An investigation into factors affecting Grade 10 and 12 students' mathematics performance in Port Moresby, Papua New Guinea (PNG)

Jerome Oko

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

School of Education
Faculty of Arts
The University of Adelaide

March, 2021

Table of Contents

Table of Contents ..... **i**
List of Tables ..... **x**
List of Figures ..... **xii**
Abstract ..... **xiv**
Declaration ..... **xvi**
Acknowledgement ..... **xvii**
Abbreviations ..... **xix**
Chapter 1: Issues in mathematics education **1**
1.1 Introduction/Background 1
1.2 Statement of the problem 2
1.2.1 Overview of Papua New Guinea education systems 2
1.2.2 Lower and upper secondary school mathematics curriculum in PNG 8
1.2.3 Content of lower secondary mathematics and its implications 10
1.2.4 Content of upper secondary mathematics and its implications 11
1.3 The status of mathematics within the PNG education system 12
1.4 Research questions 15
1.5 Aims and objectives of the study 15
1.6 Significance of the study 16
1.7 Scope and limitations of the study 18
1.8 Structure of this thesis 19
1.9 Summary 22
Chapter 2: Factors affecting students' mathematics results **24**
2.1 Introduction 24
2.2 Issues affecting mathematics achievement 24
2.3 Student-level factors 25
2.3.1 Gender differences and similarities 25
2.3.2 Parents education level 27
2.3.3 Attitude towards mathematics 29
2.3.4 Motivation toward mathematics 33
2.4 Teacher-level factors 37
2.4.1 Teacher education-level 37

2.4.2 Classroom learning environment 39
2.4.3 Quality of teaching 43
2.5 School-level factors 48
2.5.1 Instructional resources 49
2.5.2 School type 51
2.6 Theoretical framework 53
2.7 Summary 58
Chapter 3: Research Methodology 60
3.1 Introduction 60
3.1 Planning stage 60
3.1.1 Location of the study's focus 60
3.2 Research design and method 61
3.3 Ethics approval 62
3.4 Sampling and data collection 63
3.4.1 Sampling framework 63
3.4.2 Sampling selection technique 66
3.4.3 Selection of schools and participants 66
3.4.4 Instruments 67
3.4.5 Linking quantitative and qualitative research methods 69
3.5 Survey instruments 72
3.5.1 Adoption, modification and development of instruments 73
3.5.2 Development of interview questions 74
3.5.3 Final instrument of the main study 75
3.5.4 Data administration and collection 76
3.5.5 Administrating the instruments 76
3.5.6 Administrating the interviews 78
3.6 Statistical procedures and methods employed in the study 79
3.6.1 Confirmatory factor analysis 79
3.6.2 Structural equation modelling 80
3.6.3 Rasch analysis 84
3.6.3.1 Item fit statistics 87
3.6.3.2 Rating scale model, Partial credit model and Multidimensional model 88
3.6.4 Differential item functioning 90
3.6.4.1 Item threshold approach 91

3.6.4.2 Item fit approach 94
3.6.4.3 Group plot methods 94
3.6.5 Scoring 95
3.6.6 Hierarchical linear modelling (HLM) 95
3.6.7 Qualitative data 97
3.6.8 Validity and reliability 98
3.6.9 General methodological consideration 101
3.6.9.1 Aggregation and disaggregation of data 101
3.6.9.2 Preparation of data 102
3.6.9.3 Missing values 102
3.7 Summary 104
Chapter 4: Teacher and student demographic information 105
4.1 Introduction 105
4.2 The sample: descriptive information 106
4.2.1 Teacher gender 106
4.2.2 Academic qualification 107
4.2.3 Years of teaching experience 108
4.2.4 Major area of specialisation 109
4.2.5 Minor area of specialisation 110
4.2.6 School type 111
4.3 Student gender 112
4.4 Parents highest education level 114
4.5 Parents employment 116
4.6 Summary 118
Chapter 5: Grade 10 mathematics test validation 119
5.1 Introduction 119
5.2 Mathematics achievement test 119
5.3 Development of the mathematics test items 120
5.4 Results of the Rasch analysis 125
5.4.1 Overall test statistics 125
5.4.2 Individual item statistics 127
5.4.3 Differential item functioning 132
5.4.3.1 Overall test level DIF for gender 132

5.4.3.2 Item level DIF for gender 134
5.4.3.3 Item threshold approach 135
5.4.3.4 Item fit approach 138
5.4.3.5 Group plot methods 140
5.4.4 Summary of the findings of the three methods 142
5.4.5 Final instrument 144
5.4.6 Person-item map 146
5.5 Transformation to W-scores 147
5.6 Summary 148
Chapter 6: Grade 12 mathematics test validation 149
6.1 Introduction 149
6.2 Development of the mathematics test items 149
6.3 Rasch measurement model 152
6.3.1 Overall statistics 152
6.3.2 Individual statistics 156
6.3.3 Differential item functioning 158
6.3.3.1 Overall DIF level test for gender 159
6.3.3.2 Item-level DIF for gender 160
6.3.3.3 Item fit approach 161
6.3.3.4 Item threshold approach 163
6.3.3.5 Group plot method 165
6.3.4 Summary of the findings of the three methods 167
6.3.5 Final instrument 168
6.3.6 Person-item map 168
6.4 Transformation to W-scores 169
6.5 Summary 170
Chapter 7: Attitude scale validation 171
7.1 Introduction 171
7.2 Contextualisation of attitude scale 172
7.3 Attitude scale details in this study 175
7.4 CFA of the measurement model for the attitude scale 176
7.4.1 Structure analysis of the hypothesised measurement model 177
7.4.2 CFA of the hypothesised measurement model 178
7.4.3 Model fit 179

7.4.4 CFA analysis of the alternative models 179
7.5.5 CFA for the attitude alternative measurement model 181
7.4.6 Model fit 182
7.4.7 Summary of confirmatory factor analysis 183
7.5 Multidimensional Rasch analysis 183
7.5.1 Item analysis with the Partial credit model 184
7. 5.2 Person-item map 186
7.6 Transformation to W-scores 188
7.7 Comparison of the different MRM and CFA results 188
7.8 Model employed for the study 189
7.9 Summary 189
Chapter 8: Motivation scale validation 191
8.1 Introduction 191
8.2 Contextualisation of motivation scale 192
8.3 The motivation scale in this study 194
8.4 CFA of the measurement model for the motivation scale 196
8.4.1 Structure analysis of the hypothesised measurement model 196
8.4.2 CFA of the hypothesised measurement model 197
8.4.3 Model fit 198
8.4.4 Comparison of CFA of the different alternative models 199
8.4.5 CFA for the motivation hierarchical model 201
8.4.6 Model fit 202
8.5 Multidimensional Rasch analysis 203
8.5.1 Item analysis with the Partial Credit Model 204
8.5.2 Person-item map 207
8.6 Transformation to W-scores 209
8.7 Comparison of different RMM and CFA results 209
8.8 Model employed for the study 210
8.9 Summary 210
Chapter 9: Validation of teacher-level scales 212
9.1 Introduction 212
9.2 Contextualisation of teacher-level scales 212
9.2.1 Quality of teaching 212
9.2.2 Classroom environment 214

9.2.3 Instructional resources 216
9.3 CFA of the measurement models for the teacher-level scales 217
9.3.1 Structure analysis of the hypothesised measurement model 218
9.3.2 CFA of the hypothesised measurement model 219
9.3.3 Model fit 220
9.4 Rasch analysis 221
9.4.1 Item analysis with the Partial credit model 221
9.4.2 Person-item map 226
9.5 Transformation to W-scores 227
9.6 Comparison of the different RMM and CFA results 228
9.7 Model employed for the study 228
9.8 Summary 229
Chapter 10: Structural Equation Modelling: Student and teacher-level models 230
10.1 Introduction 230
10.2 Structural equation modeling 232
10.3 Mplus version 7.0 233
10.4 Models and presentation in quantitative research 234
10.4.1 Model trimming criteria 235
10.5 Constructs and variables used in the student-level analysis 235
10.5.1 The hypothesised student-level model 238
10.5.2 Student-level models results 240
10.6 Measurement model results for the combined student-level scales 240
10.6.1 Combined student-level model fit indices 241
10.6.2 Combined Grade 10 and 12 student-level structural model results 242
10.7 Grade 10 student level mesurement model results 246
10.7.1 Grade 10 student model fit indices 247
10.7.2 Grade 12 student-level structural model results 247
10.8 Grade 12 student-level measurement model results 250
10.8.1 Grade 12 student-level model fit indices 251
10.8.2 Grade 12 student-level model results of the structural model 251
10.9 Scales and variables used at the teacher-level analysis 254
10.9.1 The hypothesised teacher-level model 255
10.9.2 The final teacher-level model 257
10.9.3 Mesaurement model results for teacher-level scales 257

10.9.4 Teacher-level model fit indices……………………………………………………………………..258
10.9.5 Results of teacher-level structural model…………………………………………………………..258
10.10 Summary 260
Chapter 11: Hierarchical Linear Modelling 263
1.1 Introduction 263
11.2 Overview of Hierarchical linear modelling (HLM) 264
11.3 Application of HLM 6.08 264
11.4 Variables in the two-level model 265
11.5 The two-level model of mathematics achievement 268
11.5.1 The hypothesised model and analysis framework 269
11.6 Model building and evaluation 271
11.7 Null model 272
11.8 Combined student sample results 274
11.8.1 Analysis steps for the HLM student models 276
11.8.2 The final combined students' model 276
11.8.3 The direct effect of level -1 and level-2 predictors on mathematics results 279
11.8.4 The cross-interaction effects 280
11.8.5 Variance explained for the two-level combined model 282
11.9 The Grade 10 students' model 284
11.9.1 The direct effect of level -1 and level-2 predictors on mathematics results 287
11.9.2 Cross-interaction effects 287
11.9.3 Variance explained for the two-level Grade 10 model 290
11.10 The Grade 12 students' results 292
11.10.1 The direct effect of level -1 and level-2 predictors on mathematics results 294
11.10.2 Cross-interaction effects 295
11.10.3 Variance explained for the two-level Grade 12 model 297
11.11 Summary 298
Chapter 12: Teacher-level interview analysis 302
12.1 Introduction 302
12.2 Thematic analysis approach 303
12.3 NVivo 12 304
12.4 Emerging themes 305
12.5 Theme 1: Teaching methods 306
12.5.1 Traditional (teacher-centred) method 306

12.5.2 Student-centred method 308
12.6 Using both teaching methods 310
12.6.1 Discussion of teaching methods 313
12.7 Theme 2: Content knowledge 315
12.7.1 Discussion of content knowledge 319
12.8. Theme 3: Outcomes of the teaching methods 321
12.8.1 Discussion of the outcomes of the three teaching methods 324
12.9 Theme 4: Encouragement 325
12.9.1 Discussion of encouragement 327
12.10 Summary 328
Chapter 13: Discussion and Conclusion 330
13.1 Introduction 330
13.2 Summary of the findings 332
13.2.1 The effect of students' attitudes and motivation on their mathematics results 332
13.2.2 The effect of students' demographic variables of their mathematics results 333
13.2.3 The influence of teacher-level factors on students' mathematics results 334
13.2.4 The effect of teacher demographic variables on students' mathematics results 335
13.2.5 The interaction effect of student and teacher level factors on mathematics results…………..335
13.2.6 Qualitative interview results 336
13.3 Significance of the Study 337
13.3.1 Methodological significance 337
13.3.2 Theoretical significance 338
13.3.3 Practical significance 340
13.4 Limitations and future research 341
13.5 Recommendations 342
13.6 Conclusion 345
References 347
Appendices 387
Appendix A 387
Permission/approval documents 387
Appendix B 397
Student and teacher survey questionnaires 397
Appendix C 410
Grade 10 & 12 students mathematics tests 410

Appendix D 424
Information sheets, and consent and complaint form.. 424
Appendix E.. 429
Letters of request for permission to conduct the study.. 429
Appendix F……………………………………………………………………………………………………431
PNG MSU Grade 10 & 12 mathematics Table of Specification 2017……………………………………… 431

List of Tables

Table 2. 1 NSW dimensions of a quality-teaching model ... 44
Table 3. 1 Plan A: Summary of location, schools and participants of the study ... 65
Table 3. 2 Plan B: Summary of schools and participants of the study ... 65
Table 3. 3 Fit indices for examining the model fit of the student and teacher-level models ... 83
Table 3. 4 Reasonable item MNSQ ranges ... 88
Table 3. 5 Factors at each hierarchical level that affect Grade 10 and 12 students' mathematics results . 97
Table 4. 1 Distribution of teacher respondents by gender ... 106
Table 4. 2 Distribution of teacher academic qualification ... 107
Table 4. 3 Years of teaching experience ... 108
Table 4. 4 Major area of specialisation ... 109
Table 4. 5 Minor area of specialisation ... 110
Table 4. 6 School type in Port Moresby ... 112
Table 4. 7 Student respondents by gender ... 112
Table 4. 8 Student respondents according to mathematics type ... 114
Table 4. 9 Fathers highest education level ... 115
Table 4. 10 Mothers' highest education level ... 116
Table 4. 11 Mothers' employment ... 117
Table 4. 12 Fathers' employment ... 117
Table 5. 1 Grade 10 mathematics test table of specification (TOS) ... 124
Table 5. 2 Results of the initial analysis of the Grade 10 test items ... 126
Table 5. 3 Results of the final ConQuest run analysis of the Grade 10 test items... 128
Table 5. 4 Test level gender difference statistics ... 133
Table 5. 5 Test gender differences in ability estimates ... 133
Table 5. 6 DIF results for the 36 mathematics test items by item threshold ... 137
Table 5. 7 Test item fit statistics for gender (male and female) ... 139
Table 6. 1 Combined Grade 12 advance and general mathematics test Table of Specification (TOS) ... 151
Table 6. 2 Results of the initial analysis of the Grade 12 test items ... 154
Table 6. 3 Results of the final analysis of the Grade 12 test items... 155
Table 6. 4 Test level gender difference statistics ... 159
Table 6. 5 Gender facet estimates for 35 mathematics items for Grade 12 ... 160
Table 6. 6 Test item fit statistics for gender (male and female) ... 162
Table 6. 7 Result of DIF for the 35 mathematics test items by item threshold approach ... 164
Table 7. 1 Summary of items in the attitude scale used in the study from TIMSS 2015 and PISA 2012 . 176
Table 7. 2 Factor loading of Attitude scale ... 178
Table 7. 3 Results of fit indices for the single-factor attitude structure ... 179
Table 7. 4 Different model fit comparison for attitude towards mathematics ... 180
Table 7. 5 Results of the factor loading of the hierarchical model ... 182
Table 7. 6 Result of the model fit ... 183
Table 7. 7 Model parameter estimates of attitude towards math (n=729) ... 184
Table 7. 8 INFIT MNSQ for gender and item threshold (estimates) of gender ... 186
Table 8. 1 Summary of items in the motivation scale used in the study ... 195
Table 8. 2 Factor loading of motivation scale ... 198
Table 8. 3 Results of fit indices for the single factor motivation structure ... 199
Table 8. 4 Different model fit comparison for motivation towards mathematics ... 199

Table 8. 5 Results of the factor loading of the model ... 202
Table 8. 6 Result of the model fit ... 203
Table 8. 7 Model parameter estimates of motivation towards maths (n=729) ... 205
Table 8. 8 Model parameter estimates of motivation towards maths (n=729) ... 206
Table 8. 9 DIF Item fit (INFIT MNSQ) results for gender and Item fit threshold results for gender ... 207
Table 9. 1 Summary of items in the quality teaching scale used in the study ... 214
Table 9. 2 Summary of items in the classroom environment scale used in the study ... 216
Table 9. 3 Summary of items in the curriculum scale used in the study ... 217
Table 9. 4 Factor loadings of teacher level scales ... 219
Table 9. 5 Results of fit indices for the single teacher scales structures ... 220
Table 9. 6 Statitsical fit results for quality of teaching scale (n=41) ... 222
Table 9. 7 Statitsical fit results for classroom environment and instructional resource scales (n=41) ... 223
Table 9.8 DIF for quality of teaching for gender ... 224
Table 9.9 DIF for instructional resources and classroom environment scales ... 225
Table 10. 1 The variables used in the student-level model ... 237
Table 10. 2 Measurement model results for combined student level scales ... 241
Table 10. 3 The results of the fit indices of the final combined student-level model ... 242
Table 10. 4 Grade 10 and 12 students' estimated path model of the outcome of mathematics results ... 244
Table 10. 5 Measurement model results for Grade 10 student-level scales ... 246
Table 10. 6 Grade 10 student –level fit indices ... 247
Table 10. 7 Estimated path models of the outcome of mathematics results for Grade 10 students ... 248
Table 10. 8 Measurement model results for Grade 12 student level scales ... 250
Table 10. 9 Grade 12 student-level fit indices ... 241
Table 10. 10 Estimated path models of the outcome of mathematics results for Grade 10 students ... 252
Table 10. 11 The variables used in the teacher-level model ... 255
Table 10. 12 The results of the fit indices of the final teacher-level model ... 258
Table 10. 13 Teacher level variables with strutural model results ... 259
Table 11. 1 List of variables in the three-level model ... 267
Table 11. 2 Fully unconditional model results for 2L/HLM for the combined student model ... 275
Table 11. 3 The final model results for two levels model for combined students' mathematics results ... 277
Table 11. 4 Estimation of variance components for the final two-level model for combined students ... 283
Table 11. 5 Fully unconditional model results for the 2L/HLM for Grade 10 student model ... 284
Table 11. 6 Final results for the two-level model for Grade 10 mathematics results ... 285
Table 11. 7 Estimation of variance components for the final two-level model for Grade 10 students ... 291
Table 11. 8 The Null model results for the two-level model for Grade 12 mathematics results ... 292
Table 11. 9 The Final model results for the two-level model for Grade 12 mathematics results ... 293
Table 11. 10 Estimation of variance components for the final two-level model for Grade 12 students ... 298
Table 12. 1 Descriptive statistics of maths results for the three different teaching methods ... 311
Table 12. 2 ANOVA analysis result of the overall teaching method. ... 312
Table 12. 3 Post Hoc Tests for the each of the teaching methods for comparison ... 312
Table 12. 4 Descriptive statistics of the teachers' confidence in teaching mathematics content ... 317
Table 12. 5 The output of independent sample t-test analysis ... 318

List of Figures

Figure 1. 1 The western pyramidal education structure 1985-1999 (NDOE, 2010) 3
Figure 1.2 PNG Outcome Based Education System Structure (Apelis, 2012) 5
Figure 1.3 PNG Reformed standard based education system (NDOE, 2016) 7
Figure 1.4 Multi-level ecological model of change for students learning (Izu, O'Malley & Voight ,2014) ... 14
Figure 2. 1 TALIS Theoretical Framework (OECD, 2010 p.32) 54
Figure 2. 2 PISA theoretical framework (OECD, 2004) 56
Figure 2. 3 Proposed theoretical framework for the study (OECD, 2004a & OECD, 2010). 57
Figure 3. 1 Using mixed method design for evaluating multilevel data (Bamberger, 2012, p.14) 70
Figure 3. 2 Multilevel mixed method design (Bamberger, 2012, p.26) 71
Figure 3. 3 Integrated qualitative and quantitative linking of data 72
Figure 3. 4 Validation of scales used in the study (adapted from Ben, 2010) 100
Figure 4. 1 Distribution of teacher respondents by gender 107
Figure 4. 2 Distribution of teacher academic qualification 108
Figure 4. 3 Distribution of years of teaching experience 109
Figure 4. 4 Distribution of major area of specialisation 110
Figure 4. 5 Distribution of minor area of specialisation 111
Figure 4. 6 Distribution of school type in Port Moresby 112
Figure 4. 7 Distribution of student respondents by gender 113
Figure 4. 8 Distribution of student respondents according to mathematics type 114
Figure 4. 9 Distribution of fathers' highest education level 115
Figure 4. 10 Distribution of mothers' highest education level 116
Figure 4. 11 Distribution of mothers' employment 117
Figure 4. 12 Fathers' employment 118
Figure 5. 1 ICC for item 4 129
Figure 5. 2 Item 22 analysis and ICC 131
Figure 5. 3 ICC for Item 06 showing non-bias between male and female respondents 140
Figure 5. 4 ICC for items 27 biased in favour of male respondents 141
Figure 5. 5 ICC for items 29 biased in favour of male respondents 141
Figure 5. 6 ICC for items 31 biased in favour of female respondents 141
Figure 5. 7 Item 09 showing improved characteristics curve by category 145
Figure 5. 8 Person-item map of Grade 10 test items 147
Figure 6. 1 Item characteristics curve for item 38 157
Figure 6. 2 Item characteristics curve by category for item 32 158
Figure 6. 3 ICC by score for gender DIF of item 05 165
Figure 6. 4 ICC curve by score for gender DIF of item 08 166
Figure 6. 5 ICC by score for gender DIF of item 21 166
Figure 6. 6 ICC curve by score for gender DIF of item 25 167
Figure 6. 7 Person-item map for Grade 12 items 169
Figure 7. 1 Structure of one-factor model for attitude 177
Figure 7. 2 Structure of the hierarchical model 181
Figure 7. 3 Person-item map of attitude scale 187
Figure 8. 1 Structure of one-factor model for motivation scale 197
Figure 8. 2 Structure of the hierarchical model 201
Figure 8. 3 Person-item map of motivation scale 208

Figure 9. 1 Structure of one-factor model for teacher level factors 218
Figure 9. 2 Person-item map of teacher-level scales 227
Figure 10. 1 The hypothesised student level model 239
Figure 10. 2 The final combined student level model (n=729) 243
Figure 10. 3 The final Grade 10 student level model (n=354) 248
Figure 10. 4 The final Grade 12 student level model (n=375) 252
Figure 10. 5 The proposed teacher level model (n=41) 256
Figure 10. 6 The final teacher level model (n=41) 257
Figure 11. 1 Two-level model of mathematics results 268
Figure 11. 2 The hypothesised two-level model of mathematics results 271
Figure 11. 3 Two-level model of students' mathematics results for the final combined model 278
Figure 11. 4 Cross-level interaction effect of students in class on Ave-MHEDUL on the slope of combined students' attitude towards mathematics. 282
Figure 11. 5 Two-level model of students' mathematics results for Grade 10 sample 286
Figure 11. 6 Cross-level interaction effect of students in the class on Ave-MHEDUL on the slope of Grade 10 students' attitude towards mathematics 290
Figure 11. 7 Two-level model of students' mathematics results for the Grade 12 student model 295
Figure 11. 8 Cross-level interaction effect of students in class on Ave-MHEDUL on the slope of Grade 12 students' gender 297

Abstract

Though Papua New Guinea (PNG) has great development potential with its natural resources, the country faces challenges with its education system. One of these challenges is the poor mathematics performance of students at secondary schools. This poor performance is evident both through the significant number of students who cannot continue to Grade 11, and through the simultaneous decline in student enrolment in science related degrees at the university level. That being the case, this study aims to examine the school, teacher and student-level factors and their interrelationships that affect Grade 10 and 12 students' mathematics results.

A quantitative-dominant mixed-method approach is employed in this study, which has 729 student and 41 teacher participants. The instruments used in this study are: survey questionnaires, mathematics test, and interview questions. The survey questionnaires for students and teachers are adopted from international studies such as Trends in International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA) and are modified according to the context of this study. The mathematics test questions are adopted and modified from past Grade 10 and 12 national mathematics examination papers, respectively. These survey questionnaires and tests are validated and verified through confirmatory factor analysis and Rasch analysis. Rasch analysis scores obtained from the survey questionnaires and tests are used to examine the relationships between the independent and dependent variables at the teacher and student-levels through the use of Structural Equation Modelling, respectively. Hierarchical Linear Modelling is also employed to examine the direct effects from teacher and student-levels and cross-interaction effects between variables at these two different levels. The interview data at the teacher-level is analysed through thematic approach to capture information that might have been missed in teachers' survey questionnaires to complement the findings of quantitative (survey) data.

A number of significant factors (attitude, motivation, fathers' occupation, private schools, students' gender, teachers' with mathematics major and quality of teaching) that are influencing students' mathematics results are identified by this study, with mothers' highest education level found to be the most critical factor. The study concludes that more attention should be given to the factors identified in this study, in order to improve Grade 10 and 12 students' mathematics results. Overall, this study contributes theoretical, methodological, and practical knowledge to teaching and learning in mathematics education.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institutions without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed by candidate:

Date: 4 March, 2021

Acknowledgement

First and foremost, I want to thank God Almighty for the wisdom and knowledge in completing this thesis. I would also want to extend my gratitude to some unique individuals and organisations. Indeed, they greatly helped me achieve a milestone that would have been impossible without their invaluable support, advice and guidance. In particular, I am incredibly grateful to the following people.

Both my principal supervisors, Dr IGusti Darmawan and Professor Sivakumar Alagumalai and also co-supervisor Associate Professor Nicholas Buchdahl for their scholarly guidance and dedication in this journey. Dr IGusti Darmawan for always having the patience to guide and support the data analysis process with different statistical approaches and simultaneously giving continuous critical feedback for improvement with calmness and patience. Professor Sivakumar Alagumalai for the recommendation to obtain the Australian Awards Scholarship and devoting time to develop the proposal and the critical feedback on the validation chapters for the thesis. Associate Professor Nicholas Buchdahl for consistently reading and providing critical and yet detailed feedback for the chapters of the thesis for improvement. The supervisory team have been a great source of inspiration. Dr Owen Lindsay for the editorial service for English grammar and usage for the thesis.

The Australian Awards Scholarship and the taxpayers of Australia for the scholarships for both my Masters and PhD studies. The scholarship afforded me a great opportunity to pursue the highest degrees in my field, and more contribute to the future development of my community/country.

PNG National Department of Education (NDOE) and National Capital District secondary school principals for granting permission to collect data and also the teachers with their students who participated in the study.

Retired Catholic Archbishop of Madang/Bishop of Mendi Diocese Steven Richeirt OFM, Cap and Salesians of Don Bosco (SDB) for giving me a second chance in my education at Don Bosco Technological Institute (DBTI) when I dropped out in Grade 12. Thank you for the financial support, guidance, prayers and always believing in my potential/ability in my education.

My wife Cathy Oko, for the love, support, encouragement, understanding and sacrifice through thick and thin throughout my research journey. Thank you for your patience in making my dream come true.

Agima Ambros Rero and family (Sharon, Pheobe and Marylove) for the consistent assistance and support during the PhD study. Brother, thank you for always been there for me when I most needed it.

Cousin Roy Cornelius, brothers Pius Kalambe and Christopher Kasu for the support, encouragement, motivation and fun when it most matters (tough times). Cousin and brothers this thesis is the result of all those conversations and fun. Hope you three enjoy reading this thesis.

Daniel Mura and family for the support and prayers in the study always.

My dad Luta Oko Rama for installing discipline and encouragement to further my education though he was not educated. The dream about education that you had for me has come to reality. Mum Clare Lena Oko for your hard work towards my education journey and prayers always. This thesis is dedicated to both of you.

Finally, this thesis is dedicated to Aramirupa/Kimapolo (Yame/Raleloma)/Porane Kalipa (Porane), the mighty Mirupa tribe and the people of Kagua/Erave in Southern Highlands Province, Papua New Guinea.

Abbreviations

CFA- Confirmatory factor analysis

DIF-Differential Item Functioning

HLM-Hierarchical Linear Modelling

IRT-Item Response Theory

MRM-Multidimensional Rasch Model

MSD-Measurement Service Division

NOED-National Department of Education

OECD-Organisation for Economic Co-operation and Development

PISA-Programme for International Student Assessment

PCM- Partial Credit Model

PNG-Papua New Guinea

RMM-Rasch measurement model

SEM-Structural Equation Modelling

TALIS-Teaching and learning International Survey

TIMSS-Trends in International Mathematics and Science Study

WLE-Weighted Likelihood Estimate

Chapter 1: Issues in mathematics education

1.1 Introduction/Background

Mathematics is of central importance to modern society, as it provides a vital underpinning for economic knowledge and systems, and is significant in the physical sciences, technology, business, financial services, and many areas of information communication technology (ICT) (Brown, 2013; Gasparyan & Smirnova, 2015). Moreover, mathematics is important for biology, medicine, and many of the social sciences. Mathematics is, consequently, a significant area of knowledge for the development of a country's economy and society. Thus, many countries employ measures to improve and maintain the quality of their education system relating to mathematics education (Craft, 2018; Lim, 2018; Scheerens, Luyten,& Van Ravens, 2011; Thune, 2017). Related to these efforts, numerous academic studies have been undertaken to assist, investigate, and address the areas of vital concern in mathematics education, and to examine the causes associated with these challenges (Craft, 2018; Davis, 2008; Fafunwa, 2018). One finding emerging from these studies has been that mathematics results in many countries are declining at an alarming rate (Kennedy, Lyons,& Quinn, 2014; Wilson & Mack, 2014). This phenomenon is evident in international studies, such as Trends in International Mathematics and Science Studies (TIMSS) and the Programme for International Student Assessment (PISA), which both clearly indicate that students in many countries are performing poorly in mathematics (Mullis, Martin, Foy,& Arora, 2012; OECD., 2005). In turn, a poor standard of mathematics has potentially detrimental effects for economies and societies (Wilson & Mack, 2014). As a developing country, Papua New Guinea (PNG) continuous to face challenges with instructional resources, teachers , curriculum development and implementation in the education system.

This chapter focuses on defining the problem that this study seeks to investigate, and is organised as follows. The first section presents a statement of the problem, and outlines the issues relevant to this study. Next, the research questions, aims and objectives of the study are presented, which represent the rationale for this project. This is followed by a section defining the scope and limitations of the study, and the structure of this thesis and, finally, a summary of the chapter.

1.2 Statement of the problem

This section provides an overview of the Papua New Guinean education system, identifying the strengths and flaws in its mathematics curriculum, and the shortcomings of the education system in general. Furthermore, this section also highlights the context of the problem and the rationale for this study.

1.2.1 Overview of Papua New Guinea education systems

From the 1970s until now, PNG has implemented three education systems: the pyramidal education system used from 1970-1999; the outcome-based education (OBE) used from 2000-2014; and, from 2015 onwards, the standard based education system (SBE), which was initially introduced at the primary level and is being gradually phased in at the secondary level until the OBE system is eradicated. As such, the OBE system is still actively used today, and data for this study were collected accordingly. Each of these structures is now discussed separately, in order to provide background information on PNG education systems, and how mathematics education functions within these systems.

The pyramidal education system was adopted from the western model, following the colonial influence of Australia in the 1970s (Conroy, 2017; Le Fanu & Kelep-Malpo, 2015; Tapo, 2004). Though this education system enabled many students to enrol in elementary (students enrolled from prep to Grade 3) and primary schooling (students enrolled from Grade 4 to Grade 6), it reduced enrolments in the later years of schooling

(Crossley & Sprague, 2014; Guthrie, 2014; Matane, 1986). Through a standardised national examination in Grade 6, 10 and 12, this pyramid system (see Figure 1.1) allowed students with better academic performance, evident through their school results, to continue on to the next level of education (Leke, 2010; Tapo, 2004). As a result, the system produced and promoted quality graduates during the 1980s. As shown in Figure 1.1, high school students with high grades/marks were selected to do Grade 11 and 12 at the four National High Schools in the country (Kerevat, Aiyura, Sogeri and Passam). Most of these students then achieved better results and secured places at universities for further study (Leke, 2010; Matane, 1986). Other high school students with average and below-average grades/marks were selected for teaching, nursing and technical colleges. Though only a handful of students dropped out of Grades 10 and 12, these students still had the chance to enrol at vocational schools for training in basic skills such as cooking, welding, and mechanics or at distance education centres to improve their marks in failed subjects (Matane, 1986; Leke, 2010; Tapo, 2004), and later engaged in formal employment. Overall, this system enabled all students to be successful in their education and enter into jobs that contributed meaningfully to the economy. One of the reasons for the success of this system may be the promotion of quality in education, with less competition among students for spaces at tertiary institutions. This system and its structure are shown in Figure 1.1, which indicates the exit point stages in students' education.

