>
<td>-3</td>
<td>7</td>
<td>2</td>
<td>-7</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Y8</td>
<td>-5</td>
<td>-3</td>
<td>3</td>
<td>0</td>
<td>-7</td>
<td>6</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

(B) The dual solutions for cluster CL2 (OPTIMAL WEIGHT VECTORS)

$\eta^2 = .3756519; \delta\% = 92.50$

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Category boundaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y4</td>
<td>1.0180</td>
</tr>
<tr>
<td>Y5</td>
<td>1.0195</td>
</tr>
<tr>
<td>Y6</td>
<td>1.0195</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category boundaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
</tr>
<tr>
<td>$\tau_2$</td>
</tr>
<tr>
<td>$\tau_3$</td>
</tr>
</tbody>
</table>

Stimulus values

<table>
<thead>
<tr>
<th>Stimulus values</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
</tr>
<tr>
<td>X2</td>
</tr>
<tr>
<td>X3</td>
</tr>
<tr>
<td>X4</td>
</tr>
<tr>
<td>X5</td>
</tr>
</tbody>
</table>
Table 6-45. Solutions (parameter location vectors) for CL1 and CL2.

<table>
<thead>
<tr>
<th>Category boundary values</th>
<th>CL1*</th>
<th>CL2*</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_1 )</td>
<td>-0.6800</td>
<td>-1.1645</td>
</tr>
<tr>
<td>( \tau_2 )</td>
<td>-0.2267</td>
<td>-0.6987</td>
</tr>
<tr>
<td>( \tau_3 )</td>
<td>0.3156</td>
<td>1.0606</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stimulus values</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>1.1778</td>
<td>0.1809</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0.9554</td>
<td>-1.6303</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>-1.3600</td>
<td>1.3207</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>1.1778</td>
<td>0.1809</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>-1.3600</td>
<td>0.7504</td>
</tr>
</tbody>
</table>

| \( \eta^2 \) (eta squared) | .2251 | .1700 |
| Tr (C) (Trace of C)         | .4062 | .4062 |
| \( \delta \% \) (delta)    | 55.41 | 41.86 |

*The Pearson's product-moment correlation between these solutions was equal to zero.
A computer program (ORTHOG.BAS) for carrying out a Gram-Schmidt orthogonalization was used to orthogonalize the solution given by RM with respect to Solution I. The resulting orthogonal component of the solution yielded by RM was designated as the second solution and given the label, Solution II. Incidentally, the elements of the solution yielded by RM did not change because of orthogonalization. which meant that this solution was already orthogonal to Solution I. Solution II accounted for 41.86% of the total variance for IDM. Collectively, Solution I and Solution II accounted for 97.27% of the total variance for IDM.

The resulting two-dimensional configuration of stimuli and category boundaries is given in Figure 5. Here, Solution I formed the horizontal axis and Solution II the vertical axis. With respect to the vector of stimuli \((X_1, X_2, X_3, X_4, X_5)\), Solution I captured the response pattern \((4, 4, 1, 4, 1)\) while Solution II captured the pattern \((3, 2, 3, 3, 3)\), where the numbers in brackets refer to the labels of categories 1, 2, 3 and 4, respectively.

A remarkable aspect of the pattern of parameters in Figure 5 was that \(X_1\) and \(X_4\) were clumping in the two-dimensional space just as they did in the space in Figure 4. Similarly, the positions occupied by \(X_2\), \(X_3\) and \(X_5\) were consistent with their
Table 6-46. Dual scaling solutions from the first residual matrix (RM).

(A) First residual dominance matrix, RM

<table>
<thead>
<tr>
<th></th>
<th>$Y_1$</th>
<th>$Y_2$</th>
<th>$Y_3$</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>0.0139</td>
<td>0.0048</td>
<td>-0.3988</td>
<td>-0.2202</td>
<td>0.7655</td>
<td>0.0277</td>
<td>-0.2202</td>
<td>0.0277</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>0.0139</td>
<td>0.0048</td>
<td>-0.3988</td>
<td>-0.2202</td>
<td>0.7655</td>
<td>0.0277</td>
<td>-0.2202</td>
<td>0.0277</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>0.0139</td>
<td>0.0048</td>
<td>-0.3988</td>
<td>-0.2202</td>
<td>0.7655</td>
<td>0.0277</td>
<td>-0.2202</td>
<td>0.0277</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>-6.5284</td>
<td>-3.5095</td>
<td>3.7093</td>
<td>2.6472</td>
<td>-4.8526</td>
<td>2.9432</td>
<td>2.6472</td>
<td>2.9432</td>
</tr>
<tr>
<td>$Y_5$</td>
<td>-5.8121</td>
<td>-3.2707</td>
<td>5.3769</td>
<td>2.4066</td>
<td>-5.8590</td>
<td>5.3758</td>
<td>2.4066</td>
<td>-0.6242</td>
</tr>
<tr>
<td>$Y_6$</td>
<td>-5.8121</td>
<td>-3.2707</td>
<td>5.3769</td>
<td>2.4066</td>
<td>-5.8590</td>
<td>5.3758</td>
<td>2.4066</td>
<td>-0.6242</td>
</tr>
<tr>
<td>$Y_7$</td>
<td>-0.0571</td>
<td>-0.0189</td>
<td>1.6341</td>
<td>0.9027</td>
<td>-3.1348</td>
<td>-0.1142</td>
<td>0.9027</td>
<td>-0.1142</td>
</tr>
<tr>
<td>$Y_8$</td>
<td>-5.0958</td>
<td>-3.0319</td>
<td>7.0445</td>
<td>2.1660</td>
<td>-6.8654</td>
<td>1.8084</td>
<td>2.1660</td>
<td>1.8084</td>
</tr>
<tr>
<td>$Y_9$</td>
<td>-6.5284</td>
<td>-3.5095</td>
<td>3.7093</td>
<td>2.6472</td>
<td>-4.8526</td>
<td>2.9432</td>
<td>2.6472</td>
<td>2.9432</td>
</tr>
<tr>
<td>$Y_{10}$</td>
<td>0.0139</td>
<td>0.0048</td>
<td>-0.3988</td>
<td>-0.2202</td>
<td>0.7655</td>
<td>0.0277</td>
<td>-0.2202</td>
<td>0.0277</td>
</tr>
</tbody>
</table>

(B) Solutions from RM (OPTIMAL WEIGHT VECTORS)

$\eta^2 = 0.1699785; \ \delta% = 93.88$

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Category boundaries</th>
<th>Stimulus values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>$\tau_1$ -1.4438</td>
<td>$X_1$ 0.6079</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$\tau_2$ -0.8054</td>
<td>$X_2$ -1.4265</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$\tau_3$ 1.2604</td>
<td>$X_3$ 0.9038</td>
</tr>
<tr>
<td>$Y_4$</td>
<td></td>
<td>$X_4$ 0.6079</td>
</tr>
<tr>
<td>$Y_5$</td>
<td></td>
<td>$X_5$ 0.2958</td>
</tr>
<tr>
<td>$Y_6$</td>
<td></td>
<td>$X_6$ 0.3253</td>
</tr>
<tr>
<td>$Y_7$</td>
<td></td>
<td>$X_7$ 0.0803</td>
</tr>
<tr>
<td>$Y_8$</td>
<td></td>
<td>$X_8$ 0.3253</td>
</tr>
<tr>
<td>$Y_9$</td>
<td></td>
<td>$X_9$ 0.0803</td>
</tr>
<tr>
<td>$Y_{10}$</td>
<td></td>
<td>$X_{10}$ 0.0803</td>
</tr>
</tbody>
</table>
**Table 6-47.** Summary of solutions (parameter location vectors) for data in Table 6-3.

<table>
<thead>
<tr>
<th>Category boundary values</th>
<th>SOLUTION I*</th>
<th>SOLUTION II*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>-0.6800</td>
<td>-1.4438</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-0.2267</td>
<td>-0.8054</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>0.3156</td>
<td>1.2604</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stimulus values</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.1778</td>
<td>0.6079</td>
</tr>
<tr>
<td>$X_2$</td>
<td>0.9554</td>
<td>-1.4265</td>
</tr>
<tr>
<td>$X_3$</td>
<td>-1.3600</td>
<td>0.9038</td>
</tr>
<tr>
<td>$X_4$</td>
<td>1.1778</td>
<td>0.6079</td>
</tr>
<tr>
<td>$X_5$</td>
<td>-1.3600</td>
<td>0.2958</td>
</tr>
</tbody>
</table>

$\eta^2$ (eta squared)  
Tr(C) (Trace of C)      
$\delta$% (delta)       
$\Sigma\delta$% (cumulative delta)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.2251</td>
<td>.1700</td>
</tr>
<tr>
<td></td>
<td>.4062</td>
<td>.4062</td>
</tr>
<tr>
<td>55.41</td>
<td>41.86</td>
<td></td>
</tr>
<tr>
<td>55.41</td>
<td>97.27</td>
<td></td>
</tr>
</tbody>
</table>

*The Pearson's product-moment correlation between these solutions was equal to zero.*
relative positions in the graph (see Figure 4) used to generate the artificial data set. Thus, **Solution I** and **Solution II**, obtained by DSMASC, reproduced the structure originally built into the artificial data set. It should be noted that the positions of the subjects (respondents) vis-a-vis the positions of the parameters can be displayed in a two-dimensional space as follows. One can plot the subject scores \( y_j \) and the projection of parameter values \( \rho x_j \) in the same graph, with the understanding that the rank order of the distances between \( y_j \) and \( \rho x_j \) \((j=1, 2, ..., m)\) is a rank-2 approximation to the rank order of the elements in row \( i \) of the dominance matrix analyzed.

**Some Technical Observations.** From the analysis results of the artificial SC data set, some technical observations can be noted.

1. The behaviour of category boundaries depends on the composition or structure of a successive-categories data set. If half of the subjects have responded to a set of stimuli in one direction (of the continuum for categories) and the other half have responded differently, then the resulting successive-categories data set might fail to yield an admissible solution when analyzed by standard dual scaling.
Figure 5. A two-dimensional plot of Solution 1 versus Solution 2 for the 10-by-5 subject-by-stimulus artificial data set.

**KEY:**

Stimuli:  
D = $\mu_1 = X_1$  
E = $\mu_2 = X_2$  
F = $\mu_3 = X_3$  
G = $\mu_4 = X_4$  
H = $\mu_5 = X_5$

Category Boundaries:  
A = $\tau_1$  
B = $\tau_2$  
C = $\tau_3$

Category 1  
Category 2  
Category 3  
Category 4
2. In this study, some solutions derived from clusters were mutually orthogonal. This should not imply that dual scaling solutions are the same as clusters obtained by a traditional cluster analysis method. Cluster-derived solutions do not necessarily represent orthogonal subspaces. As such, clusters are not equivalent to factors or principal components (solutions). A cluster can be projected onto several orthogonal subspaces. In other words, a cluster may be resolved into more than one dimension. Alternatively, two or more clusters can be projected onto one dimension. Therefore, clusters do not necessarily represent the same information as dual scaling solutions.

3. Solutions derived from clusters might be interesting in their own right. The pattern captured by a solution isolates something that happens several times and consistently so in a specific way. If a solution from one cluster is independent of or correlated with a solution from another cluster, then the underlying representation of the stimuli will show either independence or dependence.

4. The discrete nature of the underlying representation as depicted by two dual scaling solutions (graphical display) could be seen by the clumping of the hypothetical stimuli in the two-dimensional space. A wide category (marked by
category boundaries) is one that has been chosen more often by subjects giving a consistent ranking of the parameters.

5. Orthogonalization is necessary if the independence of the dimensions represented by solutions is to be investigated. Also, orthogonal solutions can be easily displayed in a graphical form using Cartesian coordinate system.

SECTION III

6.4 DSMASC solution vs SDSS solutions

First, restating the difference between standard dual scaling of successive categories (SDSS) and DSMASC is necessary. In SDSS, dual scaling is carried out (on a rank-order table) without the order constraint on category boundaries. In DSMASC, category boundary values are required to be ordered so multiple dimensions are generated from mutually exclusive subsets of subjects.

