UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

The Validity of Value-Added Measures in Secondary Schools

By Vanessa Scherman

Submitted in fulfilment of the requirements for the degree of
PhD: Assessment and Quality Assurance

In the Department of Curriculum Studies
Faculty of Education
University of Pretoria
PRETORIA

March 2007

Supervisor: Prof. Sarah Howie, University of Pretoria
Co-supervisor: Prof. Tjeerd Plomp, University of Twente

Table of Contents

List of Tables **vi**
List of Figures **viii**
List of Acronyms **ix**
Summary **xi**
Acknowledgements **xiii**

CHAPTER 1 INTRODUCTION **1**

1.1 Introduction 2
1.2 South African education context 4
1.2.1 The South African education system 4
1.2.2 National revised curriculum statement 8
1.2.3 Assessment policies 12
1.3 Key concepts 16
1.3.1 Quality in education 16
1.3.2 Monitoring in education 17
1.3.3 Value-added assessment as a monitoring mechanism 19
1.4 Policies on monitoring quality in education 19
1.5 Research problem in context 23
1.6 Structure of the dissertation 26

CHAPTER 2 MONITORING QUALITY EDUCATION: REVIEW OF THE LITERATURE **27**

2.1 Introduction 27
2.2 Monitoring systems in education 28
2.2.1 Characteristics of monitoring systems 30
2.2.2 The ZEBO-project in the Netherlands 31
2.2.3 The Victorian Certificate of Education (VCE) data project in Australia 33
2.2.4 The ABC+ model in the United States of America 35
2.3 Comparison of selected monitoring systems 38
2.4 Value-added monitoring systems 42
2.4.1 Value-added monitoring systems in the United Kingdom 43
2.4.2 Value-added monitoring systems in the United States 45
2.4.3 Concluding remarks regarding value-added monitoring systems 46
2.5 Ability as a predictor of academic success 48

2.6 The relevance of skills and abilities against the backdrop of curriculum 50
2.7 Factors influencing performance 52
2.8 Conclusion 59

CHAPTER 3 CONCEPTUAL FRAMEWORK FOR THE STUDY 63

3.1 Introduction 63
3.2 School effectiveness models 64
3.3 Scheeren's model for school effectiveness research 66
3.4 Model for monitoring education in South Africa 69
3.4.1 The context 74
3.4.2 Input indicators 75
3.4.3 Process indicators 75
3.4.4 Output indicators 78
3.5 Specific research questions 79
3.6 Conclusion 85

CHAPTER 4 MIDDLE YEARS INFORMATION SYSTEM (MIDYIS): CHALLENGES AND CHANGES 86

4.1 Introduction 86
4.2 MidYIS in the United Kingdom 88
4.3 The MidYIS instruments 90
4.3.1 The MidYIS Scales 90
4.3.2 The vocabulary sub-test 91
4.3.3 The mathematics sub-test 92
4.3.4 The proof reading sub-test 93
4.3.5 The perceptual speed and accuracy sub-test 93
4.3.6 The cross-sections sub-test 94
4.3.7 The block counting sub-test 95
4.3.8 The pictures sub-test 95
4.3.9 Extended MidYIS 96
4.4 Feedback provided by the MidYIS project 97
4.4.1 Nationally standardised feedback 98
4.4.2 Band profile graphs 100
4.4.3 Predictions to Key Stage 3 and GCSE 101
4.4.4 Chances graphs 103
4.4.5 Value-added feedback 103
4.5 Evaluation of the MidYIS project and relevance for South Africa 104

4.6 Summary and adaptations to enhance MidYIS for South Africa 107
4.7 Conclusion 112

CHAPTER 5 RESEARCH DESIGN AND METHODS 117

5.1 Introduction 117
5.1.1 The development of pragmatism 118
5.1.2 Rationale for working within a pragmatist paradigm 121
5.2 Design issues 123
5.3 Methodology 130
5.3.1 Sample 131
5.3.2 Instruments 132
5.3.2.1 Assessment instrument 132
5.3.2.2 Questionnaires 132
5.3.2.3 Interview schedules 135
5.3.3 Validity issues related to the instruments 135
5.3.3.1 Validation of assessment instrument 136
5.3.3.2 Validation of the questionnaires 142
5.3.4 Data collection 143
5.3.4.1 Document analysis 144
5.3.4.2 Evaluation reports 145
5.3.4.3 National-level data collection 145
5.3.4.4 Provincial level data collection 146
5.3.4.5 School-level data collection 146
5.3.5 Data analysis 147
5.3.5.1 Analysis of documents 148
5.3.5.2 Analysis of qualitative data 149
5.3.5.3 Analysis of assessment data 151
5.3.5.4 Analysis of contextual data 158
5.3.5.5 Building an exploratory model using assessment and contextual data 159
5.3.6 Research procedures 163
5.4 Ethical issues 164
5.5 Methodological constraints 165
5.6 Conclusion 165