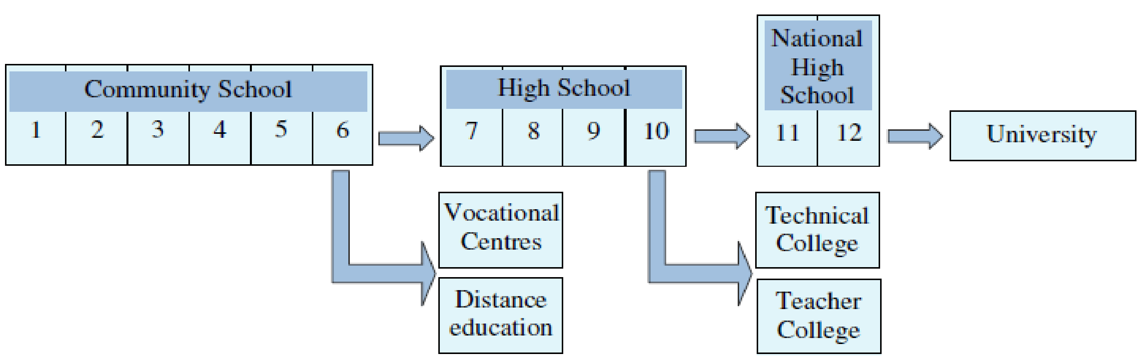


Figure 1. 1 The pyramidal education structure 1985-1999 (NDOE, 2010)

However, following PNG's population growth from approximately 4 million people in 1985 to 5.6 million in 1999 (United Nations, 2019), and with socio-economic and technological advancement, the pyramidal education system began to face numerous challenges (Howes et al., 2014; Turner, 1990). Growing numbers of students were unable to continue on to the next stages of their education (see Figure 1.2 p.5), due to poor academic performance and an increase in tuition fees. Consequently, these students engaged themselves in subsistence farming or roaming the streets of towns and cities looking for opportunities. The PNG Government therefore failed to foresee the impact that the country's economic problems with the growing population would have on its western education system model (O'Donoghue, 1994; Weeks, 1993). This situation prompted the national government through the Department of Education (NDOE) to review the pyramidal education system and introduce an alternative system in 2000 known as outcome-based education (OBE), with its first syllabus published in 2003 (Asimi, 2014). This OBE system shown in Figure 1.2 (p.5) was introduced with the rationale that all students should have access to basic education (Eldridge, Larry, Baird,& Kavanamur, 2018; Neofa, 2010), especially reading and writing education. OBE features three years' elementary schooling in local communities in local vernaculars; six years' primary schooling (Grade 4-8); and four years' secondary schooling (Grades 9-12). Accordingly, in 2003 the government opened up more elementary, primary and secondary schools (PNG NDOE, 2006). Despite this reform, however, the OBE system came to exert a negative impact on the economy, with dropouts from schools increasing at an alarming rate compared to the pyramidal education system discussed earlier, without producing quality graduates who were able to contribute meaningfully to the economy (Eldridge et al., 2018; Joskin, 2013). Based on the researcher's experience, this decline may be due to several factors, including the low quality of teaching associated with the introduction and implementation of outcome-based education, which was a system that was adopted without proper consultation. This OBE system promoted quantity rather than quality of outcomes from students in progressing to the next stages of education.

Though most high schools were upgraded to secondary school status to prepare for the increased intake of students, the number of tertiary institutions has not been increased (NDOE, 2016). Consequently, this has forced competition among students to attain high grades to secure spaces at tertiary institutions. For instance, although 18,000 students sat for Grade 12 examinations in 2015, only 6,000 students secured places in higher education institutions in PNG, due to low grades and limited spaces. The researcher was involved in the selection of students for higher education that year, and has personally witnessed the decline in the number of students entering tertiary institutions. When compared to the previous education system discussed above (Figure 1.1, p.3), the government created minimal alternative pathways for the increased number of dropouts in the system from 2003-2014 with few improvements made to the education system. As a result, this cohort of students has been marginalised by the reformed education system, which has denied them opportunities to pursue careers in science and mathematics related fields, such as technology and engineering in the later years of their education.

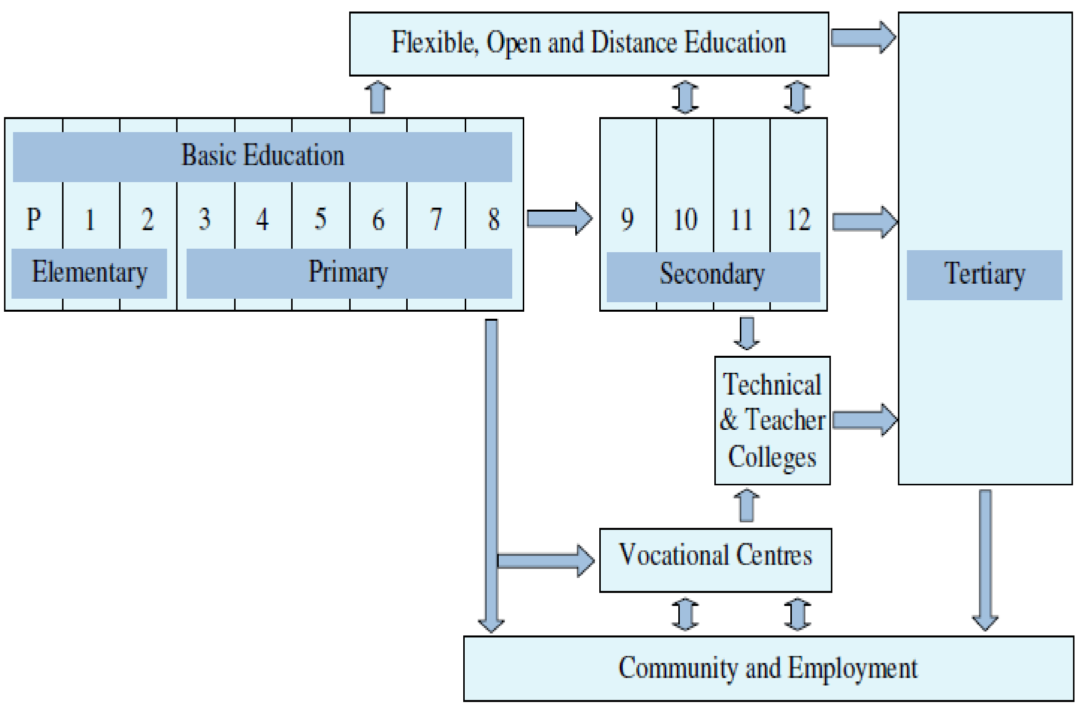


Figure 1.2 PNG Outcome Based Education System Structure (Apelis, 2012)

By 2009, the OBE system had grown to employ around 44,558 teachers, and had enrolments of 1.43 million students in 5,366 elementary schools, 3,275 primary schools, 218 secondary and high schools, and 139 vocational schools (AusAID report, 2018). As mentioned earlier, the reformed education system was intended to increase access for all children at the lower and upper secondary levels. However, this system was implemented without parallel improvements in the curriculum implementation process (i.e., teaching and learning), classrooms, learning resources, teacher training and other educational factors. Consequently, only students who performed well in their subjects were able to advance to Grade 11 or tertiary institutions, with the remainder dropping out from the system due to low grades. This indicates that the OBE system lacked proper consultation with stakeholders regarding its implications and relevance for the PNG context, and was not sufficiently focused on students' needs and achievements. The poor performance of students under the revised system raised concerns from stakeholders, teachers, parents and guidance in terms of content delivery in the syllabus, quality of teaching, instructional instruments and assessment used in the curriculum to improve students' mathematics results. Simultaneously, poor results also raised concerns about students' critical thinking levels, attitudes and motivations towards mathematics.

Given the shortcomings of the two education systems discussed above, in 2015 the government introduced a new standards-based education (SBE) structure, known as 2-6-6 (PNG NDOE, 2016). This structure (shown in Figure 1.3, p.7) provides 2 years of early childhood education, 6 years of primary education (Grades 1-6), and 6 years of high school/secondary education (Grades 7-12). The overall goal of SBE structure is for access to education to be increased: "Students will no longer be pushed out of education at Grade 8 and 10 like the two previous education systems due to more places and relevant choices of subjects will be created to cater for the additional students remaining in education" (PNG NDOE 2016, p.31). However, in the new structure, the selection of students for colleges and universities still takes place in Grade 12, with academic results still playing a crucial part in determining student placement at these institutions.

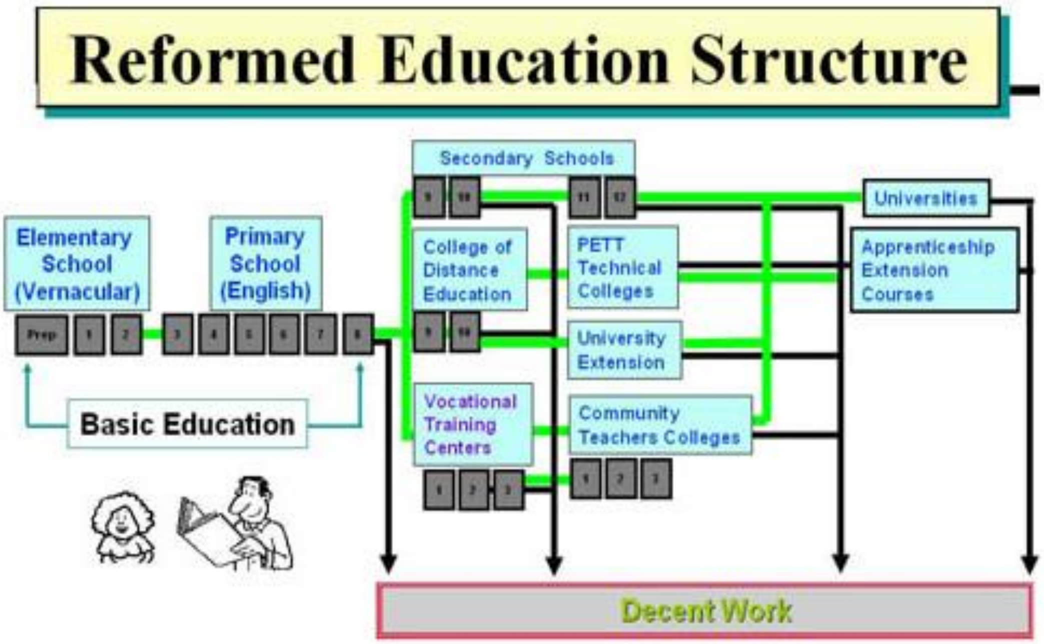


Figure 1.3 PNG reformed standards- based education system (NDOE, 2016)

With the SBE system, students are able to choose learning pathways to suit their needs, strengths and interests, while academic pathways are strengthened for those intent on tertiary studies. Vocational studies and distance education learning colleges have an increased importance, providing much needed skills for work and life while catering for increased student numbers and pathways choices (PNG NDOE, 2016). Students gain nationally certified vocational qualifications while at school and are able to mix vocational studies with more traditional subjects to suit their needs. This stable, embedded curriculum adds more relevant learning choices for students' academic, work and life skills. Under the SBE structure, social and community values and 21st century learning will be an important component of every child's learning (PNG NDOE, 2016). The OBE syllabus will be completely eradicated soon in favour of the SBE. As such, though SBE was implemented in 2015, the results of the new education system will only be seen when the outcomes of Grade 12 subjects are available. The validity and importance of the conclusions of this study is not affected, even though the education system is currently changing from OBE to SBE. This is because the same contents of the mathematics syllabus used in this study will be still used despite the change in education system. Data collected in this study accordingly reflect the results of the OBE mathematics curriculum. The next section discusses the OBE mathematics curriculum for lower and upper secondary schools.

1.2.2 Lower and upper secondary school mathematics curriculum in PNG

In PNG, 'lower secondary' refers to Grades 9 and 10, and 'upper secondary' refers to Grades 11 and 12 in a secondary school setting (PNG NDOE, 2006a; 2006b). The mathematics topics (see Appendix C) covered in both lower and upper secondary schools are tested each school year in a national examination at the end of Grades 10 and 12, respectively. The purpose behind this examination is for students to demonstrate sound basic mathematical concepts and to provide the numerical knowledge necessary for upper secondary school or for tertiary institutions (PNG NDOE, 2006; PNG NDOE, 2009). Another aim is to filter students at the Grade 10 and 12 education stages, as discussed earlier, to allow only capable students to continue on to Grade 11 and tertiary institutions. Those students who do not achieve the required marks are left out of the education system. This process of selecting students is similar to the pyramidal education system discussed earlier.

The mathematics curriculum has been developed and designed with learning outcomes for both lower and upper secondary schools (PNG NDOE, 2006). Hence, these outcomes identify the knowledge and skills that students should achieve and demonstrate in their learning at the end of Grades 10 and 12 (Joskin, 2013; Le Fanu & Kelep-Malpo, 2015; Owens, 2015).

The units/topics covered by the curriculum contain academic and practical components that emphasise students' development of mathematics (Kravia & Owens, 2014; Matang, 2002; Owens, 2015; Tapo, 2004). This reflects the rationale for the secondary curriculum in PNG, which is based around students understanding underlying background theories and applying practical aspects to their daily lives (Owens, 2014; Owens & Kaleva, 2008). For example, the counting system, barter systems, patterns in weaving or billum (bag) making, traditional calendars, and measurement systems and navigational skills are all highlighted in the curriculum Thus, students are required to develop abilities to use mathematics for reasoning and communicating in day-to-day activities such as trading, buying and selling, weighing, measuring, and estimating (PNG NDOE, 2006). Despite this objective, however, in practice these skills are only minimally developed because most school-teachers still employ traditional methods of teaching (Brownlee, Farrell, & Davis, 2012). This is evident in their delivering lessons that lack the practical aspects outlined in the syllabus (Brownlee et al., 2012; Guthrie, 2015). Without this practical component, some teachers' lessons fail to engage students' interest, which results in poor grade performance. Matang (2002) highlights that highly theoretical teaching approaches by teachers can hinder individuals' abilities to contribute meaningfully with confidence to their communities after Grades 10 and 12, because they are less competent mathematically. Additionally, some individuals do not have the confidence to further their knowledge of mathematics at universities or colleges (PNG NDOE, 2006; 2009).

Furthermore, the present curriculum lacks the capacity to bridge the gap of how the academic aspects of mathematics principles are integrated with the traditional PNG context (Kravia & Owens, 2014; Owens, 2014, 2015). The curriculum promote an increased quantity of graduates moving through the different stages of education, without focus on quality (Owens, 2015; Sukthankar, 2006). As such, teaching strategies used to disseminate syllabus content in the curriculum needs to focus more on the weaker students while simultaneously considering the brighter ones. In practice, however, teaching methods give emphasis to a small number of students (brighter students) who are progressing to the next stage of their education (Grade 11 or tertiary education). This imbalance has been a common issue for both the reformed and previous education systems, as discussed earlier, and is a situation that reflects educational leaders' leadership skills and visions for the country's education system.

1.2.3 Content of lower secondary mathematics and its implications

The Grade 10 national examination preparation curriculum covers topics in the lower secondary syllabus (see Appendix C). This study examines only the core units of this curriculum, and not the optional units and topics with the assumption that only the core units are the priority for examination preparation (PNG NDOE, 2006a). As optional units may not be covered in all schools, if they are included in this study's test, it would not be fair to students who were only taught the core units.

Through the study of mathematics content at the lower secondary level, students explore ways of solving problems using mathematical skills and processes (PNG NDOE, 2016: PNG NDOE, 2009). However, as students at this level use minimal quantitative and spatial information in problem solving and decision-making (PNG NDOE, 2006a; PNG NDOE, 2009), they do not learn to enjoy and value mathematics. Studies have shown that students who learn mathematics with quantitative and spatial information in problem solving at the lower secondary level grow more confident and motivated to think analytically and rationally, and appreciate the role of mathematics as they continue learning mathematics in higher levels (Clarke, Clarke & Sullivan,1996; Mel, 2007; Saxe & Esmonde, 2012). As a result, this comprehension and appreciation is evidently not being successfully accomplished with Grade 10 graduates in PNG. In this researcher's experience teaching in both secondary and tertiary institutions, the majority of students demonstrate low levels of analytical and rational thinking skills when using mathematics, which leads them to have a lower appreciation of mathematics. A survey carried out by the National Department of Education

(NDOE) in 2009 found that the majority of students avoid asking questions in mathematics classrooms, and instead tend to become passive learners due to low levels of motivation (PNG NDOE, 2009). These students are unable to think mathematically and are consequently less empowered to operate effectively through inquiry. Many students in PNG are unaware that the vast amount of mathematics information can be understood through being active and engaged with mathematical techniques and having a positive attitude (Clarke et al., 1996; Saxe & Esmonde, 2012), which would result in them having the confidence to apply content knowledge in practical community contexts.

1.2.4 Content of upper secondary mathematics and its implications

The units studied in general mathematics and advanced mathematics for Grades 11 and 12 are outlined in Appendix C. These units/topics in both syllabuses are designed to enable students to solve problems logically. Students acquire mathematical knowledge and skills to be applied in fields such as medicine, economics, management studies, physical sciences, and industries (PNG NDOE, 2006b; PNG NDOE, 2009). Acquiring and developing this mathematical capacity provides students with a set of skills necessary for employment in the formal and informal sectors, as well as preparing students who will continue on to higher educational institutions or other pathways.

Additionally, general mathematics units are designed for specific programmes at colleges or universities. For instance, a student enrolled in general mathematics can apply for arts courses such as accounting, social works, and human resources career pathways. On the other hand, students enrolled in advanced mathematics can apply for courses such as engineering, science foundation, and economics at universities in PNG (PNG NDOE, 2016). This choice allows students to determine which mathematics course they should enrol in to follow their career pathway (PNG NDOE, 2006a). Some of the units/topics of general and advanced mathematics are similar, differing in the level and context of the mathematics employed. For instance, functions and graphs in advanced mathematics are discussed in greater detail than they are in general mathematics, because they are a pre-requisite of topics that follow later (PNG NDOE, 2006). In general, the two kinds of mathematics courses are designed to prepare students for the type of mathematics that are involved in the next stage of the course they are enrolled in at college or university (PNG NDOE, 2006b; PNG NDOE, 2009).

In addition to the above ideas that underpin the upper secondary mathematics curriculum (NDOE, 2006a), there remain challenges confronting teachers and students regarding syllabus implementation, delivery, assessment, as well as with their motivation and attitude towards mathematics. Similarly, PISA and TIMSS studies highlight that other factors such as teaching resources, the quality of teaching, overcrowding of students in classrooms, inadequate instructional resources, poor management, time allocated to teach mathematics, and cultural aspects are also everyday obstacles faced by teachers and students. Therefore, one of the purposes of this study is to investigate some of these factors that may have effect on students' mathematics results (OECD, 2012).

1.3 The status of mathematics within the PNG education system

Mathematics plays a significant role in all aspects of everyday life (d'Ambrosio, 2016; Gravemeijer, Stephan, Julie, Lin, & Ohtani, 2017; Hackling, 2016). A robust mathematics education system, therefore, is central for a society (Nicholas, Poladian, Mack, & Wilson, 2015). As such, mathematics contributes to accelerating the development of a nation through its application in the private and public sectors, such as in industries and government departments (Mujtaba, Reiss, & Hodgson, 2014). However, a number of studies reported a decrease in students' general mathematics abilities in mathematics during the 2000s (Kennedy et al., 2014; Whannell & Tobias, 2015; Wilson & Mack, 2014). Furthermore, enrolment in mathematics has declined globally over the last 20 years, with fewer students choosing to take up mathematics as a subject. Whannell and Tobias (2015) indicate that this decline is caused by students' perception of mathematics as a difficult and boring subject. This state of affairs poses a definite problem and is a particular concern in PNG, where there is a low emphasis on teaching and learning mathematics effectively within the school system (Joskin, 2013; Lagani et al., 2010).

As a result, the PNG Government fears that the continuous decline in the performance at national mathematics examinations for Grade 10 and 12 students will have a negative effect on the future economy. With the declining mathematics performance of students in many schools, PNG cannot produce well-qualified medical doctors, engineers, statisticians, and economists who are able to contribute to the socio-economic development of the nation (Kaleebu, Gee, Watson, Jones, & Jauk, 2017; Macdonald, 2016; McCormick, 2011). An outcome of this situation would be that infrastructure developments such as roads and bridges continuously deteriorate due to inaccurate calculations that underpin their design (Lagani et al., 2010; Macdonald, 2016).

Consequently, improvement of mathematics learning and teaching to obtain better results in national examinations is a major interest for the government, school administrators, teachers and parents' alike (Cleverley, 2007; Joskin, 2013; Mel, 2007). The National Education Secretary in PNG states that national mathematics achievement levels have declined in the past decade and therefore need scrutiny (PNG NDOE, 2009; PNG NDOE, 2004; PNG NDOE, 2016; PNG DHERST, 2006a). This viewpoint is based on the annual Grade 10 and 12 students' mathematics national examinations results. Consequently, few students graduating from Grade 12 are able to enrol in universities to undertake mathematics related programs such as engineering and medicine. Simultaneously, there is a significant number of Grade 10 students who do not have a chance to continue to Grade 11 and are therefore forced out of the education system (Joskin, 2013; Le Fanu & Kelep-Malpo, 2015; Rena, 2011). The decline in students' mathematics performance may be due to different contextual factors affecting their results. Therefore, investigation of these factors underpinning the decline in students' performance is vital to address the concerns raised by the PNG National Department of Education reports.

The multilevel ecological model of change for students learning by Izu, O'Malley and Voight (2014) in Figure 1.4 shows that the students' behaviour or performance is the consequences of what happens in the classroom, school and the local community. This is because students are placed in classrooms with teachers; classrooms are within schools; and schools are within local communities. This indicates that factors at the school, classroom, and student-levels require analysis due to their impact on students' mathematics achievements. These factors include school types, attitude, motivation, quality of teaching, classroom environment, instructional resources, and gender. Therefore, the direct effect with the interrelationships or interaction effects of these factors from the different layers/levels, and how they affect Grade 10 and 12 students' mathematics results in the annual examination in Port Moresby, PNG, will be the focus of this study.

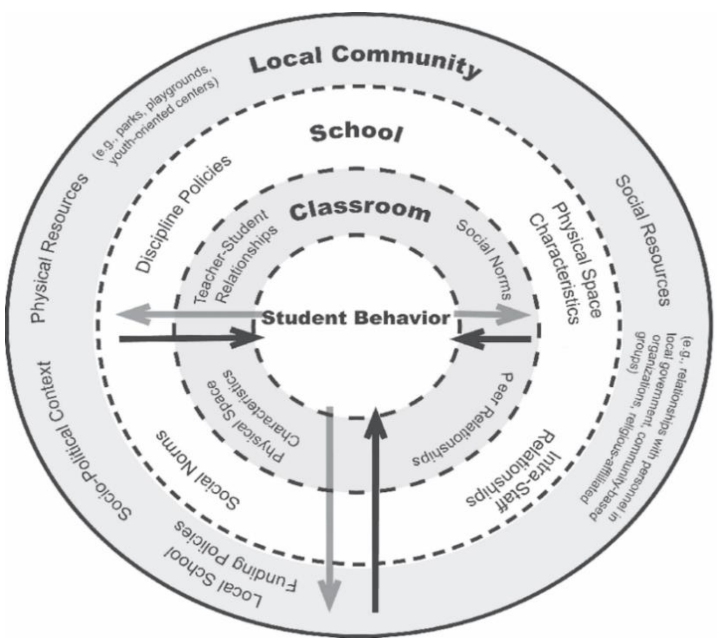


Figure 1.4 Multilevel ecological model of change for student learning (Izu, O'Malley & Voight ,2014)

1.4 Research questions

This study's research questions highlight the multilevel nature of the school, classroom and student-level factors shown in Figure 1.4 that will be examined further.

1. What are the student, teacher and school-level factors that are affecting Grade 10 and 12 students' mathematics results in Port Moresby, PNG?

2. How do the student, teacher and school-level factors interact to influence Grade 10 and 12 students' mathematics results?

3. What are the attributes that are deemed important by teachers regarding their mathematics teaching strategies and content knowledge?

In this study, student-level factors considered include demographic variables, attitude and motivation. Teacher-level factors considered are demographic variables, quality of teaching and classroom environment; and school-level factors considered are school type and instructional resources. These factors are examined for their possible effect on students' mathematics results in order to improve teaching and learning in mathematics.

1.5 Aims and objectives of the study

The aim of this study is to examine the factors that are affecting Grade 10 and 12 students' mathematics performance. Specifically, the study will focus on four primary objectives to facilitate the achievement of this aim:

a. To examine the factors/constructs at school, teacher and student-level that affect Grade 10 and 12 students' mathematics achievements/results in Port Moresby, PNG. This study will examine the impact of these factors on mathematics achievements as reflected in international studies such TIMSS, PISA, Teaching, and the Learning International Survey (TALIS). The researcher's own experience as a mathematics teacher in PNG also provides anecdotal evidence for the influence of factors at school, teacher and student levels on students' mathematics results

b. To examine the direct and indirect effect of factors at each level (school, teacher, and student).
c. To examine the interaction effects or interrelationships of the school, teacher, and student-level factors on Grade 10 and 12 students' mathematics results.
d. To examine the mathematics teachers' views on their content knowledge and teaching pedagogies that affect Grade 10 and 12 students' mathematics results.

1.6 Significance of the study

This study will significantly contribute to the discipline in three ways:

a. *Theoretical significance*

Many mathematics educational studies have been conducted with a limited or narrow scope, and are focused on only one or two factors relating to how teaching and learning of mathematics occurs. By contrast, this study focuses on a number of factors and their interrelationships that are believed to influence Grade 10 and 12 students' mathematics performance. This study is relevant because to date there have been no studies undertaken related to factors affecting students' mathematics results in PNG. Factors that are central to this study include motivation and attitude towards mathematics, the influence of the gender gap in PNG on the educational aspects of mathematics, and issues of quality of teaching and the classroom environment. In addition, PNG has not participated in international studies such as TIMSS and PISA, and the results of this study will therefore enable PNG's participation in those studies. In general, studies on mathematics education have focused on achievements in Grades 4 and 8. However, this study will focus only on Grade 10 and 12 students (ages 15-19) in secondary schools, to identify the factors affecting their results. The findings of this study will provide insights into improving students' performance and to minimise attrition rates.

b. Methodological significance

This study provides an empirical-based analysis using various quantitative techniques such as confirmatory factor analysis, Rasch modelling, structural equation modelling and multi-level analysis, to test and extend the existing research framework and model how different factors affect Grade 10 and 12 students' mathematics results. These different statistical techniques employed for student, teacher and school-level scales provide evidence from the data on the factors influencing students' mathematics performance. A comparison of the two data sources (surveys and interviews) was carried out through triangulation (Bamberger, 2012; Sandelowski, 2000), which provides a deep understanding of the processes associated with teachers and their content knowledge and teaching strategies. These research techniques and methods can be used by other researchers to conduct similar empirical research. Moreover, classroom teachers can also use the Rasch modelling techniques in their assessment practices to check the validity and reliability of the test items.

c. Practical significance

The results of the study will provide PNG education leaders and teachers with better understandings of the school, teacher and student-level factors involved in improving mathematics results. The primary data collected from Grade 10 and 12 students with their mathematics teachers in Port Moresby are analysed with quantitative and qualitative methods. These two methods examine the factors affecting Grade 10 and 12 mathematics results. Consequently, this study will offer recommendations to improve teaching and learning of mathematics, ultimately to improve Grade 10 and 12 students' performance in national examination. As this research project was conducted in Port Moresby, a city in which students from different parts of PNG are present, it has the potential to provide broader outcomes. Additionally, the findings of this study can offer teachers and educators the direct, indirect and cross-interaction effects (discussed in Chapters 10 and 11) of the school, teacher and student-level factors have on students' mathematics results.

1.7 Scope and limitations of the study

The study's focus is on the factors that affect Grade 10 and 12 students' mathematics achievements in Port Moresby, PNG. The factors examined in this study are significant for developing inferences and conclusions about students' mathematics results. Furthermore, this study employs a multilevel analysis to examine the school, teacher and student-level factors that are affecting students' mathematics results. The factors examined by this study at the student-level are attitude and motivation toward mathematics; those at the teacher-level are quality of teaching and classroom learning environment; and those at the school-level are school type and instructional resources. The study also includes investigation of the effects of demographic variables such as gender, age, academic qualification, years of teaching experience, parents' education level and teachers' major area of teaching. Outcomes at the student-level can be generally measured through results of mathematics tests, and this study operationalises a mathematics for Grade 10 and 12 students test for this purpose. Scope wise, the study uses quantitative (survey) method to collect data from 729 Grade 10 and 12 students and qualitative (interview) method for 41 of the students' mathematics teachers from 16 secondary schools in Port Moresby, PNG. A comparison of the two data sources (surveys and interviews) through triangulation provides a deep understanding of the processes associated with teachers' content knowledge and their teaching strategies.

This study acknowledges several limitations. Firstly, in this single study, it is not possible to consider all of the complex group of factors and their interrelationships affecting students' mathematics outcomes. Teaching and learning are complicated processes for teachers and students, and cannot be adequately represented by responses to a single survey. Consequently, in this study limited constructs/variables are used, even though there are many other factors that may influence students' mathematics results. The mathematics results' can also influence attitudes and motivation (two-ways) for the students. However, in this study the investigation is one-way.

However, due to time and financial constraints, this research only includes the above-mentioned variables/factors. Secondly, participants for this research were selected only from the capital city of PNG, Port Moresby, and were recruited using a stratified sampling method (discussed in Chapter 3). A greater number of research sites/areas could therefore have been covered to make this study more generalisable. However, the decision to collect data from one area was made to ensure the success and timeliness of the data collection process, as the researcher has familiarity with the location. Thirdly, the data analysed in this study are somewhat limited as they were collected at the same time and therefore have less generalisability compared to other research approaches such as longitudinal studies. Despite these limitations, however, this study has the advantage of supplying findings that are associated with relatively cheaper processes, and includes more samples than found in longitudinal designs (Ben, 2010; Creswell, 2008).

1.8 Structure of this thesis

This thesis has 13 chapters. Chapter 1 discusses the education system of Papua New Guinea, including the Grade 10 and 12 mathematics curriculum/syllabi to provide the background of the study with the statement of the problem, research questions, purpose and scope of the study.

Chapter 2 reviews previous studies on the school, teacher and student-level factors and their interrelationships that are affecting mathematics performance. This review highlights the issues in mathematics education and discusses the theoretical framework and the research design for the study. Moreover, the theoretical framework of the study is adopted from TALIS and PISA studies with an extension from the review of the previous literature.

Chapter 3 presents the research methods and design with the sampling framework and data collection techniques employed in the study. The chapter further explains the development and implementation of the survey questionnaires, test items and interview questions for the students and teachers. Additionally, this chapter discusses the process of linking the quantitative and the qualitative data at the teacher level. Finally, the theoretical aspects of the statistical procedures such as Rasch measurement model, confirmatory factor analysis, structural equation model and multilevel analysis are discussed.

Chapter 4 provides the student and teacher respondents' demographic and descriptive information. This chapter's results are important for subsequent analysis. In this chapter, descriptive analysis is carried out using the statistical software package IBM SPSS 24.

Chapters 5 and 6 discuss the table of specification and the process of developing the Grade 10 and 12 mathematics tests used in this study. The test items came from past years mathematics examination papers (permission from PNG government). Further, the chapters explain the statistical techniques for instrument verification, calibration and validation as well as scoring of the tests. The statistical software used in these two chapters is ConQuest 4.0 (ACER, 2012).

Chapters 7, 8 and 9 present the validation of the student-level scales (attitude and motivation) and teacher-level scales (quality of teaching, classroom environment and instructional resources) employed in this study. The scales items were drawn from TIMSS & PISA (permission from IEA and OECD respectively). Confirmatory factor analysis (CFA) is conducted using the Mplus 7.0 software to examine the structure of the construct of each scale. CFA is used to test a series of models to compare the models. The best fitting CFA model is used for Rasch analysis using ACER ConQuest 4.0. In this study, the Rasch model is used for verification, calibration, validation, as well for the scoring process of the scales with their items/observed variables to identify the structure. The verification of the observed variables (items) is carried out through Rasch analysis. The Item Separation Reliability and Weighted Likelihood Estimate (WLE) are used to measure the internal consistency of the scales. Finally, the scoring for each scale from the student and teacher-levels are conducted using the Weighted Likelihood Estimate (WLE) scores obtained from the scales (Woodcock, 1999). These WLE scores are then converted to W-scores and are used in the subsequent analysis of structural equation modelling and hierarchical linear modelling.