For the case of three categories, standard dual scaling of successive categories (SDSS) produces solutions, all of which satisfy the weak order constraint \( \tau_1 \leq \tau_2 \) on category boundaries. However, some solutions may need to be reflected to be \textit{admissible}. Thus, the SDSS provides a multidimensional analysis for SC data with
three categories. Naturally, the question anticipated was: Of what relevance would the DSMASC method be for a multidimensional analysis of SC data with three categories? This question was addressed by carrying out two types of analyses, one by SDDS and the other by DSMASC, on the same data set.

The hypothetical data set considered had six subjects, three categories \((m+1=3)\) labeled 1, 2 and 3, respectively and three stimuli \((n=3)\). The 6-by-3 subjects-by-stimuli data set in section (A) of Table 6-51 was converted into the subjects-by-parameters (dominance) matrix in section (B) and, subsequently, submitted to a dual scaling computer program, DUAL3RO (see Nishisato & Nishisato, 1984). The three solutions shown in section (D) of Table 6-51 were obtained.

### 6.4.1 SDSS solutions.

The raw data in (A) of Table 6-51 yielded two main solutions by the SDSS procedure. (For convenience, Solution 1 in (D) of Table 6-51 was labeled SDSS1 and Solution 2, SDSS2.) SDSS1 captured the profile, \((X_1: \mu_1 = 2, X_2: \mu_2 = 1, X_3: \mu_3 = 3)\) with Subjects 2, 3 and 4 being the main contributors. SDSS2 captured the profile, \((X_1: \mu_1 = 3, X_2: \mu_2 = 2, X_3: \mu_3 = 2)\) with Subjects 1, 5 and 6 being the main contributors. SDSS1 and SDSS2 collectively accounted for 98.99% of the total variance for 6-by-3 subjects-by-stimuli data set considered.
6.4.2 DSMASC results. The data set as used section 6.6.1 was analyzed by DSMASC (see, section (A) of Table 6-52). After applying a combination of cluster analysis and dual scaling, two solutions were extracted. These solutions, labelled Solution I and Solution II in section (D) of Table 6-52 were, for discussion, relabeled DSMASC1 and DSMASC2, respectively.

DSMASC1 captured the pattern \((X_1: \mu_1 = 2, X_2: \mu_2 = 1, X_3: \mu_3 = 3)\) or simply \((2 1 3)\), and DSMASC2 captured the pattern \((X_1: \mu_1 = 3, X_2: \mu_2 = 2, X_3: \mu_3 = 2)\) or simply \((3 2 2)\). Solution I and Solution II collectively accounted for 98.93% of the total variance for the small data set considered.

6.4.3 Vector comparisons. In dual scaling of SC data, values given to the rank-ordered stimuli \((\mu_j)\) and category boundaries \((\tau_k)\) are estimates of "location parameters" (Nishisato, 1994, p.270). This implies that the rank score information is more important than the actual values in the solution vector. As ordinal variables, the location parameters should not be treated as if they are quantitative variables with the rank scores considered as scores measured at interval scale (Kiers, 1993). In other words, a direct comparison of the elements of any two solution vectors cannot be justified.
Table 6-51. SDSS dual scaling results for a hypothetical three-category data set.

(A) Raw data

<table>
<thead>
<tr>
<th>Subjects</th>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3.</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4.</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>5.</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6.</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

(C) Dominance Matrix E

<table>
<thead>
<tr>
<th></th>
<th>-2</th>
<th>2</th>
<th>4</th>
<th>0</th>
<th>-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>2</td>
<td>-2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>4</td>
<td>-2</td>
<td>2</td>
<td>4</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

*The Total Variance to Be Accounted for = .4958

(B) Rank Order Table

<table>
<thead>
<tr>
<th>Subjects</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3.</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4.</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5.</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6.</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

(D) Solutions

Solution 1 (SDSS1): Optimal Weight Vectors

<table>
<thead>
<tr>
<th>SUBJECT SCORES</th>
<th>PARAMETER VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 1.3425</td>
<td>r₁ -1.1606</td>
</tr>
<tr>
<td>2. 0.3551</td>
<td>r₂ 0.9105</td>
</tr>
<tr>
<td>3. 0.3551</td>
<td>μ₁ 1.4458</td>
</tr>
<tr>
<td>4. 0.3551</td>
<td>μ₂ -0.5002</td>
</tr>
<tr>
<td>5. 1.3425</td>
<td>μ₃ -0.6954</td>
</tr>
<tr>
<td>6. 1.4202</td>
<td></td>
</tr>
</tbody>
</table>

Statistics: $\eta^2 = 0.26667$, $\delta$(partial) = 53.82, $\sum\delta% = 53.82$

Subject profile represented: 2 1 3

Solution 2 (SDSS2): Optimal Weight Vectors

<table>
<thead>
<tr>
<th>SUBJECT SCORES</th>
<th>PARAMETER VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 1.3425</td>
<td>r₁ -1.1606</td>
</tr>
<tr>
<td>2. 0.3551</td>
<td>r₂ 0.9105</td>
</tr>
<tr>
<td>3. 0.3551</td>
<td>μ₁ 1.4458</td>
</tr>
<tr>
<td>4. 0.3551</td>
<td>μ₂ -0.5002</td>
</tr>
<tr>
<td>5. 1.3425</td>
<td>μ₃ -0.6954</td>
</tr>
<tr>
<td>6. 1.4202</td>
<td></td>
</tr>
</tbody>
</table>

Statistics: $\eta^2 = 0.22396$, $\delta$(partial) = 45.17, $\sum\delta% = 98.99$

Subject profile represented: 3 2 2

Solution 3* (SDSS3): Optimal Weight Vectors

<table>
<thead>
<tr>
<th>SUBJECT SCORES</th>
<th>PARAMETER VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. -0.9234</td>
<td>r₁ -1.3576</td>
</tr>
<tr>
<td>2. -0.3289</td>
<td>r₂ -0.9916</td>
</tr>
<tr>
<td>3. -0.3289</td>
<td>μ₁ 0.3438</td>
</tr>
<tr>
<td>4. -0.3289</td>
<td>μ₂ 1.1523</td>
</tr>
<tr>
<td>5. -0.9234</td>
<td>μ₃ 0.8531</td>
</tr>
<tr>
<td>6. 1.9925</td>
<td></td>
</tr>
</tbody>
</table>

Statistics: $\eta^2 = 0.0050$, $\delta$(partial) = 1.0, $\sum\delta% = 100.00$

Subject profile represented: 3 3 3

*This solution has been reflected.
Such a comparison would only be of interest if a model-fitting procedure involves means vectors and Sum of Squares and Cross Products (SSCP) as the matrix for decomposition, which is not so in dual scaling. These considerations limited the scope of the comparisons made between SDSS and DSMASC solutions. These solutions were compared according to the patterns they captured and the amount of the total variance for which they accounted.

**SDSS1 and DSMASC1.** When the elements of the SDSS1 vector (rank-order of stimuli (μj) and category boundaries (τk)) was translated, the pattern captured for the set of stimuli \{X_1, X_2, X_3\} was (2 1 3); this solution accounted for 53.82% of the total variance for the data set considered. The same results were observed for DSMASC1 but this solution was associated with 53.27% of the total variance. (See section (D) of Tables 6-51 & 6-52.)

**SDSS2 and DSMASC2.** The pattern captured by SDSS2 for the set of stimuli \{X_1, X_2, X_3\} was (3 2 2), and DSMASC2 was (3 2 1) as in section (D) of Tables 6-51 & 6-52. The two solutions, SDSS2 and DSMASC2, were similar only to the extent that each gave the same classification for X_1 and X_2. These solutions in question were different with respect to their classification of the stimulus, X_3. Both patterns, (3 2 2) and (3 2 1) were verified to be in the raw data. The profile (3 2 2), captured by SDSS2, represented Subject 3 in the raw data set; the profile (3 2 1), captured
by DSMASC2, represented Subject 1 and Subject 5 in the raw data set. Furthermore, the contribution of SDSS2 to the total variance was 45.10% and the contribution for DSMASC2 was 45.66%.

For the data analyst, the need for decision and action may be compelling enough to press the question: Which of the two solutions, SDSS2 and DSMASC2 is a better representation of the data? This question was addressed by recognizing the fact that both solutions were important in their own rights, at least, to the extent that each captured a legitimate pattern in the raw data.

The other question to consider is: If both solutions are equally important, is there a geometric relationship between them? This question was addressed by calculating the angle between the two vectors since both SDSS and DSMASC solution vectors have an arbitrary zero origin. Given a common zero origin, the angles between the pairs of vectors, [SDSS1 & DSMASC1] and [SDSS2 & DSMASC2] were calculated. To explain how this was done, the following definitions were made: SDSS1 = u1; DSMASC1 = v1; SDSS2 = u2; DSMASC2 = v2; the angle between SDSS1 and DSMASC1 = φ1; the angle between SDSS2 and DSMASC2 = φ2.
Table 6-52. DSMASC results for a dominance matrix in (B) of Table 6-51.

(A) Case listing of Cluster membership.

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Cluster</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>.260</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>.000</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>.000</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>.000</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>.260</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3.485</td>
</tr>
</tbody>
</table>

(B) Cluster matrix, CL1

```
-2 2 0 -4 4
-2 2 0 -4 4
-2 2 0 -4 4
```

Optimal weight vectors (Solutions)

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Category boundaries</td>
</tr>
<tr>
<td></td>
<td>Stimulus values</td>
</tr>
<tr>
<td>1</td>
<td>-1.0000</td>
</tr>
<tr>
<td>2</td>
<td>-1.0000</td>
</tr>
<tr>
<td>3</td>
<td>-1.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.0000</td>
</tr>
<tr>
<td>5</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Statistics: \( \eta^2 = 5 \), \( \delta\% = 100 \)

(C) Residual matrix, RM

```
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.4</td>
<td>2 4</td>
<td>-8</td>
<td>-3.2</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0 0</td>
<td>-0.0001</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0 0</td>
<td>-0.0001</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0 0</td>
<td>-0.0001</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>-2.4</td>
<td>2 4</td>
<td>-8</td>
<td>-3.2</td>
<td></td>
</tr>
<tr>
<td>-3.4</td>
<td>1 4</td>
<td>-2</td>
<td>-2.2</td>
<td></td>
</tr>
</tbody>
</table>
```

Optimal weight vectors (Solutions)

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Category Boundaries</td>
</tr>
<tr>
<td></td>
<td>Stimulus Values</td>
</tr>
<tr>
<td>1</td>
<td>1.4489</td>
</tr>
<tr>
<td>2</td>
<td>-0.0000</td>
</tr>
<tr>
<td>3</td>
<td>-0.0000</td>
</tr>
<tr>
<td>4</td>
<td>-0.0000</td>
</tr>
<tr>
<td>5</td>
<td>1.4489</td>
</tr>
<tr>
<td>6</td>
<td>1.3422</td>
</tr>
</tbody>
</table>

Statistics: \( \eta^2 = .2264 \), \( \delta\% = 97.72 \)

(D) DSMASC optimal vectors (Solutions)

```
DSMASC1 | DSMASC2

<table>
<thead>
<tr>
<th>Category Boundaries</th>
<th>Category Boundaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>T₁ -0.7071</td>
<td>T₁ -0.7071</td>
</tr>
<tr>
<td>T₂ 0.7071</td>
<td>T₂ 0.7745</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stimulus Values</th>
<th>Stimulus Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>µ₁ 0.0000</td>
<td>µ₁ 1.4829</td>
</tr>
<tr>
<td>µ₂ -1.4142</td>
<td>µ₂ -0.1813</td>
</tr>
<tr>
<td>µ₃ 1.4142</td>
<td>µ₃ -1.0710</td>
</tr>
<tr>
<td>η̂ .26416</td>
<td>η̂ 2.264</td>
</tr>
<tr>
<td>tr (c) .4959</td>
<td>tr (c) .4959</td>
</tr>
<tr>
<td>δ% 53.27</td>
<td>δ% 45.66</td>
</tr>
<tr>
<td>Σδ% 53.27</td>
<td>Σδ% 98.93</td>
</tr>
</tbody>
</table>
```

Subject profiles represented:

```
2 1 3
3 2 1
```
The calculations showed that \( \cos \phi_1 = 0.970791 \) and \( \cos \phi_2 = 0.971287 \).