CHAPTER 6 THE CONTENT–RELATED VALIDITY OF THE MIDYIS ASSESSMENT 167

6.1 Introduction 167
6.2 Exploring the curriculum validity of the MidYIS assessment 170

6.2.1 Perspectives from National and Provincial Departments of Education 171
6.2.2 The language learning area 177
6.2.3 The mathematics learning area 184
6.2.4 Exploring possible suggestions for the revision of MidYIS 191
6.3 Exploring the content validity of the MidYIS assessment 194
6.4 Conclusion 197

CHAPTER 7 THE CONSTRUCT-RELATED VALIDITY AND RELIABILITY OF THE MIDYIS ASSESSMENT 201

7.1 Introduction 201
7.2 Participant characteristics 205
7.3 Elaborating on construct validity 206
7.3.1 Investigating construct validity by means of Rasch analysis 207
7.3.1.1 Vocabulary sub-test 208
7.3.1.2 Mathematics sub-test 211
7.3.1.3 Proof reading sub-test 213
7.3.1.4 Perceptual speed and accuracy sub-test 217
7.3.1.5 Cross-sections sub-test 220
7.3.1.6 Block counting sub-test 222
7.3.1.7 Pictures sub-test 225
7.3.2 Conclusions drawn from the Rasch analyses 228
7.4 Exploring the reliability of the MidYIS assessment 230
7.5 Exploring relationships between MidYIS scores and academic achievement 232
7.6 Conclusion 238

CHAPTER 8 FACTORS THAT INFLUENCE PERFORMANCE ON MIDYIS 240

8.1 Introduction 240
8.2 Preparation for model building 242
8.2.1 Identifying variables to be explored with multilevel analyses 242
8.2.2 Approach to model building 246
8.3 The results of the multilevel analyses 248
8.3.1 The null model 248
8.3.2 The learner, classroom, school-level model 250
8.4 Conclusion 256

CHAPTER 9 CONCLUSIONS AND RECOMMENDATIONS **260**

9.1 Introduction 260
9.2 Summary of the research 261
9.3 Discussion and reflection 278
9.3.1 Methodological reflections 278
9.3.2 Reflection on this research in of light school effectiveness research 280
9.3.3 Contribution to scientific and practical knowledge 283
9.4 Recommendations 288
9.4.1 Recommendations and issues regarding monitoring systems 288
9.4.2 Recommendations and issues regarding policy 289
9.4.3 Recommendations for further research 290
9.5 The role of feedback and intervention 292
9.6 Conclusion 299

REFERENCES **301**

Appendix A: Description of the MidYIS sub-tests **334**
Appendix B: Description of constructs included in the learner questionnaire **337**
Appendix C: Description of constructs included in the educator questionnaire **340**
Appendix D: Description of the constructs in the principal questionnaire **344**
Appendix E: Audit trail documents **348**
Appendix F: Evaluation report guidelines **350**
Appendix G: Summary of reports from the language and mathematics specialists **352**
Appendix H: Diagramatic representation of the research procedures undertaken **360**
Appendix I: Letters of consent **362**
Appendix J: Assessment framework for mathematics **369**
Appendix K: Complete list of ability factors **376**
Appendix L: Rasch and correlation analyses **380**
Appendix M: Multilevel analyses **382**
Appendix N: Ethical clearance and language editing **385**