Chapter 10 discusses the Combined, Grade 10 and Grade 12 student models separately from each other to compare the results. The combined model is used because the same survey questionnaires are employed for both Grades 10 and 12. Moreover, in the combined model, the variable grade is used as a control variable to compare the differences between Grade 10 and 12 students. In the student-level models, factors are examined for interrelationships and their impact on students' mathematics results. The teacher-level model is also discussed separately to the student-level models. Teacher-level factors are also examined for the interrelationships and their impact on students' mathematics performance. This analysis was conducted using the Mplus 7.0 software (Muthén & Muthén, 1998), with the results discussed in this chapter.

Chapter 11 presents the results of Hierarchical Linear Models (HLM) for Combined, Grade 10 and Grade 12 student models with the teacher-levels. The HLM examines the cross-level interaction effects and the relationship between the predictor and the outcome variable (mathematics results) using the statistical software HLM 6.08.

Chapter 12 discusses the teacher-level interview data. A thematic approach is employed to analyse the data using the NVivo software (Hart & Achterman, 2017), to organise and analyse the data. However, this qualitative interview data has been converted to quantitative data for triangulation purposes using statistical techniques such as ANOVA and t-tests to compare the different teaching methods and confidence level of the teachers in mathematics content knowledge. Additionally, the themes identified from the interview data are discussed accordingly.

Chapter 13 summaries and discusses the main findings of the study. The limitations and issues pertaining to the study are also discussed. The study's significance and conclusions are highlighted with some recommendations for further consideration based on the results of the study.

1.9 Summary

The chapter has provided the background and the context for this study. It has presented the issues and problems in mathematics education that will be investigated for Grades 10 and 12 mathematics results in Port Moresby, PNG. In order to highlight the context of these issues, the chapter provided a background of the PNG education system and the Grade 10 and 12 mathematics curriculum. Further, the chapter has contextualised the issue within the PNG economy, underpinning the importance of a robust mathematics education system for the country. The extent of the problem, research questions, and the study's scope and limitations were also presented. This chapter has additionally provided discussion relating to the relevance of this study to the research area, and an overview of the factors affecting mathematics results through discussion of the study's aims/objectives, significance and an outline of the structure of the thesis. Next, in Chapter 2 (the literature review) the school, teacher and student-level issues/factors identified and discussed in this chapter will be elaborated in more detail.

Chapter 2: Factors affecting students' mathematics results

2.1 Introduction

This chapter reviews the literature related to factors affecting students' mathematics achievements, highlighting the arguments by different studies. Its purpose is to identify the factors affecting mathematics achievements, and their relationships to each other. Moreover, the literature review in this chapter synthesises what has been written about students' mathematics performance, in terms of the different factors influencing results at the student, teacher, and school-level. Specific issues/factors are identified and examined under these three levels with reference to previous studies.

2.2 Issues affecting mathematics achievement

Declining and poor mathematics achievement has been an alarming issue for schools all over the world (Hine, 2016; Mullis et al., 2012; Murray, 2011). Since the 1990s mathematics education researchers have been investigating the possible causes of the problem. Many researchers see declining mathematics performance as a serious issue because of the impacts that mathematics has on the broader economy. One obvious problem is enrolment of students who are ineligible to progress to the next year level of schooling. This enrolment process in many developing countries education systems, particularly in PNG, is associated with social chaos and instability that is associated with the selection process, which contributes to overall poor performance in mathematics (Joskin, 2013; Mel, 2007; Neofa, 2010). Decades of research into poor mathematics results has produced a number of theories related to likely causes, which can be organised into problems at three levels: the school, teacher and student-levels. Factors at these three levels are the basis of the theoretical framework of this study.

Studies have examined one or two of the factors that may affect students' mathematics results, however, this chapter consolidates the findings of many studies on factors affecting mathematics results. This literature review is divided into three major sections. The first section starts with a discussion of student-level factors (gender difference, attitude and motivation), examining effects on mathematics learning and students' outcomes. The second section provides a review of teacher-level factors (quality of teaching and classroom environment), and finally, the third section examines school-level factors (instructional resources and school type). This chapter then concludes with a summary, highlighting the key points of the review.

2.3 Student-level factors

This section reviews and discusses different studies/research conducted on the student-level factors that affect mathematics results.

2.3.1 Gender differences and similarities

Many studies have examined the impact that gender has on students' performances in mathematics (Brown & Kanyongo, 2010; Leder, Forgasz, & Solar, 1996; Sarouphim & Chartouny, 2017; Sukthankar, 1995). This is due to the widening gap in developing countries between boys' and girls' mathematics achievements. To address this gender gap issue in mathematics education, researchers have examined the differences and similarities between male and female students. Some of these studies have indicated gender gaps that favour boys in academic achievements. For instance, a quantitative study in China by Zhu, Kaiser, and Cai (2018) found that boys performed significantly better than girls in mathematics over the last decade. A recent study by Reilly, Neumann, and Andrews (2019) also reveals similar results.

Explanations for the gender gap between boys and girls have focused on different issues/factors, such as the role of academic attitudes, achievements, emotions, learning, motivation and lower self-belief hindering performance in mathematics (Brown & Kanyongo, 2010; Sarouphim & Chartouny, 2017). In general, girls' lower performances in mathematics are associated with both internal and external contextual factors. The drop in girls' performances is related to perceptions that their classrooms are unattractive, uncomfortable and hostile (Innabi & Dodeen, 2018; Leder et al., 1996; Preti et al., 2013). As a result, Buckley (2016) reports that female participation rates in mathematics are decreasing in some countries, due to the lower levels of confidence that girls have in completing their mathematics tasks. For instance, Buckley claims that in Australia the gap in mathematics confidence between females and males is larger than in other countries like Singapore and Canada due to less confidence and engagement. He further argues that this is a consequence of some girls having less confidence in solving mathematics problems that involve certain levels of reasoning. In light of this situation, Goldman and Penner (2016) examine females' and males' preferences in mathematics, identifying that girls are more likely to easily give up mathematics when they are less interested and have no belief in their abilities. A similar study by Law (2018) finds that girls are weaker in solving mathematics problems that involve higher-order thinking, and that teacher and peer support are important for girls' performance in mathematics. Another important aspect identified by Young (1995) was boys' and girls' attitudes towards mathematics and science, with boys having more favourable attitudes towards these subjects than girls. Besides this, emotions towards mathematics were studied by Frenzel, Pekrun, and Goetz (2007) who found that girls experienced less enjoyment and pride and more shame than boys, while boys experienced less anxiety and less hopelessness towards mathematics than girls.

On the other hand, recent research indicates that girls can perform better academically and overcome the gender gap (Goldman & Penner 2016; Stewart et al 2017; Fredricks, Hofkens, Wang, Mortenson, & Scott, 2018; Uwineza, Rubagiza, Hakizimana, & Uwamahoro, 2018). Similarly, PISA studies have shown a narrowing gap in mathematics achievement for males and females. For example, recent studies by Barry

(2019) in Saudi Arabia shows that females significantly outperformed male students in both mathematics and science results, because girls were given a higher level of classroom support in order to help them to achieve their learning goals (Barry, 2019; Reilly et al., 2019). Other studies have shown that, although biological differences exist between the sexes, these differences do not affect academic performance (Ben, 2010; Gherasim, Butnaru, & Mairean, 2013). Moreover, several studies have shown that performance in mathematics is not gender-dependent, and that intelligence is not correlated with gender (Barry, 2019; Ben, 2010; Reilly et al., 2019). These studies also affirm that females can perform as well as (or better) than their male counterparts in mathematics. In many countries, the gap between male and female students' mathematics performance closes with an increased support of the latter's attitude, emotions, and self-belief (Delaney & Devereux, 2019; Li, Zhang, Liu, & Hao, 2018; Sarouphim & Chartouny, 2017; Yoo, 2018). The results of these studies confirm that there are no significant biological differences underpinning either achievement or attitudes toward mathematics. Thus, it is important for this study to examine the role of gender in the creation of the gap in mathematics achievements of male and female students.

2.3.2 Parents' education level

Parents' level of education is one of the most important factors for students' educational achievements. There is evidence that parents' education level affects students' academic achievement in mathematics (Baliyan, Rao, & Baliyan, 2012; Khan, Iqbal, & Tasneem, 2015; Oyerinde, 2014). Parents who are educated raise children to have better self-perceptions in their academic abilities and engage them to develop a positive attitude towards learning experiences (Baliyan et al., 2012; Khan et al., 2015; Oyerinde, 2014). According to Dekar (2016), parents with higher education levels take more responsibility for their child's academic abilities, and they also have higher expectations for their children. They have confidence in their children to earn good grades in their academic subjects, which leads to children building their own confidence in their academic abilities, and being more likely to succeed (Dekar, 2016; Khan et al., 2015).

Studies by Musgrave (2000) and Dekar (2016) found that students that come from an educated family background work actively in their studies to follow the examples of their parents.

Moreover, there is a positive correlation between the level of parental education and student academic performance (Abuya, Mutisya, & Ngware, 2015; Baliyan et al., 2012; Tan, Zhou, & Li, 2020). According to Baliyan et al. (2012) students with parents who finished primary, senior school or university performed exceptionally well. These results reflect the ability of parents to support their children's schoolwork. For instance, in households with educated parents, parents and students are likely to interact regarding homework, with parents assisting with difficult exercises. A recent study by Tan et al. (2020) in China found that a higher maternal education level is related with more children's books at home and more extracurricular activities, which were then related to better performance in mathematics. On the other hand, students with parents that have lower education levels tend to perform worse than students whose parents have high education levels (Abuya et al., 2015; Elliott, Bachman, & Henry, 2020; Faize & Dahar, 2011). Similarly, Okumu, Nakajjo, and Isoke (2008) found that high academic attainment of a mother and father significantly reduces the chance of primary school dropout and improves student performance. For instance, an educated mother is more likely to devote time to her children's education than her uneducated counterparts. Faize and Dahar (2011) highlight that educated mothers are more effective in helping their children in academic work through monitoring and supervision of their children's academic progress. For fathers, higher academic success it is attributed to the fact that educated fathers are also interested in their children, and thus, they would be willing to spend more time to assist them with academic problems. Several studies examining the academic outcomes of students showed that parent's level of education made a significant contribution to achievements (Abuya et al., 2015; Baliyan et al., 2012; Dekar, 2016; Okumu et al., 2008; Oyerinde, 2014; Sarı & Cenkseven, 2008). Additionally, parents educated to a certain level would aspire to see their children advance further than they did educationally.

2.3.3 Attitude towards mathematics

In mathematics education literature, there has been much debate over the definition of *attitude*. The multidimensional nature of the concept has led to multiple definitions. Kiwanuka et al. (2017, p. 3) define *attitude* as “a disposition towards an aspect of mathematics that has been acquired by an individual through his or her beliefs and experiences, but which could be changed”. A recent study by Yavuz Mumcu and Cansız Aktaş (2020, p. 5) refers to *attitude* as “an individual’s tendency to have a positive or negative judgment of a sign, an object, or an individual around him/her”. However, in the context of this study, attitude is defined as a “ liking or disliking of mathematics, a tendency to engage in or avoid mathematical activities, a belief that one is good or bad at mathematics, and a belief that mathematics is useful or useless” (Neale, 1969, p. 632). Despite the ambiguity in the definition of attitude from other researchers, over the last 25 years there have been a large number of studies that document the development of students’ attitudes (Farooq & Shah, 2008; Kiwanuka et al., 2017).

Attitudes towards mathematics are the extent to which learners hold positive or negative feelings towards mathematics and their perception of its relevance, value and difficulty (Farooq & Shah, 2008; Kiwanuka, Van Damme, Anumendem, Van Den Noortgate, & Namusisi, 2015; Tok, 2015). Several studies have also shown a correlation between students’ attitudes towards mathematics and their achievements in the subject (Chen et al., 2018; Kiwanuka et al., 2017; Yavuz Mumcu & Cansız Aktaş, 2020). This relationship is highlighted in a study by Benken, Ramirez, Li, and Wetendorf (2015), which concludes that attitude toward mathematics is one of the most significant factors influencing participation in the subject. For instance, a study by Aytekin and Isiksal-Bostan (2019) in Turkey has shown that, when compared with their female counterparts, male students tend to have more negative attitudes towards the use of technology in mathematics lessons. This kind of difference in attitude tends to continuously become more negative as students move from primary to secondary school. Pitsia, Biggart, and Karakolidis (2017) highlight that many students fail to excel in mathematics due to an inadequacy of self-belief that can ensure the development of a positive attitude. Consequently, student performance in mathematics drops as they view and believe that mathematics is an exceptionally difficult and boring subject (Iqbal, Mirza, & Shams, 2017; Kiss, 2018; Soni & Kumari, 2017). Another study by Shin, Slater, and Ortiz (2017) on reading and mathematics in Korea and the USA highlights that this kind of negative attitude is a hindrance to students' achieving better results. The factor which has the biggest impact on mathematics performance is negative student attitudes. From the author's teaching experience and literature, it is believed that students with positive attitude in a particular subject perform well and obtain better academic results

There are several elements related to the formation and development of a student's attitude towards mathematics. In studies by Yavuz Mumcu and Cansız Aktaş (2020, p. 4) and Kiwanuka et al. (2015, pp. 1-2) attitude is discussed in the context of three components:- *cognitive, affective* and *behavioural*. These three components also comprise the attitude construct in the context of this study. The affective component refers to the liking-disliking aspect of attitude that changes from student to student and cannot be explained by facts (Yavuz Mumcu & Cansız Aktaş, 2020, p. 5). The cognitive component relates to students' knowledge and possession of confidence in their ability to learn and perform well in mathematics tasks (Kiwanuka et al., 2017; Yavuz Mumcu & Cansız Aktaş, 2020). The behavioural component is associated with the value that individual students place on the usefulness of mathematics, which enables him/her to understand the importance of mathematics for everyday use and later in life (Kiwanuka et al., 2017). Together, these three dimensions comprising the attitude construct interact and manifest themselves in the establishment of positive student attitudes in mathematics. The chapter now turns to discuss each component in the context of this study.

The *affective* component (*liking* mathematics) changes students' attitudes as they show interest in learning and performing mathematics tasks (Dowker, Cheriton, Horton, & Mark, 2019; Kiwanuka et al., 2017; Yavuz Mumcu & Cansız Aktaş, 2020). Students like mathematics when they are intrinsically motivated, and, as a result, learn mathematics with interest and enjoyment (Mullis et al., 2012; Stephens, Landeros, Perkins, & Tang, 2016). This is evident through the strong relationship shown in TIMSS studies between students' liking of mathematics and their achievements in the subject. For instance, a study in Indonesia by Oktiningrum, Zulkardi, and Hartono (2016) reveals that students who like learning mathematics had a higher average in mathematics achievements than those who do not like learning mathematics. On the other hand, students who do not like learning mathematics had the lowest mathematics achievements (Mullis et al., 2012; Thomson et al., 2012). This demonstrates that students who like and have positive attitudes towards mathematics are good at it, are more likely to enjoy learning it as a subject and so perform better overall. Further, studies have also shown a correlation between students' mathematics achievement and their liking and enjoyment of the subject. (Fan, Xu, Cai, He, & Fan, 2017; Özsoy & Ataman, 2017; Yavuz Mumcu & Cansız Aktaş, 2020). This correlation between attitude and achievement is highlighted in a multilevel study by Kiwanuka et al. (2017) and Benken et al. (2015), who highlight that when students like mathematics, they start to enjoy it and become more engaged in performing mathematics activities. As a result, these studies report that the affective component has a significant influence on the students' attitudes towards participation in mathematics. Such a finding emphasises the importance of further identification and clarification of the other influential elements of attitude, such as the *behavioural component (importance/value)* that impacts mathematics achievement.

The *behavioural* component *(importance/valuing)* relates to students' attitudes about the importance and usefulness of mathematics (Russo et al., 2020; Siahaan & de Mello, 2020; Yavuz Mumcu & Cansız Aktaş, 2020). The results of international studies have consistently shown a strong relationship between students valuing mathematics and their achievements in the subject (Mullis et al., 2012; PISA, 2013; Thomson et al., 2012). The degree to which the students value mathematics partly determines their performance in the subject (Siahaan & de Mello, 2020; Yavuz Mumcu & Cansız Aktaş, 2020). Students who do not value mathematics as a significant subject for furthering their studies and future career may therefore lack the motivation to continue studying it (Ahmed & Mudrey, 2019; Daher, 2020). Results from several studies have shown that students who said they valued mathematics typically had higher achievements than students who did not value mathematics (Gaspard et al., 2015; Harris et al., 2015; Thomson et al., 2012). For instance, findings from TIMSS studies indicate that, on average, Grade 8 students in Singapore who placed a high value on mathematics obtained better results. Even though some eighth-grade students do not especially enjoy learning mathematics, they do appreciate the value of the subject (Mullis et al., 2012, p. 329). This kind of attitude in turn prevents the decline of mathematics achievements, as valuing mathematics is an important subject for success. It is evident that, when students start to value mathematics and appreciate its benefits, their approaches towards mathematics change (Gaspard et al., 2015). In short, more positive attitudes towards mathematics improves students' mathematics results.

The cognitive component *(self-belief*) refers to the relationship that exists between students' academic self-confidence and positive attitude to perform better in academic tasks (Yavuz Mumcu & Cansız Aktaş, 2020). Over the last three decades, studies in educational research have discussed the relationship between self-belief and achievement. For example, a meta-analysis of 39 studies by Multon, Brown, and Lent (1991) reveals a positive and statistically significant relationship between self-efficacy beliefs and academic performance. Self-belief reflects students' convictions and positive attitudes that they can successfully perform a given academic task at designated levels (Susanti & Rosyida, 2020). For example, a study by Mack, Stanich, and Goldman (2019) in Chemistry reveals that higher levels of self-belief promote positive attitudes, which leads to better performance. On the other hand, students who have low levels of mathematics self-belief may perform poorly in mathematics, even despite their abilities, due to their negative attitude (PISA, 2013; Rutherford, Liu, Lam, & Schenke, 2020). In other words, in order to accomplish mathematical tasks successfully, students must have belief in their own mathematics ability. Students who have low self-belief are less likely to regulate their achievement behaviours or be motivated to engage in learning (Eklöf, 2007; PISA, 2013).

Strong self-belief encourages positive attitudes for students to engage with instruction and show persistence, effort, and attentiveness with a feeling that it is possible to succeed (PISA, 2013; Reed, Kirschner, & Jolles, 2015). Self-belief tends to improve mathematics performance through intrinsic motivation, and reduces students' anxiety towards the subject (Pitsia et al., 2017; Reed et al., 2015). For instance, according to PISA 2012 studies, Iceland saw one of the greatest improvements in students' self-belief to learn mathematics, and one of the largest reductions in anxiety towards mathematics among countries that participated in both PISA 2003 and PISA 2012. This reflects the positive attributes of the self-belief (cognitive) component of the attitude construct. Together, the three components (*affective, cognitive and behavioural*) comprising the attitude scale used in this study have been shown to have an impact on students' mathematics achievements.

2.3.4 Motivation towards mathematics

Motivation is a critical element in the learning process of any subject or course. However, motivating learners to engage in learning tasks/activities is a constant challenge, particularly for mathematics teachers (Durksen et al., 2017; Herges, Duffied, Martin, & Wageman, 2017; Kim, 2007), and it is a complex phenomenon that needs investigation to develop improved learning and instruction techniques. Recognition of this fact has prompted education researchers to conduct studies to investigate the underlying aspects attached to motivation in learning mathematics (Hashim Ali, 2013; Herges et al., 2017; Kim, Park, Cozart, & Lee, 2015; Seals, 2018). One such study, by Durksen et al. (2017) found that motivation is important because it enables learners to have an interest in pursuing learning goals. Further, the authors state that, for the student, a purposeful learning process is difficult to generate and sustain without motivation. This indicates that motivation is a crucial element in sustaining students' interest and promotes positive attitudes that improve results (Ahmed & Mudrey, 2019; Park, Gunderson, Tsukayama, Levine, & Beilock, 2016). However, recent studies have found that students' motivation in mathematics learning decreases as they progress to the upper years of their schooling, due to the increasing difficulty of the topics they study (Ng, Liu, & Wang, 2016; Seals, 2018; Yu & Singh, 2018). Based on the author's teaching experience, consistent motivation is required for students to be engaged in learning mathematics at all levels. Consequently, students find enjoyment and are actively solving the mathematics problems. In this study, motivation is discussed in terms of: a) *interest*, b) *engagement,* and c) *achievement goals.* This approach is consistent with existing international studies, though many studies differentiate motivation into both intrinsic and extrinsic aspects.

There are several reasons to study *interest,* particularly its importance in relation to learning. When one is interested, there are opportunities to learn and, likewise, when one learns, interest flourishes. According to Hidi and Renninger (2006, p. 111), "the level of a person's interest has repeatedly been found to be a powerful influence on learning". Interest is thus an important motivational factor (Ufer, Rach, & Kosiol, 2017) and, seen from the teacher perspective, engaging students is a constituent of good mathematics teaching (Ufer et al., 2017; Wilson, Cooney, & Stinson, 2005). The development of interest is one of the aims of mathematics teaching in the classroom, and teaching should develop students interest towards mathematics. Nyman (2017, p. 16) states that "interest can be visible to an observer as expressed by engagement directed towards something and is helpful in a classroom context". Another study by Frenzel, Goetz, Pekrun, and Watt (2010, p. 509) concluded that “contemporary approaches define interest as a motivational variable that refers to an individual’s engagement with particular classes of objects and activities”. As such, interested students are focused and aware of learning new knowledge. This focus and awareness is associated with what happens in the learning process of new mathematical knowledge in which students are engaged (Norton & Zhang, 2013; Nyman, 2017) and how certain tasks can be structured to be more interesting and engaging (Durksen et al., 2017; PISA, 2013). For instance, a study by Barkley (2018) indicates that teachers structuring lessons using flexible strategies can gauge the student interest level that needs to be engaged to solve a mathematical problem. This approach is significant for students’ motivation to engage in learning (Watt, Carmichael, & Callingham, 2017; Wilson, 2011). A study by Samuelsson (2011) shows a statistically significant correlation between interest, engagement, and test results, which underpins the important role of motivation in relation to learning. Similarly, empirical findings by Baumert and Schnabel (1998) show a relationship between interest, engagement and mathematics achievement. Hence, this study also stresses that motivation is a driving force in learning. Motivation is enhanced when students have an interest and active engagement in a mathematics learning goal.

Students’ *engagement* in learning is driven by their motivation. When students form connections in learning, they experience a sense of achievement, accomplishment and satisfaction. As students enjoy learning tasks, they become more engaged and motivated in the classroom. The structure of the classroom and the style of teaching conducted there therefore motivates students. For instance, Seals (2018) and Durksen et al. (2017) highlight that structure in terms of clear expectations from teachers is influential in maintaining student motivation and engagement. This is because the clarity of the lesson content and the teaching pedagogy arouse interest and motivate students to be engaged in the learning process. Motivation in learning activities underpins engagement and leads to improved achievements (Cai & Liem, 2017; Kim, 2007; Pantziara & Philippou, 2015). This is associated with students’ intrinsic motivation levels and may not be by extrinsic motivation levels. For example, a study by Mujtaba et al. (2014) of Year 8 and 10 students in London found that, though some students may have lower motivation levels, they still engage in learning, due to their high expectations for success, even though many of them do not enjoy mathematics. Similarly, Kim (2007) and Kim et al. (2015) stress that successful students believe in their own abilities to be successful and have the desire to be involved in learning activities. As such, increasing students' motivation can lead them to be fully engaged in tasks, promoting improvement in performances (Durksen et al., 2017; Mujtaba et al., 2014; Seals, 2018). This notion is also supported by the researcher's teaching experience: increased student motivation increases mathematics skills engagement, and that leads to increased self-confidence and improved results. Despite this, however, several studies have found that students in some countries lack motivation, engagement, interest and desire to succeed in mathematics (Kim et al., 2015; Mujtaba et al., 2014; Seals, 2018). Inspiring students to be motivated and engaged is not a straightforward process but needs time and consistent, everyday practice by teachers, as it is linked to students' attitudes, as discussed by Kim (2007). Accordingly, motivation to learn can be promoted and enhanced through the active involvement of students.

Students' *achievement goals* are one of the most important components of motivation in mathematics (Gherasim et al., 2013; OECD, 2005). This component relates to why a learner engages in specific achievement-related behaviours. "Achievement goals include the idea of mastery, because students with mastery goals tend to focus on developing competence, and are more likely to persist in learning activities, have greater interest in their classes, and seek help" (Simzar, Domina, & Tran, 2016, p. 3). This allows individuals to have greater motivation to reach their goals, even if they are less talented than their classmates, they can succeed more than those talented students because they are capable of setting goals for themselves and remain focused on achieving them (Duckworth et al., 2011; Eccles and Wigfield, 2002). The motivation to achieve goals prompts them to compete with others and directs them to pursue work they perceive to be valuable (Gherasim et al., 2013; Herges et al., 2017). This drive may come from an internal or external source. Overall, students with high achievement goals tend to do better in mathematics (Gherasim et al., 2013; Herges et al., 2017; Simzar et al., 2016). These students seem to possess higher self-esteem and stronger cognitive flexibility, which leads them making a greater effort in mathematics study, because they have higher autonomy and an internalised achievement motivation (Herges et al., 2017). A study by Ross (2008) found that students who are highly motivated to achieve goals are often autonomous individuals who believe that they can affect their learning in positive ways and solve problems. They keep their work spaces organised, have a sense of duty and obligation in learning mathematics, devote great effort toward achieving success (Cleary & Kitsantas, 2017), and regulate their behaviour to achieve their goals (Schenke, Lam, Conley, & Karabenick, 2015).

2.4 Teacher-level factors

This section reviews and discusses different studies and research projects that have been conducted on the teacher-level factors that affect students' mathematics results.

2.4.1 Teacher education-level

Teacher education level (highest educational degree obtained by a teacher) is one of the main attributes of teacher quality that has gained attention in the literature, and has been the focus of many investigations. However, the results of existing studies examining the relationship between student achievement and teacher education level have found contradicting outcomes, with some studies suggesting a positive relationship and others suggesting that there is no relationship at all (Kosgei, Mise, Odera, & Ayugi, 2013). For instance, a recent study by Tella (2017) finds that the relationship between teacher education level and student achievement is still unclear, while other studies suggest no relationship, and still others suggest a small, positive relationship (Tella, 2017; Zhang, 2008). However, recent studies have focused on whether teachers with a master's degrees or greater have a significantly greater impact on student achievement (Horn & Jang, 2017; Lee, 2018; Sauceda, 2017). These studies found that a teacher having a master's degree in science or education is a statistically insignificant predictor of student achievement in science. By contrast, a study by Baliyan et al. (2012) discovered that advanced degrees such as master's degrees in science or education have a direct influence on student achievement in science.

Other studies emphasize that a teacher cannot be determined to be qualified simply by checking his or her education level and years of experience. This is because teachers influence students through their interactions in the classroom, and teacher education level and experience only represent a portion of the ability to manage the classroom efficiently and to promote student achievement (Kosgei et al., 2013; Sauceda, 2017). Despite that, teacher experience and teacher education level are related to teacher quality. Several studies investigating the relationship between teacher education level and student achievement showed statistically significant positive results (Horn & Jang, 2017; Lee, 2018). In these studies, students were found to have learned more mathematics when their math teachers had an additional degree in mathematics (Kosgei et al., 2013; Lee, 2018; Zhang, 2008). However, only mathematics teachers who have master's degrees in mathematics will significantly improve students' mathematics achievements; and mathematics teachers with a master's degree in a subject other than mathematics will not have great impact on their students' mathematics achievements. In contrast, Zhang (2008) highlighted that an advanced degree in science or education significantly and positively influenced student science achievement. Based on the author's experience, teachers with advanced degrees in mathematics or education are more likely to exhibit positive teaching behaviours, which in turn is a characteristic associated with better student performance in mathematics

2.4.2 Classroom learning environment

The classroom-learning environment refers "to the social, psychological, and pedagogical contexts in which learning occurs and affect students' achievement" (Yang, 2015, p. 251). In the past decade, there has been a remarkable growth of research in investigating classroom-learning environments in both Western and Eastern countries, such as Australia, the USA, Singapore and South Korea (Yang, 2015). This is due to the impact that learning environments have had on students' mathematics and science results in TIMSS and PISA. For example, the TIMSS 2015 study found a positive relationship between classroom environment and student achievements. This is because an effective learning classroom environment has also been found to increase student motivation and participation (Hogan, Thompson, Sellar, & Lingard, 2018; Malik & Rizvi, 2018; Wang, Yin, Lu, & Zhang, 2017). Therefore, in the context of this study, the classroom learning environment is discussed in respect to a) its role as an innovative learning environment, b) teachers' role in the classroom, and c) the teacher-student interaction/relationship. Together, these aspects can promote a conducive learning environment for both teachers and students in the classroom. Each of these aspects is now discussed in turn.

Innovative and positive *learning environment*s are significant because they determine the quality of education received by students (Deieso & Fraser, 2019; Hogan et al., 2018). An innovative and positive learning environment is a system comprising of the physical involvement, the organizational objectives, and the characteristics of teachers and students (Schmidt & Cagran, 2006). According to Shmis, Ustinova, and Chugunov (2019), there is a significant relationship between innovative and positive learning environments and improved achievements through student engagement and behaviour. Therefore, the need for teachers to maintain positive environments in the classroom has been continuously emphasised in schools.

A study by Borich (1996) two decades ago identifies three types of classroom learning environment that teachers still utilise today in order to be innovative in their teaching: a) competitive, b) co-operative, and; c) individualistic classrooms. Firstly, in competitive classrooms, students are allowed to compete with each other while teachers lead or guide, with small amount of autonomy allowed for students (Howe, Hennessy, Mercer, Vrikki, & Wheatley, 2019). This is because, though classrooms led by teachers enhance students' achievements, competition can also be motivating for students. Secondly, co-operative classrooms are those in which classrooms student co-operation is emphasised, along with the teacher's timely intervention to guide them towards learning goals (Byrne & Prendeville, 2019). These classrooms effectively develop crucial co-operative learning skills that can enhance student achievements. Thirdly, individualistic classrooms are those in which students work on their own with little intervention by the teacher (Aji & Khan, 2019). The focus in these classrooms is on students developing their own independent learning skills. This allows them "to monitor and evaluate their own learning and reflect personal learning experiences in different authentic environments and social contexts" (Brown & Campione, 1996 p. 439). For instance, a study by Ampad and Danso (2018) discovered that this learning style assists students to undergo meaningful learning experiences and perform at their best. The three types of classroom highlighted by highlighted by Borich (1996) offer alternative innovative pathways for students to enjoy and engage in the learning process.

The *teacher's role* in the classroom is important for effective student learning to take place. Teachers play a key role in creating an atmosphere that is conducive to understanding how different students learn and apply different teaching methods. This allows students to use their own strategies in identifying ways to solve problems (Istance & Paniagua, 2018; Turner, 2017). To ensure success, teachers facilitate access to mathematics problems for students in three ways, as enumerated by Turner (2017, p. 17):

a) by allowing independent learning (allowing students to solve problems using their own strategies),
b) by presenting a range of problems at a varying degree of difficulty, guaranteeing that each student can solve at least one or more problems correctly, and
c) by considering student expectations.

Several other studies also strongly confirm that these three approaches allow students to learn better (Gherasim et al., 2013; Tshewang, Chandra, & Yeh, 2017; Turner, 2017). The independent learning method mentioned above is encouraged by many researchers in mathematics education because it allows students to construct their own ideas and strategies in learning. Besides this, Turner (2017) also highlights that teachers need to become facilitators so that students can become owners of their own learning. This facilitation process ensures that students participate in lessons and are engaged with the mathematics being presented. A study by Tshewang et al. (2017) in Bhutan found that one of the best ways to assist students to develop mathematics skills is through process-oriented activities in a comfortable classroom environment that encourages students to think, probe, communicate and discuss. For this to happen, Tshewang et al. (2017) and Gherasim et al. (2013) state that teachers are the constant variable that ignites students' interest in mathematics. In an inquiry-based classroom, discussions focus on students' thinking and attempts to understand relationships among qualities and concepts in relation to different algorithms. This means teachers provide positive feedback, listen and respond to students' questions, and are empathetic to students' needs by creating a supportive environment (Wang et al., 2017). A study by Shernoff, Ruzek, and Sinha (2017) supports Wang et al. (2017) by emphasising that environmental support is significantly related to learning. This could include the frequency with which the teacher assists student learning, as well as the frequency at which the teacher asks students to express their opinions.