Consequently, \( \phi_1 \) equals 0.2423 radians and \( \phi_2 \) equals 0.2402 radians. Clearly, it seems that DSMASC solutions could be mapped onto SDSS solutions by a transformation involving a rigid rotation of the former through approximately 0.24 radians about an arbitrary zero origin.

It was found that a rotational transformation could map SDSS solutions onto DSMASC solutions. This is interesting because it engenders the idea that DSMASC solutions represent some consistent property of the data that can be represented in different perspectives. When more than three successive categories are used in data collection, DSMASC may be used to provide a multidimensional perspective of the data. As Hartwig (1979) asserts, when operating within an exploratory mode of data analysis, the analyst must be open to the possibility of several alternative, but equally legitimate, structures in the data. He argues that this openness is best helped when the analyst does not place excessive trust in numeric summaries of data, but uses visual displays of the data as well. A multidimensional representation would make visual displays more informative than not.
SECTION IV

6.5 A practical application of DSMASC

6.5.1 Introduction. This section presents an exploration of the multidimensionality of a real data set using the DSMASC procedure proposed in this study. As a general technique of data analysis, dual scaling maps simultaneously the columns and the rows of a two-way table onto points on a scale. (The column variables of a dominance matrix $E$ are stimuli and ordered category boundaries.) Such a scale or a linear continuum is usually implied when we say that a person has more education or that a crime is more serious, yet, when pressed, we may acknowledge that the criteria used in each case have little, if anything, in common. It seems likely that many opinions are multidimensional and that they cannot be represented on a single continuum as observed by Thurstone (1929). Also, it seems that, sometimes, the continuum implied in a “more” or “less” rating may be conceptual, that it does not necessarily have the physical existence of a yardstick as observed by Stevens (1951). These observations could well apply to educational issues.

When a successive categories data set is analyzed by dual scaling, a question likely to be raised is whether a single solution or more solutions are needed to explain
that data set adequately. As concerns educational issues, the question might be, are educational phenomena necessarily homogeneous with respect to the positions that students, teachers or parents have on various issues or not? This question might be addressed by a multidimensional analysis of the specific items considered. In the present study, the proposal is to evaluate the success of DSMASC by using it to obtain at least two admissible solutions from a real data set. The process of evaluating DSMASC also includes an attempt to interpret the resulting solutions.

6.5.2 Description of real data sets. Data for a group of students in a Secondary School in Malawi\(^4\) were used to illustrate DSMASC. A full description of these data sets is given in Appendix 1A. and the data matrices in Appendix 1B. Although the data set is a 72-by-10 matrix, logically, it consists of two 72-by-5 matrices because only five stimuli (test formats) were considered. Therefore, the matrix in Appendix 1B consists of two data matrices, which for analyses were labeled STUDENT1 and STUDENT2.

In brief, Student Data had been realized as follows. Seventy-two students rated five different testing methods with respect to how good the methods were in

\(^4\) This data set was part of a larger set of data collected by Howard Mzumara in 1993 for a doctoral research project at OISE.
measuring a student’s ability, and a student’s overall achievement. The students rated five testing methods using the set of categories \{Very poor, Poor, Average, Good, Very Good\}. These categories had been assigned sequential numbers from 1 to 5, with \textit{Very Poor} assigned 1 and \textit{Very Good} assigned 5.

6.5.3 Analysis of Student Data. The DSMASC procedures described in section 3.6.1, were followed in analyzing the real data sets. This section contains the results for the two 72-by-5 matrices for the five testing methods. The \text{STUDENT1} data matrix was concerned with “one’s ability as a test taker” and the \text{STUDENT2} data matrix with “how well a test evaluates learning.”

\textbf{Solutions for \text{STUDENT1} data set.} The \text{STUDENT1} data set had been generated by the students’ responses to the item: \textit{Rate your ability as a test taker in each of the following testing methods}. A list of testing methods was rated by students using the set of categories \{Very Poor, Poor, Average, Good, Very Good\}. The first solution (\textbf{Solution 1}) for \text{STUDENT1} data set was admissible; this solution accounted for 63.03\% of the total variance. (See Table 6.61.) Since \textbf{Solution 1} left approximately one-third of the total variance unaccounted for, a second solution was obtained as described the next paragraph.
The residual matrix, RM1 yielded a solution that was not admissible (see Table 6.62). Consequently, RM1 was subjected to a cluster analysis using SPSSPC+ Quickcluster computer program. Initially, two clusters, labeled CL1 and CL2 were obtained. Neither of these two clusters yielded an admissible solution. RM1 was then resubmitted to cluster analysis and three clusters, labeled CL1, CL2 and CL3 were obtained. Of these three clusters, only CL2 yielded an admissible solution (see Table 6-63). This solution (for CL2) was subsequently orthogonalized with respect to Solution I. A computer program written in QBASIC programming language for this study (see ORTHOG.BAS in Appendix 4) was used to carry out a Gram-Schmidt orthogonalization of the two solutions. Also, this program returned the percentage of the total variance attributed to Solution II. The two solutions extracted by DSMASC were recorded in Table 6-64.

Solution I may be roughly interpreted as representing a continuum of the students’ confidence in their abilities to do well in multiple-choice (F) and Short-Answer-sentence (H) testing methods. As can be seen in Figure 6, multiple-choice (F) and Short-Answer (H) formats were classified in the category labeled Good. The students’ expectations of performance in Fill-in-the-Blank (G), True/False (E) and Essay (I) testing formats were classified in the category labeled Average. These results compare with observations made by Mzumara (1995) concerning other
students from the same school. According to the results of analysis by Mzumara (1995), students reported having more confidence in their abilities to do better in multiple-choice tests than in Constructed-Response (Essay) tests. He acknowledged that a student’s confidence in certain testing formats had something to do with his/her beliefs about the formats. Further, Mzumara explained that the students had expressed the belief that a multiple-choice testing format affords one a chance to guess the correct answer to an item, provides clear multiple-choice item stems and removes any threat of penalty for spelling mistakes. These perceptions, Mzumara adds, may have been the reason that students had a greater sense of security and self-confidence, hence their rating of multiple-choice format in the category labeled Good.

According to Solution II in Figure 6, Fill-in-the-Blank (G) and True/False (E) formats were classified as very good, Essay (I) format as average, multiple-choice (F) and Short-Answer-sentence (H) as Very Poor. Solution II may roughly represent a continuum from less to more exposure to (or experience with) the various testing formats as a dimension. At the lower end of the continuum would be ‘very poor exposure’ and at the upper end would be ‘very good’ exposure.
According to the results of analysis by Mzumara (1995), students reported having more confidence in their abilities to do better in multiple-choice tests than in Constructed-Response (Essay) tests. He acknowledged that a student's confidence in certain testing formats had something to do with his/her beliefs about the formats. Further, Mzumara explained that the students had expressed the belief that a multiple-choice testing format affords one a chance to guess the correct answer to an item, provides clear multiple-choice item stems and removes any threat of penalty for spelling mistakes. These perceptions, Mzumara adds, may have been the reason that students had a greater sense of security and self-confidence, hence their rating of multiple-choice format in the category labeled Good.

According to Solution II in Figure 6. Fill-in-the-Blank (G) and True/False (E) formats were classified as very good, Essay (I) format as average, multiple-choice (F) and Short-Answer-sentence (H) as Very Poor. Solution II may roughly represent a continuum from less to more exposure to (or experience with) the various testing formats as a dimension. At the lower end of the continuum would be ‘very poor exposure’ and at the upper end would be ‘very good’ exposure.
Table 6-61. First Solution for STUDENT1 data set.

<table>
<thead>
<tr>
<th>STUDENT1: Solution 1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>-1.8990</td>
<td></td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-1.2685</td>
<td></td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>0.3646</td>
<td></td>
</tr>
<tr>
<td>$\tau_4$</td>
<td>1.6780</td>
<td></td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>-0.0618</td>
<td></td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.5489</td>
<td></td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>-0.0629</td>
<td></td>
</tr>
<tr>
<td>$\mu_4$</td>
<td>0.7257</td>
<td></td>
</tr>
<tr>
<td>$\mu_5$</td>
<td>-0.0250</td>
<td></td>
</tr>
</tbody>
</table>

$\delta\% = 63.03.$

Table 6-62. Solution from RM1 for STUDENT1 data set.

<table>
<thead>
<tr>
<th>STUDENT1: Solution from RM1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>-0.2884*</td>
<td></td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-0.4639*</td>
<td></td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-0.4427</td>
<td></td>
</tr>
<tr>
<td>$\tau_4$</td>
<td>-0.3133</td>
<td></td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>1.5844</td>
<td></td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>-0.1162</td>
<td></td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>1.8589</td>
<td></td>
</tr>
<tr>
<td>$\mu_4$</td>
<td>-0.2872</td>
<td></td>
</tr>
<tr>
<td>$\mu_5$</td>
<td>-1.5316</td>
<td></td>
</tr>
</tbody>
</table>

*These values do not satisfy the order constraint on $\tau_1$ and $\tau_2$. 
Table 6-63. Solutions for clusters CL1, CL2, CL3 for RM1

<table>
<thead>
<tr>
<th></th>
<th>CL1</th>
<th>CL2</th>
<th>CL3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>0.0827*</td>
<td>-0.4954</td>
<td>-0.1783*</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-0.0115*</td>
<td>-0.4487</td>
<td>-0.5283*</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-0.3939*</td>
<td>-0.4311</td>
<td>-0.2883*</td>
</tr>
<tr>
<td>$\tau_4$</td>
<td>-0.3168*</td>
<td>-0.0303</td>
<td>-0.4322*</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>1.9886</td>
<td>1.3295</td>
<td>1.4294</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>-0.2330</td>
<td>-0.9038</td>
<td>0.0849</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>-1.2235</td>
<td>2.2210</td>
<td>1.7867</td>
</tr>
<tr>
<td>$\mu_4$</td>
<td>1.3237</td>
<td>-0.8196</td>
<td>-0.0943</td>
</tr>
<tr>
<td>$\mu_5$</td>
<td>-1.2163</td>
<td>-0.4217</td>
<td>-1.7797</td>
</tr>
</tbody>
</table>

* These values do not satisfy the order constraint on category boundaries.

Table 6-64. Two admissible solutions for STUDENT1 data set.

<table>
<thead>
<tr>
<th></th>
<th>Solution I</th>
<th>Solution II</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1$</td>
<td>-1.8990</td>
<td>-0.4955</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-1.2685</td>
<td>-0.4487</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>0.3646</td>
<td>-0.4311</td>
</tr>
<tr>
<td>$\tau_4$</td>
<td>1.6780</td>
<td>-0.0302</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>-0.0618</td>
<td>1.3295</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.5489</td>
<td>-0.9038</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>-0.0629</td>
<td>2.2210</td>
</tr>
<tr>
<td>$\mu_4$</td>
<td>0.7257</td>
<td>-0.8196</td>
</tr>
<tr>
<td>$\mu_5$</td>
<td>-0.0250</td>
<td>-0.4217</td>
</tr>
<tr>
<td>$\text{tr}(C)$</td>
<td>0.3995</td>
<td>0.3995</td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>0.2518</td>
<td>0.5018</td>
</tr>
<tr>
<td>$\delta%$</td>
<td>63.03</td>
<td>10.53</td>
</tr>
</tbody>
</table>
Figure 6. Two-dimensional Plot of Solution I vs Solution II for STUDENT I Data Set.

