http://researchspace.auckland.ac.nz

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author’s right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.
MAXIMIZING INFORMATION: APPLICATIONS OF IDEAL POINT MODELING AND INNOVATIVE ITEM DESIGN TO PERSONALITY MEASUREMENT

Heidi Vanessa Leeson

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Education,

The University of Auckland, 2008
ABSTRACT

Recent research has challenged the way in which personality and attitude constructs are measured. Alternatives have been offered as to how non-cognitive responses are modeled, the mode of delivery used when administering such scales, and the impact of technology in measuring personality. Thus, the major purpose of the studies in this thesis concerns two interrelated issues of personality research, namely the way personality responses are best modeled, and the most optimal mode by which personality items are presented and associated modal issues. Three studies are presented. First, recent developments using an ideal point approach to scale construction are outlined, and an empirical study compares modeling personality items based on an ideal point approach (generalized graded unfolding model; GGUM) and a dominance approach (graded response model: GRM). Second, an extensive review of literature pertaining to the mode effect when transferring paper-and-pencil measures to screen was conducted, in addition to a review of the various types of computerized and innovative items and their associated psychometric information. Finally, nine innovative items were developed using various multimedia features (e.g., video, graphics, and audio) to ascertain the advantages of these methods to present items constructed to elicit response behavior underlying ideal point approaches, namely, typical response behavior.

It was found that the dominance IRT model continued to produce superior model-data fit for most items, more attention needs to be placed on developing principles for constructing ideal point type items, the web-based version supplied 20% more construct information than the paper version, and innovative items seem to
provide more data-model fit for students with lower personality attributes. While the
innovative items may require more initial outlay in terms of time and development
costs, they have the capacity to provide more information regarding test-takers’
personality levels, potentially using fewer items.
I would like to thank the numerous people, students, and schools who have assisted or participated in the research contained in this thesis. I am especially grateful for the support and guidance of my supervisor Professor John Hattie. Thank you for the invigorating debates on psychometric issues and statistical approaches, and your never-ending enthusiasm for my work. Your feedback, advice, and assistance have made this thesis not only possible, but an exciting adventure. I would also like to thank my second supervisor Associate Professor Mike Townsend. Thanks also to the fabulous team at Project asTTle, you are truly my ‘work family’ and I have appreciated the incredible support (and humour) that you have all given me throughout this journey. Thank you also Jana and Satomi, you have been great fellow PhD students, and I have appreciated the support from both of you.

I acknowledge the financial support from the University of Auckland, whose scholarship has meant that I could undertake this commitment financially. Special thanks also to the Ministry of Education, who provided the significant funding required for the innovative item development conducted in this thesis. Thank you for making it possible to undertake that project.

At a personal level, a special thanks to my family for providing me with the motivation and desire to achieve this goal, you will never know your impact. Extra special thanks to my Mum for deciding that I should learn to read when I was 3 years old, thank you for introducing to me the worlds created from text in a book.

Lastly, I would like to acknowledge and dedicate this thesis to Andrew, Ben and Lilly. You are everything.
LIST OF TABLES

Table 1 Sample items from the Academic Self-Worth Scale 39
Table 2 Total variance explained in the EFA for factors whose eigenvalues exceed 1.00 50
Table 3 Sub-Scale Characteristics 50
Table 4 GGUM and GRM Average Item Parameter Estimations 52
Table 5 Means, Standard Deviations, and Frequencies of Chi-Square to $df$ Ratio in the Calibration Sample 54
Table 6 Number of Fit Plots Providing Best-Fit to Sub-Scales 55
Table 7 The twelve fonts compared in Bernard, Mills, Peterson, et al. (2001) study 77
Table 8 Summary of the main human and technological findings 113
Table 9 Descriptive statistics and cross-mode correlations for the RSPS – Progress items 133
Table 10 Item factor loadings, absolute and incremental fit indices for web-based and paper-and-pencil questionnaires 136
Table 11 Descriptive statistics and difficulty and discrimination item parameters for paper-and-pencil and web-based versions 137
Table 12 Frequencies and Percentages for Reasons for Preferred Mode Choice 144
Table 13 Frequencies and Percentages for the Mode that Best Reflected their Self-Perceptions of Reading Progress 146
Table 14 Frequencies and Percentages for Likes and Dislikes of the Web-based Questionnaire 147
Table 15 Frequencies and Percentages for Participants Attitudes Regarding Different Features of the Web-based Questionnaire 148
### LIST OF FIGURES