The quality *of interaction* between teachers and students in the classroom plays a major role in motivating students' learning (Hu, Fan, Wu, LoCasale-Crouch, & Song, 2020; Pianta et al., 2020). In the context of mathematics, interest in the subject can be enhanced through teacher-student interactions in the classroom, particularly when a teacher responds through reinforcing and scaffolding at critical moments in teaching and learning. A study by Durksen et al. (2017) reveals that teacher-student interaction is important for motivation in mathematics to promote both confidence and connectedness. Further, a similar study by Brandenberger, Hagenauer, and Hascher (2018) finds that motivation for students is significantly increased when there is dialogue and effective communication with teachers to establish good learning relationships. Many improvements in student learning are provided through teacher interactions with students, and so reversing the declining trend of mathematics results globally necessitates focus on the quality of these interactions. The "quality of interaction between a student and a teacher that conjoins affective and cognitive realms in the process of aiming for mathematical learning" (Durksen et al., 2017, p. 177). According to Varhol, Drageset, and Hansen (2020) and a study by Yumiati and Kusumah (2019), the interaction between teachers and students is centered on two basic aspects, especially when teachers provide support to students. They are:

a) Emotional support of students that focuses on a positive climate dimension where the aim is on assessing relationships, positive effects, positive communications, and respect
b) Classroom organisation that includes the dimension of a positive climate used to assess the overall level of positivity between a teacher and students;

These two aspects that encourage high levels of student motivation in mathematics are primarily associated with a positive teacher-student relationship experienced within a supportive environment. A study by Hirshberg, Flook, Enright, and Davidson (2020) on classroom organisation found that teacher-student interactions are crucial for students' motivation to be engaged and to build confidence in mathematics. Students' willingness to try different solutions and be persistent is promoted in an appropriate classroom climate that allows for continuous interaction (Los, 2016; Shernoff et al., 2017; Yu & Singh, 2018). Elsewhere, other studies by Watt et al. (2017) and Tshewang et al. (2017) highlight that engagement is known to be associated with positive school outcomes, and is influenced by environmental factors such as students and teachers perceiving their classroom environments to be favourable. Similarly, Istance and Paniagua (2018) argue that positive outcome occurs when teachers create conducive environments that allow students to be comfortable in learning and enable them to improve in mathematics through satisfactory communication with teachers. The three aspects discussed above explain the classroom-learning environment in the context of this study.

2.4.3 Quality of teaching

The concept of quality of teaching has been a core issue debated in educational circles in many countries (Akyuz, Dixon, & Stephan, 2013; Escribano, Treviño, Nussbaum, Irribarra, & Carrasco, 2020; Kanzow & Wiegand, 2020). As such, education systems in many countries, including Papua New Guinea have undergone significant changes in many aspects of quality of teaching. These changes to improve teaching methods and academic results include curriculum reform, changes to textbooks, new teaching practices/approaches, and a focus on teacher retention rates (Pennell, 2020). However, there is still an inconsistency in teachers' roles in the practical aspects of teaching, which has effects on students' critical thinking levels. For an education system to be successful, the dimension of quality of teaching is important in driving better student academic outcomes/results.

In this study, the Quality Teaching Model (QTM) developed and implemented by the New South Wales (NSW) education system is adopted as a benchmark for pedagogic practice to clarify the concept *quality of teaching*. This model is chosen due to the association of comprehensive teacher attributes that has the potential to promote teaching and learning. Despite attempts through in-service and teacher training to encourage teachers to employ the QTM concept, it remains little utilised by them (Prieto-Rodriguez, Gore, & Holmes, 2016; Sakarneh, Paterson, & Minichiello, 2016). The QTM provides a theoretical framework that can be used to investigate the effectiveness and quality of teachers' approaches, practices and behaviour (Green, Eady, & Andersen, 2018; Kyriakides, Creemers, & Panayiotou, 2018). The model centres on the challenging issues of intellectual quality and equitable student outcomes. According to the NSW Department of Education and Training (2003), there are three main features of teacher classroom practices that are associated with improving student-learning outcomes: 1) Intellectual quality (content knowledge) that enables and promotes high-standard teaching, 2) Quality learning environments that ensure effective teaching, 3) Significance of students' work that teaching must develop and make clear to them.

The Australian Institute of Teaching and School Leadership (AITSL) (2011) argues that teaching strategies that focus on these three dimensions have a positive influence on student learning. With these dimensions, teachers engage, challenge and lead students to better educational outcomes (Green et al., 2018; Kyriakides et al., 2018; Sakarneh, 2014). Each of the three dimensions are comprised of six elements (refer to Table 2.1) that further define and clarify the nature and function of the model in terms of classroom behaviours and practices.

Table 2. 1 NSW dimensions of a quality-teaching model

	Intellectual quality	**Quality learning environment**	**Significance**
Elements	deep knowledge	explicit quality criteria	background knowledge
	deep understanding	engagement	cultural knowledge
	problematic knowledge	high expectations	knowledge integration
	higher-order thinking	social support	inclusivity
	Metalanguage	students' self-regulation	connectedness
	substantive communication	student direction	narrative

The three dimensions highlighted in Table 2.1 are inter-related and interactive in nature (Green et al., 2018). The framework shows that the intellectual quality of teachers is at the centre of the QTM because teachers' involvement in the process of teaching and learning is measured through their effectiveness to facilitate quality learning outcomes for students (Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Sakarneh, 2014; van de Grift, 2014). These three dimensions and their constituent elements are now discussed in further detail.

First, the dimension of *intellectual quality* requires teachers to select and organise the essential knowledge, skills, and values from the syllabus for easy understanding by students (Green et al., 2018; Kyriakides et al., 2018; Sakarneh, 2014). This enables teachers to have deep content knowledge and assist students to progress in their work. Teachers with stronger content knowledge are more likely to respond appropriately to students' mathematical ideas and make fewer mathematical or language errors during instruction. The converse of this relationship also appears to be true: that is, a lack of content knowledge seems to limit a teacher's instruction capabilities. For instance, a study by Charalambous, Hill, Chin, and McGinn (2019) found that when teachers with weak content knowledge departed from their instructional materials, they tended to obscure or distort the mathematics concepts that are expected to be learnt by the students. This is because they chose to add inappropriate mathematical representations to their instruction.

However, several studies reveal that experienced teachers can impart these concepts effectively through interactive teaching strategies (Ambrosetti, Capeness, Kriewaldt, & Rorrison, 2018; Chao, Chen, Star, & Dede, 2016; Green et al., 2018), even if they have only limited exposure to content knowledge and general 21st century instructional processes of planning and delivery (Ambrosetti et al., 2018; Chao et al., 2016). On the other hand, fresh graduates may possess a depth of knowledge with the subject, curriculum content and up-to-date methods of teaching, but have only limited experience in imparting knowledge to students

(Ambrosetti et al., 2018; Chao et al., 2016). These two kinds of teachers can influence the educational outcomes of students. According to Green et al. (2018), both types of teachers have the ability to develop students' deep and critical understanding of the selected knowledge, understandings, skills and values to form connections among them. Teachers are the driving force facilitating students' active engagement in the learning process through prerequisite skills and strategies (Fauth et al., 2014; Kyriakides et al., 2018). Therefore, intellectual quality elements such as problem knowledge, higher-order thinking, analysis of instructions and deeper thinking in teaching are crucial for the critical dissemination of information (Akyuz et al., 2013; Fauth et al., 2014; Polikoff & Porter, 2014). One of the most effective teaching strategies is to connect teachers' and students' knowledge to real-world, practical situations both inside and outside of school (Heikkinen, Wilkinson, Aspfors, & Bristol, 2018; Hogan et al., 2018). This is encouraged through teachers' effective pedagogy and deep knowledge of learners (Heikkinen et al., 2018; Hogan et al., 2018). Consequently, teaching with a high level of intellectual quality is more contextually connected and enables students to engage with tasks in more complex, deeper, subtle, and sophisticated ways.

Quality learning environment is the second dimension of the NSW quality-teaching model. In a high-quality learning environment, good results are expected by teachers from their students (Stronge, 2018). Teachers explain clearly the expected quality of work students need to produce and present students with some choices of learning activities (Istance & Paniagua, 2018; van Vendeloo et al., 2018). Teachers and students embrace mutual respect, with teachers encouraging students to have an interest in their work to enhance their learning. However, in some schools, teachers are required to work in overcrowded classrooms in which they need to keep constant control of student behaviour to avoid disruptions. This situation hinders teachers' abilities to engage students in actual learning, and forces teachers to be preoccupied with concerns about behaviour. A balanced and equal interaction between teachers and students can instead enable teachers to spend more time and effort facilitating learning rather than regulating student behaviour

(Mueller, Mylonas, & Schumacher, 2018; Stronge, 2018). Student-centred teaching and learning processes, rather than traditional teacher-centred approaches, are encouraged because they allow students to take ownership of their learning (Kyriakides et al., 2018). The process of teaching is a robust tool that teachers use for students to acquire knowledge and develop their skills (Akyuz et al., 2013; Winheller, Hattie, & Brown, 2013). Enhancement of learning through the cooperation of teachers and students in the classroom boosts productivity and leads to better results. This is because such an environment interrogates and supports all students and so encourages autonomy and enthusiasm (Mueller et al., 2018; Stronge, 2018).

Furthermore, it is necessary to provide effective links between new lessons and students' knowledge of the world. By including viewpoints and understandings from other cultures in lessons (Mueller et al., 2018; Stronge, 2018), teachers demonstrate that different types of knowledge are valued and are a legitimate part of students' learning. Similarly, teachers explaining the purpose of learning a particular topic with the knowledge and its associated skills (Fauth et al., 2014; Sakarneh, 2014) is important because it makes learning relevant to students and the world in which they live. In such an environment, teachers ensure all students are included, and have input into the teaching-learning process (Polikoff & Porter, 2014; Sakarneh, 2014; van de Grift, 2014), and students can appreciate the reasons for studying a topic.

Finally, Fleming, Gibson, Anderson, and Martin (2020, p. 5) define that *significance,* the third dimension of quality teaching, as the pedagogy that assists students to make their learning meaningful and important. This is due to this pedagogy's clear connections with students' prior knowledge and identities. The relationship is built with contexts outside of the classroom, and with different ways of knowing or cultural perspectives (Gore & Rosser, 2020; Moloney & Xu, 2018; Sakarneh, 2014). As stated by the NSW Department of Education and Training (2003 p.14), "to achieve high quality learning outcomes for each student, students need to see why, and to understand that, their learning matters". The significance of students' learning depends on the connections among their peers as individuals and social beings (Fleming et al., 2020; Sakarneh, 2014). Moreover, the kind of work they are involved in and the contexts of this work are paramount. Teachers should make connections and link their lessons to students' prior knowledge, cultural backgrounds and communities. A study by Buchanan (2017) highlights that such an approach can allow interaction between students and teachers in differing fields of knowledge. The connection between teachers and students from the combination of the former's content knowledge and their latter's cognitive, social and cultural backgrounds (Moloney & Xu, 2018; Sakarneh, 2014) can facilitate effective instruction.

This broad NSW quality teaching model is used due to its applicability to the general principles in teaching at schools in PNG. This model has assisted to inform the factors teachers experience in their teaching in the classroom. This quality teaching is defined by teachers' content knowledge and their teaching pedagogies as stipulated in the theoretical framework in Figure 2.3 p.56. These two factors are associated with the dimensions highlighted in Table 2.1 p. 44 of the NSW quality teaching model which are applicable at PNG schools. Thus, the quality of teaching in this study is measure by instruments developed to specifically measure teachers' content knowledge and their teaching strategies as enshrined in the theoretical framework in Figure 2.3 p.56.

2.5 School-level factors

This section reviews and discusses studies and research projects that have been conducted into the school-level factors that affect students' mathematics achievements.

2.5.1 Instructional resources

Instructional resources refer to objects or devices that assist teachers to make learning meaningful to learners. Instructional resources can be classified into two types (Aina & Adekanye, 2013; Ode, 2014). The first type are visual resources made up of reading and non-reading resources such as textbooks, teachers' guides, pictures, maps, charts and graphs. The second type consists of audio-visual resources that are made up of both electronically operated and non-electronically operated resources. These resources include laptop computers, audio tape recorders, video tape recorders and slide projectors. These resources offer a variety of learning experiences to meet different teaching and learning challenges both individually or in combination (Fan et al., 2017; Sabah & Hammouri, 2010). A study by Olayinka (2016) in Nigeria found that instructional resources that increase learning are of vital importance to the teaching of any subject in the school curriculum. He further stresses that the use of instructional resources can assist students to be more engaged and motivated in learning. For instance, a teacher who enthusiastically makes use of appropriate instructional resources to supplement teaching enhances and promotes students' innovative and creative thinking. Thus, mathematics results in many countries are reflections of the quality of the instructional resources utilised by teachers, and their use by students at schools (Fan et al., 2017; Olayinka, 2016). This is because such resources are major sources of teachers' planning and practice, and function to bring different discourses together. Fan et al. (2017) stress that the use of instructional resources is beneficial for teachers' learning as well. Instructional resources can contribute to the professional development of teachers' and guide them to teach effectively and with confidence. A study by Sabah and Hammouri (2010) highlights that well-designed, instructional resources can increase participation in learning for both students and teachers. Teachers use textbooks and teachers' guides to assist them to differentiate teaching styles and in understanding the curriculum. A similar sentiment is shared by Clement et al. (2012) who affirm that teaching guides and syllabi assist teachers with new ways to teach and simultaneously ensure that students learn what is intended for them at school.

Despite these beneficial qualities, however, many schools in developing countries have limited access to instructional resources (Evans & Popova, 2015; Musau & Migosi, 2014). In such context teachers are therefore hampered in their lesson preparation activities and abilities to provide guidance to students to reach their full potential (Lake, 2010). For instance, in the researcher's experience teaching in PNG secondary schools, often there is only one mathematics textbook or teachers' guide to share among teachers, and sometimes there is no textbook at all, so teachers are forced to borrow them from other nearby schools. Similarly, a hierarchical linear model study by Sabah and Hammouri (2010), shows that a shortage of instructional resources negatively affected the mathematics achievements of Jordanian eighth graders. This kind of scenario hinders effective learning and teaching between students and teachers, and consequently leads to poor performance by students (Olayinka, 2016). Use of instructional resources in implementing a curriculum is therefore crucial for the better transmission of information.

The use of instructional resources assists students to intensively explore, experiment, create and interact with the learning environment (Olayinka, 2016; Tuomi, 2006; Wei, Clifton, & Roberts, 2011). When more instructional resources are provided to students, a more effective learning environment is enabled (Tuomi, 2006), with these resources acting as instruments that more constructively facilitate teaching and learning. Instructional resources are therefore necessary ingredients in the development and implementation of a mathematics curriculum (Olayinka, 2016; Wei et al., 2011; Zakharov, Tsheko, & Carnoy, 2016). The aim of these resources is to increase the effectiveness of teaching mathematics as a means of preparing students for their future responsibilities as adults (Tuomi, 2006; Wei et al., 2011; Zakharov et al., 2016). At the same time, these resources may influence teachers' beliefs about how they teach mathematics and how this may influence students' learning.

2.5.2 School type

Different types of schools namely public, private and Catholic schools each have their own unique style of education. Several studies show that the type of education system a school offers has effects on the academic achievements of its students (Belfield & Levin, 2015; Berends & Waddington, 2018; Singh & Sarkar, 2015; Urquiola, 2016). This influence is caused by factors such as tuition fees, school environments and the quality of teaching and learning resources provided by schools. The three types of schools (public, private and Catholic) and their impacts on students' achievements are now discussed.

Public schools are supported and fully funded by governments. In some countries, the government also subsidises tuition fees and expects enrolment figures to increase to give equal opportunity to children (Fryer Jr, 2016; Kozlowski, 2016). Some researchers in education argue that this puts more pressure on principals and teachers to accept more students despite sometimes poor conditions of the learning environment (Abdulkadiroğlu, Pathak, & Walters, 2018; Fryer Jr, 2016). This leads to having large class sizes that usually impact negatively on students' learning processes (Hallinan & Kubitschek, 2012; Lefebvre, Merrigan, & Verstraete, 2011). For example, in this author's experience of teaching in Papua New Guinea, a teacher teaches 60 to 70 students in a class, making the effective delivery of lessons very difficult. Such poor learning environments affect academic outcomes and result in low quality education (Berends & Waddington, 2018; Hallinan & Kubitschek, 2012; Lefebvre et al., 2011).

On the other hand, private schools are market oriented and focused on students and their parents as customers. They are operated as businesses, as well as being concerned with the outcomes of the education they deliver. Stevans and Sessions (2000) highlight that the private school industry thrives in the USA. The continued success of private education is based upon the social status associated with a private school education, and the perception that private schools provide a higher quality `product' than public schools (Grimes, 1994; Stevans & Sessions, 2000). For instance, according to a 1999 study by the US Center for Education Statistics (US Department of Education, 1999 cited in Stevans & Sessions, 2000), at that time about 50% of the elementary and secondary school student population attended private institutions, and this percentage has been increasing over the previous past 25 years. This is because private schools charge school fees according to the quality of education they provide to their clients/students. These schools have conducive learning environments, and high-quality teachers with strong experience who are equipped with content knowledge and teaching pedagogy (Lefebvre et al., 2011). In addition, a comparative study by Siddiqui and Gorard (2017) in Pakistan showed that, compared to public schools, private school class sizes are manageable, therefore enabling for teachers to facilitate classes with one-on-one effective interactions and discussion. These different aspects of learning enable students to produce better academic outcomes (Siddiqui & Gorard, 2017).

Finally, Catholic schools are a different type of school from the two types discussed above. Their education system is centred on forming students to become good and honest citizens (Fleming, Lavertu, & Crawford, 2018; Miserandino, 2019), and they place more emphasis on the integral human development aspect (Fleming et al., 2018; Miserandino, 2019). This means students' overall discipline and commitment towards their studies is emphasised by schools and their teachers. Coleman (2018) stresses that there is a mutual understanding among parents, students and teachers about the outcomes and expectations of each party. Students are encouraged to work hard either in their studies or in other beneficial activities in their lives. Many Catholic school teachers are specialised; they have experience in the subjects they teach and are committed to raising students' academic standards (Agirdag, Driessen, & Merry, 2017; Pollock & Mindzak, 2018; Starkey & Rymarz, 2018). In addition, a study by Shouse (2018) asserts that such a school environment is conducive to successful learning, and fosters an atmosphere of friendliness among teachers, students and parents who collaborate and work together as a community. For instance, in this author's teaching experience in PNG Catholic schools, their learning environments are organised and systematised in all aspects of learning, and students perform well in their end of year national examinations. This suggests that Catholic schools' organisational approaches towards teaching and learning have positive impacts on students' academic results. The three different types of schools discussed above each have an impact on their students' academic achievements depending on their systematic approach to students' leaning. Overall, well-organised schools with experienced teachers and conducive learning environments support better academic outcomes.

2.6 Theoretical framework

The proposed research model tested in this study is an integration of the theoretical framework of the Teaching and Learning International Survey (TALIS) and the Programme for International Student Assessment (PISA). The integration is possible because the factors identified in the literature are best explained and represented by these two international studies (TALIS and PISA) theoretical frameworks. As such, the factors/constructs tested in this study's research framework are based on established and authoritative knowledge from the above literature review.

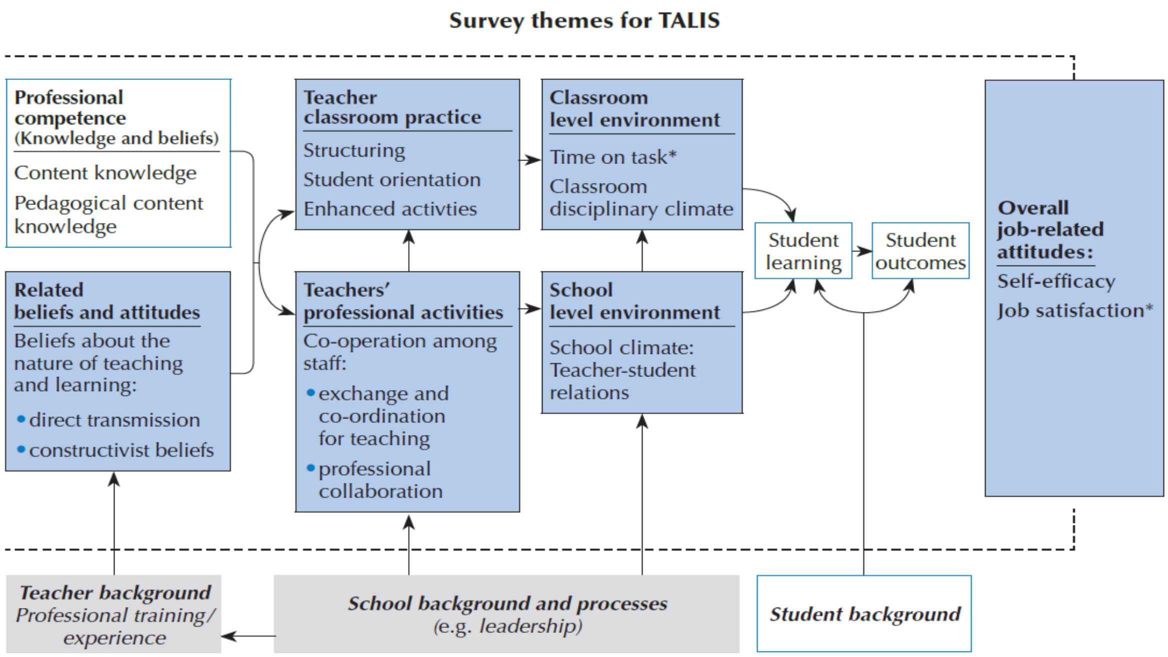


Figure 2.1 TALIS Theoretical Framework (OECD, 2010 p.32)

The TALIS theoretical framework (Figure 2.1) highlights that the quality of the learning environment is the most important causal factor for student leaning outcomes (OECD, 2010 p.32). The *learning environment* includes both the school and classroom levels. It is evident from the TALIS framework that the classroom-learning environment has a greater influence on student learning. This implies that factors at the classroom level are significant in the learning processes of students.

The quality of learning environment at the classroom level is determined by classroom practices. These practices are performed by teachers, chiefly through the process of giving instruction to students. This makes teachers one of the most important factors in creating a quality-learning environment, as they have the responsibility of implementing practices. In other words, teachers' professional activities are significant due to their positive impact on student learning in the classroom. Furthermore, the TALIS framework (Figure 2.1) highlights the importance of teachers' professional development as it equips and develops their skills and knowledge. Professional development allows teachers to co-operate and share ideas to improve their practices and attributes, such as beliefs and attitudes towards teaching (OECD, 2009; 2010). Consequently, such development opportunities enable teachers to contribute positively to student learning outcomes in the classroom. The framework in Figure 2.1 highlights that professional competencies (such as knowledge) and attributes (such as attitudes and beliefs of teachers) can affect learning at the classroom level. In turn, this can also affect student learning and other outcomes. This means that students' attributes such as attitude and belief in the ability to learn are influenced by teacher competences and other characteristics such as teacher engagement in professional activities (OECD, 2009; 2010).

The PISA framework emphasises mathematical literacy and the three main domains of mathematical knowledge: content, context and competencies. This means that mathematical tasks need to be tied to the context of problems in the real world. Further, for a student to solve a problem, he/she must have a degree of mastery over the relevant mathematical content, and the problem needs to be solved by following the proper steps/processes. However, in order to successfully implement these processes, students need a certain level of competency. The PISA framework elaborates that mathematical literacy is centred on the abilities of students to examine, reason, and identify concepts. This is because they process, develop, solve and explain mathematical problems in different ways. Assessment should therefore be patterned on real world situations, to apply the skills learnt in the classroom (OECD, 2004a).

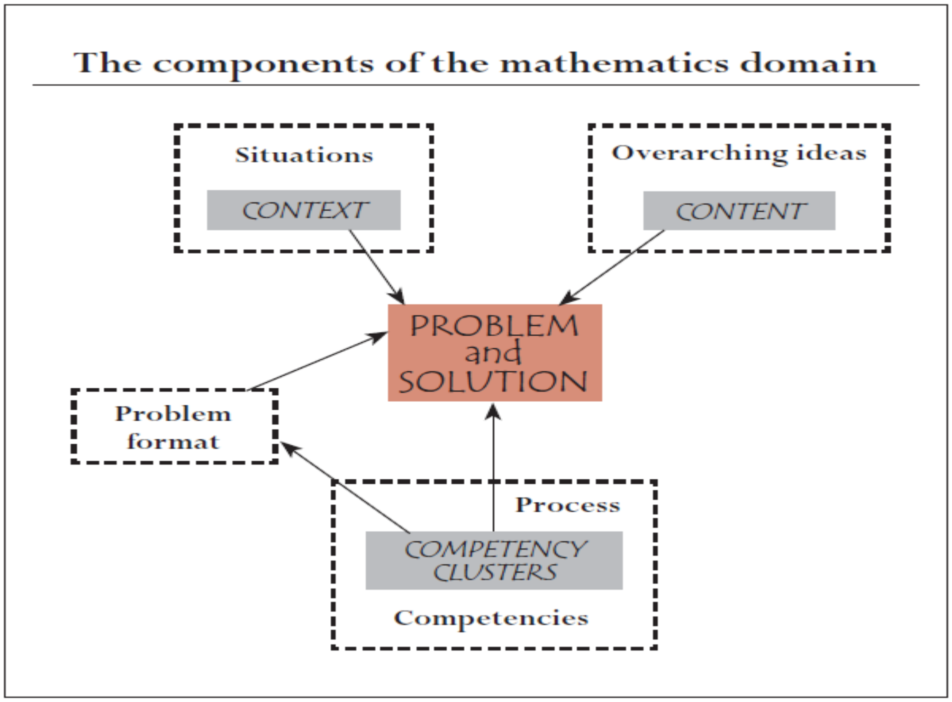


Figure 2. 2 PISA theoretical framework (OECD, 2004 p.18)

This framework highlights that attitudes and emotions (such as self-belief, eagerness, interest and desire to accomplish tasks) and relevance are important prerequisites for mathematics literacy (OECD, 2004a). Students must have some degree of positive attitude and emotions towards mathematics, even though these constructs may not be components in the definition of *mathematical literacy*. In practice, the framework's implication is that literacy can be applied to practice in the real world by students that have some degree of self-confidence, curiosity, feeling of interest and relevance, and the desire to engage with and understand problems that contain mathematical components. As such the PISA framework (Figure 2.2) underpins and reflects this study's student level factors such as attitude and motivation (reviewed in the literature) that affects students' mathematics results.

The proposed theoretical framework of this study shown in Figure 2.3 is derived from this chapter's literature review of the different factors affecting students' mathematics achievements. This framework is integrated with constructs from the school and teacher-level factors highlighted by TALIS, and student-level factors from PISA. These student-level factors include attitude and motivation towards mathematics and demographic variables. Furthermore, teacher-level factors include quality of teaching and classroom environment with demographic variables and school-level factors include instructional resources and school type.

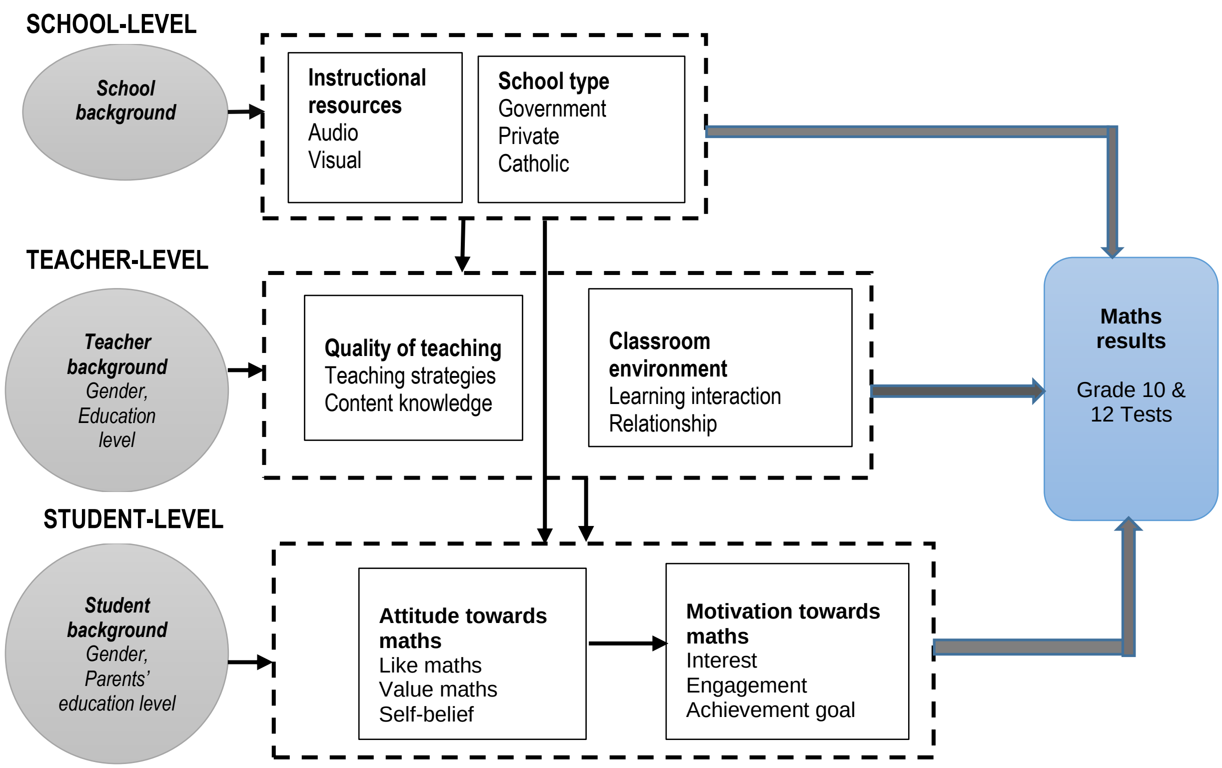


Figure 2. 3 Proposed theoretical framework for the study (OECD, 2004 & OECD, 2010)

This study's framework underpins the aim of the study to examine the interrelationship of these factors and their influence on students' mathematics achievements. The effect of the demographic variables on the factors at the three different levels on mathematics achievements are also examined. These variables with the factors/scales at each level (school, teacher and student-levels) may directly or indirectly affect mathematics results. Moreover, as shown in the proposed theoretical framework in Figure 2.3, the scales with demographic variables from one level may have a cross-interaction relationship with another level. For instance, the instructional resources scale from the school-level may have a cross-interaction with the attitude scale at the student-level that may affect the Grade 10 and 12 students' mathematics results. As such, survey questionnaires for the scales with the demographic variables at each level and the Grades 10 and 12 mathematics tests (see Figure 2.3) are examined to see whether the proposed relationships in the theoretical framework exist.

2.7 Summary

This chapter has reviewed the relevant literature and existing studies covering the factors tested for in this study. The chapter began by examining issues in mathematics education, discussing the factors that affect mathematics results at the student, teacher and school-levels, to provide a conceptual background. At the student level, factors such as gender difference, attitude and motivation were discussed. Quality of teaching and classroom environment were the factors examined at the teacher level. Finally, instructional resources and school types were the factors discussed at the school level. These factors were discussed in the context of how they affect students' mathematics results, and it was noted that the factors at the three different levels are interrelated and have relationships with each other, either directly or indirectly, that influence students' results.

Next, this chapter discussed the theoretical framework adopted in this study and its rationale. The model in this study proposes that attitude, motivation, quality of teaching, classroom environment and instructional resources affect students' mathematics achievements. Demographic factors at the student and teacher levels such as gender, years of teaching experience, school type parents' employment and highest education, similarly have an influence on results. However, for better results, the instruments that measure the factors/constructs reviewed in the literature still need to be validated in the context of this study through data collection and analysis. Therefore, the next chapter discuss the methods and techniques employed to develop the instruments to collect and analyse data for the factors identified that seem to affect Grade 10 and 12 mathematics results.