**Stimuli:**
- E=μ₁ = True/False
- F=μ₂ = Multiple-Choice
- G=μ₃ = Fill-in-the-Blank
- H=μ₄ = Short Answer (sentence)
- I=μ₅ = Essay (paragraph)

**KEY:**

**Category Boundaries:**

- Very Poor
  - A=τ₁
- Poor
  - B=τ₂
- Average
  - C=τ₃
- Good
  - D=τ₄
- Very Good
This observation was similar to Mzumara’s (1995) observation that students generally had more exposure to the Essay (I) format than the multiple-choice (F) format.

**Results for STUDENT2 Data Set.** STUDENT2 data set represented students’ responses to the item: “Based on your testing experiences, please rate these testing methods according to how well they can evaluate student learning.” The students evaluated this item using the set of categories {Very Poor, Poor, Average, Well, Very Well}. The first solution extracted from STUDENT2, labeled **Solution I** in Table 6.7 was admissible. The residual (RM) matrix for STUDENT2 data also yielded an admissible solution, labeled **Solution II** in Table 6.7. Because of this, cluster analysis was not carried out on RM. Also, **Solution II** was orthogonal to **Solution I** and, therefore, no orthogonalization was done. A graphical display of **Solution I** and **Solution II** for STUDENT2 data is shown in Figure 7.

What **Solution I** in Figure 7 represents is that the students believed the following test methods evaluate learning and/or measure student knowledge after studying well: multiple-choice (F), Short-Answer (H), Fill-in-the-Blank (G) and Essay (I). Only the True/False (E) format was given an average rating in this respect.
Table 6-7. Two Solutions for STUDENT2 Data Set.

<table>
<thead>
<tr>
<th></th>
<th>Solution I</th>
<th>Solution II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_1 )</td>
<td>-1.9447</td>
<td>-0.3019</td>
</tr>
<tr>
<td>( \tau_2 )</td>
<td>-1.2799</td>
<td>-0.2798</td>
</tr>
<tr>
<td>( \tau_3 )</td>
<td>0.0306</td>
<td>-0.1052</td>
</tr>
<tr>
<td>( \tau_4 )</td>
<td>1.3806</td>
<td>0.2159</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>-0.4929</td>
<td>1.1376</td>
</tr>
<tr>
<td>( \mu_2 )</td>
<td>0.6026</td>
<td>1.1984</td>
</tr>
<tr>
<td>( \mu_3 )</td>
<td>0.3177</td>
<td>-0.1793</td>
</tr>
<tr>
<td>( \mu_4 )</td>
<td>0.6401</td>
<td>0.6528</td>
</tr>
<tr>
<td>( \mu_5 )</td>
<td>0.7458</td>
<td>-2.3631</td>
</tr>
<tr>
<td>( tr(C) )</td>
<td>0.4036</td>
<td>0.4036</td>
</tr>
<tr>
<td>( \eta^2 )</td>
<td>0.2407</td>
<td>0.0493</td>
</tr>
<tr>
<td>( \delta% )</td>
<td>59.63</td>
<td>12.22</td>
</tr>
</tbody>
</table>
What is suggested by Solution II in Figure 7 is that the students gave the Essay (1) format a Very Poor rating but gave the True/False, multiple-choice and Short-Answer formats a Very Well rating. The students gave the Fill-in-the-Blank format a Poor rating. Considering that Solution I accounted for 63% of the total variance, Solution II seems to represent only a minor dimension that accounts for approximately 11% of the total variance. An interpretation of Solution II should therefore be attempted cautiously since a unidimensional rating scale was used to collect the data.

A review of Mzumara’s (1995) findings revealed that many students reported experiencing more anxiety for Essay (I) tests than for multiple-choice tests. A high level of anxiety for a test format was partly attributed to a lack of confidence in one’s ability to do well in it. Perhaps this was the reason that the Essay test format had been rated at the Very Poor end of the continuum by the students in Solution II. Further comments by Mzumara (1995) suggest that the ability to self-express in writing, budget time properly in preparing a written response and interpret ambiguous or vague questions may be associated with high anxiety and low self-confidence.
Figure 7. Two-dimensional Plot of Solution I vs Solution II for STUDENT2 Data Set.

KEY:

Stimuli:
E = μ₁ = True/False
F = μ₂ = Multiple-Choice
G = μ₃ = Fill-in-the-Blank
H = μ₄ = Short Answer (sentence)
I = μ₅ = Essay (paragraph)

Category Boundaries:
A = τ₁ Very Poor
B = τ₂ Poor
C = τ₃ Average
D = τ₄ Well

Very Well
less ambiguous than Essay (I) items. Two was the belief that a multiple-choice (F) format demands fewer reasoning/thinking skills/strategies (such as association and elimination). Therefore, in Solution II, the Multiple-choice (F), True/False (E) and Short-Answer (H) formats may have been given a Very Well rating because the students had more confidence in their ability to do well in these test formats. One could speculate on several reasons why such a level of confidence in the Multiple-choice (F) format would have expressed.

First, the students might have perceived Multiple-choice (F) items as less complex and less ambiguous than Essay (I) formats, therefore their lack of confidence in doing well in these test formats. Second, the students expressing more confidence in the Multiple-choice format might have believed this format demands fewer reasoning/thinking skills/strategies (such as association and elimination) than the Essay (I) format. Third, the students expressing more confidence in the Multiple-choice format might have felt comfortable about the response alternatives in the Multiple-choice (F) format; the response alternatives provide clues to the correct answer and the opportunity to guess the correct answer to an item.
7.0 Introduction

In this study, the problem of how to provide multidimensional analysis for successive categories data (a type of categorical data) by dual scaling was addressed. Central to this problem was how to derive admissible solutions in multiple dimensions. The problem of interpretation posed by solutions that violate the order constraint on category boundaries was discussed. The nature of this problem was illustrated using cases of data with three successive categories as well as cases with more than three successive categories. The results of the study showed that the first dual scaling solution (a parameter location vector) may not always satisfy the order constraint on category boundaries. Furthermore, the results of the study showed that a solution that violates the order constraint could be associated with a substantive eigenvalue. In such a case, DSMASC was applied to the data set in question leading to the extraction of admissible solutions. The results of the study also showed that multidimensionality of structure could be the
reason that some cases of data sets did not yield admissible first solutions. Finally, the study illustrated the power of combining cluster analysis with dual scaling in a multidimensional analysis. More often than not, clusters yielded solutions in which the order constraint was maintained. The proposed multidimensional method was illustrated with analyses of artificial and real data examples. The present chapter consists of sections about (1) a summary of the thesis, (2) a discussion of the findings, and (3) a concluding remark.

7.1 Summary of the study

Chapter 1 provided an introduction to dual scaling of dominance data with particular reference to rating data. This chapter also briefly discussed other approaches to multidimensional analysis in dual scaling of rating data and against this backdrop introduced the research problem. Chapter 2 discussed the standard dual scaling procedure for rating data. This included brief presentations on how the appropriate input data matrix is prepared and a description of the objective function or criterion for optimization, Guttman's (1941) squared-correlation ratio. It was noted that standard dual scaling for rating data (SDSS) does not always provide a multidimensional analysis in cases with more than two category boundaries. Chapter 3 presented a short review on points of view analysis (PVA)
as it relates to multidimensional analysis. This chapter also gave a background to the problem of dual scaling of successive categories data. Chapter 4 provided an overview of cluster analysis methods. One traditional cluster analysis method (SPSSPC+ Quickcluster program) was selected for use in the study. Chapter 5 described procedures of DSMASC and the computer programs used in the study. Chapter 6 presented the results of analysis for artificial and real data sets. Chapter 7 discusses findings in the study and presents conclusions.

7.2 Discussion of findings

7.2.1 Order Constraint on Category Boundaries. In some situations, it might be difficult or impossible to collect scalar data (e.g., academic evaluation, opinion surveys, psychological profiles and political forecasts, to name a few). For this reason, the variables of interest are usually classified into two or more categories to measure a particular characteristic or attitude. Rather than assigning a simple, equally spaced numeric scaling of say 1, 2, 3, 4 to categories {Poor, Average, Good, Excellent}, it would be better if the scaling implicit in the data were estimated and used. Optimal solutions derived by dual scaling for successive categories data fit such a scaling profile.
For successive categories data, dual scaling provides least-squares type estimates for subjects scores and parameter weights (stimuli and category boundaries). To be admissible, a solution (a parameter location vector) must satisfy the order constraint on category boundaries. Any $m$ category boundary values ($\tau_j, j=1,2,\ldots,m$) assigned by dual scaling are required to satisfy at least the weak order constraint, $\tau_1 \leq \tau_2 \leq \tau_3 \ldots \leq \tau_m$. This order constraint is maintained within subjects in the dominance matrix, $E$. When this matrix is subjected to an Eckart-Young, or singular value type of analysis, a person space that could be inspected for individual differences is established. Also, established is a corresponding stimulus space. Now, if all individuals were alike, there would be only one dimension in the person space. For individuals who differ with respect to their perceptions of the stimuli, there would be two or more dimensions. Differences in individual perceptions of the stimuli as to the multidimensional space could result in clusters of individuals. In a dual scaling solution derived from each cluster, the subject scores are positive due to the forced consistency of responses on the category boundaries (order constraint). If some subject scores are negative, it means that the judgemental orders of the parameters are reversed as compared with that of the optimum solution. For example, if a cluster yields a solution that shows the order of parameters as

$$\tau_2 > \tau_1 > \mu_1 > \mu_2 > \mu_3,$$
then subjects with negative scores would tend to show the reversed order as

$$\mu_2 > \mu_1 > \tau_2 > \tau_1.$$ 

Nishisato (1994) stated that it is almost always the case that the first solution satisfies the order constraint. In Chapter 6, an example was presented in which the first solution did not satisfy the order constraint, one of his rare cases. The overall results suggest that the problem of interpreting dual scaling solutions of successive categories data may not be confined to solutions that violate the order constraint but also to solutions that satisfy only the weak order constraint. For example, in some solutions, category boundaries might be the same value, that is, $\tau_j = \tau_{j-1}$. This would imply that the category bounded by $\tau_j$ and $\tau_{j-1}$ was not chosen by any consistent subjects. However, this implication was challenged by some results in Chapter 2 (see Table 3-1): All subjects consistently used a category (category 2), yet, this category turned out to be diminished (see Figure 1 on page 39) as depicted by the parameter location vector. This example represented the hypothetical situation in which half the subjects responded to the judgemental task in one way and the remaining subjects in another way. In the results graphically illustrated in Figure 1 (see page 39), one stimulus and two category boundaries appear to be clumped at the same point on a common scale.
The implications of these findings are apparent. A solution showing an extremely narrow or a diminished category (in which category boundary values coincide or nearly do so) may occur for two possible reasons: (1) Very few consistent subjects have chosen the category in question, and (2) a multidimensional structure possibly exists. Nishisato (1980b) suggested that when a successive categories data set consists of two groups of subject that have responded in different ways, the two solutions that best describe the two groups should satisfy the order constraint on category boundaries. This could require two separate analyses in which a group of subjects with small weights is identified and analyzed separately to obtain an optimal solution for the group. Nishisato (1980b) was cautious to note that it was unlikely that one could classify subjects into clear-cut and exclusive subgroups. The same point can be made for the three clustering techniques proposed by Nishisato (1986c). In the method proposed in this study (DSMASC), the traditional cluster analysis method employed offered a way to divide the subjects into subgroups with similar points of view about the parameters. The results of the study all show that cluster-derived solutions, within the framework of DSMASC, can satisfy the order constraint on category boundaries. Thus, the proposed method provides one way to handle not only solutions that violate the order constraint but also solutions that satisfy only the weak order constraint.
7.2.2 Orthogonality of solutions. In the present study, solutions were extracted from a dominance matrix one at a time. A feature of the proposed method (DSMASC) was that a quantification matrix (for decomposition) that did not yield an admissible solution was subjected to cluster analysis. Solutions yielded by clusters were inspected for admissibility, then mutually orthogonalized and each solution's contribution to the total variance (in matrix E) was calculated. Thus, once an admissible solution was determined, the space orthogonal to this solution was inspected for clusters by clustering the residual of a quantification matrix. Orthogonality suggested that solutions (patterns of stimuli) representing the perceptions of different groups of individuals were independent. This is the reason that the cluster-derived solutions were orthogonalized before calculating their respective contribution to the total variance for E (dominance matrix).