| Figure 1 | Fit Plots for the GGUM and GRM (Item 11 of the Reflectivity sub-scale): (a) Strongly Disagree Option (GRM), (b) Strongly Disagree Option (GGUM), (c) Disagree Option (GRM), (d) Disagree Option (GGUM), (e) Neither Disagree Nor Agree Option (GRM), (f) Neither Disagree Nor Agree Option (GGUM), (g) Agree Option (GRM), (h) Agree Option (GGUM), (i) Strongly Agree Option (GRM), (j) Strongly Agree Option (GGUM). | 56 |
| Figure 2 | Fit Plots for the GGUM and GRM (Item 8 of the Self-Presented sub-scale): (a) Strongly Disagree Option (GRM), (b) Strongly Disagree Option (GGUM), (c) Disagree Option (GRM), (d) Disagree Option (GGUM), (e) Neither Disagree Nor Agree Option (GRM), (f) Neither Disagree Nor Agree Option (GGUM), (g) Agree Option (GRM), (h) Agree Option (GGUM), (i) Strongly Agree Option (GRM), (j) Strongly Agree Option (GGUM) | 59 |
| Figure 3 | An example of a figural response item (Zenisky & Sireci, 2002). | 92 |
| Figure 4 | An example of a figural response item (Martinez & Jenkins, 1993). | 93 |
| Figure 5 | An example of a drag-and-drop item presenting the Periodic Table of the Elements with missing elements (Zenisky & Sireci, 2002). | 94 |
| Figure 6 | A drag-and-drop item requiring test-takers to select and drag their chart title choices to their correct position on the work area of the item (Microsoft Corporation, 2003). | 95 |
| Figure 7 | An example of a graphical modeling item format (Bennett, Morley, Quardt, & Rock, 2000). | 96 |
| Figure 8 | An example of a drop-and-connect item format (Jodoin, 2003). | 97 |
| Figure 9 | An example of a specifying relationships item type (Zenisky & Sireci, 2002). | 98 |
| Figure 10 | An example of a create-a-tree item type (Zenisky & Sireci, 2002). | 99 |
| Figure 11 | An example of an ETS capturing frame item (Bennett, Goodman, Hessinger, Kahn, Ligget, Marshall, & Zack, 1999). | 100 |
| Figure 12 | An example of an ETS capturing frame item (Bennett, Goodman, Hessinger, Kahn, Ligget, Marshall, & Zack, 1999). | 101 |
Figure 13  An example of a multiple selection item (Shotland, Alliger, & Sales, 1998).

Figure 14  An example of an ETS analyzing situation item (Bennett, Goodman, Hessinger, Kahn, Ligget, Marshall, & Zack, 1999).

Figure 15  Example of a GRE essay item format whereby test-takers are required to type a response into a text box that has basic editing functions available (cut, paste, undo) (Microsoft Corporation, 2003).

Figure 16  Example of an ETS generating examples item (Bennett & Rock, 1998).

Figure 17  An example of an ETS document-analysis item that asks the test-taker to analyze and compare the persuasive techniques adopted in wartime propaganda (Bennett, Goodman, Hessinger, Kahn, Ligget, Marshall, & Zack, 1999).

Figure 18  An example of a text-editing item (Breland, 1999).

Figure 19  Item information for the nine RSPS Progress items in both paper-and-pencil and web-based versions.

Figure 20  Average IRT information across theta levels for paper-and-pencil and web-based versions of the RSPS Progress sub-scale
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................. I

ACKNOWLEDGEMENTS .......................................................................................... III

LIST OF TABLES .................................................................................................... IV

LIST OF FIGURES ................................................................................................. V

TABLE OF CONTENTS ......................................................................................... VII

CHAPTER ONE .......................................................................................................... 1

INTRODUCTION ...................................................................................................... 1

Purpose of the Research ......................................................................................... 5

Significance of the Research ................................................................................... 6

Design of the Research ........................................................................................... 9

CHAPTER TWO ....................................................................................................... 11

LITERATURE REVIEW ......................................................................................... 11

Self-worth Strategies ............................................................................................. 12

Self-handicapping ............................................................................................. 12

Active self-handicapping ............................................................................... 13

Self-presented self-handicapping ................................................................. 14

Self-presented affective self-handicapping ................................................... 15

Defensive Pessimism ........................................................................................ 15

Active Defensive Expectations ..................................................................... 16