Chapter 3: Research Methodology

3.1 Introduction

The investigation into factors affecting Grade 10 and 12 students' mathematics performance is the first of its kind to be conducted in Port Moresby, Papua New Guinea (PNG). Hence, a suitable method and design are needed to secure reliable data for meaningful findings and recommendations. This chapter discusses the research methods and design that are employed in this study to collect, analyse and interpret the data. It begins with a discussion of the planning stage, which provides a background for the selection of the research topic, question, design and methods. Next, this chapter includes discussion of this study's sampling framework, respondents, and the process of data collection. The different items/scales employed in the study are then described, with reference to the development and validation of the instruments (mathematics tests, survey items and interview questionnaires). This chapter then continues with a discussion of the data analysis method employed to provide the statistical procedures/steps on the data collected, and how they were treated and interpreted. Finally, this chapter then concludes with a summary of important points discussed.

3.1 Planning stage

This section highlights the significant tasks that were undertaken before the collection of data by this study.

3.1.1 Location of the study's focus

The focus of the study was developed from a substantial review of the literature about problems in mathematics education (factors affecting mathematics results). This was motivated by the researcher's observation teaching mathematics in PNG at both secondary schools and tertiary institutions. Another motivational factor was the annual decline in Grades 10 and Grade 12 students' mathematics results during selections for higher education. Related to this was, the issue (discussed in Chapter 1) that few Grade 12 students are entering tertiary institutions to undertake engineering and science courses that involve mathematics. The researcher reviewed relevant articles and books on the topic of mathematics education from Asia, Europe, United States, New Zealand and Australia. These countries/regions were selected because most of the issues in mathematics education highlighted are also relevant to the PNG context. The issues include quality of teaching, classroom environment, instructional resources, attitude and motivation towards mathematics; all of which affect the mathematics achievements of students (Cheng, 2017; Drent, Meelissen, & van der Kleij, 2013; Ingvarson, Meiers, & Beavis, 2005; Vanlaar et al., 2016).

As a result of this review in Chapter 2, several issues for mathematics education in Port Moresby, PNG, became apparent. These issues include the motivations and attitudes of students towards mathematics, which affect their results. One objective of this study is to contribute to the pool of still-fragmented knowledge about mathematics education in PNG. Furthermore, it is hoped that this study can serve to open more research opportunities in the area of mathematics education in Pacific Island countries such as PNG, Fiji, and the Solomon Islands, where there are currently few or no publications about students' attitudes, motivations, learning environment and teaching quality in mathematics learning. In addition, the findings of this study may also facilitate improvements in mathematics results in secondary schools, as well as tertiary institutions in PNG.

3.2 Research design and method

The study employs a quantitative-dominant, mixed-method research design, with both psychometric and multilevel methods that provide detailed measurement and structural information to all scales used to understand problems (Creswell & Clark, 2007; Onwuegbuzie & Combs, 2011; Onwuegbuzie & Leech, 2005; OECD,2017; TIMSS & PIRLS, 2016). A mixed-method research design was chosen because quantitative methods are useful in identifying the nature of the problem, and use of qualitative methods increases understanding of the underlying phenomenon (Creswell, 2002; Creswell & Creswell, 2017; Onwuegbuzie & Combs, 2011). Using quantitative approaches provide valuable information (through rigorous treatment of data) to address the research problem through explanation of relationships between the school, teacher, and student level factors (Creswell, 2008; Creswell & Clark, 2007; Onwuegbuzie & Leech, 2006). Additionally, a qualitative research method is used to collect additional data to support the study's third research question, 'What attributes, are deemed important by teachers regarding their mathematics teaching approaches and content knowledge?' Moreover, a qualitative method is used to provide more in-depth understandings of the collected quantitative data relating to quality of teaching (teachers' teaching methods and content knowledge) in facilitating students' mathematics learning (Onwuegbuzie & Combs, 2011; Onwuegbuzie & Leech, 2005). The integration of both methods was effective in providing comprehensive understandings of respondents' situations. This study acknowledges the nested nature of educational data that promotes quantitative methods. However, both research methods were employed in order to obtain methodologically sound data and to allow triangulation of data/information to understand better context and status of mathematics education in the sample from PNG (Creswell, 2008; Onwuegbuzie & Combs, 2011).

3.3 Ethics approval

Prior to the administration of this study, it was necessary to obtain ethical research approval from the University of Adelaide's Human Research and Ethics Committee (UAHREC). The UAHREC granted approval for this study to proceed on 14 July 2017 (Ethics Approval No H-2017-133- see Appendix A). The committee's approval was granted on seven conditions to which the researcher must conform: 1) every participant be provided with an information sheet about the study; 2) every participant must read, sign and return the consent form to the researcher to participate in the study; 3) consent from parents/guardians was to be obtained for participants below the age of 18 years; 4) the identity of every participant was to be kept confidential when conducting survey questionnaires and interviews; 5) teachers' employment and students' academic standing were not to be affected in any way; 6) teaching and learning were not to be disturbed; and 7) participation was voluntary and participants were free to discontinue at any time. Consideration of these conditions was important in carrying out the study. The conditions were made clear to participants through the participants consent letter before administration of the survey and interview to clarify doubts pertaining to ethics requirements.

Since the study was to be employed in Port Moresby, approval was likewise sought from the PNG National Department of Education (NDOE) research division, and secondary schools' secretary for education in Port Moresby, respectively. Permission to conduct the study in schools in PNG was granted on 30 August 2017 (see Appendix A) by the Papua New Guinea National Department of Education research division. Permission to collect data from secondary schools in Port Moresby was approved on 14 August 2017 by the secretary (see Appendix A). The principals of the secondary schools in Port Moresby requested reports of the results of the data analysis as part of their approval of the study. The necessary permissions were therefore obtained prior to undertaking this study.

3.4 Sampling and data collection

This section discusses how to select data from a population and how much data to collect and how often it can be collected in a subset of data points to identify patterns and trends of a research undertaken.

3.4.1 Sampling framework

The sampling framework used in this study was patterned on large-scale international studies such as Trends in International Maths and Science Study (TIMSS), Programme for International Student Assessment's (PISA) technical manuals and associated literature. Initially, the plan was to collect data from all parts (provinces) of PNG. However, because of the logistical difficulties facing the researcher, two plans for data collection (Plans A & B) were proposed in this study. Plan A was to collect data in four regions of PNG, while Plan B was to collect data in the capital city of PNG, Port Moresby. To facilitate generalisability of findings, Plan A (Table 2, p. 65) was proposed to collect data in the four different regions of PNG (to collect data in schools in one or two provinces to represent the region). The four regions of PNG; namely the Highlands, the New Guinea Islands, the Momase and the Southern were selected for data collection. In those regions, four in the New Guinea Islands, five in the Momase, nine in the Southern, and twelve secondary schools in the Highlands regions were selected to represent PNG. This plan was appropriate for the research topic under study and participants were to be representative of the sample from the provinces. The rationale for this selection was to obtain a balanced representation of the sample group with respect to the proportion of schools in the regions and gender (Roiser & Roo cited in Keeves, 1992). The different sample sizes for Plan A and Plan B were proposed due to the challenges faced by the researcher in accessing the schools, as the data collection was close to Grade 10 and 12 students national examination scheduled time (August to October). Back-up schools for each region were also identified in reserve, in the event that selected schools chose to withdraw from participating. Other secondary schools in each region were not able to be included due to financial constraints and the remoteness of the schools, which limited access to participants. Due to logistical, budgetary, and time constraints, however, Plan A was not used by this study.

Plan B (Table 31, p. 65), which proposed data collection from 17 secondary schools in Port Moresby, the capital city of PNG, was deemed feasible and was ultimately carried out because of the lesser amount of logistical costs and time in data collection. The schools were selected in proportion to the type of school (private, church and government), as well as in proportion to gender in Grades 10 and 12. Three backup schools in Port Moresby, were also identified as reserve in case the selected schools withdrew from participation. The reserve schools were; Kwikila Secondary School, St Charles Luwanga Secondary and Christian Academy Secondary schools. All secondary schools that participated in this study were recognised by the National Department of Education (NDOE) in PNG.

Table 3. 1 Plan A: Summary of location, schools and participants of the study

	SCHOOLS		**SCHOOL TYPE**			**GRADE 10**			**GRADE 12 (GEN MAT)**			**GRADE 12 (ADV MAT)**			**TRS**
REGIONS	TOTAL	SAMPLE	CHRC	PRIVATE	GOVNMT	MALE	FEMALE	TOTAL	MALE	FEMALE	TOTAL	MALE	FEMALE	TOTAL	
HIGHLANDS	42	12	2	1	9	51	43	94	19	16	35	13	9	22	24
SOUTHERN	37	9	2	1	6	26	39	65	16	18	34	10	14	24	21
MOMASE	27	5	2	1	2	59	39	98	15	13	28	12	8	20	18
NGI	24	4	1	1	2	25	38	63	12	13	25	7	9	16	13
TOTAL	**130**	**30**	**7**	**4**	**19**	**161**	**159**	320	62	60	122	42	40	82	76
TOTAL PARTICIPANTS		**524 (students) + 76 (teachers) = 600**													

Table 3. 2 Plan B: Summary of schools and participants of the study

SCHOOLS			**SCHOOL TYPE**			**GRADE 10**			**GR. 12 (ADV MAT)**			**GR. 12 (GEN MAT)**			
NO		**SCHOOL CODE**	**CHURCH**	**PRIVATE**	**GOVNMT**	**MALE**	**FEMALE**	**TOTAL**	**MALE**	**FEMALE**	**TOTAL**	**MALE**	**FEMALE**	**TOTAL**	**TEACHERS**
1	Della Sale Secondary School	635	1			21		21	9		9	11		11	3
2	Don Bosco Tech Secondary School	604		1		20		20	10		10	13		15	2
3	Caritas Tech Secondary School	606		1									15	15	2
4	Marianvile Secondary School	636	1				25	25		7	7		7	7	3
5	Kopkop College	733		1		8	10	18	1	1	2	1	1	2	3
6	Jubilee Secondary School	673	1			5	7	12	5	5	10	7	5	12	3
7	Gerehu Secondary School	603			1	17	15	32	8	6	14	7	10	17	3
8	Waigani Christian Academy	483		1		10	9	19	8	4	12	7	8	15	3
9	Tokarara Secondary School	604			1	20	18	38	7	5	12	4	3	7	3
10	Gordons Secondary School	601			1	12	16	28	5	4	9	5	7	12	2
11	Kilakila Secondary School	602			1	15	15	30	12	14	26	13	12	25	3
12	Paradise Secondary School	202		1		8	6	14	7	7	14				3
13	Laloki Secondary School	603			1	13	11	24	9	7	16	8	12	20	3
14	Badiyagua Secondary School	600			1	17	17	34	14	12	26	15	13	28	2
15	St Joseph International School	775		1		9	11	20							1
16	St Charles Luwanga Secondary School	638				9	10	19	10	14	26				2
	TOTAL		3	6	6	184	170	**354**	105	86	**191**	91	93	**184**	**41**
TOTAL PARTICIPANTS			**729 (students) + 41 (teachers) = 770**												

3.4.2 Sampling selection technique

This study applied a stratified random sampling technique (Creswell, 2008; Joncas & Foy, 2011). This technique allowed the researcher to arrange and divide the population of the schools in Port Moresby into groups, or strata, which shared common characteristics. For instance, schools were arranged within specific geographic regions, school type (private, government, church) in each region, and participants' gender group. This technique ensured a balanced representation of each school and gender in the selected sample (Joncas & Foy, 2011). A stratified sampling technique can mitigate an imbalance in the characteristics of a sample, such as gender, in different schools (boys, girls, and coeducational schools) (Barreiro & Albandoz, 2001; Sandelowski, 2000). Further, this technique ensures that the stratum desired is represented in the sample in proportion to that which exists in the population (Roiser & Roos cited in Keeves, 1992). Joncas and Foy (2011) highlight that this technique improves the effectiveness of sample design and makes survey estimate more reliable. Additionally, stratified sampling ensures proportional representation of specific groups of schools in the sample (Barreiro & Albandoz, 2001; Joncas & Foy, 2011; Siniscalco & Auriat, 2012).

3.4.3 Selection of schools and participants

Prior to data collection, it was important for this study to clarify the population to be analysed. The population of the study was defined as students formally enrolled in Grades 10 and 12 at secondary schools in Port Moresby, and their teachers. This definition was possible because mathematics is a core compulsory subject studied by students of the specified cohort. After that, the researcher fully utilised plan B, which was to collect data in Port Moresby, based on Roiser and Roos' (cited in Keeves, 1992) stratified random sampling technique. The primary data collected was from 354 Grade 10 students, 375 Grade 12 students (both advanced and general mathematics), and 41 mathematics teachers from the different secondary schools in Port Moresby. Of the 41 teachers, 19 participated in the interview session. The reason for choosing only teachers for interviews is that they should be aware of how students learn in the classroom due to their regular involvement with students. The genders in Grades 10 and 12 were in proportion to males and females in Port Moresby within the selected type of schools. The reason for selecting 16 schools in each region was based purely on the amount of research work that was scheduled, the availability of the schools and financial considerations. In order to carry out the data collection procedure, class lists were obtained from each of the schools to indicate the number of students selected. In each school, sampled data were collected from an intact classroom to randomly select the respondents (Roiser & Roo cited in Keeves, 1992). This was possible through the stratification process involved at another stage of selecting the sample population, though there were challenges faced by the researcher in accessing the schools.

3.4.4 Instruments

Mathematics test items, survey questionnaires, and open-ended interviews for teacher participants were used to collect data for this study. Within these broad categories, five instruments were used, namely: a Grade 10 maths test, a Grade 12 maths test, student survey questionnaires, teacher survey questionnaires, and a teacher interview template. The survey questionnaires were designed to answer Research Question 1, ‘What are the student, teacher and school-level factors that are influencing mathematics achievements of Grade 10 and 12 mathematics students in Port Moresby, PNG?’ Besides that, the open-ended interviews were intended to answer Research Question 3, ‘What are the attributes that are deemed important by teachers regarding their teaching strategies and content knowledge?’ The survey questionnaires for students were designed to gauge attitudes and motivation towards mathematics, as well as students’ background/demographic variables. Furthermore, survey questionnaires for teachers were designed to measure quality of teaching, classroom environment and instructional resources, as well as teachers’ background/demographic variables. The items for both teachers and students were adopted and modified from the existing instruments from Trends in International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA) to suit the context of the research site. In this study, few of the PISA and TIMSS instruments are combined to obtain a comprehensive and relevant responses in PNG context for older age group (Grade 10 and 12 students). Mathematics achievements were measured through the Grade 10 and 12 mathematics test items. The open-ended interview questions were developed to obtain responses from mathematics teachers, and interviews were conducted to capture a rich amount of information and to deeply understand teachers' responses that had been provided on the survey questionnaires. Additionally, open-ended interviews supported the quantitative data obtained from the survey questionnaires, allowing more in-depth explanations and understandings of the research problems (Creswell, 2008; Creswell & Clark, 2007).

After collecting the data, four processes were employed:

1. A validation of the test items and survey questionnaires measuring the factors affecting mathematics achievements.
2. An investigation of factors and underlying structures affecting Grade 10 and 12 students' mathematics achievements at the school, teacher, and student-levels.
3. An examination of the relationships between the factors at school, teacher and student levels that may affect students' mathematics achievements.
4. The triangulation of qualitative (interview) and quantitative (survey) data at the teacher-level to confirm the findings (Sandelowski, 2000), in order to compare the perspectives of teachers on their teaching strategies and content knowledge.

3.4.5 Linking quantitative and qualitative research methods

A multilevel-mixed method design was adopted in this study (Bamberger, 2012; Harrison, Reilly, & Creswell, 2020). This design approach involves multiple levels that are used to collect data and draw conclusions about each of these levels, in order to present a clearer understanding of the overall findings of the study. This design can be employed because the two quantitative research questions; 1) 'What are the student, teacher and school-level factors that are affecting Grade 10 and 12 students' mathematics results in Port Moresby, PNG?' and 2). 'What are the relationships among student, teacher and school-level factors which affect Grades 10 and 12 students' mathematics achievement?' are combined and investigated simultaneously with the qualitative research question, 3).'What are the attributes that are deemed important by teachers regarding their mathematics teaching strategies and content knowledge?' The three research questions with their approaches can be combined in multilevel mixed method design by triangulation, as shown in the below diagram (Figure 3.1)

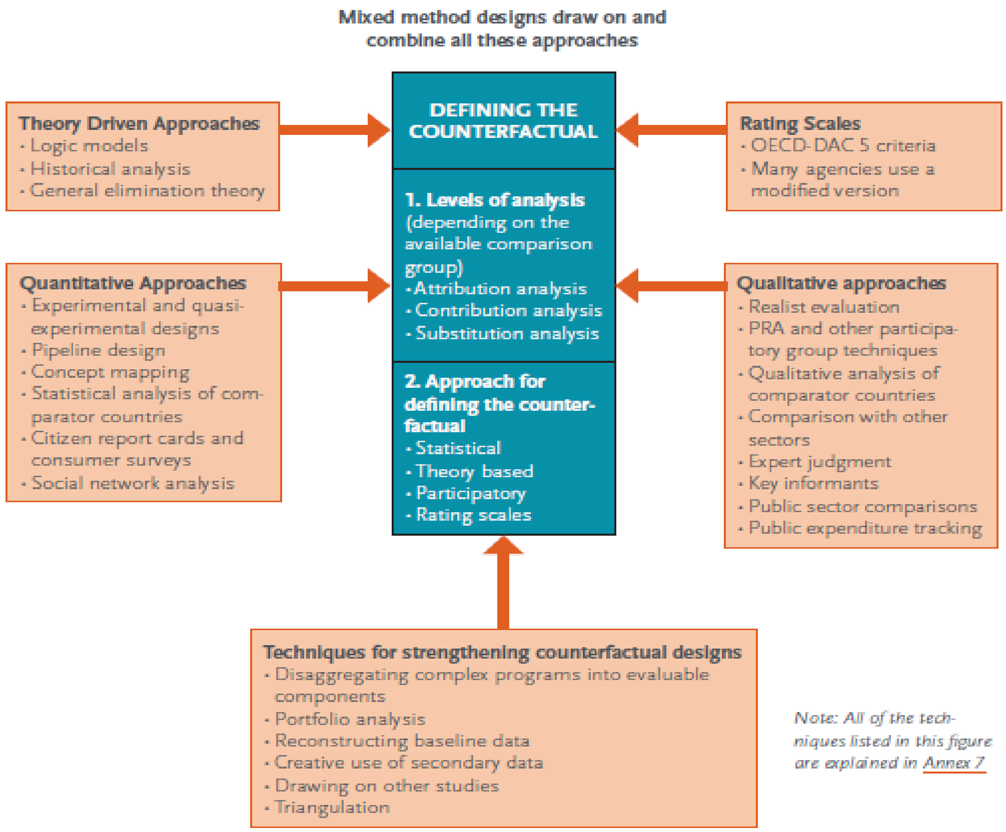


Figure 3. 1 Using mixed-method design for evaluating multilevel data

In this study, qualitative and quantitative techniques are employed at different levels of cluster (Bamberger, 2012; Harrison et al., 2020) from quantitative data sources (surveys) and qualitative data sources (interviews) to answer the research questions. A comparison of the two data sources (surveys and interviews) was carried out through triangulation (Bamberger, 2012; Sandelowski, 2000), which provides a deep understanding of the processes associated with teachers and their teaching (see Figure 3. 2).

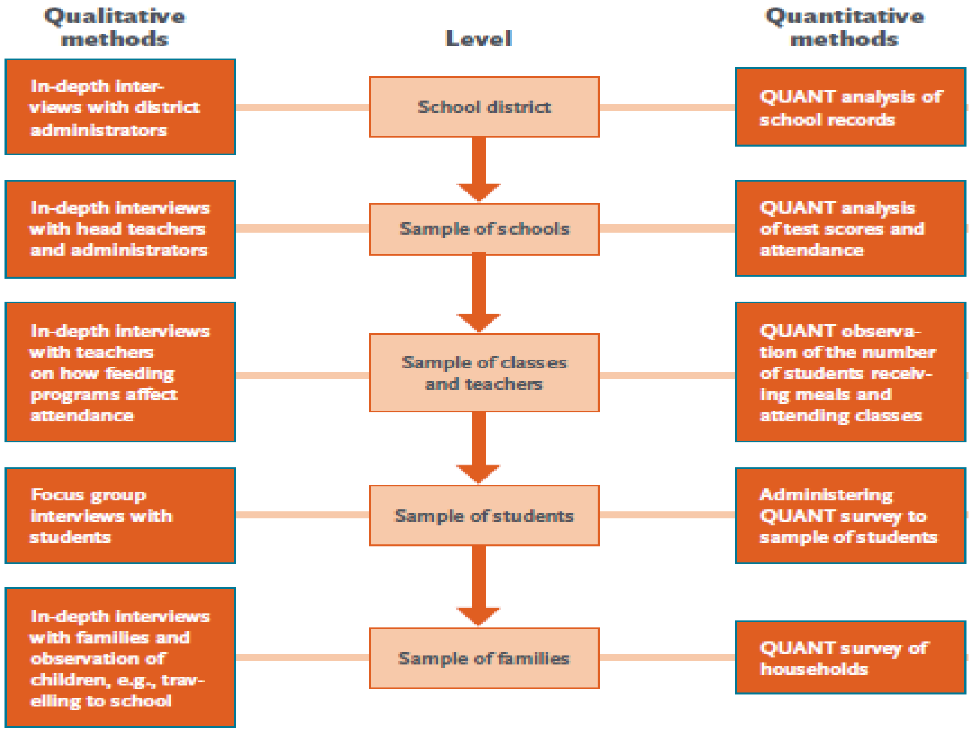


Figure 3. 2 Multilevel mixed method design (Bamberger, 2012, p.26)

The above diagram (Figure 3.2) highlights the importance of teachers' roles in student learning. The researcher compared the perspectives of teachers on their teaching strategies and content knowledge with the quality of schooling from surveys and interviews. Bamberger (2012) states that triangulation of data can bring strength to the researcher's conclusions or identify areas for further work. This is because a single method can never satisfactorily shed light on a phenomenon. Using numerous methods can therefore assist to facilitate deeper understanding (Bamberger, 2012; Harrison et al., 2020; Sandelowski, 2000).

In this study, the researcher integrated the mixed-method design for evaluating multilevel data (Figure 3.1 p.70) and multilevel mixed method design (Figure 3.2) by Bamberger (2012) to develop a diagram (Figure 3.3 p.72) that provides a clear explanation of the process of linking the quantitative and qualitative data. Figure 3.3 indicates that quantitative data were collected at school, teachers and student levels, while qualitative data were collected at the teacher level, with both sets of data then triangulated. As the researcher's hypothesis (based on experience from teaching and extant literature) is that students' mathematics achievements are affected at the teacher-level in Port Moresby, both quantitative and qualitative data sets were gathered and analysed to provide a portrait of teacher competencies. The qualitative data analysis through a thematic approach takes a perspective that is more exploratory for mathematics teachers, which alongside the quantitative data derived through statistical treatment provides balanced findings at the teacher-level.

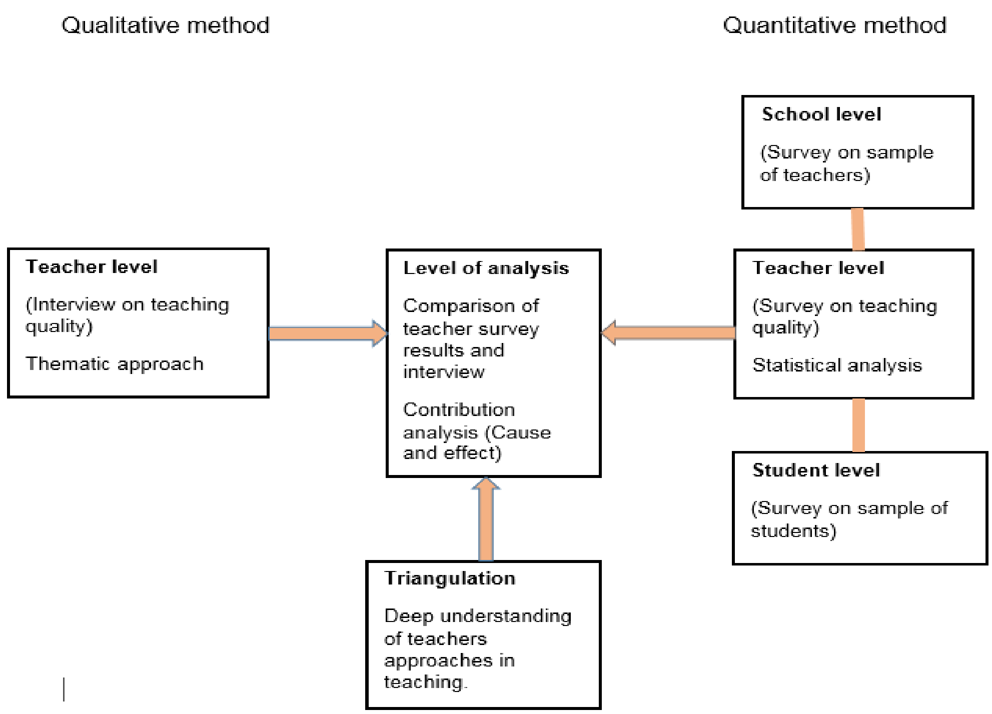


Figure 3. 3 Integrating and linking qualitative and quantitative data (Bamberger, 2012)

3.5 Survey instruments

The researcher requested approval for the use of survey instruments of TIMSS and PISA from the International Association for the Evaluation of Educational Achievement (IEA) and Organisation for

Economic Co-operation and Development (OECD), respectively. The directors granted permission for the request to use the survey instruments in this study (see Appendix A). Survey instruments in this study were therefore mostly adopted with modifications in consultation with the project supervisors. The aims, research question, and factors examined in this study were used as primary bases in adopting, modifying or developing both the survey and interview instruments. The application of the items in the research context was taken into consideration for the suitability and usefulness of the instruments. As a result, items that were insignificant were dropped in the final instruments.

Mathematics test items were developed in reference to the Grade 10 and 12 mathematics curricula in PNG. The units/topics in the syllabus were utilised by formulating a table of specification for this study (discussed in Chapters 5 and 6) according to the distribution of the domains (knowledge, comprehension and higher order) stipulated in PNG Grade 10 and 12 table of specification (see Appendix F). Grades 10 and 12 mathematics items were drawn from past examination papers of the PNG curriculum. The following sub-sections provide an overview of the development of the survey questionnaire/test items used in the study.

3.5.1 Adoption, modification and development of instruments

The scales of attitude and motivation towards mathematics, quality teaching, classroom environment and instructional resources were adopted, modified, and developed from the existing instruments, with the literature used as a guide. Adoption, modification and development of the scales was required to ensure that participants responded to the items with clarity within the time frame. Accordingly, the researcher provided a draft of the survey items to an experienced teacher in PNG to examine and make suggestions and comments based on the context of the research site. Further to this, the items were also provided to fellow researchers and faculty members from the School of Education at the University of Adelaide. This procedure ensured that survey items had a clarity in language, brevity, clear format and structure, and were applicable for student and teacher respondents. After that process, the student questionnaires were given to one Grade 10 and one Grade 12 student from PNG studying in South Australia as a trial for them to answer. This was timed in order to evaluate any difficulties that may have arisen when students in PNG were responding to the items. After their responses, the researcher increased the timing and adjusted the content of the questionnaires, accordingly.

The Grades 10 and 12 test items were adopted and developed from PNG's past standardised examination papers. Adoption and development of the items required certain steps to ensure that the participants responded to the items with clarity within the given time frame. The researcher provided a draft of the test to Associate Professor Nicholas Buchdahl from School of Mathematical Sciences (the University of Adelaide) for a content check and suggestions. After adjustments, the test items were again given as a trial to the same Grade 10 and Grade 12 students from PNG studying in South Australia. A similar procedure applied to the survey items explained above was repeated for the test items. After finalising these procedures, the study was ready to be conducted.

3.5.2 Development of interview questions

The interview questions were developed following basic guidelines, similar to the construction of the survey instruments (Archibald, 2016; Bamberger, 2012; Creswell, 2008). Each question was constructed with reference to the topic and the purpose of the study. Specifically, the questions were constructed to fill gaps and to compare respondents' views in relation to the survey questionnaires. Accordingly, the brevity and clarity of the instrument was prioritised. Biased and negative wordings that may have influenced teachers' responses were avoided. The questions developed were then discussed with the project's supervisory team, and further assessed by three experienced teachers from School of Education at the University of

Adelaide**.** Feedback from these researchers related to designing the questions on quality of teaching aspects adapted from the teaching quality model in NSW schools (discussed in Chapter 2), and how the researcher would engage teachers to truthfully express their feelings towards teaching mathematics. Consideration of validity and reliability were paramount for interview questions for the teacher participants. In-depth interpretation of the quantitative data obtained from the questionnaires was further enhanced and obtained through the development of interview questions (Archibald, 2016; Creswell, 2008).

3.5.3 Final instrument of the main study

After all relevant modification and revisions, the instruments used in the study were named '*Student Mathematics Examination', 'Student Questionnaire'* and *'Teacher Questionnaire'*. The Student Questionnaire contained three sections. The first section, Section A contained six general information items about demographics (see Appendix B). Next, Section B consisted of 20 items, relating to attitude towards mathematics learning while Section C contained 20 items related to motivation towards mathematics. The Mathematics Examination consisted of 40 items each for Grade 10 and Grade 12 students, respectively (see Appendix C). Both advanced and general Grade 12 mathematics students attempted the same items, because some topics that they study are similar. For instance, as quadratic equation is a topic taught to both groups, was included in the examination. This was also convenient for the researcher and the schools in terms of scheduling the research. The examination items were developed from a Table of Specification (TOS) that highlighted the similar topics that were studied at both levels (see Appendix F). The Teacher Questionnaire consisted of four sections (see Appendix B). Section A covered five general information items, Section B consisted of 15 items on quality teaching, Section C covered 13 items on classroom environment, and Section D consisted of 12 items on school curriculum.

Student questionnaires, examination items and teacher questionnaires were prepared separately. Cover letters were attached to each questionnaire to explain the purpose and the importance of the study to the participants, and institutions and departments were contacted to ensure ethics and appropriate administration of the instruments.

3.5.4 Data administration and collection

After obtaining the necessary approval/permissions from the NDOE, the provincial education secretary and the participating schools' principals, the survey was deployed to the respondents of Port Moresby, PNG. The survey was conducted between 25 July 2017 and 6 October 2017. There is no significant difference on the amount of mathematics lessons the students learnt. Since the data was collected towards the end, closer to their national examination at the same time/months. The following subsections describe the processes of the administration of survey and the collection of the data.

3.5.5 Administrating the instruments

A total of sixteen secondary schools (refer to Table 3.2 p. 64) that included public, private and agency (Catholic) schools agreed to participate in this study. Three schools (Port Moresby Grammar School, Port Moresby National High School and Port Moresby International School) declined to participate in the study, based on scheduling and other unspecified reasons. These three schools were consequently replaced by the reserve schools (Kwikila Secondary School, Christian Academy and St Charles Luwanga Secondary School) following the replacement process used by Keeves, 1992. The principals / deputy principals and mathematics department heads were contacted to arrange convenient times to conduct the survey through permission letters (see Appendix E). Though scheduling for schools visitations to administer instruments/tests was organised systematically, with one school per day or two schools per day to avoid conflicts, this was not always possible in practice, with some schools rescheduling the survey a few times, and others changing the schedule upon the researcher's arrival to conduct the survey. This situation delayed the deployment of the survey.