In some numerical examples not reported in this study, it was observed that orthogonalization sometimes caused cluster-derived solutions that only satisfied the weak order constraint to violate the order constraint. This happened in solutions with very narrow categories (whose upper and lower bounds are very close to each other). Thus, the order constraint on category boundaries may be maintained within a cluster but not within the multidimensional space. An interesting question is whether to ignore a cluster-derived solution that, when
orthogonalized. violates the order constraint or attempt to interpret it. This should
be left to the data analyst to decide. On the one hand, if the purpose of the analysis
is to develop a classification scheme for the individuals, then cluster-derived
solutions would be interpreted given what is perceived to be represented by the
clusters. On the other hand, if the purpose of analysis is to contrast a solution from
one cluster with a solution from a different cluster of individuals, then mutual
orthogonalization of these solutions would be necessary. Cluster-derived solutions
that remain admissible after orthogonalization would be evidence that the order
constraint on category boundaries can be maintained in a multidimensional space.

7.2.3 Total Variance to be Accounted For. The squared correlation ratio (\(\eta^2\)) is
associated with an eigenvalue in an ordinary principal component analysis (PCA).
In dual scaling, the squared correlation ratio represents the dominance or
importance of solutions that are successively extracted. For a given matrix \(E\), the
total variance is a statistic that reflects variations due to row effects (subjects),
column effects (parameters) and their interactions. When the first solution is
associated with a small amount of the total variance, the scaling result is not
necessarily poor or disappointing. In the present study, the unexplained percentage
of the total variance served as a clue for a possible multidimensionality of \(E\).
As the results of this study suggest, investigating matrix $E$ whose first solution accounts for little of the total variance could lead to a multidimensional analysis of the data. Some results showed that a solution associated with a large amount of the total variance may not necessarily be admissible. The standard dual scaling output in Table 6-3 verified this point. The first solution for this data set was not admissible, yet, it was associated with 64.22% of the total variance. Still, had the first solution been admissible, ignoring the remaining 35.78% of the total variance would be unwise.

### 7.3 Potential areas of application

The main purpose of the present study was to outline a data analytic framework in which a combination of dual scaling and cluster analysis could be used to investigate multidimensionality in successive categories data. Overall, the results of the study showed that the order constraint on category boundaries can be maintained in the multidimensional space. This was illustrated using both an artificial and a real data set. What the results showed is clear: if the first dual scaling solution does not satisfy the order constraint on category boundaries or explains only a small percentage of the total variance, then cluster analysis can be used to partition the rows of $E$, leading to the extraction of multiple solutions.
As it is known, dual scaling involves the process of reciprocal averaging to map either the rows or the columns of data to points on a common scale. Excluding a single subject from a data set can drastically change the pattern of weights in the parameter location vector. In fact, a weight assigned to a row or a column may fall near or far from the point it is supposed to characterize. This phenomenon was more profound when subjects with very different response patterns were analyzed in the same matrix.

For a solution that violates the order constraint on category boundaries, the elements of the solution vector are estimates of only those data points that the vector is nearest. Therefore, a parameter location vector (solution) that violates the order constraint on category boundaries may not be an appropriate descriptor of the columns of $E$. The results of analysis presented in Chapter Six seem to support the idea that clusters of the rows of matrix $E$ can yield admissible solutions that realistically characterize the ranking of stimuli by the subjects. Considering these observations, several potential areas of application can be noted.

1. The DSMASC method could be applied in many areas of social research. The data set from students ratings of testing methods presented in Chapter Six was one example. In other areas of social research, for example, in the investigation of
beliefs or opinions about crime the criteria used by individuals in rating the seriousness of crimes may not be apparent in a single solution. This is because *seriousness of criminal offences* demands that the rater analyses the dimensions or attributes along which seriousness can vary before giving a rating. For instance, a crime some individuals consider serious might not be considered so by other individuals. Furthermore, the term seriousness as it relates to a crime might not mean the same thing to individuals who espouse different values. By taking advantage of the proposed DSMASC method (a combination of cluster analysis and dual scaling), the different patterns of responses that relate to different dimensions of seriousness might be revealed. Thus, DSMASC would provide a multidimensional analysis of both the person space and the stimulus space. In the analysis of questionnaire data, this technique could show how items rank within a cluster and what dimensions these might suggest. Interesting hypotheses concerning the dimensions used by subjects in their decisions to place stimuli (items) into various categories could emerge.

2. Another potential application of DSMASC is in the development of rating scales. When constructing a scale, our intention is to define a clear ordering of response levels used by respondents in parallel with the underlying variable. Sometimes, however, people respond differently from the way intended.
Alternatively, a researcher may have given more categories that people can distinguish or people may respond using different dimensions. One thing to be concerned about is how to check that people have responded according to the intended ordering. The DSMASC method can be used to check, in more than one dimension, if category boundary values satisfy the order constraint.

3. Dual scaling of SC data, in the framework of DSMASC, can be used in the refinement of rating scales. As Low (1988) has pointed out, we often ignore the working of the structure of our rating scale. According to Low (1988), a rating scale employed to survey the motivation of Lithuanian entrepreneurs used the categories \{Very Important, Somewhat Important, Not Important, Less Important, Least Important\}. An initial analysis was conducted with the assumption that the scale was hierarchical, but this analysis failed to yield clear conclusions. Inspection of the scale definition and the data suggested that the respondents used only the first three categories and subsequent analysis with the last two categories removed was productive. If Low's (1988) data were analyzed by dual scaling, the categories without responses would be narrow in the first solution. However, a category may also be narrow for another reason: the responses in a category may have come from subjects with opposing patterns of ranking. In any case, the method proposed in the present study would reveal that Low's (1988) last two
categories were redundant. In addition, the proposed method could be used to investigate multidimensionality when the percentage of the total variance explained by the first solution was small.

4. Another example is a study (Loewen, 1988) involving the construction of a rating questionnaire. Here, dual scaling of successive categories data was used to discover that two distinct groups of subjects (younger and older groups) had used different parts of the continuum spanned by ordered categories. Although difference in age was accepted as an explanation for the observed patterns of responses (solutions), exploring a second dimension to see if the individuals still clustered according to their age groups could have been done. Such a process of analysis could benefit from the proposed DSMASC method. Using this method, clusters of individuals could be identified and then submitted to dual scaling. As the results of the present study have shown, clusters most of the time yield solutions in which category boundaries are nicely separated while satisfying the order constraint. Finally, the application of dual scaling to adult education research (Abbey, 1984) has suggested that multidimensional analysis is necessary to explain interesting aspects of judgemental (rating) data. It is contended that a multidimensional representation of rating data can generate interesting ideas and hypotheses in the sphere of educational research.
7.4 Significance and Limitations of the study

This study makes a modest but important contribution to the methodological literature overall and dual scaling in particular. The approach adopted is to cluster the subjects into homogeneous subgroups, find a separate solution for each cluster. when an initial solution is not admissible. This approach is an extension of standard dual scaling of successive categories data. Although other approaches toward a multidimensional analysis of successive categories data have been proposed before, none of those approaches have based their methodologies on the extraction of multiple solutions through the clustering a dominance matrix.

By focussing on successive categories data, the present study makes a contribution to data theory. The finding in this study puts to question the usual practice of evaluating components only by the amount of the total variance explained. In an illustration of the proposed method using an artificial data set, a solution may clearly account for a large amount of the total variance and yet be inadmissible.

As to a possible area of application, the proposed method can help data analysts and practitioners who find themselves constructing rating questionnaires or
analyzing successive categories data to make sense out of their data. The proposed method offers an alternative way to explore the multidimensionality of the ranking of stimuli and category boundaries by respondents. Although the present study has attempted something by sorting subjects into homogeneous groups using the so-called dominance matrix, it has the following limitations:

1. First, in the proposed method (DSMASC) is an extension of dual scaling of successive categories data. This method deals with the difficulty created by estimates of category boundaries that violate the assumed unidimensional ordering. However, the method involves a two-step procedure in which dual scaling and cluster analyses are carried out separately. Incidentally, in the two-step procedure, the data analyst must evaluate the results from a dominance matrix to decide the sequence of the steps. This is a limitation of the DSMASC method.

2. Second, only the K-means cluster analysis method (McQueen, 1967), was employed in this study. It is no secret that in the realm of cluster analysis, different clusters can be obtained depending on the method of cluster analysis employed. In other words, different sub-matrices of the dominance matrix can lead to different solutions. The absence of evidence on the sensitivity of the proposed method (DSMASC) to different methods of clustering is a limitation to note.
7.5 Concluding remark

Considering this study as a whole, it seems satisfactory that DSMASC provides a means of examining multidimensionality in successive categories data. The study revealed that dual scaling of the so-called dominance matrix may not necessarily lead to an admissible solution. However, the study showed that clusters of the rows of a dominance matrix could yield admissible solutions when a whole dominance matrix had failed to do so. Like every other study, the present study raised several questions that may require further investigation:

1. First, one direction for future research should be the comparison of results obtained by different cluster analysis methods. This could advance the scope of the present study in providing evidence on the stability of DSMASC representations.

2. Second, although DSMASC was shown to provide a two-dimensional representation of SC data with category boundary estimates numerically sequenced as required, dual scaling and clustering were applied in separate steps. In the future, a simultaneous procedure should be developed.

3. Another issue to ponder is that, historically, dual scaling and related
methods have been conceived of as essentially exploratory, structure-seeking methods. However, some researchers have begun to formulate some of these methods through confirmatory, model-based methods for the analysis of categorical data. (See, e.g., Millones, 1991; Goodman, 1986). The present study did not address the question of the appropriate number of dimensions necessary to represent the structure in a successive categories data set. The number of dimensions to extract has traditionally been addressed via measures of percent of variance accounted for. However, approaches that follow Goodman's RC association model (see, e.g., Goodman, 1986) provide maximum likelihood estimates which are typically equivalent to those of dual scaling and provide likelihood ratio tests for the number of dimensions. Clearly, this is one direction to go in future research.

4. Finally, in the present study, the dominance matrix is clustered when either the dominance matrix yields either an inadmissible solution or a solution that accounts for a small amount of the total variance. Future research could consider the unconditional clustering of the dominance matrix and investigate the properties of the multidimensional results. This may be a good starting point for further development and application of the proposed method.
References


APPENDIX 1A

Description of real data examples

I. Description: The two 72 by 5 subjects-by-stimuli data matrices in Appendix B, named STUDENT1 and STUDENT2, respectively, were part of a larger set of data collected by Howard Mzumara for his thesis research project at the Ontario Institute for Studies in Education, affiliated to the University of Toronto in March 1993. The data sets considered here came from a pool of data sets that had been collected but not analyzed or included in the provider’s report. The data collection instrument was an adapted questionnaire for student opinions about methods of educational testing (see, Mzumara, 1995). The respondents were two classes of Grade 12 students in a secondary school in Malawi. The columns of each data matrix represent five testing methods; the columns represent the seventy-two students; an element of a matrix is a sequential number assigned to the five successive categories used rate the five testing methods in the columns of each data matrix. The sample of questionnaire items concerning the two data matrices is reproduced below. The actual data matrices are shown in Appendix 1B. Note that the two data sets are for the same group of students.

II. Questions (stimuli)

1. Rate your ability as a test taker in each of the following methods of testing.
   E. True/False test format.
   F. Multiple-choice test format.
   G. Fill-in-the-blank test format.
   H. Short Answer (sentence +) test format.
   I. Essay (paragraph +) test format.