List of Tables

Table 1.1 Information on learners, educators and schools in 2003 6
Table 1.2 Trends in the purpose of assessment 15
Table 1.3 Policy initiatives aimed at monitoring the quality of education 21
Table 2.1 Components and indicators of the ZEBO-project 33
Table 2.2 Data elements of the ABC+ model 37
Table 2.3 Characteristics of the monitoring systems 39
Table 3.1 Overview of indicators and variables included in the conceptual model 73
Table 4.1 Typology of indicators for education monitored by CEM 87
Table 4.2 Age group of learners participating in the MidYIS project 89
Table 4.3 Correlations between MidYIS assessments and Key Stage 3 examinations 97
Table 4.4 Reliability coefficients for the UK, Year 7/ 8 assessment 97
Table 4.5 Reliability coefficients for the UK, Year 9 assessment 98
Table 4.6 GCSE grades and equivalent MidYIS scores 102
Table 4.7 Characteristics of the MidYIS monitoring system 105
Table 5.1 Comparison between post-positivism, constructivism, and pragmatism 120
Table 5.2 Summary of research questions, sources and data instruments 144
Table 6.1 Skills as indicated by the Provincial Department Education officials 176
Table 6.2 Outcomes in the languages learning area 177
Table 6.3 Proposed overlap between the language learning area and MidYIS 182
Table 6.4 Outcomes in the mathematics learning area 185
Table 6.5 Accessibility of mathematics items 188
Table 6.6 Proposed overlap between the mathematics learning area and MidYIS 190
Table 6.7 Summary of ability factors associated with abilities or aptitude assessments 196
Table 7.1 Age and gender distribution of participating learners 205
Table 7.2 Home language of learners who participated 206
Table 7.3 Population group of learners who participated 206
Table 7.4 Initial statistics for vocabulary sub-test 209
Table 7.5 Final statistics for vocabulary sub-test 209
Table 7.6 Initial statistics for mathematics sub-test 211
Table 7.7 Final statistics for mathematics sub-test 212
Table 7.8 Initial statistics for proof reading 1 sub-test 214
Table 7.9 Final statistics for proof reading 1 sub-test 214
Table 7.10 Initial statistics for proof reading 2 sub-test 216
Table 7.11 Final statistics for proof reading 2 sub-test 216
Table 7.12 Initial statistics for perceptual speed and accuracy sub-test 218
Table 7.13 Final statistics for perceptual speed and accuracy sub-test 218

Table 7.14 Initial statistics for cross-sections sub-test 220
Table 7.15 Final statistics for cross-sections sub-test 220
Table 7.16 Initial statistics for block counting sub-test 223
Table 7.17 Final statistics for block counting sub-test 223
Table 7.18 Initial statistics for pictures sub-test 226
Table 7.19 Final statistics for pictures sub-test 226
Table 7.20 Reliability analysis and standard error of measurement per scale 231
Table 7.21 Reliability analysis per scale and population groups of learners 232
Table 7.22 Correlations between the revised MidYIS scales and school mathematics 234
Table 7.23 Correlations between the revised MidYIS scales and school English 236
Table 8.1 Description and significant correlations with MidYIS total 244
Table 8.2 The intercept-only model 249
Table 8.3 Variance explained at the learner, classroom and school-level 250
Table 8.4 Progression in model building 252
Table 8.5 Proportion of variance explained by consecutive models for language 255
Table 9.1 A comparison of the ZEBO-project, VCE data project, ABC+ model and MidYIS 270

List of Figures

Figure 3.1 School effectiveness model as developed by Scheerens (1990) 67
Figure 3.2 Conceptual framework for monitoring education in South Africa 72
Figure 3.3 Criteria for evaluating quality of measurement used in monitoring systems 80
Figure 3.4 Components included for study 84
Figure 4.1 The scales and sub-tests of the MidYIS assessment 90
Figure 4.2 Example from the vocabulary sub-test 92
Figure 4.3 Example from the mathematics sub-test 92
Figure 4.4 Example item from the proof reading sub-test 93
Figure 4.5 Example item from the perceptual speed and accuracy sub-test 94
Figure 4.6 Example item from the cross-sections sub-test 94
Figure 4.7 Example item from the block counting sub-test 95
Figure 4.8 Example item from the pictures sub-test 96
Figure 4.9 Standardised scores 99
Figure 4.10 MidYIS bands represented on a normal distribution 100
Figure 4.11 Band profile graphs 101
Figure 4.12 Predictions to GCSE subjects 102
Figure 4.13 Learner-level chances graph for English 103
Figure 4.14 Value-added analysis 104
Figure 5.1 A mixed method process model 129
Figure 5.2 Iterative model of the different components of qualitative data 150
Figure 5.3 Example of an item characteristic curve 152
Figure 6.1 Extension of the criteria for evaluating quality of measurement 168
Figure 7.1 Extension of the criteria for evaluating quality of measurement 202
Figure 7.2 Item and person map for the vocabulary sub-test 210
Figure 7.3 Item and person map for the mathematics sub-test 213
Figure 7.4 Item and person map for the proof reading 1 sub-test 215
Figure 7.5 Item and person map for the proof reading 2 sub-test 217
Figure 7.6 Item and person map for the perceptual speed and accuracy sub-test 219
Figure 7.7 Item and person map for the cross-sections sub-test 222
Figure 7.8 Item and person map for the block counting sub-test 225
Figure 7.9 Item and person map for the pictures sub-test 228
Figure 8.1 Proposed model for the multilevel analyses 246
Figure 9.1 Overview of the main research, specific research and sub-research questions 267
Figure 9.2 Conceptual framework for monitoring education in South Africa 287
Figure 9.3 Conceptual framework for this study on monitoring the quality of education 295
Figure 9.4 School-based model for improvement 296
Figure 9.5 Evaluation model for school-based improvement 297