Participation complaint forms, information sheets and consent forms were distributed to the selected Grade 10 and 12 students and their mathematics teachers. Prior to surveys and examinations, instructions were given to both students and teachers verbally, as well as being provided in written form. Students undertook the mathematics examination for one hour under the researcher's supervision. After that, they were given 30 minutes to response to the survey questionnaire providing general information, and response to attitude and motivation towards learning mathematics sections, with the assistance of their mathematics teachers. In some schools, mathematics teachers agreed for the researcher to conduct the mathematics examination and survey during their double mathematics period (80 minutes) in a single day, while others opted for two days. However, in some schools mathematics teachers used time with permission from other subject teachers such as creative arts to conduct the survey and examination. The researcher administered the test with the survey and collected the examination answer sheets along with the survey questionnaires in the classroom. The survey questionnaires were not allowed to be taken out of the classroom, and so students were provided sufficient time to complete all the questionnaire questions. Despite this, three schools (Don Bosco Technical School, St Charles Luwanga Secondary School and Tokarara Secondary School) shortened the survey period from 40 minutes to 30 minutes, and some students consequently did not complete all survey items. The researcher was present during every survey to answer any queries and doubts pertinent to the instrument, and to ensure that efficient collection of the questionnaires was carried out.

Teacher surveys were distributed with the questionnaire, collecting general information and information about teacher quality and classroom environment. Teachers were asked to respond to the items at a convenient place in their non-contact time at school, and return it at the end of the school day. Many teachers completed the survey questionnaire and gave it to the researcher towards the end of the day, while others were collected the next day. Two teachers did not return the questionnaires at all, despite allowances made for them to complete the questionnaire and return it. In reaching the respondents, the researcher used public transport and walked between schools.

3.5.6 Administrating the interviews

Mathematics teachers of Grade 10 and 12 students were selected according to the classes that participated in mathematics examinations and responded to survey questionnaires. Teachers were purposefully selected with a mixture of experience, from expert and novice, to ensure that a balance of views and opinions was received (Creswell, 2002; Creswell & Clark, 2007). Interviews were scheduled during teachers' non-contact periods. Once the participant appointments were made, interview questions were provided to teachers, in order to obtain as honest and detailed answers as possible. Interviews were then conducted in the 16 schools, with 21 teacher participants. Participants were informed that the whole interview was audio-recorded. Prior to the interview, the researcher explained to participants the purpose, importance, and confidentiality of the interview (Creswell, 2008). After that, the researcher began to ask questions using the interview protocol. Consideration of questioning strategies was sought and observed to elicit the appropriate responses. Questions were rephrased and examples were highlighted relating to scenarios for the participants to more easily understand the questions. At the end of the interview, the researcher thanked the participants, and assurance of the confidentiality of their responses was emphasised.

3.6 Statistical procedures and methods employed in the study

This section, presents a discussion of the different statistical procedures and criteria used to validate, verify and calibrate (the process used to ensure accuracy) the instruments (survey questionnaires and test items) used in the study. These procedures are employed in order to obtain useful and meaningful results from the analysis for the factors under examination.

3.6.1 Confirmatory factor analysis

This study is centred on the constructs (latent variables) that represent the factors investigated for Grade 10 and 12 mathematics achievements. Constructs used in this study such as attitude, motivation, quality of teaching, classroom environment and instructional resources are latent/unobserved variables. Hence, confirmatory factor analysis (CFA) is conducted to examine the construct validity of these scales. The underlying latent variables/scales are represented by a number of observed variables. MacCallum (1995) suggests three approaches of conducting CFA: the strictly confirmatory approach, the model generation approach and the model comparison approach. In the strictly confirmatory approach, the model is built and examined to observe whether the data fits the model well (Lye, 2016; Marsh, Guo, Dicke, Parker, & Craven, 2020; Schaaps, 2019). A plausible model is accepted if the model fits the data; while if the model does not fit the data well, it is not accepted, and further analysis is not employed (Brown, 2014; Schaaps, 2019). For the model generation approach, a model is specified and evaluated, and the modification indices are used to improve the fit of the model: while for “the model comparison approach, a number of alternative a priori models are specified to fit the same set of data.”(Stewart, 2001, p.76 cited in Ben, 2010 p.81). Each of these alternative models specifies the theoretical assumptions regarding the possible relationships proposed in the literature review. This means the researcher has already hypothesised that the observed variables have a relationship with the latent constructs that are under study. These alternative models of CFA are evaluated and compared to see which model fits the data better based on the fit indices results for meaningful interpretation (Chawla & Saxena, 2016; Marsh et al., 2020). In this study, the model comparison approach is used to discover a model that is more consistent with the data (Curtis, 2005). This is because the strictly confirmatory approach and model generation approaches have weaknesses. "The strictly confirmatory approach is overly rigid and model generation approaches modification made to the model may not be meaningful and substantively justifiable" (Lye, 2016, p. 146).

According to Lye (2016, p. 147) the scales and constructs can be structured into four different models: single factor model, orthogonal factor model, correlated factor model, and hierarchical factor model. First, the single factor model examines to what extent the observed variables load into the single factor. Second, the orthogonal factor model presumes that, they are uncorrelated, while third, the correlated factor model assumes that factors of the constructs are loaded on correlated factors. Finally, the hierarchical model suggests that there are first-and second-order factors in the model. The first-order factors are loaded onto a single second-order factor in the model (Darmawan, 2003). In this study, these four different models are employed. However, among the four different models examined, only the best model that fit the data was selected for Rasch analysis.

3.6.2 Structural equation modelling

Structural Equation Modelling (SEM) is a comprehensive statistical approach that enables the estimation of multiple and interrelated dependence relationships simultaneously (Martynova, West, & Liu, 2017; Ramlall, 2016; Sivakumar, 2014). According to Darmawan (2003, p. 82), SEM is useful when "one dependent variable becomes an independent variable in subsequent dependence relationships". As such, SEM ensures that the sample data support the theoretical model, and allows researchers to be attentive to the use of numerous observed variables (Ramlall, 2016; Schumacker & Lomax, 2004). Hence, the main reasons for using this statistical analytical approach are: a) to examine the measurement model of the construct's validity, and b) to examine the complex and multiple relationships between the independent and dependent variables (.i.e. structural model).

Determination of the measurement model's fit with the data is assessed by various fit indices. The *measurement model* examines the relationships between the latent variables and observable variables (Hailaya, 2014; Khine, 2013; Wang & Wang, 2019). This can be indicated by the magnitude of the factor loading (outcome from factor analysis). Several studies state that there is no particular cut-off rule for choosing the appropriate values of the factor loading (Barrett, 2007; Cangur & Ercan, 2015; Diamantopoulos, Siguaw, & Siguaw, 2000; Hooper, Coughlan, & Mullen, 2008; Linacre, 2002), because the meaning of the factor loading magnitude varies by research context and the different sample sizes used. In practice, values greater than +0.5 or -0.5 are applicable (Cangur & Ercan, 2015). However, in this study the minimum acceptable factor loading value is set at +0.4 (Cangur & Ercan, 2015; Linacre, 2002), though researchers have suggested that a value of 0.3 can be used for multiple-choice test and right and wrong responses or dichotomous data (Hailaya, 2014; Hooper et al., 2008; Linacre, 2002). The factor loading establishes validity and reliability, which are crucial in the measurement model. This can allow the structural model to influence the relationships among the latent variables because the measurement model is acceptable (Ben, 2010; Hailaya, 2014; Khine, 2013; Wang & Wang, 2019).

The *structural model* is the relationship between the latent variables. Determination of the structural model's fit with the data is assessed by various fit indices (Barrett, 2007; Schumacker & Lomax, 2004). This means that model fit indices ensure that a hypothesised model is consistent with the data (Barrett, 2007; Khine, 2013; Martynova et al., 2017). The purpose of assessing model fit is to decide the level to which the model is consistent with the data collected (Diamantopoulos et al., 2000; Hooper et al., 2008; Pal Pandi, Paranitharan, & Jeyathilagar, 2018). Moreover, "these indices assess how well the sample covariances (positive or negative direction) are reproduced by the covariances predicted from the parameter estimates" (Hailaya, 2014, p. 79). Researchers have identified several model fit indices, which have been used in previous studies (Barrett, 2007; Cangur & Ercan, 2015; Hooper et al., 2008; Schumacker & Lomax, 2004). However, in this study only the applicable analysis indices are reported. These indices include chi-square statistics (x^2), ratio of chi-square to its degree of freedom, standard root mean square residual (SRMR), root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker Lewis index (TLI). Each of these indices are discussed below.

The chi-square (x^2) is an index that explains precise and accurate fit and assesses the ideal model to fit the empirical data (Cangur & Ercan, 2015; Cheung & Rensvold, 2002; Diamantopoulos et al., 2000). However, x^2 index fit is not used because it is sensitive to sample size, instead X^2/DF is employed in this study (Cheung & Rensvold, 2002; Hooper et al., 2008; Cangur & Ercan, 2015; Diamantopoulos et al., 2000). The RMSEA is one of the most informative criteria in SEM because it takes into account the error of approximation in the population (Byrne,2010; Cangur & Ercan, 2015; Hooper et al., 2008). The SRMR is an "index of the average of standardised residuals between the observed and the hypothesised covariance (direction) matrices" (Cangur & Ercan, 2015, p. 156). SRMR is easy to interpret due to its standardised nature, and it is relatively sample size independent (Cangur & Ercan, 2015). CFI is used because it is one of the most popularly reported fit indices and it is least influenced by sample (Diamantopoulos et al., 2000; Kenny & McCoach, 2003). The CFI index compares the performance of the model with the performance of the baseline (null or independence) model (Cangur & Ercan, 2015; Cheung & Rensvold, 2002; Hooper et al., 2008). Due to the baseline model, it assumes that there is zero correlation between all observed variables (Barrett, 2007; Cheung & Rensvold, 2002).

The X^2/DF is the chi-square divided by the degree of freedom (Diamantopoulos et al., 2000; Hailaya, 2014). The model fits the data well with the smaller the value of this index (X^2/DF). There is no clear-cut guideline as to which value is the maximum for a model to be accepted (Kline, 2014). An RMSEA value less than 0.05 indicates a good fit (Diamantopoulos et al., 2000; Schumacker & Lomax, 2004) and values between 0.05 and under 0.08 are of reasonable fit (Cangur & Ercan, 2015; Kenny & McCoach, 2003; Schumacker & Lomax, 2004), while values between 0.08 and 0.10 indicate fit that is neither good nor bad, and values greater than 0.10 indicate poor fit (Barrett, 2007; Hooper et al., 2008; Schumacker & Lomax, 2004). In this study, a RMSEA value under 0.06 is used for model fit (Cangur & Ercan, 2015; Diamantopoulos et al., 2000; Schumacker & Lomax, 2004). The value of SRMR ranges from 0 to 1.0 (Hooper et al., 2008; Schumacker & Lomax, 2004), with the acceptable fit being values smaller than 0.10, and an indicator of a good fit being values less than 0.05 (Cangur & Ercan, 2015; Hooper et al., 2008; Schumacker & Lomax, 2004). The perfect fit value in SRMR is 0 (Hooper et al., 2008). SRMR values less than 0.05 are understood in this study to be an acceptable model fit. CFI values between 0.8-0.9 are indicators of good fit for that index (Cangur & Ercan, 2015; Kenny & McCoach, 2003), with a value larger than 0.95 considered to be a good fit (Cangur & Ercan, 2015; Cheung & Rensvold, 2002; Hooper et al., 2008). The TLI is another incremental fit index that is similar to CFI, except that it is not normed. Similar to the CFI index, a higher TLI value indicates a better model (Byrne, 1998). In this study, a model with a TLI value of close of 0.90 (or close to 0.90) is considered acceptable.

Table 3. 3 Fit indices for examining the model fit of the student and teacher-level models

Indices	Acceptable cut-off scores for model fit
X^2/DF	The smaller the better
CFI	Close to or 0.90 for acceptance
TLI	Close to or 0.90 for acceptance
RMSEA	0.05-0.08 or below: not more than 0.10 is still acceptable
SRMR	Values less than 0.08 is acceptable

There are five sequential steps that characterise the application of SEM (Ben, 2010; Diamantopoulos et al., 2000; Ramlall, 2016). The first step is the *model specification*. This step permits the researcher to use equations or graphical forms (e.g., building a path diagram of casual relations) to demonstrate the theoretical model (Ramlall, 2016). *Identification* is the second step. The objective of this step is for the researcher to use the observed data to possibly determine or estimate the model. This could assist the researcher to obtain special sets of parameters of the model specified (Ben, 2010; Khine, 2013; Ramlall, 2016). The third step is *estimation*. This step uses complex methods such as covariance (direction) and correlation (strength) to estimate the proposed model, through computing the parameters with estimation measures such as multiple regression (Khine, 2013; Ramlall, 2016). The fourth step is *testing* the model fit for consistency with the data through checking the indices that guarantee good model fit. The final step is *re-specification* of the model if the researcher intends to improve the model. Otherwise, if the model fit is sound, the results are interpreted (Diamantopoulos et al., 2000; Khine, 2013; Ramlall, 2016). In this study, these steps in SEM serve to guide the researcher in validating and confirming the theoretical model.

Finally, SEM analysis is employed to gauge an overview of the data gathered (Guo, Zhang, Zhao, Wang, & Xi, 2018; Khine, 2013), of how the factors interact at student, teacher and school levels and how they might affect Grade 10 and 12 students' mathematics achievements. SEM analysis is used after confirming results with CFA and validating and re-confirming the variables with the partial credit model (Khine, 2013; Ramlall, 2016). The latent variables that have the best fit during CFA are examined using Rasch analysis.

3.6.3 Rasch analysis

A CFA examination of the factor structure of a set of items can reveal evidence of the underlying dimensions of the items. However, Keeves and Masters (1999) argue that conducting CFA analysis alone is not sufficient to examine the extent of the structures of a scale. Hence, the data can be further examined through Rasch analysis for person and item fit. In this study, Rasch analysis is used to verify that the scope of the structures of the scales confirmed in the CFA analysis fit the Rasch model. Another reason for using Rasch analysis is for scoring purposes in subsequent analysis, such as structural equation modelling and hierarchical linear modelling.

The Rasch model is based on the Item Response Theory (IRT) and is referred to as one-parameter IRT. Three different models can be identified through IRT (Alagumalai & Curtis, 2005; Reise, 2000). The three-parameter model assumes that the participants' abilities are affected by three characteristics of an item, which are its difficulty, discrimination and pseudo-guessing (Alagumalai & Curtis, 2005; Reise, 2000). In the two-parameter model, pseudo-guessing is removed from the three-parameter model, and only difficulty and discrimination are present (Alagumalai & Curtis, 2005). Then, after removing item discrimination from the two-parameter model, the one-parameter model remains (Alagumalai & Curtis, 2005; Hambleton & Swaminathan, 2013; Reise, 2000). The one-parameter model is a Rasch model and is used in this study because it satisfies the basic assumption of measurement compared to two and three parameter models (Alagumalai & Curtis, 2005; Alagumalai, Curtis, & Hungi, 2005; Reise, 2000). As indicated by Alagumailai and Curtis (2005), the one-parameter model is used in a strict and absolute sense to model the data. Rasch models such as the Rating Scale Model (RSM), the Partial Credit Model (PCM), and the multidimensional model are used to establish the measurement properties of the scales. Alagumalai, Curtis and Hungi (2005) highlight that using Rasch models improves measurement properties and can further develop interval scales, equate tests, detect item bias, bring items and persons to a common scale and produce estimate errors for each individual person and items instead of the instrument as a whole. These estimates are independent for the individuals responding to the items, as well as for the items included in the survey questionnaire and test items.

A strict requirement of unidimensionality for a variable in Rasch model needs to be observed in the analysis. In constructing a variable, the many different elements and properties of the variable are identified and are plotted into a single number line (Andrich, 1988; Andrich & Marais, 2019). The construction of a unidimensional variable is carried out through this process. Andrich and Marais (2019) further state that once the unidimensional variable is constructed, the difference can be compared to better understand the measurement properties. This means that the estimates of parameters will be meaningful when each scale and test items of the survey contributes to measure a single underlying trait (Bond & Fox, 2015).

Moreover, the Rasch model was used to investigate the psychometric properties of the constructs/test for validation (Bond & Fox, 2015; Boone & Scantlebury, 2006; Yao, Bull, Khng, & Rahim, 2018). This analysis highlights the test-takers' abilities and item difficulty with responses (Bond, Fox, & Lacey, 2007), which can be measured on a common measurement scale (Bond & Fox, 2015). This means that the estimation of respondent's ability or item difficulty is possible by converting raw scores to logits (the unit of measurement that results when the Rasch model is used to transform raw scores obtained from ordinal data to log odds ratios on the common interval scale) on an objective measurement scale (Bond et al., 2007). For this reason, using the Rasch model, it is possible to sum up all the marks of each item in the test, or sum up all the codes for each item in a survey (Bond et al., 2007; Korbel, 2015). Bond et al. (2007) emphasise that this allows comparison of person ability and item difficulty estimates and provides important opportunities for inferences to be drawn from the test. In addition, Bond et al. (2007) state that Rasch modelling of an individual's performance on a test can be interpreted with respect to how likely it is that the person would successfully demonstrate a skill or an understanding of a concept as assessed by a particular item, in addition to how well the person performed compared to his or her peers.

In summary, the Rasch model plays a significant role in the process of instrument development, by ensuring content and construct validity of the instrument (Bond & Fox, 2015; Boone & Scantlebury, 2006). The Rasch Model is considered a rigorous approach for constructing scales that conform with the principles of objective measurement (Bond & Fox, 2013; Bond et al., 2007). Therefore, the usefulness of Rasch model has been widely recognized in developing objective measurements to obtain valid data through the assessment of item fit statistics criteria.

3.6.3.1 Item fit statistics

In this study, residual-based fit statistics (values that represent the difference between Rasch model theoretical expectations and actual performance) are used to judge and accept the test items (Bond & Fox, 2015; Gómez, Arias, Verdugo, & Navas, 2012). Fit statistics focus on two aspects of fit namely, item infit and outfit (Tejada, Luque, Rojas, & Moreno, 2011; Yao et al., 2018).The infit and outfit statistics are mean square (MNSQ) and standardised mean square (ZEMP). Hence, infit is an information-weighted fit statistic, and, is more responsive to unexpected behaviour influencing responses to items near to the respondent's ability (Bond et al., 2007; Tejada et al., 2011; Wu, Tam, & Jen, 2016). On the other hand, outfit is an outlier-sensitive fit statistic, more responsive to unexpected behaviour influencing responses to items away from the respondent's ability (Bond et al., 2007; Tejada et al., 2011). Further, infit statistics provide comparatively more weight to the performance of the respondent closer to the item value (Bond & Fox, 2015; Tejada et al., 2011). By contrast, outfit statistics are subject to the influence of outlying scores because they are not weighted. Finally, both infit and outfit statistics are robust to large sample size variations (Bond et al., 2007; Wu & Adams, 2007). In this study, infit statistics are employed because of their advantage over outfit statitics, with weight to the performance of the test taker.

Bond and Fox (2015) state that the unstandardised form of fit statistics is reported as mean square (MNQS) and the standardised form is the t statistic. This MNSQ is the mean or average value of the squared residuals of the item. In this study, the acceptable values of the MNSQ are placed in the interval between 0.7 and 1.30, where 1 is the ideal for mathematics test items (Wu et al., 2016). There are no hard rules for cut-off scores, though Skrodal (2010) suggests an infit MNSQ range of 0.6-1.40 as reasonable for data collected from surveys. The different MNSQ ranges are shown in Table 3.4. In ZEMP (t value), 0 means that the model satisfactorily envisaged the observed data, and an interval between -2 and 2 specify acceptable fit (Bond et al., 2007; Wu et al., 2016).

Table 3. 4 Reasonable item MNSQ ranges

Type of test	Range
Multiple-choice test (high stakes)	0.80-1.20
Multiple-choice test (run of the mill)	0.70-1.30
Rating-scale (Likert /survey)	0.60-1.40
Clinical observation	0.50-1.70
Judged (where agreement is encouraged)	0.40-1.20

In this study, items that abide to the criteria and regulations mentioned above and satisfy the measurement attributes and fit indicators/items are preserved, while those that fail to conform to the requirements are deleted/removed, as they violate the measurement criteria (Boone & Scantlebury, 2006; Oliveira, Fernandes, & Sisto, 2014; Wu et al., 2016).

3.6.3.2 Rating scale model, Partial credit model and Multidimensional model

The two polytomous models, namely the rating scale model (RSM) (Andrich, 1978) and the partial credit model (PCM) (Masters, 1982), are used in the Rasch measurement model (RMM). In the RSM, response categories are scored in a way that the total score for all items constitutes a rating of the respondents on a latent scale (Abou El-Komboz, Zeileis, & Strobl, 2014; Andersen, 1997; Oliveira et al., 2014). "The RSM can be used for ratings of two or more categories and when the interval between each response category on a particular scale is to remain the same for all items in an instrument such as the questionnaire" (Andrich, 1999 cited in Lye, 2016 p.157), ensuring that the highest score is the first or the last category (Andersen, 1997; Gómez et al., 2012; Oliveira et al., 2014; Wang & Wu, 2011). Thus, the phrasing of the response categories must reflect a scaling of the responses: e.g., *""always," "sometimes," "rarely," "never."* Andersen (1997, p. 67) argues that "the main assumption for the rating scale model, apart from being a polytomous Rasch model, is that scoring of the response categories must be equidistant, i.e., their values must increase by a constant". This argument is supported by Andrich (1978), who also specifies this condition. On the other hand, with PCM the interval between each response category on a particular scale is different (Eggert & Bögeholz, 2010, p. 235). The concept behind this model is that partial success on item responses are expressed by partial credits and are hierarchical (Eggert & Bögeholz, 2010; Gomez, 2008; Wu et al., 2016). This implies that a response that is given a higher partial credit is better than a response that is assigned a lower partial credit. Hence, in this study PCM was employed to analyse the items for all the scale response categories because in RSM the units between each response category may be unequal although the categories are generally ordered (Blais et al.,2011).

The multidimensional model is useful to examine scales that contain a hierarchical structure (Paek & Cai, 2014; Reckase, 2009). Wu et al. (2016) indicate that, in the past, the analysis of scales that contain several subscales was conducted by either applying a unidimensional model to each separately or by ignoring the multidimensionality and treating the scales as unidimensional. Wu et al. (2016) argue that both of these methods have limitations, and thus a joint, multidimensional calibration is preferred. It is important to note the multidimensional models in this study have not violated the requirements of unidimensionailty (Paek & Cai, 2014; Wu et al., 2016). In fact, in multidimensional analysis, each subscale is unidimensional (Paek &

Cai, 2014; Reckase, 2009). As noted by Reckase (2009) the multidimensional test enables the examination of several subscales in which each measure related but distinct latent dimensions.

3.6.4 Differential item functioning

Differential item functioning (DIF) analysis was conducted to confirm the fairness and equity of the test items and survey questionnaires due to the different gender groups surveyed (male and female). DIF analysis was utilised because it allows the researcher to determine if individual items or groups of items function differently for specified groups by looking at both the item level difference and the group difference on the items (Abou El-Komboz et al., 2014; Bond & Fox, 2015; Wu et al., 2016). The respondents with the same ability level (both males and females) may have a different probability of a correct response due to items having different difficulties. This provides evidence of different probabilities or likelihoods of success on an item when matched on the ability or interest of different groups (Bansilal, 2015; Bond & Fox, 2015; Hagquist & Andrich, 2015). Therefore, "in order to preserve the unidimensionality trait of the construct under measurement, an important aspect of Rasch analysis was the examination of the presence of DIF in the various items" (Bansilal, 2015, p. 6). This procedure allows the researcher to determine if certain items were problematic when item bias is suspected, or if there is a response difference to certain types of items for gender subgroups. In this study, DIF analysis was employed to investigate whether individual items have similar psychometric properties among males and females who are from the same population (Abou El-Komboz et al., 2014; Guilera, Gómez-Benito, & Hidalgo, 2010; Hagquist & Andrich, 2015). The rationale for this is to confirm the item bias if different test-takers with equal ability who are from different subgroups do not have the same chance of success on an item (Brodersen et al., 2007).

DIF analysis investigates the quality of a measurement instrument (Bond & Fox, 2015; Bond et al., 2007; Wu et al., 2016). The principle underlying the detection of DIF is to investigate whether there is invariance in the person and item plots across different test situations (Bond et al., 2007; Brodersen et al., 2007; Wu et al., 2016). The procedure for detecting DIF items is necessary in the process of examining the fairness of a test and identifying the problematic items for revision or elimination before the administration of the test (Abou El-Komboz et al., 2014; Hagquist & Andrich, 2015; Wu et al., 2016). If item difficulties of a group are plotted against another group, the slope that describes the invariance of the item difficulties should be equal to one (Abou El-Komboz et al., 2014; Bond & Fox, 2015; Wu et al., 2016). Three methods of identifying DIF in the RMM were carried out in this study: the item threshold approach, item fit approach, and group plot methods to detect measurement invariance. These methods are now discussed in turn.

3.6.4.1 Item threshold approach

The item threshold approach is the first approach used to identify DIF in this study, and a common procedure used for evaluating DIF within the context of the RMM (Abd-El-Fattah, Al-Sinani, El Shourbagi, & Fakhroo, 2014; Hungi, 2005; Wu et al., 2016). This approach focuses on the difference between the threshold values (difficulty levels) of item in sub-groups. If the difference in the item threshold values is noticeably large, this implies that the item is particularly difficult for members of one of the groups being compared (Abd-El-Fattah et al., 2014; Hagquist & Andrich, 2015). This is not because of their different levels of the underlying latent trait, but due to other factors probably related to being members of that group (Abd-El-Fattah et al., 2014; Le, 2006; Strobl, Kopf, & Zeileis, 2015). With the item threshold approach, an item found to be more difficult for a group than the other items in a test is considered biased against that group (Abd-El-Fattah et al., 2014; Strobl et al., 2015). Those items with the largest differences in scale value are the suspect items.

A biased item can be detected by the difference of the threshold values (difficulty levels) of the items in the two groups. This is because the difficulty of an item (d parameter) is estimated separately for each group

(Hungi, 2005; Meade & Fetzer, 2009; Scheuneman & Bleistein, 1999). Scheuneman and Bleistein (1999, p. 231) highlight that the difference in item difficulty between groups can be calculated with t statistics with the given formula:

$$t_i = \frac{(d_1 - d_2)}{\sqrt{SE_{i1}^2 + SE_{i2}^2}}$$

where: SE represents the standard error of d.

d_1 item difficulty value for group 1

d_2 item difficulty value for group 2

This formula can predict whether DIF is present when the t-value is large (Scheuneman & Bleistein, 1999). For example, in Table 5.6 p.135 in Chapter 5 all t-values or the standardised difference in item threshold are calculated following this calculation for item 01:

$$t_i = \frac{(-0.107 - 0.107)}{\sqrt{0.12^2 + 0.12^2}} = \frac{-0.214}{\sqrt{0.0288}} = \frac{-0.214}{0.169} = -1.27$$

Further, if there is a noticeable difference in the t-value, that particular item is recognized as being more difficult for a certain group than another group (Hungi, 2005; Scheuneman & Bleistein, 1999). As mentioned earlier, this is not due to the groups' different performance levels, but is a result of other factors related to the different features of the group members (Andrich & Hagquist, 2015; Hungi, 2005; Le, 2006). Both the unexpectedly difficult items and unexpectedly easy items for a specific group can be identified with this process. Bond and Fox (2015) point out that, in a high-stakes test, if item threshold difference is greater than 0.5 logits, this can be used as a criterion for detecting the DIF. This view is supported by Hungi (2005), who stresses that the absolute value of the difference of ± 0.5 logits indicates the difficulty difference of two sub-groups of one-year of school learning. The formula shown below is utilised.

$$-0.05 < \mathrm{d}_1 - \mathrm{d}_2 < 0.50$$

where: d_1= the item's difficulty value in group 1, and

d_2 = the item's difficulty value in group 2

Hungi (2005,p.146) states that the specified range value between -2.00 to +2.00 is the acceptable standardized difference of item difficulty: "Items whose differences in standardised item threshold between any of the groups fall outside a predetermined range do not confirm to the model and can be identified biased items". The formula below illustrates this scenario:

$$-2 < \mathrm{std}(\mathrm{d}_{1-}\mathrm{d}_2) < 2.00$$

where: std = standardized difference of item difficulty

d_1= the item's difficulty value in group 1, and

d_2 = the item's difficulty value in group 2

The standardised difference in item threshold is important to test the premise that the difference in difficulty between males and females is statistically significant. This process is carried out through checking a set of criteria of a range of values of the standardised difference in item threshold among genders, as specified by Adams and Khoo (1993; cited in Hungi, 2005).

Furthermore, an item can be flagged as a DIF item when the absolute difference of item threshold for an item is greater than 0.25 logits (Abd-El-Fattah et al., 2014; Le, 2006). This absolute difference value is equal to approximately half of the school year to learn a distinct content area. However, when the difference in the item threshold value of an item between male and female respondents is outside the ±0.25 logit range, such a difference may cause significant concern related to item DIF (Abd-El-Fattah et al., 2014; Le, 2006).

3.6.4.2 Item fit approach

The item fit approach is the second approach in the Rasch model used to detect DIF. This approach investigates whether the items have equal discrimination power, allowing the infit mean square value for all items within the acceptable range (Hungi, 2005; Scheuneman & Bleistein, 1999). However, if the INFIT MNSQ values of the items are outside the range, the items can be assumed to be biased. This demonstrates that the items cannot equally discriminate in all different sub-groups (Bond et al., 2007; Hungi, 2005). In other words, non-biased items would fit the model in each group (Scheuneman & Bleistein, 1999).The acceptable range of item fit is between 0.77 to 1.30 for test items, and 0.60 to1.40 for Likert scales (Bond & Fox, 2015; Hagquist & Andrich, 2015; Hungi, 2005). This INFIT MNSQ range is useful to identify whether all items are satisfactorily fitting to the models when comparing subgroups (Hungi, 2005).

3.6.4.3 Group plot methods

The group plot method is the final approach for Rasch model, and it compares the area of the item characteristic curves (ICC) of difference in sub-groups. According to Alagumalai et al. (2005) and Wu et al. (2016), for the ICC, the gradient/slope of the ICC is positive when the probability (p) is 0.5. This is because p=0.5 describes the latent trait for item location for the specific group. The comparison of the area of the ICC of difference in two groups is equivalent to the difference between the item difficulties (Andrich & Hagquist, 2015; Hagquist & Andrich, 2015; Wu et al., 2016). This is because item difficulty is only one parameter of RMM. It can be concluded that the results of the group plot method will be similar to those of the item threshold approach discussed earlier. Results of the group plot method use graphs to demonstrate the difference between the two groups.

3.6.5 Scoring

This study uses Weighted Likelihood Estimate (WLE) to score respondents' answers to the test items and survey questionnaires. The rationale for using WLE over other transformation methods such as maximum likelihood estimate (MLE) is its advantage in minimising the estimate bias (Linacre, 2009; Wu et al., 2016). The WLE is presented in the form of Logits, and reflects respondents' ability levels in choosing a certain category in the test items and survey questionnaires (Magis & Verhelst, 2017; Warm, 1989). In theory, respondents who choose a more difficult category should obtain a higher average WLE than those who do not choose the category (Magis & Verhelst, 2017; Warm, 1989; Woodcock, 1999; Wu et al., 2016). The ability estimate of the respondents for the questionnaires for each scale and test items are generated through Rasch analysis using the ACER ConQuest 4.0 software. The WLE values are transformed into W-scores to eliminate negative values, and the need for decimal points (Benson, Beaujean, Donohue, & Ward, 2016; Linacre, 2009; Woodcock, 1999). Microsoft Excel is used for this transformation using the formula W=9.1024 x (WLE logits) + c, where c is a constant term, selected to eliminate the negative values (Benson et al., 2016; Woodcock, 1999). The formula [W=9.1024 x (WLE logits) + c] is adopted from international studies such as the Programme for International Assessment (PISA) and is used in this study. The constant term used in this study is 500, which is also consistent with PISA. The W-scores are used in subsequent chapters for structural equation modelling and hierarchical linear modelling.

3.6.6 Hierarchical linear modelling (HLM)

HLM is a statistical technique used in this study to analyse hierarchical, nested or cluster data at the school, teacher and student-levels to be consistent with other studies (Polly et al., 2018; Zlatkin-Troitschanskaia et al., 2016). This technique simultaneously investigates relationships within and between hierarchical levels of grouped data levels (Kanyongo & Ayieko, 2017; Singh, 2016; Woltman, Feldstain, MacKay, & Rocchi, 2012). HLM is used because it is more efficient in estimating for variance among variables at different levels. It also provides direct effects from various levels and cross-interaction effects between variables at different levels (Garson, 2013a; Mertens, Pugliese, & Recker, 2017; Polly et al., 2018). The HLM approach assumes only one dependent variable at the individual level of the hierarchy to be predicted by a number of independent variables at different levels (Liou & Jessie Ho, 2018; Polly et al., 2018; Thien, Razak, Keeves, & Darmawan, 2016).