2. Based on your testing experiences, please rate these methods according to how well they can evaluate student learning.
   J. True/False
   K. Multiple-choice
   L. Fill-in-the-blank
   M. Short Answer (sentence +)
   N. Essay (paragraph +)

III. Successive categories (assigned numbers in brackets)

1. (1) Very poor, (2) Poor, (3) Average, (4) Good, (5) Very Good
2. (1) Very poor, (2) Poor, (3) Average, (4) Well, (5) Very Well
APPENDIX 1B. Two 72 by 15 subjects-by-questions (stimuli) data matrices for students' responses.

**TEST FORMATS**

<table>
<thead>
<tr>
<th>SUBJECTS</th>
<th>STUDENT1</th>
<th>STUDENT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>E F G H I</td>
<td>J K L M N</td>
<td></td>
</tr>
</tbody>
</table>

1. 1 4 4 4 5 1 2 2 3 5
2. 3 3 3 5 3 3 3 4 5 5
3. 4 4 3 5 4 4 5 3 4 5
4. 3 4 3 3 3 4 5 4 5 3
5. 3 3 4 2 1 5 4 4 5 2
6. 3 4 4 4 4 3 4 4 4 4
7. 3 3 4 4 3 3 3 4 4 4
8. 4 5 4 5 3 3 4 4 5 4
9. 3 4 3 4 5 2 4 4 4 5
10. 5 4 5 4 3 3 4 3 5 3
11. 3 4 4 3 3 2 3 4 2 5
12. 5 4 5 4 3 4 5 5 4 4
13. 3 4 3 3 3 3 4 4 3 3
14. 1 4 1 3 3 3 5 3 5 4
15. 2 1 4 5 3 1 2 4 5 3
16. 4 4 3 5 4 4 4 4 5 4
17. 3 4 3 4 5 2 4 5 3 5
18. 2 1 4 5 3 1 2 4 5 3
19. 3 3 4 5 2 5 4 5 5 4
20. 1 5 2 5 4 1 5 4 4 5
21. 4 3 3 4 3 3 4 5 5 4
22. 3 4 3 3 4 5 4 4 5 4
23. 3 4 4 4 3 3 2 4 4 5
24. 4 4 3 4 3 5 5 5 4 5
25. 4 4 3 4 3 3 4 4 4 4
26. 4 3 3 4 3 4 5 4 5 3
27. 4 4 3 4 3 2 3 4 5 3
28. 4 5 3 4 3 3 2 3 4 2
29. 5 4 3 5 3 1 3 5 4 5
30. 4 3 3 3 4 3 3 4 4 5
31. 5 3 2 4 3 1 4 3 4 5
32. 4 3 4 3 4 3 3 4 5 5
33. 3 4 3 4 3 3 4 3 3 5
34.  4 3 3 4 4
35.  3 5 2 4 4
36.  2 2 1 4 3
37.  2 5 3 4 3
38.  3 4 4 4 4
39.  2 3 3 4 4
40.  4 3 5 4 3
41.  4 5 4 5 3
42.  5 4 5 4 3
43.  3 4 3 4 3
44.  4 3 5 3 5
45.  2 2 1 4 3
46.  3 3 3 3 3
47.  3 4 3 3 3
48.  3 3 3 3 3
49.  3 3 3 3 3
50.  3 4 3 3 4
51.  3 3 3 5 3
52.  2 4 2 3 2
53.  3 3 3 3 3
54.  4 3 5 4 5
55.  3 3 3 3 3
56.  3 4 3 2 2
57.  3 4 3 4 5
58.  3 4 3 3 3
59.  3 4 3 4 3
60.  1 3 1 3 4
61.  3 4 3 4 4
62.  5 4 1 5 1
63.  4 5 5 5 4
64.  3 4 3 3 3
65.  3 5 4 5 2
66.  4 5 4 3 3
67.  3 3 3 4 2
68.  3 3 4 4 3
69.  3 4 3 4 5
70.  5 4 5 4 3
71.  2 2 1 4 3
72.  3 5 5 4 3
APPENDIX 2. **DSOFE.BAS**, a program for analyzing a dominance matrix.

' dual scaling program applied to dominance matrix for Whole data set
CLS
DIM G(72), R1(15), B(15), C(15, 15), Y(72)
DIM FD(72, 15), DN(72), DG(72), D2(72)
DIM EM(50, 50), WT(5, 72)
DIM ADAT(50, 50), D3(50), FADAT(50, 100), AADAT(40, 15), AA(75, 75)
DIM AAM(50, 50), AM(100, 50), FDT(75, 75)
PRINT "*******************************************************************"
PRINT "* DUAL SCALING OF DOMINANCE MATRIX *"
PRINT "*******************************************************************"
PRINT
6500 INPUT "FILE NAME for DOMINANCE MATRIX...> ", FILENAME$
OPEN FILENAME$ FOR INPUT AS #1
INPUT "NUMBER OF QUESTIONS? >"; NUMVAR
INPUT "NUMBER OF OBSERVATIONS? >"; NUMOBS
INPUT "NUMBER OF SUCCESSIVE CATEGORIES? >"; NCAT
NCOL = NUMVAR + NCAT - 1
FOR I = 1 TO NUMOBS
FOR J = 1 TO NCOL
INPUT #1, FD(I, J)
PRINT FD(I, J);
NEXT J
PRINT
NEXT I
CLOSE #1
NR = NUMOBS
NC = NUMVAR
FORMWS$ = "##.####"
FORMHS$ = "##.##
##.##
##.##
##.##
##.##
##.##
##.##
##.##
##.##
##.##

JS = NCAT - 1
LK = JS * (JS - 1) / 2 + NC * (NC - 1) / 2
LA = JS + NC
'Begins construction of diagonal
'T = 2 * (JS * NC + (JS * (JS - 1) / 2) + (NC * (NC - 1) / 2))
T = (JS + NC) * (JS + NC - 1)
FOR I = 1 TO NR 'matrices for ROW & COLUMN TOTALS
DN(I) = T 'ROW TOTALS
NEXT I
S1 = (NC + JS - 1) * NR
FOR I = 1 TO JS
DG(I) = S1 'Tau's COLUMN TOTALS
NEXT I
S1 = (JS + NC - 1) * NR
L = JS + NC
FOR I = NCAT TO L
DG(I) = S1 'Mu's COLUMN TOTALS
NEXT I
FT = NR * (JS + NC) * (JS + NC - 1)
NC = JS + NC
6800 'END
PRINT
PRINT
PRINT "*** NUMBER OF RESPONSES FOR ROWS ***"
PRINT
FORM1$ = "######## ######## ######## ######## ######## ######## ########
######## ######## ######## ######## ########
FOR I = 1 TO NR
PRINT USING FORM1$: DN(I);
NEXT I
PRINT
PRINT
PRINT "*** NUMBER OF RESPONSES FOR COLUMNS ***"
PRINT
FORM2$ = "######## ######## ######## ######## ######## ######## ########
######## ######## ######## ######## ########
FOR I = 1 TO NC
PRINT USING FORM2$: DG(I);
NEXT I
PRINT : PRINT
PRINT "*** MATRIX C FOR EIGENEQUATION ***"
PRINT
CHK1 = 0!
CHK2 = 0!
FOR I = 1 TO NR
CHK1 = CHK1 + DN(I)  'vector of ROW TOTALS
NEXT I
FOR I = 1 TO NC
CHK2 = CHK2 + DG(I)  'vector of COLUMN TOTALS
NEXT I
DCHK = ABS(CHK1 - CHK2)  'checks that GRAND TOTAL is correct
IF DCHK > .5 THEN 7405
FOR I = 1 TO NR
FOR J = 1 TO NC
PRINT FD(I, J);  ' Prints matrix E (dominance matrix)
NEXT J
NEXT I
FOR I = 1 TO NC
IF DG(I) < .5 THEN D2(I) = 0!
IF DG(I) >= .5 THEN D2(I) = 1! / SQR(DG(I))  ' D2(I) = D-1/2
NEXT I
INPUT "TYPE NAME OF OUTPUT FILE for SUBJECT WEIGHTS> ";
SCORE$
OPEN SCORE$ FOR OUTPUT AS #4
INPUT "TYPE NAME OF OUTPUT FILE for PARAMATER WEIGHTS> ";
SOLUTE$
OPEN SOLUTE$ FOR OUTPUT AS #5
7100  FOR I = 1 TO NC
FOR J = 1 TO NC
C(I, J) = 0!
FOR K = 1 TO NR
IF DN(K) < .1 THEN DR = 0!
IF DN(K) > .1 THEN DR = 1! / DN(K)  ' DR = Dn-1
C(I, J) = C(I, J) + D2(I) * FD(K, I) * FD(K, J) * D2(J) * DR
C(J, I) = C(I, J)
' This expression: C = D-1/2 F' Dn-1 F D-1/2
NEXT K
NEXT J
NEXT I
7200 TRACE = 0!
FOR I = 1 TO NC
TRACE = TRACE + C(I, I)
NEXT I
PRINT
PRINT
PRINT "** TRACE OF C = ": TRACE; " **"
PRINT
'Iterative power method to find the maximum value and the
'corresponding eigenvector. first generate a trial vector
'for initial run
NROOT = 0
SSP = 0!
T = NC - 1
STP = 2! / T
ITER = 1
G(1) = 1!
B(1) = 1!
PRINT
PRINT " *** THE TRIAL VECTOR *** " 'bo
PRINT
FOR I = 2 TO NC
   KK = I - 1
   G(I) = G(KK) - STP
   B(I) = G(I)
   PRINT B(I);
NEXT I
NROOT = NROOT + 1
PRINT
7300 PRINT
PRINT "** SOLUTION ": NROOT: "**
7310 FOR I = 1 TO NC
   R1(I) = 0!
   FOR J = 1 TO NC
      R1(I) = R1(I) + C(I, J) * G(J)   'Rbo = b1
   NEXT J
   NEXT I
   S1 = ABS(R1(I))
   FOR I = 2 TO NC     'search for largest element, k1
      Z = ABS(R1(I))
      IF S1 > Z THEN 7330
      S1 = Z
   7330 NEXT I
   7350 FOR I = 1 TO NC     'divide Rbo with k1, i.e. b1/k1
G(I) = R1(I) / S1
NEXT I
CR = 0!
FOR I = 1 TO NC
DFR = ABS(B(I) - G(I))
IF DFR < .00005 THEN 7360 'if each element of bk is stable
GOTO 7370 'if each element of bk not stable
7360 NEXT I
GOTO 7400 'if bs stable proceed on to 7400
7370 FOR I = 1 TO NC
B(I) = G(I) 'replace contents of e.g., b1 with b2
NEXT I 'and iterate
ITER = ITER + 1
IF ITER < 1000 THEN 7310
PRINT "**** NO CONVERGENCE -- RUN ABORTED *****"
GOTO 7700
'**CALCULATION OF OPTIMAL AND ORTHOGONAL WEIGHT VECTORS
'**FOR THE ROWS AND THE COLUMNS OF THE DATA MATRIX
'**A STOPPING CRITERION
7400 IF S1 <= 1! AND S1 >= 0! THEN 7410 'ensure standardized values
0<=S1<=1
7405 PRINT "***** CHECK INPUT DATA FOR ERROR **"
GOTO 7700
7410 S2 = 0! 'calculating principal components
FOR I = 1 TO NC
S2 = S2 + G(I) ^ 2 'calculates b'b
NEXT I
FOR I = 1 TO NC
G(I) = G(I) / SQR(S2) 'computes the ratio b/SQRT(b'b)
NEXT I
CRT = SQR(FT)
FOR I = 1 TO NC
G(I) = G(I) * CRT 'computes SQRT(FT)*b/SQRT(b'b)=w
NEXT I 'Or SQRT(FT/b'b)*b = w
INPUT "TYPE NAME OF OUTPUT FILE for DUAL SCALING RESULTS >":
DUAL$
OPEN DUAL$ FOR OUTPUT AS #6
SROOT = SQR(S1) 'eta = SQRT(largest absolute element of bs)
PRINT #6,
PRINT #6, "SQUARED CORRELATION RATIO = "; S1
PRINT #6, "MAXIMUM CORRELATION = "; SROOT
SSP = SSP + 100! * S1 / TRACE
PRINT #6, "DELTA(TOTAL VARIANCE ACCOUNTED FOR) BY"; NROOT: " SOLUTION(S) = "; SSP; " (PER CENT)"
PRINT #6, "NUMBER OF ITERATIONS = "; ITER
PRINT #6, 7500 FOR I = 1 TO NC
B(I) = G(I) * D2(I) * -1 'computes w.D-1/2 = x
WT(3, I) = B(I) * SROOT 'w.D-1/2.eta
WT(4, I) = G(I) * SROOT 'w.eta
NEXT I
TR = 1! / SROOT '1/eta
FOR I = 1 TO NR
Y(I) = 0!
IF DN(I) < .5 THEN TT = 0!
IF DN(I) >= .5 THEN TT = 1! / DN(I)
FOR J = 1 TO NC
Y(I) = Y(I) + TR * FD(I, J) * B(J) * TT 'optimal score vector:
Y(I) = (1/ETA)FxDn-1
NEXT J
NEXT I
FOR I = 1 TO NR 'rescales all the results in terms of
WT(1, I) = Y(I) * SROOT 'the relative size of each dimension
NEXT I 'i.e., (1/ETA).F.x.Dn-1.ETA
PRINT #6.
PRINT #6, " OPTIMAL/LOCALLY OPTIMAL WEIGHT VECTORS
(UNWEIGHTED, WEIGHTED BY ETA) FOR" PRINT #6, "----------------------------------------------------------"
PRINT #6, " ** ROWS **  ** COLUMNS **"
PRINT #6, "----------------------------------------------------------"
PRINT #6, " CATEGORY BOUNDARIES"
FORM10$ = "### #######.###### "### #######.######"
FOR I = 1 TO NCAT - 1
PRINT #6, USING FORM10$; I; Y(I); I; B(I)
PRINT #4, USING FORMW$; Y(I)
PRINT #5, USING FORMW$; B(I)
NEXT I
PRINT #6,
PRINT #6, "STIMULUS VALUES"
FOR I = NCAT TO NC
    J = I - (NCAT - 1)
    PRINT #6, USING FORM10$; I; Y(I); J; B(I)
    PRINT #4, USING FORMW$; Y(I)
    PRINT #5, USING FORMW$; B(I)
NEXT I
IF NR = NC THEN 7700
    K = NC + 1
    FORM11$ = "### ########.####"
    FOR I = K TO NR
        PRINT #6, USING FORM11$; I; Y(I)
        PRINT #4. USING FORMW$; Y(I)
    NEXT I
    PRINT #6, "-----------------------------------------------"
CLOSE
7600 PRINT "**** END OF ANALYSIS ****"
7700 END
APPENDIX 3.  ORTHOG.BAS, a program for Gram-Schmidt orthogonalization.