List of Acronyms

ABC+ - Attitudinal/Behavioural/Cognitive Indicators plus Context
Alis - A-level Information System
CASS - Continuous Assessment
CEA - Centre for Evaluation and Assessment
CEM - Curriculum, Evaluation and Management Centre
DAT - Differential Aptitude Test
FET - Further Education and Training
GAT - General Achievement Test
GCSE - General Certificate for Secondary Education
GDE - Gauteng Department of Education
GET - General Education and Training Band
GSAT - General Scholastic Aptitude Test Battery (GSAT)
HET - Higher Education and Training
HLM - Hierarchical Linear Models
HSRC – Human Sciences Research Council
IQMS - Integrated Quality Management System
LEA - Local Education Authorities
JAT – Junior Aptitude Test
MLA - Monitoring Learning Achievement
MidYIS - Middle Years Information System
NAPTOSA – National Professional Teachers' Organisation of South Africa
NFER - National Foundation for Educational Research
OBE - Outcomes-Based Education
OFSTED - Gauteng Department of Education Office for Standards in Education
PARIS - Predictions and Reporting Interactive Software
PIPS - Performance Indicators at Primary School
PIRLS - Progress in International Reading Literacy Study
QAIT/MACRO - Quality, Appropriateness, Incentive, Time of instruction/Meaningful goals, Attention to academic focus, Coordination, Recruitment and training, Organisation
QUAN - Quantitative Research
QUAL - Qualitative Research
QUASE - Quantitative Analysis for Self Evaluation
RNCS - Revised National Curriculum Statement

SACMEQ - Southern Africa Consortium for Monitoring Educational Quality

SER – School Effectiveness Research

SGB – School-Governing Body

SSAIS - Senior South African Individual Scale (SSAIS)

SASSIS - South African Secondary School Information System

SAT - Senior Aptitude Test

SATIS - Student Attitudes Information System

SE - Systemic Evaluation

SITES - Second International Technology in Education Study

TAD - Test of Developed Ability

TIMSS - Third International Mathematics and Science Study

UK - United Kingdom

USA - United States of America

VCE - Victorian Certificate of Education

WAIS - South African Wechsler Adult Intelligence Scale

WISC - Wechsler Intelligence Scale for Children

WSE - Whole School Evaluation

ZEBO - Self-Evaluation in Primary Schools

Summary

The issue of quality education is a critical topic of discussion, for South Africa facing the challenge of implementation amidst a plethora of progressive policies. This research project is undertaken in collaboration with the Curriculum, Evaluation, and Management Centre (CEM) at Durham University in the United Kingdom. The Middle Years Information System (MidYIS) project was developed with the aim of providing schools with information on how learners would perform at the end of two national examinations namely Key Stage 3 and General Certificate in Secondary Education, in addition to providing value-added information. The purpose of the research reported here is to describe the procedures undertaken to explore the feasibility of implementing the MidYIS system in the South African context.