One of this study's research questions is: *What school, teacher and student factors influence Grade 10 and 12 students' mathematics results in Port Moresby, PNG?* This question involves a hierarchy with three levels (discussed in Chapter 1) to clarify the idea of HLM. At the highest level of the hierarchy, the macro level (Level 3), are school related variables, such as instructional resources and school type (Darmawan & Keeves, 2009). Situated at the middle level of the hierarchy, the meso level (Level 2), are classroom variables, such as quality of teaching (e.g., teachers' teaching pedagogy and content knowledge), classroom environment, and years of teaching experience (Woltman et al., 2012; Zlatkin-Troitschanskaia et al., 2016). Level 2 variables are nested within Level 3 groups and are impacted by Level 3 variables (Zlatkin-Troitschanskaia et al., 2016). For example, schools (Level 3) that are in remote geographic locations (Level 3 variable) had smaller class sizes (Level 2) than classes in metropolitan areas. This can affect the quality of personal attention paid to each student, and behaviour in the classroom (Level 2 variables). The student-level variables such as attitude and motivation towards mathematics are situated at the micro level (Level 1) (Zlatkin-Troitschanskaia et al., 2016). Variables at the lowest level of the hierarchy (Level 1) are nested within Level 2 groups and share in common the impact of Level 2 variables (Shen, 2016; Shoraka, Arnold, Kim, Salinitri, & Kromrey, 2016). In summary, students (Level 1) are situated within classrooms (Level 2) that are located within schools (Level 3) (see Table 3.5 for factors under each level). The outcome variable of mathematics achievement was measured at the micro level (Level 1). This was because in HLM, the outcome variable of interest is always situated at the lowest level of the hierarchy.

Table 3. 5 Factors at each hierarchical level that affect Grade 10 and 12 students' mathematics results

Hierarchical level	Example of hierarchical level	Example variables
Level 3	School Level	Instructional resources, School type
Level 2	Classroom Level	Quality of teaching, Classroom environment
Level 1	Students Level	Attitude towards mathematics, Motivation towards mathematics

Aldous, Darmawan and Keeves, (2009) indicate that HLM 6.0 can analysis two and three level models (as shown in Table 3.5). The choice of an appropriate model and the variables to be examined at the two or three levels chosen requires consideration of the situation under investigation and the independent variables likely to influence the criterion and nesting (Aldous, Darmawan & Keeves, 2009). Misspecification of the variables included in an analysis of HLM can seriously distort the results obtained (Garson, 2013; Woltman et al., 2012). Therefore, prior to conducting an HLM analysis, background interaction effects between predictor variables should be accounted for, and sufficient amounts of within-and between-level variance at all levels of the hierarchy should be ensured (Mohammadpour & Shekarchizadeh, 2015; Shen, 2016; Woltman et al., 2012).

3.6.7 Qualitative data

Qualitative data were obtained through teacher interviews. These data were used to support/complement quantitative data findings at the teacher level to gauge how the factors reviewed in Chapter 2 affect Grade 10 and 12 students' mathematics results. In addition, the interview analysis can capture information that might have been missed in teacher surveys about their teaching strategies and content knowledge, and can facilitate an in-depth exploration of the attributes and factors at the teacher-level that are affecting Grade 10 and 12 students' mathematics results. Using a single data collection technique such as the survey questionnaires used in this study (a quantitative approach), alone can never adequately shed light on a phenomenon (Bamberger, 2012; Cresswell, 2008; Onwuegbuzie & Combs, 2011; Sandelowski, 2000).

Supplementing this method by using interviews (a qualitative approach) can therefore assist in providing a deeper understanding of the factors affecting mathematics results for Grade 10 and 12 students. Moreover, the triangulation of the two methods (quantitative and qualitative approaches) facilities validation of data through cross verification of two data sources (survey and interview) (Bamberger, 2012; Onwuegbuzie & Combs, 2006; Sandelowski, 2000). The consistency of the findings are tested through both the survey and interview and increases chances to control and assess some of the factors that are influencing Grade 10 and 12 students' mathematics results (Bamberger, 2012; Onwuegbuzie & Combs, 2006; Sandelowski, 2000). The triangulation process balances out each method and provides a richer and better understanding of the results (Bamberger, 2012; Creswell & Creswell, 2017; Onwuegbuzie & Combs, 2011).

A thematic approach is employed in the qualitative section of this study (Chapter 12). This approach is used because it pinpoints, examines and records patterns (themes) within data. Themes are patterns across data sets that become the categories for analysis, and are important in describing a phenomenon associated with a specific research question. "Thematic analysis can be an essentialist or realist method, which reports experiences, meanings and the reality of participants" (Braun & Clarke, 2006, p. 81). Therefore, thematic analysis works both to reflect reality and to unpick or unravel the surface of 'reality'. Transcriptions of interviews were analysed using the NVivo analysis software.

3.6.8 Validity and reliability

In a quantitative-dominant research project such as this study, issues of validity and reliability are important. This is because "reliability and validity are the scale construction counterparts of precision and accuracy for research" (Dawis, 2000 p.87). A valid scale demonstrates that items derived from a certain concept really measure the concept, and a reliable scale is free from random error, therefore issues of validity and reliability are of importance (Cordier et al., 2018; Dawis, 2000; 2016; Mertens, Pugliese, & Recker, 2016; Mohamad, Sulaiman, Sern, & Salleh, 2015; Yasar & Cogenli, 2014).

Validity is the most important criterion in examining the effectiveness of the instruments (Cordier et al., 2018; Mohajan, 2017; Mohamad et al., 2015; Yasar & Cogenli, 2014). In this study, face, content and construct validity were carried out as shown in Figure 3.4 (p.98) to validate the instruments. Face validity was perform by experts at different stages of instrument development to determine the suitability of the instrument. For content validity, the instrument was checked by experts to ensure that it measured what it was intended to be measured. The construct validity of the instruments was carried out through CFA to examine the underlying structure of the scale, while Rasch analysis was used to verify if the items from the scale structures scale fit the Rasch models.

Reliability is another important criterion used to examine the consistency of an measure of the instrument (Creswell & Clark, 2007; Mohajan, 2017; Mohamad et al., 2015; Yasin, Yunus, Rus, Ahmad, & Rahim, 2015). Two test for assessing the reliability of a scale that are often used are: test-retest reliability and internal consistency. Test-retest reliability requires using the measure on a group of respondents at one time, then using it again on the same group of respondents later. After that, the researcher examines the test-retest correlation between the two sets of scores (Enkavi et al., 2019; Leppink & Pérez-Fuster, 2017). The higher the test-retest reliability the more reliable the scale. The internal consistency measure is based on the correlations between different items on the same scale (or the same subscale on a larger scale), and measures whether several items that propose to measure the same general construct produce similar scores. Cronbach’s coefficient alpha is the most commonly used statistical test to measure internal consistency (Kim & Abraham, 2017; Leppink & Pérez-Fuster, 2017). A higher Cronbach’s alpha indicates a more reliable scale, with values ranging from 0 to 1.

In this study, the reliability of the instrument was examined through two types of reliability test provided by Rasch analysis: item separation reliability and person separation reliability (Boone, Staver, & Yale, 2014). Person separation reliability indicates the scope of the separation caused by the instrument on the persons included in the data (Boone et al., 2014; Wright & Stone, 1999). According to Lye (2016, p. 169) item separation reliability refers to the "replicability of item location on the logits scale if these items are given to another group of people with similar size and who behave in the same way as this sample". High person or item values indicate less errors or precise estimates of the person or item (Wright & Stone, 1999; Wu et al., 2016).

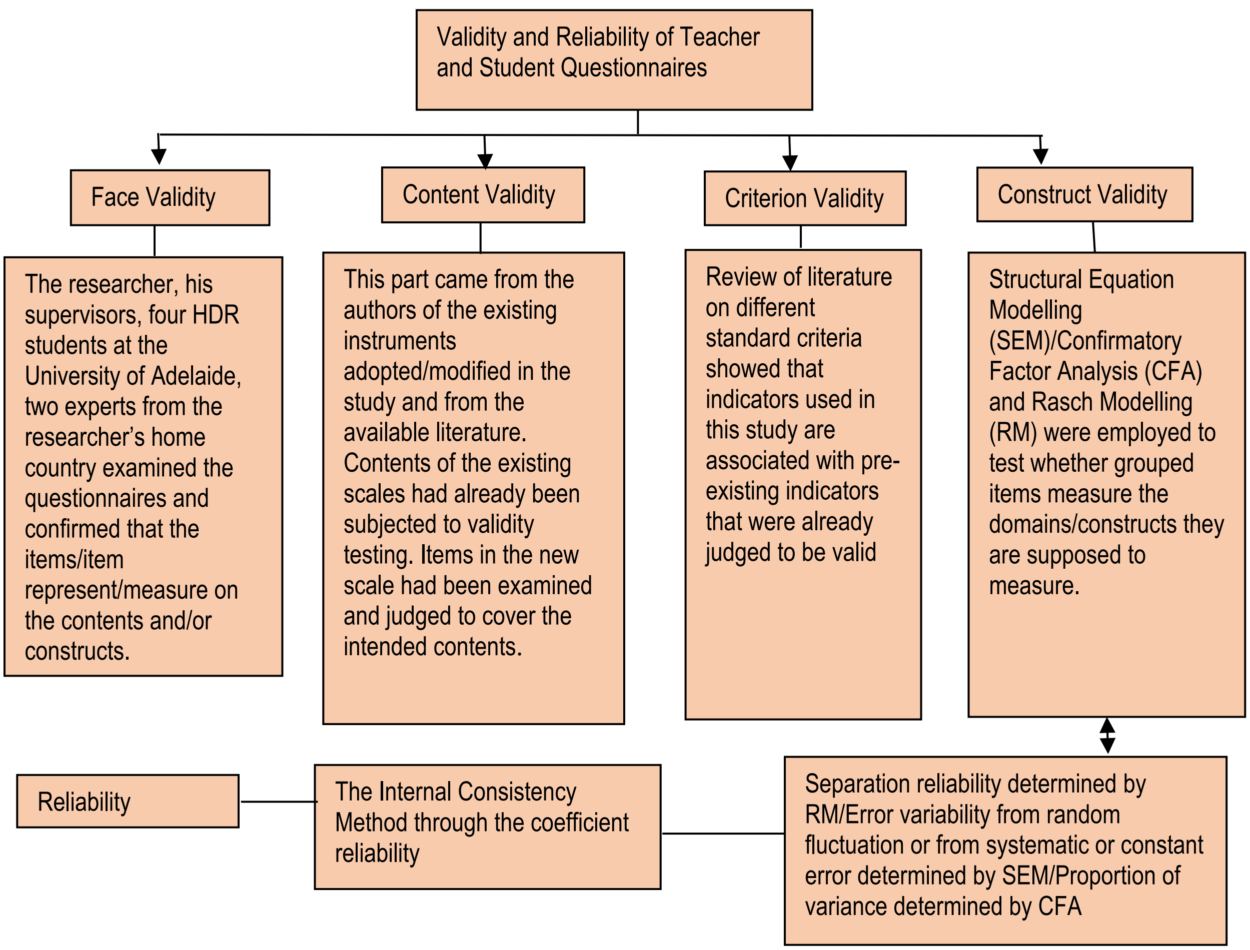


Figure 3. 4 Validation of scales used in the study (adopted from Ben, 2010)

3.6.9 General methodological considerations

This section of the thesis discusses the general statistical methods employed to organise and clean the data before the actual analysis using the different statistical software packages.

3.6.9.1 Aggregation and disaggregation of data

Survey data were collected at the student, teacher and school-levels, producing information at these three different levels. Two methods, aggregation and disaggregation of data were used to explain the differences in the three-levels; through both of these methods may produce incorrect estimates of the effects among the constructs.

The aggregation of data combines data at a lower (student) level to perform calculations to obtain data at higher (teacher) level units (Pourghebleh & Navimipour, 2017; Williamson, 2017). Consequently, in this process the dependent variable can lose information, which may lead to bias in the estimates of standard error of means and regression weights (Pourghebleh & Navimipour, 2017; Williamson, 2017). Aggregation can cause four errors: a) it can prevent examination of cross-level variables in the model; b) when the original data structure is neglected, there is the possibility of inappropriate tests of significance being employed; c) there can be a change in meaning when aggregating variables from higher levels and not directly to the lower level units; and d) correlation between higher-level variables cannot be used to make claims about lower- level relations, as this could introduce ecological fallacies.

Disaggregation on the other hand is the reverse process i.e., the breaking up of aggregate data to lower level details. In other words, when data is disaggregated, a few data values from a small number of higher-levels are loosened into many more values for many more lower-level units (Budhwani & De, 2017; Nguyen et al., 2017). As a result, the assumption of the independence of observation fails to apply, as the proper sample size for these data is the number of high-level units (Nguyen et al., 2017). This could lead to researchers finding many significant results that are completely false. Variations between levels can result in incorrect estimates of the effects, and further lead to a more serious risk of providing inappropriate and incorrect conclusions (Budhwani & De, 2017). Therefore, in this study a hierarchical linear modeling (HLM) approach is employed in order to account for the differences in the levels.

3.6.9.2 Preparation of data

The entering and processing of survey data obtained in Port Moresby started immediately after data was collected in each school, and ended when data collection finished in October, 2017. Microsoft Excel was used to enter and organise student and teacher data in a raw form. Teacher interviews were transcribed and typed into a Microsoft Word document at the end of November 2017. Raw Excel data for the Grade 10 and 12 examination results and survey items was imported into IBM SPSS 25 for descriptive statistics analysis and other statistical analysis. Data in the SPSS file format was then converted to a tab delimited (dat) file format to carry out CFA, Rasch analysis and SEM. After that, the teacher and student-level demographic variables in the SPSS file format were classified into an appropriate numerical form for use in the statistical software packages employed in the study. For instance, numbers were assigned for demographic variables such as gender (i.e., male 0 and female 1) and schools (i.e., private 0 and catholic 1). The checking of data for errors was conducted prior to further analysis, using the above-discussed screening and cleaning procedures. Any error detected through this process was corrected accordingly. This process was repeated until all data were free from errors.

3.6.9.3 Missing values

Any research, project needs to contend with the problem of missing data, and it is, therefore, important to consider how to deal with missing values in the data analysis phase. The three techniques used to deal with missing values in this study include listwise deletion, pairwise deletion and the imputation method. These three techniques each have advantages and disadvantages. In the listwise deletion technique, any individual in a data set is deleted from analysis if they are missing data on any variable in the analysis (Howell, 2019; Nassiri, Lovik, Molenberghs, & Verbeke, 2018). One of the disadvantages of this technique is that a large amount of missing data (that leads to a loss of sample size) can provide problems for multivariate data analysis (Nassiri et al., 2018). Pairwise deletion occurs when the statistical procedure uses cases that contain some missing data (Howell, 2019; Nassiri et al., 2018; Shi, Lee, Fairchild, & Maydeu-Olivares, 2020). This procedure does not include a particular variable when it has a missing value, but still does include it when it contains non-missing values (Howell, 2019; Shi et al., 2020). This technique allows the use of more data. Finally, missing values can also be solved by employing the imputation methods (Shi et al., 2020). These methods are carried out using mean or regression substitutions to get a complete set of data. Mean substitutions may reduce the variance of the distribution and distort the distribution of data (Rosenthal, 2017; Xiao, Bruner, Dai, Guo, & Hanlon, 2019). However, they are best used when the missing data are relatively low and the relationships among the variables are relatively strong.

In this study, missing data at student and teacher-levels were less than 1% for an individual case or observation. None of the variables were missing more than 1% of data. However, the missing values were replaced with “99” for analysis purposes in the SPSS file. The Rasch model has the advantage of analysing and calculating the estimates of ability and difficulty values based on the available data scores even when missing values of the variables exist in the data set.

3.7 Summary

This chapter discussed how this study evolved, how data was collected, and the procedures that were used to analyse and interpret data. The initial planning of this study evolved from the researcher's experience and observations. Since the study has adopted a quantitative-dominant, mixed-method approach, qualitative data is employed to support and fill in gaps missed in quantitative data analysis. As such, this chapter also discussed how the quantitative and qualitative data of this study were linked through triangulation at the teacher-level. The participants of the study were Grade 10 and 12 students in Port Moresby, selected through a stratified sampling method for data collection methods, including as a mathematics test, surveys, and interviews. The instruments used were modified and adopted from IEA and OECD with the directors' permission, to develop teacher and student questionnaires on constructs such as attitude, motivation, quality of teaching and classroom environment. Statistical procedures, including the Rasch model and confirmatory factor analysis, which are used to validate the items/tests in subsequent chapters, were also discussed, as well as the structural equation model (SEM) and hierarchical linear model (HLM). A structural model was used to test the relationships among the constructs in the study that are predicted by the measurement model (observed variables). Finally, HLM was utilised to determine the nested nature of the data from the school, teacher, and student levels, in order to evaluate the effects, that they may have on students' mathematics results. Three specialist software programs, ACER ConQuest 4.0, Mplus 7.0 and HLM 6.08 are used in the analysis. The next chapter examines the demographic information of participants and discusses the descriptive statistics and frequency distributions of the data.

Chapter 4: Teacher and student demographic information

4.1 Introduction

In the previous three chapters, discussions were focused on how the three-level (school, teacher and student) factors influence Grade 10 and 12 students' mathematics results. However, beginning from this chapter onwards, the school-level data have been disaggregated with teacher-level data, and discussions will only be focused on the teacher and student-levels. The reason for the teacher and student-levels serving as this study's theoretical framework, instead of the three-level model (as outlined in Chapter 2), is that even though school type appears as one of the categorical variables; six of the 16 schools had fewer than three teacher participants. This makes it impossible for the HLM 6.08 software to estimate the coefficient for the subsequent chapter, hierarchical linear modelling (see Chapter 11), as the software can only analyse three or more participants in each school.

In Chapter 3, the methods employed in the study were described and discussed. In this chapter, the demographic information/variables of the teacher and student respondents are discussed. The demographic variables examined at the teacher-level are gender, academic qualifications, school type, years of teaching and major/minor area subject specialisation. Student-level factors including gender, type of mathematics, parents' education level and employment are also examined for clarity purposes. This approach is based on the theoretical model that has been developed from the insights of previous studies (see Chapter 2). It is important to generate descriptive information from the dataset to provide a profile of the samples according to the demographic factors/variables examined in this study before carrying out subsequent analysis. This step can provide a complete picture of the data for each of the factors, allowing for proper interpretation of the relevant results.

This chapter describes the sample in terms of the distribution of the following: teacher and student gender, the academic qualifications of teachers, their years of teaching experience, school type, and students' parents highest education level and employment.

4.2 The sample: descriptive information

In this section, descriptive information of the teacher sample is discussed.

4.2.1 Teacher gender

In this study, gender is one of the factors investigated, and it is important to present this distinction in the sample. Table 4.1 shows the distribution of the teachers by gender.

Table 4. 1 Distribution of teacher respondents by gender

Gender	Frequency	Percent
Male	21	52.1%
Female	20	48.8%
Total	41	100%

Table 4.1 shows that there are 21 male and 20 female mathematics teachers in the sample. A graphical representation of Table 4.1 is also provided in Figure 4.1 to provide a visual summary of the teacher sample distribution.

A questionnaire was administered to the mathematics teachers discussed above to obtain the data required for this study, and the raw data became part of the study. These data contain teachers' demographic information and data for each of the scales in the questionnaire intended for teacher participants.

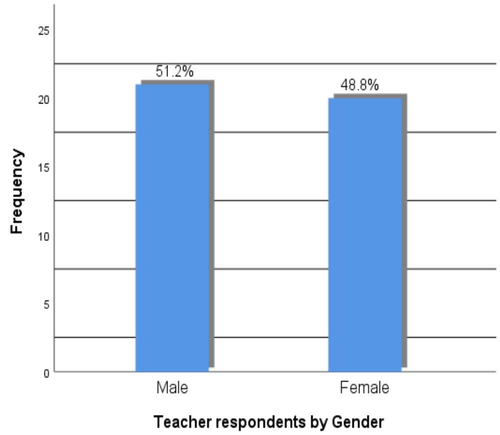


Figure 4. 1 Distribution of teacher respondents by gender

4.2.2 Academic qualification

Academic qualification of teachers is one of the important factors considered in this study that might have an influence on students' mathematics results. In the PNG context, to teach Grade 10 and 12 mathematics teachers must finish an undergraduate degree in education focusing on mathematics. This is reflected in Table 4.2, which shows that a relatively high frequency of 30 teachers (more than 73%) have a bachelor's degree, and less than 30% have a diploma. Teachers have the option to complete a postgraduate degree such as Master's or PhD, but for unclear reasons this rarely occurs. The results of the level of education of the teachers who participated in the study are shown in Table 4.2.

Table 4. 2 Distribution of teacher academic qualification

Highest Academic Qualification	**Frequency**	**Percent**
Bachelor's Degree	30	73.2 %
Diploma	11	26.8 %
Postgraduate	0	0 %
Total	41	100%

For clarity, this information is represented on a graph in Figure 4.2.

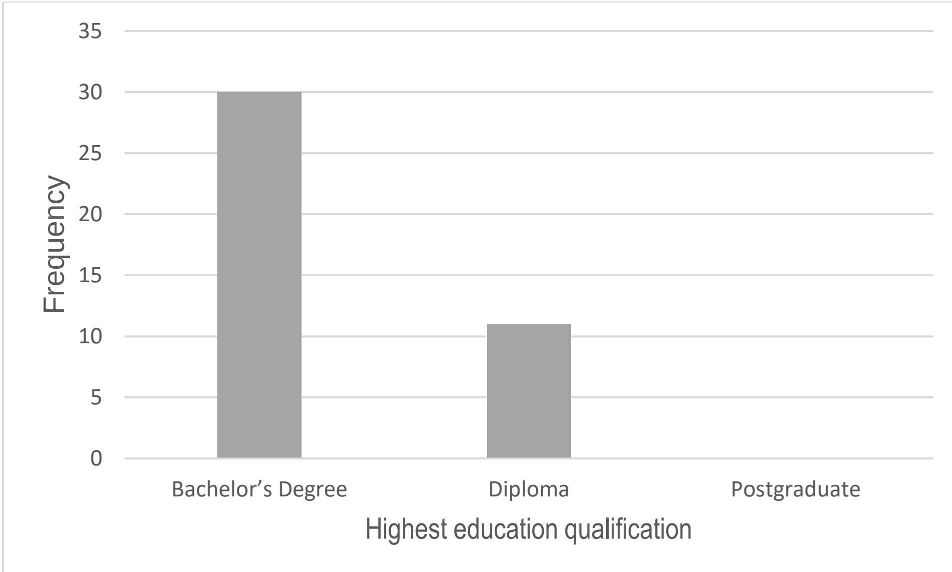


Figure 4. 2 Distribution of teacher academic qualification

4.2.3 Years of teaching experience

The number of years teachers have been teaching is examined according to its impact on students' mathematics results. Teachers' teaching experience is set at 10-year increments, because the teachers' responses on the questionnaire item had a wide range. The teachers' responses are tabulated in Table 4.3.

Table 4. 3 Years of teaching experience

Years of experience	Frequency	Percentage
1-10yrs	24	58.5%
11-20 yrs	3	7.3%
21-30yrs	11	26.8%
Over 30 yrs	3	7.3%
Total	41	100%

It is evident from Table 4.3 that 1-10 years teaching experience had the highest frequency of 24 teachers (58.5%), followed by 11 teachers (26.8%) with 21-30 years teaching experience as the second highest frequency. Three teachers (7.3%) each had 11-20 years and over 30 years' teaching experience, respectively. The bar graph shown in Figure 4.3 clearly shows the distribution of teachers based on their length of teaching experience.

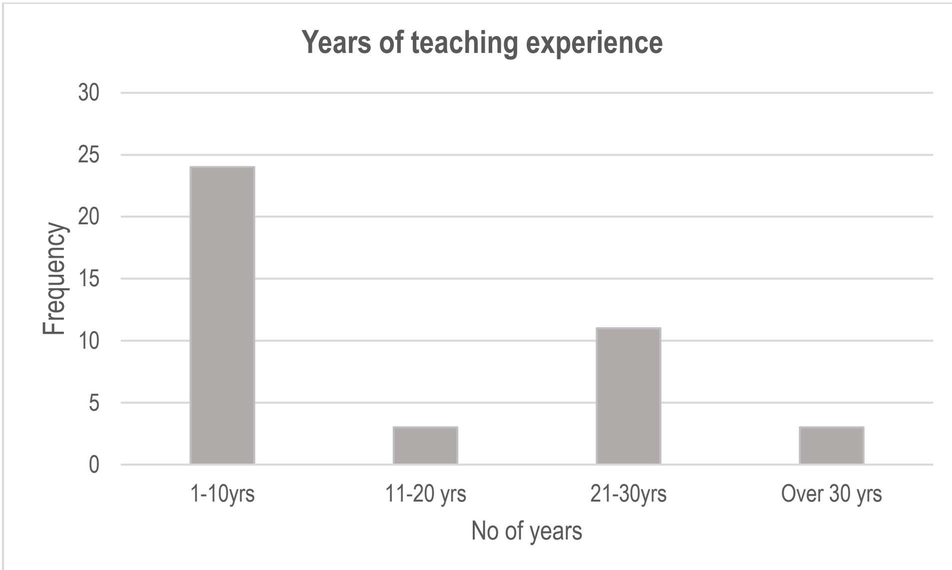


Figure 4. 3 Distribution of years of teaching experience

4.2.4 Major area of specialisation

Table 4.4 shows that teachers with a mathematics major had a relatively high frequency, with 18 teachers (43.9%) having mathematics as their specialised area of teaching. The second-highest major was science, with a relative frequency of 13 teachers (31.7%). About 22% of teachers majored in technology, while 2.4% majored in social science. This information is also reflected in the bar graph of Figure 4.4.

Table 4. 4 Major area of specialisation

Major area of specialization	Frequency	Percent
Mathematics	18	43.9%
Science	13	31.7%
Technology	9	21.9%
Social Science	1	2.4%
Total	41	100.0%

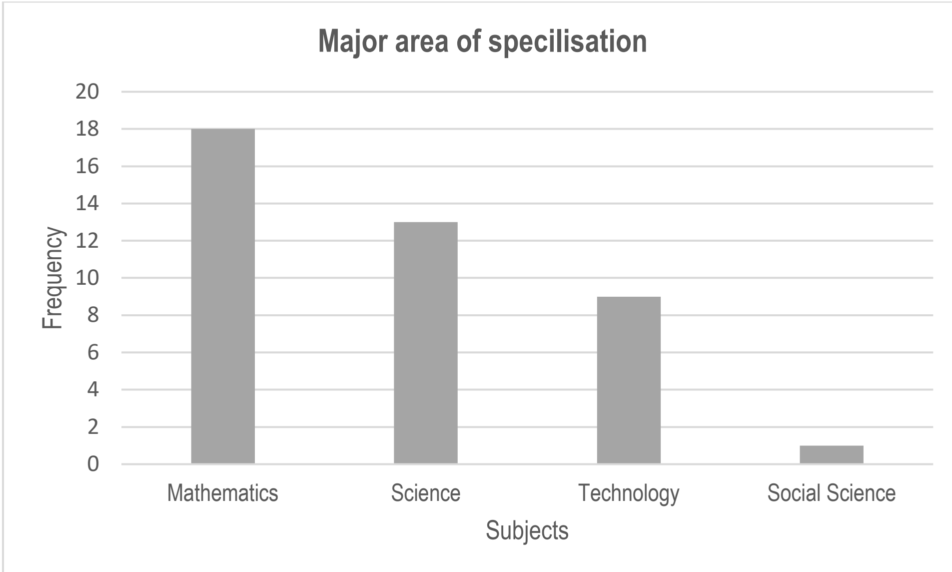


Figure 4. 4 Distribution of major area of specialisation

4.2.5 Minor area of specialisation

Table 4.5 shows the teachers' minor areas of specialisation. The table shows that a relatively high frequency of 14 teachers (34.1%) had mathematics as their minor specialised area during their university studies. Technology as the minor subject had the second-highest frequency of 10 teachers (24.4%). The reminder of teachers had minor specialisation in science (21.9%) and social science (14.6%). This information is further demonstrated in the bar graph in Figure 4.5.

Table 4. 5 Minor area of specialisation

Minor area of specialisation	**Frequency**	**Percent**
Math	14	34.1%
Science	9	21.9%
Technology	10	24.4%
Social Science	6	14.6%
Total	41	100%

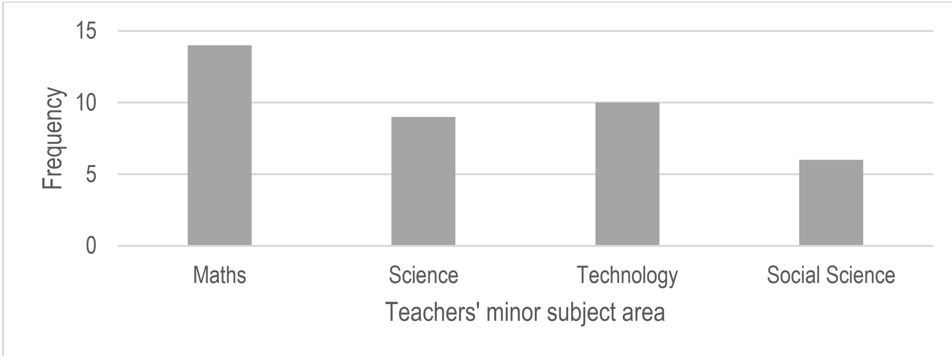


Figure 4. 5 Distribution of minor area of specialization

4.2.6 School type

School type as a school-level variable at the secondary schools in Port Moresby is discussed in this section as it was disaggregated into the teacher-level, as discussed above. As shown in Table 4.6, the types of schools included in the study are public, private and Catholic schools. Public schools are administered by the PNG government, with teachers paid by the Teaching Service Commission (TSC) of the National Department of Education (NDOE). Catholic schools are administered by different Catholic religious groups, with their teachers also paid by the TSC. Their students' tuition and boarding fees are also subsidized by the government. Private schools are owned by individuals, on-government organizations or churches. These schools pay their own teachers directly, with parents paying students' full tuition fees. Despite these differences, all schools implement the same mathematics curriculum developed by the PNG NDOE, and the same national mathematics examination at the end of Grades 10 and 12.

A shown in Table 4.6, the majority of the teachers participants (more than 43%) come from Catholic schools, more than 37% of respondents taught in public schools; and the remainder of about 18% of participants are from private schools in Port Moresby. A graph (Figure 4.6) also represents this information visually.