CLS
PRINT
PRINT "THIS PROGRAM ORTHOGONALIZES SOLUTIONS AND COMPUTES DELTA% FOR EACH"
1000 DIM OPT(15, 15), OTP(15, 15), PROD(15, 15), BH(15, 15), DELTA(15)
     DIM FDT(72, 50), FD(72, 50), H(72, 50), C(15, 15)
PRINT
INPUT "NAME OF DISK FILE OPTIMAL SCORES...">"; OPTDAT$ 
OPEN OPTDAT$ FOR INPUT AS #1
INPUT "NUMBER OF STIMULI (ITEMS)> "; M
INPUT "NUMBER OF OBSERVATIONS IN WHOLE DATA SET > ";
ROWS
INPUT "NUMBER OF successive CATEGORIES "; NCAT
INPUT "NUMBER OF cluster solutions from <DSMASC3.BAS> run "; NC
INPUT "FILE FOR OUTPUT..."; GRAMST$ 
INPUT "NAME OF FILE WITH MATRIX C"; MATCS$ 
NR = M + NCAT - 1
FOR J = 1 TO NC
FOR I = 1 TO NR 
INPUT #1, OPT(I, J)
NEXT I
NEXT J
CLOSE #1
GOSUB 2000 'ORTHOGONALIZES SOLUTIONS AND CALCULATES DELTA% FOR EACH
FOR V = 1 TO NC
FOR X = V + 1 TO NC
IF DELTA(V) < DELTA(X) THEN
GOSUB 8000 'SORTS VECTORS BY DELTA
GOSUB 2000
END IF
NEXT X
NEXT V
GOSUB 7600
END
2000 OPEN GRAMST$ FOR APPEND AS #2
PRINT
PRINT #2, "CORRELATED SOLUTIONS"
PRINT #2.
FORTH$ = "#.##### #.####### #.##### #.#####"
FOR I = 1 TO NC
PRINT #2, " SOLUTION "; I;
NEXT I
PRINT #2,
FOR I = 1 TO NR
FOR J = 1 TO NC
PRINT #2. USING FORTH$: OPT(I, J);
NEXT J
PRINT #2,
NEXT I
PRINT
FOR K = 1 TO NC - 1
'Find the inner product
FOR J = K TO NC
SUM = 0!
FOR I = 1 TO NR
SUM = SUM + OPT(I, K) * OPT(I, J)
NEXT I
PROD(K, J) = SUM
NEXT J
'Iterate to Gram-Schmidt solutions
FOR J = K TO NC - 1
FOR I = 1 TO NR
OPT(I, J + 1) = OPT(I, J + 1) - PROD(K, J + 1) * OPT(I, K) / PROD(K, K)
NEXT I
NEXT J
IF K = NC - 1 THEN
SUM = 0!
FOR I = 1 TO NR
SUM = SUM + OPT(I, NC) * OPT(I, NC) 'inner prod of last vector
NEXT I
PROD(NC, NC) = SUM
END IF
NEXT K
PRINT #2, "ORTHOGONAL SOLUTIONS"
PRINT #2.
FOR I = 1 TO NC
PRINT #2, " SOLUTION "; I;
NEXT I
PRINT #2,
FOR I = 1 TO NR
FOR J = 1 TO NC
PRINT #2, USING FORTH$; OPT(I, J);
NEXT J
PRINT #2,
NEXT I
FORMLS$ = "\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\####\#
NEXT I
CLOSE #3
FOR K = 1 TO NR
OTP(L, K) = OPT(K, L) 'transposes theta
NEXT K
FOR J = 1 TO NR
SUM1 = 0!
SUM3 = 0!
FOR K = 1 TO NR
SUM1 = SUM1 + OTP(L, K) * H(K, J) 'Theta'.C
NEXT K
BH(L, J) = SUM1
NEXT J
FOR K = 1 TO NR
SUM3 = SUM3 + BH(L, K) * OPT(K, L)
NEXT K
PRINT #2,
PRINT "ETA SQUARED = "; SUM3
PRINT #2, "ETA SQUARED = "; SUM3 'the squared correlation-ratio
SSP = 100! * SUM3 / TRACE
PRINT "TRACE = "; TRACE
PRINT #2. "TRACE OF C = "; TRACE
PRINT #2, "% DELTA FOR SOLUTION "; L; "IS = "; SSP; "%"
PRINT #2, "ETA = "; SQR(SUM3)
PRINT #2.
PRINT #2, "PARAMETERS": " "; "** OPTIMAL SCORES **"
PRINT #2, "---------------------------------------"
PRINT #2,
PRINT #2. "CATEGORY BOUNDARY VALUES"
PRINT #2,
FOR I = 1 TO NCAT - 1
PRINT #2, USING FORMP$; I; OPT(I, L)
NEXT I
PRINT #2,
PRINT #2, " STIMULUS VALUES "
PRINT #2,
FOR I = NCAT TO NR
PRINT #2, USING FORMP$; I - NCAT + 1; OPT(I, L)
NEXT I
PRINT #2, "-----------------------------"
TOTAL = TOTAL + SSP
DELTA(L) = SSP
NEXT L
CLOSE #2
RETURN

7600 OPEN GRAMST$ FOR APPEND AS #2
PRINT #2, "TOTAL INFORMATION EXPLAINED BY ALL SOLUTIONS = "; TOTAL; "%
PRINT #2,
PRINT #2, " **** END OF ANALYSIS ****"
PRINT " **** END OF ANALYSIS ****"
CLOSE #2
PRINT "THE MULTIPLE SOLUTIONS IDENTIFIED ARE IN THE FILE ":
GRAMST$
CLOSE
7700 RETURN

8000 OPEN OPTDAT$ FOR INPUT AS #1
NR = M + NCAT - 1
FOR J = 1 TO NC
FOR I = 1 TO NR
INPUT #1, OPT(I, J)
NEXT I
NEXT J
CLOSE #1
FOR J = 1 TO NC
FOR K = J + 1 TO NC
IF DELTA(J) < DELTA(K) THEN
SWAP DELTA(J), DELTA(K)
FOR I = 1 TO NR
SWAP OPT(I, J), OPT(I, K)
NEXT I
END IF
NEXT K
NEXT J
RETURN
APPENDIX 4. DESIGN.BAS, program converts a dominance matrix into rank-order matrix.