Self-Presented Defensive Expectations ......................................................... 16

Reflectivity ........................................................................................................ 16

Response Behaviors ......................................................................................... 17

Measurement Theory ......................................................................................... 19

Item Response Theory (IRT) ............................................................................ 19

Graded Response Model (GRM)................................................................ 20

Ideal Point Item Response Theory ............................................................. 21

Generalized Graded Unfolding Model (GGUM) ............................................ 22

Applications of Ideal Point and Item Response Approaches to Personality
Items .................................................................................................................. 23

Reading Self-efficacy ........................................................................................ 26

Innovative Item Design - Framework ............................................................ 27

Multimedia Testing ........................................................................................... 31

Equivalence Testing ......................................................................................... 32

Efficiency Testing .............................................................................................. 35

CHAPTER THREE ................................................................................................... 37
STUDY ONE: COMPARISON MODEL-DATA FIT FROM IRT AND IDEAL POINT IRT MODELS TO AN IDEAL POINT SCALE ............................................. 37

Method .................................................................................................................. 38
Participants ........................................................................................................ 38
Scale Development ........................................................................................... 38
Instrument ...................................................................................................... 38
Self-handicapping items ................................................................................ 39
Defensive pessimism items ........................................................................... 41
Procedure .......................................................................................................... 42
Statistical Analysis ............................................................................................ 44
Cross-validation ............................................................................................ 44
Exploratory Factor Analysis of the Academic Self-Worth Scale .......... 45
IRT Item Parameter Estimation ..................................................................... 45
Model-Data Fit .............................................................................................. 46

Results ................................................................................................................... 49
Exploratory Factor Analysis of the Academic Self-Worth Scale .......... 49
IRT Item Parameter Estimation ........................................................................ 52
Model-Data Fit .................................................................................................. 54

Discussion ............................................................................................................. 62

CHAPTER FOUR ...................................................................................................... 66

STUDY TWO: A LITERATURE REVIEW OF HUMAN, TECHNOLOGICAL AND ITEM DESIGN ISSUES IN COMPUTERIZED TESTING ........................... 66

Participant Issues ................................................................................................. 67
Race, Ethnicity, and Gender ............................................................................. 67
Cognitive Processing ......................................................................................... 68
Ability ............................................................................................................... 69
Familiarity with Computers ............................................................................ 71
Computer Anxiety ............................................................................................. 73

User Interface – Legibility .................................................................................. 74
Screen Size and Resolution ............................................................................. 74
Font Characteristics ........................................................................................ 76
Line Length ....................................................................................................... 79
Number of Lines ............................................................................................... 82
Interline Spacing ............................................................................................... 83
White Space ...................................................................................................... 84

User Interface – Interactive ................................................................................ 86
Scrolling ............................................................................................................ 86
Item Review ...................................................................................................... 87
Item Presentation ................................................................................................. 90

Computerized Item Design Types ................................................................. 91
Item Format and Response Actions ............................................................... 91
Mouse-based Response Action ................................................................... 91
Figural Response Items .................................................................................. 91
CHAPTER FIVE ........................................................................................................ 121

STUDY THREE: WEB-BASED AND PAPER-AND-PENCIL VERSIONS OF 
THE READER SELF-PERCEPTION SCALE: A COMPARISON OF 
MEASUREMENT EFFICIENCY AND PARTICIPANTS’ PERCEPTIONS...... 121

Method ................................................................................................................ 123
  Participants........................................................................................................ 123
  Design............................................................................................................. 123
Measures ............................................................................................................ 124
  Reader Self-Perception Scale (RSPS) – Paper-and-Pencil ......................... 124
  Reader Self-Perception Scale (RSPS) – Web-based .................................... 125
  RSPS – Progress sub-scale modifications ................................................... 126
  Post-assessment feedback measure ............................................................ 128
Procedure .......................................................................................................... 129
  Paper-and-pencil administration ................................................................. 130
  Web administration ...................................................................................... 130

Results .............................................................................................................. 131
Data Analysis .................................................................................................. 131
  Descriptive Statistics ................................................................................... 131
  Confirmatory Factor Analysis .................................................................... 134
  Item Parameter Statistics ........................................................................... 136
  Efficiency Analysis ..................................................................................... 137
  Item Level ................................................................................................ 138
  Test Level ................................................................................................ 139
  Mode Preferences ....................................................................................... 142

Discussion ...................................................................................................... 149

CHAPTER SIX ..................................................................................................... 157