The research was guided by two main research questions. The first research main research question is ***how appropriate is the Middle Years Information System (MidYIS) as a monitoring system in the South African context?*** The word "appropriate" here interrogates the suitability of the MidYIS system for South Africa looking specifically at validity and reliability issues. This non-experimental study used a mixed methods design, rooted in pragmatism, to explore validity and reliability issues of using MidYIS as a possible monitoring system that would provide a balanced view of the school's contribution to academic gains made by learners. The sample included in the study ranged from National Department of Education officials (two officials from curriculum and assessment), Provincial Department of Education officials (one mathematics specialist, one language specialist and one specialist from the Gauteng Department of Education Office for Standards in Education), specialists in the field of language, mathematics, and psychology as well as 11 schools. In particular content-related validity (including curriculum validity), construct-related validity, and predictive validity were examined while inferences drawn with regard to reliability were done by means of internal consistency reliability. From a curriculum perspective for content-related validity, it was found that there was moderate curriculum validity for language while inferences drawn for mathematics were substantially stronger. For content-related validity from a psychometric perspective, it was found that there was overlap between the domain of developed abilities and the MidYIS assessment. Construct-related validity was explored by means of Rasch analysis and it was found that items in the MidYIS assessment tend to form well-defined constructs. Predictive validity was explored by means of correlation analysis between the MidYIS assessment and school-based results in language and mathematics. The analysis shows that it could be possible to use the MidYIS assessment for prediction purposes. However, additional research would be needed to explore this facet of validity further with a larger sample and using standardised school-based results. The MidYIS assessment was found to be reliable for the sample as a whole as well as for population groups within the sample.

The second main research question extends the first research question. If MidYIS is valid, with South African adaptations, and reliable, then what factors on a school, classroom, and learner-level could have an effect on learner performance. Thus, the second main research question is ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system?***

In order to explore factors, multilevel analysis was undertaken on the various levels within the school system namely the principals, mathematics and language educators, as well as learners who completed questionnaires. It was found that four learner-level factors (with whom learners live, mother's level of education, importance of mathematics and importance of English), one educator level factor (challenges to assessment due to lack of in-service training) and two school-level factors (educators make use of monitoring systems and encouraging academic achievement) seem to have an effect on the performance of learners.

Key words: *school effectiveness, school improvement, monitoring, quality education, monitoring systems, factors influencing achievement, construct-related validity, content-related validity, curriculum validity, test-curriculum-overlap, predictive validity, reliability, Rasch analysis, multilevel analysis, mixed methods, pragmatism*

Acknowledgements

I am thankful to God for providing me with this special opportunity. He provided every step of the way whether sending someone upon my path to encourage me or giving me the quiet assurance that He is present and will make a way.

I would like to extend my gratitude and appreciation to Prof Sarah Howie and Prof Tjeerd Plomp who guided my along this Safari. It has been a long journey not without its heartaches and rewards. You have shown me what it means to dig a little deeper, to stretch the boundaries a little further, to explore a little more even when time seemed too short and the ideas too big.

To the most precious gift of all Victor Scherman, my husband, you have been a continued source of support. I am extremely thankful for your hugs or reassurance and kind words of encouragement when things just seemed to be too much or the road just too long to travel.

To Tina Lopes for walking this path called PhD with me. Your friendship has meant the world to me.

To my Family, thank you for understanding when I regrettably declined the opportunity of spending time with you.

To Elsie Venter, Hayley Barnes and Liz Archer, your friendship and ongoing support is greatly appreciated.

To the Staff at the Centre for Evaluation and Assessment (CEA), you have continually supported and encouraged me throughout my studies. Thank you for your patience when my thoughts were elsewhere.

To the Staff in the Department of Curriculum Studies for the words of encouragement and interest in my progress, especially Prof William Fraser for his continued support.

To Prof Jonathan Jansen for believing in me and for the words of encouragement given.

Thank you to the National Research Foundation from whom I received a grant holder's doctoral bursary.

To John Barnard and Hans Luyten who helped with some of the technical aspects of the analysis included in the pages to follow. Thank you for your willingness to assist and guide me.

To Pieter van der Merwe who assisted with the language editing. The dissertation is a better product thanks to your patient intervention.

To Jeannie Beukes from the Faculty of Education Administration, who provided substantial administrative support during the examination process. Thank you for your time and effort.

To the Curriculum, Evaluation and Management Centre (CEM) for allowing the use of their instruments. Also to the staff at the CEM centre, who assisted with queries and helped me to understand the rationale behind the use of the monitoring systems in the context of the United Kingdom.

To the specialists in the fields of education and psychology. Thank you for your time and effort.

A special word of thanks to the National and Provincial Department of Education officials who participated in this research.

Finally but by far not the least, to the 11 schools, educators, parents and learners who participated in the study, thank you for allowing me into your schools and classrooms. This research would not have been possible without you.