'PROGRAM CONVERTS RATINGS INTO DOMINANCE MATRIX OR RANK ORDERS
CLS
DIM FD(50, 15)
DIM EM(50, 50)
DIM ADAT(50, 110), D3(50), FADAT(50, 160), AADAT(50, 15), AA(50, 50)
DIM AAM(110, 50), AM(110, 15), FDT(50, 50)
PRINT "*********************************************************************"
PRINT "* THIS PROGRAM CONVERTS RATINGS INTO RANK ORDERS *"
PRINT "*********************************************************************"
PRINT
6500 INPUT "ENTER DISK FILE NAME> ", FILENAMES$ OPEN FILENAMES$ FOR INPUT AS #1
INPUT "NUMBER OF QUESTIONS? >"; NUMVAR
INPUT "NUMBER OF OBSERVATIONS? >"; NUMOBS
INPUT "NUMBER OF SUCCESSIVE CATEGORIES? >"; NCAT
INPUT "OUTPUT RANK ORDER DATA FILE >"; RANK$
PRINT "*** INPUT MATRIX ***"
FOR I = 1 TO NUMOBS
FOR J = 1 TO NUMVAR
INPUT #1, FD(I, J)
PRINT FD(I, J);
NEXT J
PRINT
NEXT I
CLOSE #1
GOSUB 5000
NR = NUMOBS
NC = NUMVAR
FORMH$ = "##.#### ##.#### ##.#### ##.#### ##.#### ##.#### ##.#### ##.####
##.#### ##.#### ##.#### ##.#### ##.####
JS = NCAT - 1
LK = JS * (JS - 1) / 2 + NC * (NC - 1) / 2
LA = JS + NC
FOR I = 1 TO NR 'Begins constructing first F component
M = 0
FOR J = 1 TO NC
K = FD(I, J)
L = K - 1
IF L = 0 THEN 6600
FOR KK = 1 TO L
M = M + 1
EM(I, M) = -1!
NEXT KK
IF L = JS THEN 6610
6600 FOR KKK = K TO JS
M = M + 1
EM(I, M) = 1!
NEXT KKK
6610 NEXT J
6620 NEXT I
KL = JS * NC
FOR I = 1 TO NR
FOR J = 1 TO KL
'PRINT EM(I, J); 'Prints out F matrix
NEXT J
NEXT I
'BEGINNING OF CONSTRUCTION OF MATRIX Fa
OPEN "TRANS.RES" FOR OUTPUT AS #2
FOR P = 1 TO NR 'Begin creating matrix Fa
N = 1
FOR J = NC TO 1 STEP -1
FOR I = 1 TO J - 1
Q = I + N
IF FD(P, I) > FD(P, Q) THEN D3(I) = 1!
IF FD(P, I) < FD(P, Q) THEN D3(I) = -1!
IF FD(P, I) = FD(P, Q) THEN D3(I) = 0!
'PRINT "(": I; ";": ; Q; ";")"; "REP"; FD(P, I); ";-"; FD(P, Q); ";="; D3(I)
PRINT #2, D3(I);
NEXT I
N = N + 1
NEXT J
NEXT P
CLOSE #2
' The design matrix Fa *** (in TEMPORARY file TRANS.RES)
FOR I = 1 TO NR
FOR J = 1 TO JS * (JS - 1) / 2  'Inputting tau's
ADAT(I, J) = -1!
NEXT J
NEXT I
LM = JS * (JS - 1) / 2
LN = JS * (JS - 1) / 2 + NC * (NC - 1) / 2
OPEN "TRANS.RES" FOR INPUT AS #2
FOR I = 1 TO NR
FOR J = LM + 1 TO LN
INPUT #2, ADAT(I, J)  'Inputting mu's
NEXT J
NEXT I
OPEN "FDAT.RES" FOR OUTPUT AS #3
FOR I = 1 TO NR
FOR J = 1 TO LN
PRINT #3, ADAT(I, J);
'PRINT ADAT(I, J);  'This prints the full matrix Fa
NEXT J
NEXT I
CLOSE #3
CLOSE #2
OPEN "FDAT.RES" FOR INPUT AS #3
'Construct F & Fa combined
LFA = (JS * NC) + (JS * (JS - 1) / 2) + (NC * (NC - 1) / 2)
LF = JS * NC + 1
FOR I = 1 TO NR
FOR J = 1 TO LF - 1
FADAT(I, J) = EM(I, J)  ' Inputting F
NEXT J
NEXT I
FOR I = 1 TO NR
FOR J = LF TO LFA
INPUT #3, FADAT(I, J)  ' Inputting Fa
NEXT J
NEXT I
"** COMBINED MATRIX F & Fa **"
FOR I = 1 TO NR
FOR J = 1 TO LFA
PRINT FADAT(I, J); 'Prints combined F,Fa
NEXT J
PRINT
NEXT I
GOSUB 5000
'begins construction of diagonal
'T = 2 * (JS * NC + (JS * (JS - 1) / 2) + (NC * (NC - 1) / 2))
T = (JS + NC) * (JS + NC - 1)
FOR I = 1 TO JS * NC 'creates Design matrix A with all entries as 0's
FOR J = 1 TO JS + NC
AA(I, J) = 0!
NEXT J
NEXT I
I = 1
N = JS + NC
FOR K = NCAT TO N 'begins construction of design matrix A
FOR J = 1 TO JS
AA(I, J) = 1! 'Inserts 1's
AA(I, K) = -1! 'Inserts -1's
I = I + 1
NEXT J
NEXT K
FOR I = 1 TO LFA 'JS * (JS - 1) / 2 + NC * (NC - 1) / 2 'creates Aa with all entries as 0's
FOR J = 1 TO JS + NC
AAM(I, J) = 0!
NEXT J
NEXT I
P = 1
N = 1 'Inserts 1's & -1's for tau's (JS)
FOR J = JS - 1 TO 1 STEP -1
FOR I = 1 TO J
Q = I + N
AAM(P, I) = 1!
AAM(P, Q) = -1!
P = P + 1
NEXT I
N = N + 1
NEXT J
N = 1
FOR J = JS + NC - 1 TO JS STEP -1  ' Inserts 1's & -1's for mu's (NC)
FOR I = JS + 1 TO J
Q = I + N
AAM(P, I) = 1!
AAM(P, Q) = -1!
P = P + 1
NEXT I
N = N + 1
NEXT J
OPEN "ADAT.RES" FOR OUTPUT AS #4
' *** MATRIX Aa *** (FOR THE DESIGN MATRIX A)
FOR I = 1 TO JS * (JS - 1) / 2 + NC * (NC - 1) / 2
FOR J = 1 TO JS + NC
PRINT #4, AAM(I, J);
'PRINT AAM(I, J):  ' Prints partitioned design matrix Aa
NEXT J
PRINT #4,
NEXT I
CLOSE #4
OPEN "ADAT.RES" FOR INPUT AS #4
FOR I = 1 TO JS * NC
FOR J = 1 TO JS + NC
AM(I, J) = AA(I, J)  ' Reading in A portion
NEXT J
NEXT I
FOR I = JS * NC + 1 TO LFA 'JS * NC + JS * (JS - 1) / 2 + NC * (NC - 1) / 2
FOR J = 1 TO JS + NC
INPUT #4, AM(I, J)  ' Reading in Aa portion
NEXT J
NEXT I
CLOSE #4
PRINT  ' Multiplying [F,Fa] with [A,Aa]
PRINT "*** DESIGN MATRIX A, aA (combined) ***" 
PRINT
FOR I = 1 TO LFA 'JS * NC + JS * (JS - 1) / 2 + NC * (NC - 1) / 2
FOR J = 1 TO JS + NC
PRINT AM(I, J);
NEXT J
PRINT
NEXT I
PRINT
GOSUB 5000
PRINT "*** DOMINANCE MATRIX E *** "
PRINT
FOR I = 1 TO NR
FOR J = 1 TO JS + NC
FD(I, J) = 0!
SUM = 0!
FOR K = 1 TO JS * NC + JS * (JS - 1) / 2 + NC * (NC - 1) / 2
SUM = SUM + FADAT(I, K) * AM(K, J)
NEXT K
FD(I, J) = SUM
PRINT FD(I, J); 'Prints matrix E
NEXT J
PRINT
NEXT I
PRINT
GOSUB 5000
OPEN RANK$ FOR OUTPUT AS #6
FOR I = 1 TO NR
FOR J = 1 TO NC + JS
FDT(I, J) = (NUMVAR + NCAT - FD(I, J)) / 2
PRINT FDT(I, J); ' Prints RANK ORDER TABLE
PRINT #6, FDT(I, J);
NEXT J
PRINT
PRINT #6,
NEXT I
CLOSE
END
5000 'HOLD SCREEN
PRINT
PRINT "HIT ANY KEY TO RETURN TO CONTINUE>"
PRINT
WHILE INKEY$ = ""
WEND
RETURN
APPENDIX 5.  RESIDUAL.BAS program calculates \( \eta^2 \), y-score & residual matrix.

* This program calculates \( \eta^2 \), y-score and residual E
CLS
DIM G(80), R1(80), B(80), E1(80, 15), Y(80)
DIM FD(80, 20), DN(80), DG(15), D2(80)
DIM EOBS(80, 15), M(80)
PRINT "*******************************************************************************"
PRINT "* ETA\(^2\), Y-SCORE and RESIDUAL DOM. MATRIX E  *
PRINT "*******************************************************************************"
PRINT
6500 INPUT "FILE NAME for DOMINANCE MATRIX...=> ", FILENAME$
OPEN FILENAME$ FOR INPUT AS #1
INPUT "NUMBER OF QUESTIONS? >"; NUMVAR
INPUT "NUMBER OF OBSERVATIONS? >"; NUMOBS
INPUT "NUMBER OF SUCCESSIVE CATEGORIES? >"; NCAT
INPUT "FILE CONTAINING ...OPTIMAL WEIGHTS..."; OPTIM$
OPEN OPTIM$ FOR INPUT AS #3
NCOL = NUMVAR + NCAT - 1
FOR I = 1 TO NUMOBS
FOR J = 1 TO NCOL
INPUT #1, FD(I, J)
PRINT FD(I, J);
NEXT J
PRINT
NEXT I
CLOSE #1
FOR I = 1 TO NCOL
INPUT #3, B(I)
NEXT I
CLOSE #3
NR = NUMOBS
NC = NUMVAR
FORMW$ = "##.####"
FORMH$ = "##.#### ##.#### ##.#### ##.#### ##.#### ##.#### ##.#### ##.####
##.#### ##.#### ##.#### ##.#### ##.#### ##.#### ##.####
JS = NCAT - 1
LK = JS * (JS - 1) / 2 + NC * (NC - 1) / 2
LA = JS + NC
'Begins construction of diagonal
'T = 2 * (JS * NC + (JS * (JS - 1) / 2) + (NC * (NC - 1) / 2))

T = (JS + NC) * (JS + NC - 1)
FOR I = 1 TO NR 'matrices for ROW & COLUMN TOTALS
DN(I) = T 'ROW TOTALS
NEXT I
S1 = (NC + JS - 1) * NR
FOR I = 1 TO JS
DG(I) = S1 'Tau's COLUMN TOTALS
NEXT I
S1 = (JS + NC - 1) * NR
L = JS + NC
FOR I = NCAT TO L
DG(I) = S1 'Mu's COLUMN TOTALS
NEXT I
FT = NUMOBS * (NUMVAR + NCAT - 1) * (NUMVAR + NCAT - 2)
NC = JS + NC

6800 'PRINT
PRINT
PRINT "*** NUMBER OF RESPONSES FOR ROWS ***"
PRINT
FORM1$ = "######## ######## ######## ######## ######## ######## ########
######## ######## ######## ######## ######## ########
FOR I = 1 TO NR
PRINT USING FORM1$; DN(I);
NEXT I
PRINT
PRINT
PRINT "*** NUMBER OF RESPONSES FOR COLUMNS ***"
PRINT
FORM2$ = "######## ######## ######## ######## ######## ########
######## ######## ######## ######## ########
FOR I = 1 TO NC
PRINT USING FORM2$; DG(I);
NEXT I
PRINT : PRINT
CHK1 = 0!
CHK2 = 0!
FOR I = 1 TO NR
CHK1 = CHK1 + DN(I) 'vector of ROW TOTALS
NEXT I
FOR I = 1 TO NC
CHK2 = CHK2 + DG(I) 'vector of COLUMN TOTALS
NEXT I
DCHK = ABS(CHK1 - CHK2) 'checks that GRAND TOTAL is correct
IF DCHK > .5 THEN 7700
FOR I = 1 TO NC
IF DG(I) < .5 THEN D2(I) = 0!
IF DG(I) >= .5 THEN D2(I) = 1! / DG(I) ' D2(I) = DC-1
NEXT I
 INPUT "TYPE NAME OF OUTPUT FILE for SUBJECT SCORES> ": SCORES$
OPEN SCORE$ FOR OUTPUT AS #4
INPUT "TYPE NAME OF OUTPUT FILE for RESIDUAL E": SOLUTE$
OPEN SOLUTE$ FOR OUTPUT AS #5
K1 = 0!
7100 FOR I = 1 TO NC
   K2 = 0!
   FOR J = 1 TO NC
      FOR K = 1 TO NR
         IF DN(K) < .1 THEN DR = 0!
         IF DN(K) > .1 THEN DR = 1! / DN(K) ' DR = Dn-1
         K1 = K1 + B(I) * FD(K, I) * DR * FD(K, J) * B(J) ' K1 = THETA.E'.Dn-1.
   NEXT K
   K2 = K2 + B(J) * DG(J) * B(J)
NEXT J
NEXT I
FOR K = 1 TO NR
   Y(K) = 0!
NEXT I
FOR J = 1 TO NC
   Y(K) = Y(K) + DR * FD(K, J) * B(J)
NEXT J
NEXT K
PRINT "...."; K1; "***"; K2; "***"; K1 / K2
FOR K = 1 TO NR
Y(K) = (1 / SQR(K1 / K2)) * Y(K)
PRINT #4, USING FORMWS; Y(K)
NEXT K

7200 TT = 0!
FOR I = 1 TO NR
FOR J = 1 TO NC
EOBS(I, J) = SQR(K1 / K2) * DN(I) * Y(I) * B(J) * DG(J) / FT' residual dominance matrix
NEXT J
NEXT I
FOR I = 1 TO NR
FOR J = 1 TO NC
E1(I, J) = FD(I, J) - EOBS(I, J)
PRINT #5, USING FORMHS; E1(I, J);
NEXT J
PRINT #5,
NEXT I
CLOSE

7600 PRINT " **** END OF ANALYSIS ****"
7700 END