Table 4. 6 School type in Port Moresby

School Type	Frequency	Percent
Public	6	37.5%
Private	3	18.8%
Catholic	7	43.7%
Total	16	100%

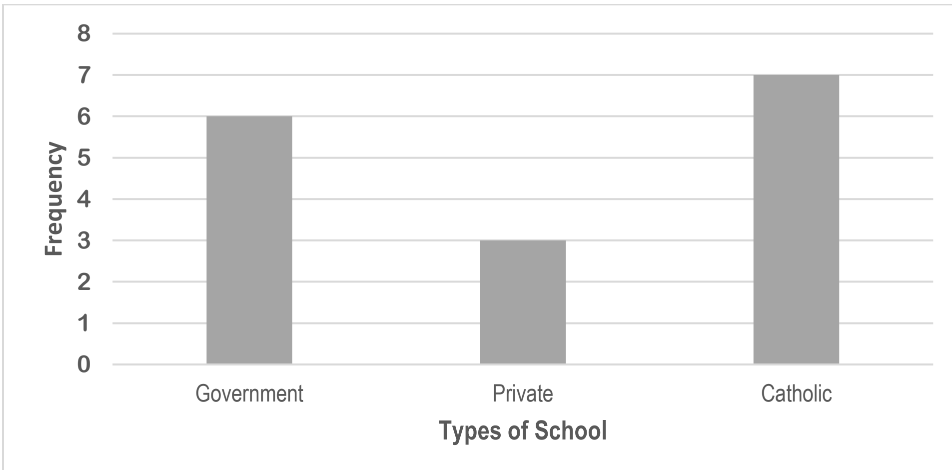


Figure 4. 6 Distribution of school type in Port Moresby

4.3 Student gender

Students' gender is one of the factors examined in this study and it is important to present the distribution of this factor. It can be observed in Table 4.7 that there is a relatively high frequency of 385 male (52.8 %) and 344 (47.2%) female respondents. There are about 5.6% more male than female participants because more male students are enrolled in mathematics at the different schools. A graphical representation of these data is provided in Figure 4.7

Table 4. 7 Student respondents by gender

Student Gender	Frequency	Percent
Male	385	52.8%
Female	344	47.2%
Total	729	100%

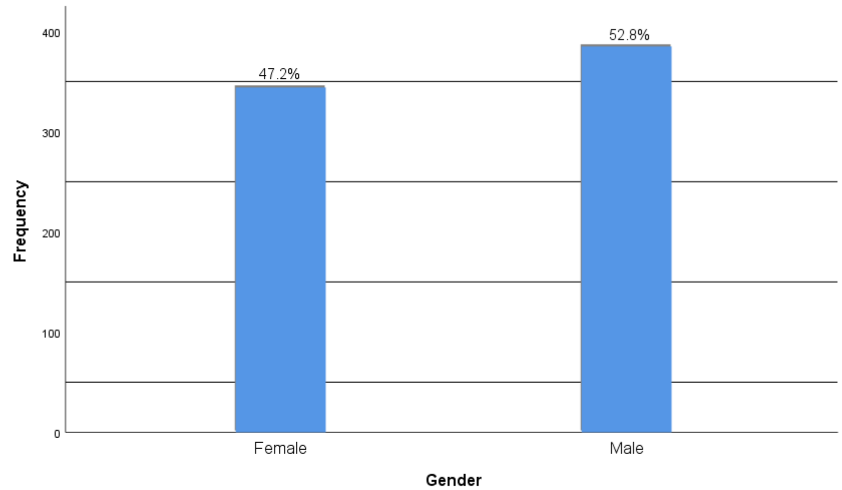


Figure 4. 7 Distribution of student respondents by gender

The student participants in this study came from different school levels and classes: Grade 10 and Grade 12 general and advanced mathematics classes as discussed in Chapter 3. Table 4.8 shows the breakdown of how many student participants there are in each schooling level. Table 4.8 further shows that the highest frequency of students, 375 students, are from the combined Grade 12 general and advanced mathematics students. The number of Grade 10 students who participated in this study were 354 students. The difference between the number of male and female student participants in each of the cohorts is shown in Table 4.8 and Figure 4.8. It is clear from Figure 4.8 that Grade 10 males make up 25.24 % of the total number participants in the study, compared to 23.32% for Grade 10 females. Furthermore, the number of Grade 12 male participants is 27.57 % (Advanced maths 14.54% + General maths 13.03%), which is higher than the number of Grade 12 female participants, at 23.87% (Advanced maths 11.65 % + General maths 12.21 %). The gap between male and female participants in both grades is relatively minor, indicating that the data was collected in proportion to gender.

Table 4. 8 Student respondents according to mathematics type

Math type	Frequency	Percent
Grade 10	354	48.6%
Gr12 General maths	191	26.2%
Gr12 Advance maths	184	25.2%
Total	729	100%

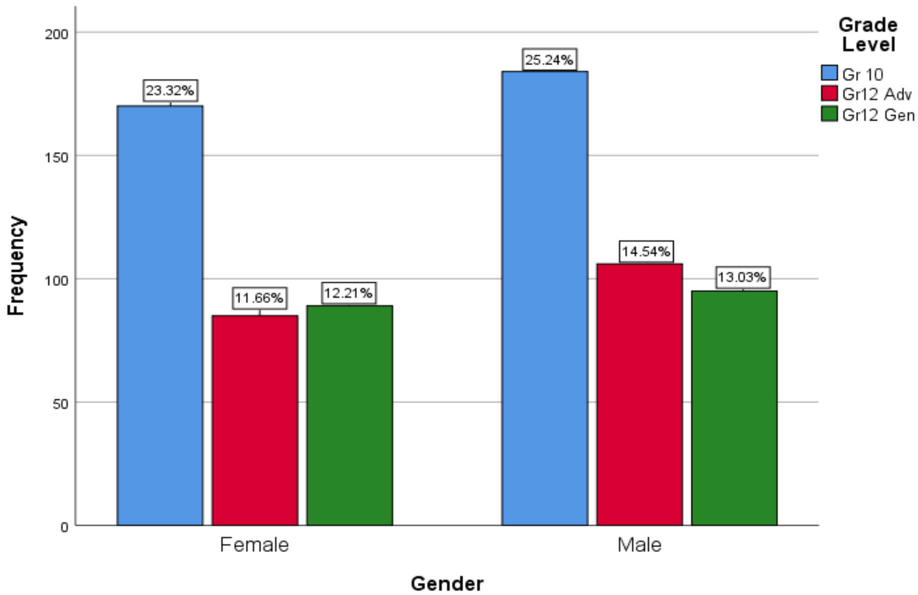


Figure 4. 8 Distribution of student respondents according to mathematics type

4.4 Parents highest education level

Table 4.9 shows that the highest frequency of 360 (over 49%) of students' fathers have highest qualification being certificate level. This reflects the fact that most students' fathers have limited access to higher levels of education due to their financial situations and the status of the economy. This is consistent with the reports of the 2050 vision statement of PNG. Moreover, about 174 (more than 23%) of students' fathers have a diploma and 173 (more than 23%) have a bachelor's degree; a similar distribution. Postgraduate levels of education, such as Masters and PhDs are achieved by 3% of students' fathers. Figure 4.9 shows a clearer picture of fathers' qualification levels.

Table 4. 9 Fathers' highest education level

F_Highest Education level	Frequency	Percent
Postgraduate (Masters/PhD)	22	3.0 %
Bachelor's Degree	173	23.7 %
Diploma	174	23.9 %
Certificate	360	49.4 %
Total	729	100 %

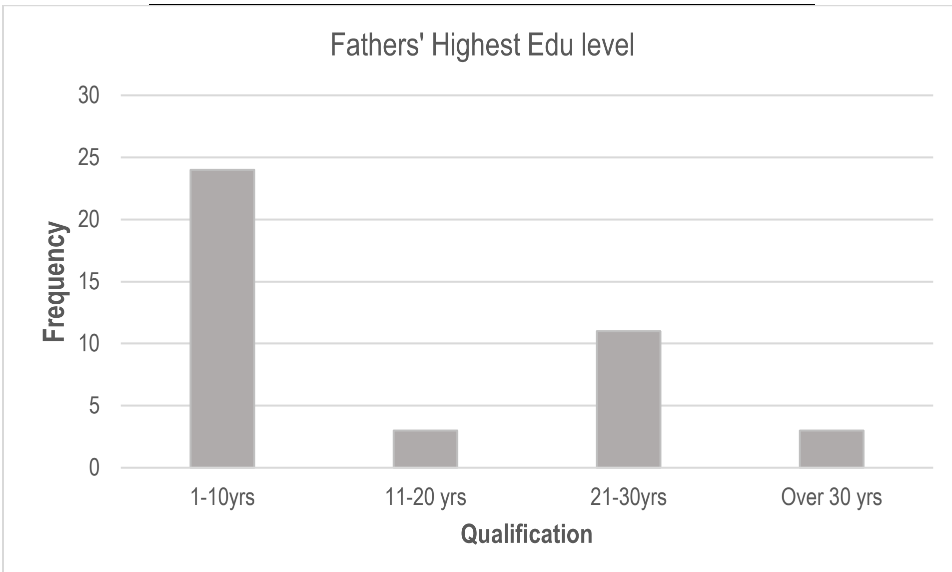


Figure 4. 9 Distribution of fathers' highest education level

Table 4.10 shows that over 44% of the students' mothers have their highest qualification at the certificate level. This reflects the fact that most of the students mothers' are less educated due to cultural aspects that affect women in PNG (Spark, 2010; PNG NDOE, 2006). This trend continues today, with women being less likely to move to higher education levels, or never being educated in schools at all. This issue is also highlighted in the 2050 vision statement of PNG. More than 38% of participants' mothers are educated up to the diploma level, and over 15% achieved a bachelor's degree, with only 1.4% of the mothers completing a postgraduate level of education. Figure 4.10 shows a clearer picture of the mothers' qualification levels.

Table 4. 10 Mothers' highest education level

M_Highest Education level	Frequency	Percent
Postgraduate (Master/PhD)	10	1.4%
Bachelor's Degree	113	15.5%
Diploma	281	38.5%
Certificate	321	44.0%
Total	729	100%

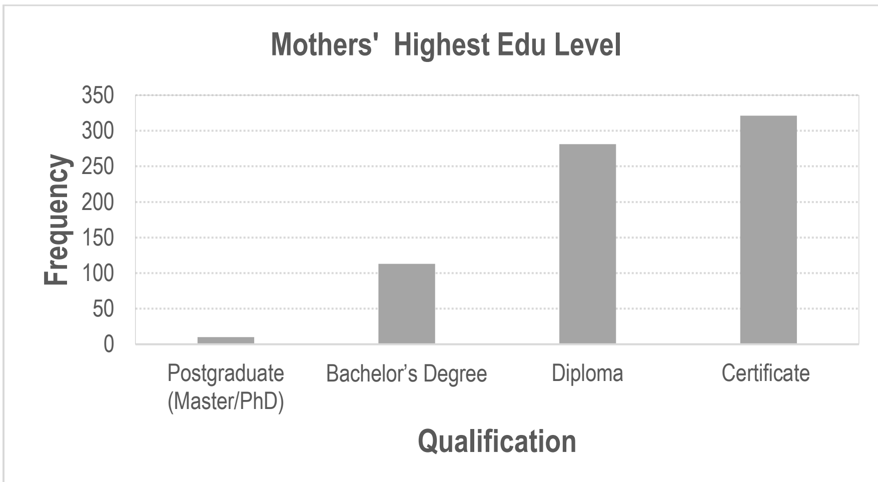


Figure 4. 10 Distribution of mothers' highest education level

4.5 Parents employment

As shown in Table 4.11, the relative highest frequency of 483 of the students' mothers (more than 66%) of the mothers are employed in informal sector work, such as subsistence farming or self-employment, and about 246 of the mothers (over 33%) have formal employment, such as being teachers or nurses. This means that most of the students' mothers are engaged in informal employment; a result that reflects the actual lifestyle for women/mothers in Port Moresby. Figure 4.11 clearly demonstrates the distribution of mothers' employment.

Table 4. 11 Mothers' employment

Mothers Employment	Frequency	Percent
Informal	483	66.3%
Formal	246	33.7%
Total	729	100%

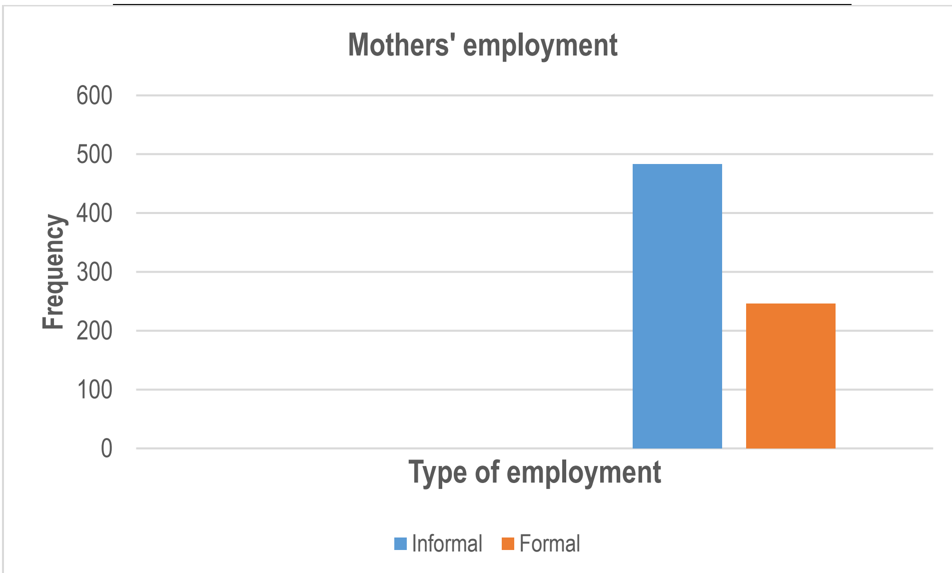


Figure 4. 11 Distribution of mothers' employment

As shown in Table 4.12, a relatively higher frequency of 491 of the students' fathers (more than 67 %) are engaged in informal employment, such subsistence farming or street sales, while 238 (over 32 %) are engaged in formal employment, such as being lawyers, accountants or teachers. This means that less than 50% of the students' fathers are engaged in the formal employment sector. This information is also clearly represented in Figure 4.12.

Table 4. 12 Fathers' employment

Fathers' Employment	Frequency	Percent
Informal	238	32.6%
Formal	491	67.4%
Total	729	100 %

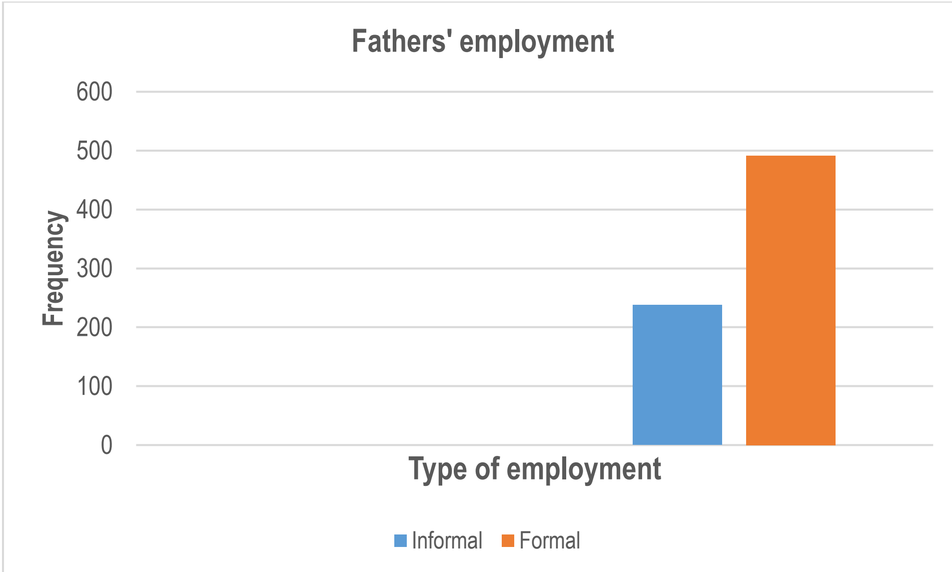


Figure 4. 12 Fathers' employment

4.6 Summary

The chapter discussed the descriptive information of the teachers and students who participated in the study. This study's student respondents were Grade 10 and 12 at secondary schools in Port Moresby, and the number of male students were roughly in proportion to the number of female students participating. The students' parents' education level, employment, and level of schooling achieved are also analysed and reported. The gender of students' teachers' is also in proportion. The teachers' years of experience teaching mathematics, their major and minor subjects, and their qualification levels are additionally analysed and reported to present a clearer understanding of the demographic information obtained from the data. The results of this demographic information for both student and teacher-levels provide bases for analysis in the subsequent validation chapters for tests and the survey items for the different scales.

Chapter 5: Grade 10 mathematics test validation

5.1 Introduction

The chief focus of this chapter is to examine the Grade 10 mathematics test items that were administered to students in Port Moresby. Students' results on this test are influenced by both student and teacher-level factors, which will be discussed in the later chapters. To answer this study's research questions relating to the factors affecting students' mathematics learning, a mathematics test was developed. As such, the instrument (mathematics test questions) were developed and was administered to Grade 10 students at 16 secondary schools in Port Moresby, PNG. This test was analysed using the Rasch measurement model (RMM) to ensure that it is reliable, valid and objective. This chapter presents and discusses the analysis process of the Grade 10 mathematics test. The analysis process for the Grade 12 mathematics test is discussed separately, in Chapter 6 of this study.

This chapter begins by explaining the concept of mathematics achievement and the development of the test items. The next section explains the process of adoption and construction of the Table of Specification (TOS) for the mathematics test in this study. The test item analysis process and statistical results are then discussed using the RMM, with the chapter then ending with a summary, highlighting the key ideas discussed.

5.2 Mathematics achievement test

An achievement test is an examination designed to assess how much knowledge an individual person has in a certain content area or set of content areas as a result of teaching (McMillan, 2012; Oermann & Gaberson, 2016). It is an instrument administered to an individual or group as a stimulus to elicit certain desired or expected responses that represent his/her ability (Oermann & Gaberson, 2016). Every measuring instrument, including the Grade 10 test used in this study, is expected to possess certain characteristics such as relevance, difficulty level and comprehensiveness. Any test, and indeed any evaluation instrument must satisfy the criteria of reliability, validity as well as objectivity (McMillan, 2012; Oermann & Gaberson, 2016).

This chapter discusses and provides a comprehensive description of the mathematics achievement test administered to the Grade 10 students in Port Moresby, including analysis of the test results. The 40 mathematics test items that were developed are consistent with the goal of covering mathematics topics comprehensively. This test was designed in accordance with the Grade 10 national examination Table of Specification (TOS) and the syllabus for lower secondary schools (Grade 9 and 10) obtained from the Measurement Service Division (MSD) of the PNG National Department of Education (see Appendix F) . After data collection, student responses to the test items were analysed through the Rasch model, as discussed in Chapter 3. This analysis was carried out to determine the test-takers' ability levels and the item difficulty based on their responses (Bond et al., 2007; Glynn, 2012). Subsequently, the estimation of the test-takers' ability levels or item difficulty is possible by converting raw scores to logits on an objective measurement scale (Bond & Fox, 2015). ConQuest version 4.0 (Wu, Adams, Wilson & Haldane, 2015) generated both latent estimates and plausible values, by adding all the marks for each item in the test. The comparison of personal ability and item difficulty estimates provides important information for inferences to be drawn from test achievements as discussed in Chapter 3.

5.3 Development of the mathematics test items

In this study, the Grade 10 mathematics national examination TOS for 2017 and Grade 9 and 10 syllabuses from MSD (see Appendix F) were used to guide instrument development and interpretation. The rationale for using these two documents was to ensure that the TOS of this study was in accordance with the TOS of MSD. Hence, the Grade 9 and 10 syllabuses are chosen because their units and topics are indicated on the Measurement Services Division's TOS for the Grade 10 examination. As such, this study also uses the same syllabus for the selection of the topics with the belief that it comprehensively covers the aspects pertaining to mathematics. The TOS defines the different domains, describes the scope of the assessment, and specifies the structure of the test and the preferred distribution of items with their weighting (marks) (Anderson & Morgan, 2008).

The new TOS in this study is designed and aligned in coordination with the syllabus and the TOS for Grade 10 from MSD, encompasses the 2050 vision. This TOS assisted in developing the mathematics test to assess students' knowledge in mathematics and their understanding of the application of that knowledge. A TOS is an important document that guides test development, analysis, and report writing (Anderson & Morgan, 2008), and it specifies the proportion of items in a test that address the various aspects of a curriculum. The TOS in this study is developed to encompass three key points highlighted by Anderson and Morgan (2008, p. 12) in their discussion of developing national examinations in PNG and other developing countries. These three key points are:

a) The proportion of items within a content area that assess different skills. For example, this includes the inclusion of different content areas such as number, measurement, space, and patterns;
b) The proportion of test items that address different cognitive processing skills such as simple recall, comprehension and application; and
c) The proportion of test items devoted to stimulus texts of different kinds in reading mathematics tables, charts, and diagrams.

Finally, the TOS for mathematics used in this study was utilised to ensure that the items were developed to meet the guidelines laid out in PNG's national examination.

As mentioned above, this study patterned the process of selecting test items from the TOS and syllabus coupled with the researcher's own experiences in teaching mathematics in secondary schools in PNG. The reason for this methodological approach is because of the unavailability of an extant framework that describes the process of constructing national examination items from MSD in PNG. However, the process was modified to suit the context of the study. Test items were elaborated and selected from past Grade 10 national examination questions, according to the different cognitive domains and content areas, to ensure that the participants were responding to appropriate mathematics questions. The Grade 10 national examination items were sourced from different years: (1999, 1995 and 2000), and were utilised according to their availability, coverage of the topics in the units and different cognitive developmental domains. Of the 60 probable test items provided by these examination questions, 40 items were then selected based on the topics covered from Grade 9 to the date the researcher conducted the Grade 10 mathematics test. This approach was intended to avoid selecting test items that the students had not yet covered in the classroom. More importantly, as highlighted earlier, the researcher's experience teaching in secondary school in PNG also played a major role in the item selection process. Another essential element that contributed to selecting the test items was the response time allocated for each question. These two elements were considered crucial to achieve the objectives of the mathematics test, and to mitigate potential difficulties. After analysis, the initial selection and approval for the final 40 mathematics test items was undertaken, with these items then evaluated by reviewers for suggestions on content area, grade appropriateness, coverage of the cognitive domains specified and item format. The purpose of this review was to ensure the alignment and accuracy of the test items, as discussed in Chapter 3.

The TOS from MSD distributed the cognitive domains according to the Grade 10 mathematics national examination in 2017. Simple recall questions comprised 20% of items; comprehensive questions consisted of 50%, and application, analysis, synthesis, evaluation questions 30% (NDOE, 2017). Similarly, the TOS for this study also adopted the same percentage distribution as the TOS from MSD for the cognitive domains (see Table 5.1). In the same way, the estimated percentage of content of the test items covered in this study's TOS was derived from each unit(s) from both Grades 9 and 10. Furthermore, the researcher ensured that the content percentage of scores used in the TOS for this study and the estimated percentage of content of test items from the TOS of MSD, were approximately equal to one another (NDOE, 2017). These techniques were employed to ensure that the test items in this study covered approximately the same content and cognitive domains for alignment and uniformity purposes.

The Grade 10 mathematics national examination in PNG consists of three different types of test items: multiple choice, short answers and calculation. However, in this study, 40 multiple choice test items were used that highlighted both the content area in the syllabus and the different cognitive domains. The reasons for this choice were: a) for easy item analysis, b) for quick marking, c) in consideration of the time constraint involved in conducting the test at schools, and d) the cost involved in printing the test questions. Each test item required students to select one of four options (A, B, C, and D), with the examination lasting an hour total. The test-takers marked their responses by shading their best answers on the sheet provided. As mentioned earlier, the mathematics test used is a reflection of the TOS for this study, and was developed with reference to the MSD's TOS and syllabus, and past examination papers. Table 5.1 shows the TOS used in this study, including the distribution of content/topic area and cognitive development domains for the 40 mathematics tests items.

Table 5. 1 Grade 10 mathematics test table of specification (TOS)

Content Area		Cognitive Domain				Percentage	
Unit (s)	Topic (s)	Knowledge 20%	Comprehension 50%	Higher Order 30 %	Score	Score %	MSD Content %
9.1 Mathematics in Our Community	Measurement		40	19	8	20%	18.5%
	Rates		14, 26				
	Ratio & Proportion		6	31			
	Number & Operation	18,17					
9.2 Pattern of Change	Equations		5,34,35	1	6	15%	13.5%
	Directed Numbers	3					
	Patterns of Change			33			
9.3 Working with Data	Representing Data			24,25	4	10%	12.5%
	Central Tendency and Spread	8	9				
9.4 Design in 2D and 3D Geometry	Surface area and Volume		21	22	4	10%	14.5%
	Deductive Reasoning			39			
	Design in 2D & 3D Geometry		15				
10.1 Managing your Money	Spending Money		2	20	5	12.5%	13%
	Managing your Money		10				
	Earning Money		28	37			
10.2 Functions and Graphs	Basic Algebra		13,11,38		8	20%	17.5%
	Indices & Scientific Notation	12, 7	36				
	Simultaneous Equations		16	32			
10.3 Trigonometric Functions	Trigonometric Applications	23,4	29	30	5	12.5%	10.5%
	Pythagoras Theorem		27				
	Total	8	20	12	40	100%	100 %
	Researcher Set Goal %	20%	50%	30%			

5.4 Results of the Rasch analysis

Analysis of the mathematics test was carried out using the responses from the 354 Grade 10 students from the 16 secondary schools surveyed in Port Moresby. The item fit statistics in the Rasch measurement model (RMM) discussed in Chapter 3 were utilised to evaluate each individual person's response to the test. Analysis was conducted using ConQuest, version 4.0 (Wu, Adams, Wilson & Haldane, 2015) to assist and detect discrepancies between the RMM prescriptions and data, to verify the utility of test analysis. The results of this analysis were determined through statistics from the 40 mathematics test items, which were then analysed and presented under key points to identify items that were misfitting to the RMM.

5.4.1 Overall test statistics

The initial run of the 40 multiple-choice results are shown in Table 5.2. These results show a summary of the fit statistics such as estimates, INFIT MNSQ, t-value and the discrimination index for each item and person obtained through ConQuest (Wu, Adams, Wilson & Haldane, 2015). As illustrated in Table 5.2, the INFIT MNSQ ranged from 0.99 to 1.11, and the t -value from -3.4 to 3.1.This indicates that the INFIT MNSQ results for the 40 items are within the acceptable INFIT MNSQ range of 0.70 to 1.30 (Bond & Fox, 2015; Bond et al., 2007). However, among these items, three fitted perfectly, with an INFIT MNSQ value of 1.0, while the remainder are over-fitting (INFIT MNSQ value less than 1) and under-fitting (INFIT MNSQ value greater than 1) though otherwise provided sufficient information. Linacre (2002) verifies this result, by highlighting that items with an INFIT MNSQ within a range of 0.70 and 1.30 provide adequate information without overfitting and misfitting the model, which demonstrates adequate fit to the theoretical item characteristics curve. As a result, the items fit the model and students with average ability levels from the sample are responding correctly. Moreover, as shown in Table 5.2, the INFIT MNSQ are within the acceptable range of 0.70 to 1.30, showing high respondent participation. The OUTFIT MNSQ is not reported in this study because it can determine unanticipated responses from participants; when it is high, this can be ascribed to noise related to a low number of respondents due to the influence outliers (Wu & Adams, 2007; Wu et al., 2016). Additionally, the t-values of four items (see Table 5.2) were out of the accepted maximum and minimum values of 2 and -2, and so were removed due to violation of the RMM rule. More importantly, the t-values played a significant role in the deletion of items, as they represented how closely the data fits the RMM (Wu & Adams, 2007; Wu et al., 2016)

Table 5. 2 Results of initial ConQuest run analysis of the Grade 10 test items

Item No	Estimate	Error	INFIT MNSQ	ZSTD (t)	Pt Bis for correct answer	Item No	Estimate	Error	INFIT MNSQ	ZSTD (t)	Pt Bis for correct answer
Item 01	-1.13	0.12	0.98	-0.4	0.27	Item 21	0.24	0.12	1.07	1.6	0.11
Item 02	0.71	0.12	0.94	-1.0	0.35	Item 22	1.36	0.15	0.97	-0.3	0.25
Item 03	0.05	0.11	1.11	3.1	0.02	Item 23	-0.62	0.11	0.98	-0.7	0.28
Item 04	0.01	0.11	1.11	3.0	0.04	Item 24	-0.21	0.11	0.93	-2.3	0.39
Item 05	-0.49	0.11	1.02	0.6	0.24	Item 25	-0.50	0.11	0.92	-2.9	0.42
Item 06	0.66	0.13	1.02	0.3	0.20	Item 26	1.08	0.14	1.00	0.0	0.19
Item 07	-0.69	0.11	0.95	-1.5	0.33	Item 27	-0.18	0.11	1.05	1.6	0.14
Item 08	-0.83	0.11	1.00	0.1	0.26	Item 28	-0.14	0.11	0.9	-3.4	0.43
Item 09	0.23	0.11	0.91	-2.3	0.43	Item 29	1.35	0.13	1.06	0.6	0.06
Item 10	0.29	0.12	1.02	0.5	0.21	Item 30	0.80	0.13	1.00	0.1	0.21
Item 11	0.73	0.12	1.01	0.2	0.21	Item 31	-0.72	0.12	0.99	-0.3	0.28
Item 12	-1.39	0.12	0.98	-0.4	0.27	Item 32	-0.12	0.12	1.03	0.9	0.20
Item 13	-1.00	0.12	0.96	-0.9	0.31	Item 33	0.43	0.12	0.92	-1.6	0.38
Item 14	-0.56	0.11	0.94	-2.1	0.37	Item 34	-0.08	0.12	1.00	0.1	0.27
Item 15	0.49	0.12	0.97	-0.6	0.30	Item 35	-0.24	0.12	0.96	-1.2	0.33
Item 16	-0.00	0.12	1.04	1.0	0.18	Item 36	-0.17	0.12	1.00	0.1	0.25
Item 17	-1.05	0.12	0.99	-0.2	0.26	Item 37	0.31	0.12	1.02	0.5	0.20
Item 18	0.34	0.12	0.94	-1.3	0.34	Item 38	0.53	0.13	1.08	1.4	0.09
Item 19	-0.51	0.11	1.06	1.9	0.13	Item 39	0.35	0.12	1.01	0.3	0.20
Item 20	0.16	0.11	0.98	-0.6	0.29	Item 40	0.48*	0.13	1.10	1.8	0.07

*Separation Reliability = 0.97 Chi-square test of parameter equality = 1075.85, df = 35, Sig Level = 0.00

Item separation reliability in RMM provides an analytical tool for the evaluation of the successful development of a test, and to monitor its continuing utility (Wright & Stone, 1999). The range of the separation reliability index is from 0 to 1 .0 (Bond & Fox, 2015; Wright & Stone, 1999), and the statistic of the separation reliability index for this analysis is very good, with a result of 0.97 (see Table 5.2). This result indicates that the items had more precise measurement and reliability (Hailaya, 2014; Wright & Stone, 1999).

Moreover, Table 5.2 shows the range of item difficulty of the 40 calibrated test items was from -1.39 to 1.36 logits, and is connected to a standard error of 0.11 to 0.15 logits. These items show difficulty index measures, identify the difficulty level of an item and classify the level of an item as easy, moderate, or hard (Zainuri et al., 2016). The difficult level of the items is discussed in the latter part of this analysis, with the item-person map in Section 5.4.6.

Furthermore, the results of the point biserial index (r_{pb}) of the items in Table 5.2 ranges from 0.07 to 0.43. This result shows that the items discriminated and differentiated among respondents with different ability levels, and implies that the items indicated a relationship between the respondents' performances on the given item (correct or incorrect) and the respondents' scores on the overall test (Quaigrain & Arhin, 2017; Wu & Adams, 2007; Wu et al., 2016). It was also evident that high-ability examinees responded to the items more frequently than lower-ability examinees (Adedoyin & Mokobi, 2013; Wu & Adams, 2007).

5.4.2 Individual item statistics

The Rasch analysis indicated that the scoring key of the items were working as required by the model, and contributed to the measurement of respondents' abilities. Next, graphical analysis provided by the item characteristic curve (ICC) (see Figure 5.1) was used to check item functioning. Each of the 40 items is checked for the item fit statistics, ICC, as well as the functioning of the categories. Based on these diagnostic procedures, items that were not abiding by RMM rules were identified and removed one by one. Further, Table 5.2 indicates that the initial analysis identified items 03, 04, 09,14,24,25 and 28 as having misfit statistics (t-values), outside of the recommended limits of -2 to 2. Some of the lower values (negative) may show possible redundancy in the respondents' answers (lack of expected stochastic fit or violation of local item independence). Further, the statistical fit criteria discussed in Chapter 3 determined how the data fits the RMM (Wright & Stone, 1999; Wu et al., 2016). Simultaneously, this criterion, the ICC and category probability (see Figure 5.1) for each item were checked and the data was calibrated. With this statistical evidence, the researcher decided to omit items 03, 04, 25, and 28 individually, for the next calibrations. As shown in Table 5.3, this calibration process resulted in improved fit statistics for the misfit items, and at the same time the statistical fit for other items improved and some regressed slightly, but there was no longer any significant misfit.

Table 5. 3. Results of final ConQuest run analysis of the Grade 10 test items

Item #	**Estimate**	**Error**	**INFIT MNSQ**	**ZSTD (t)**	**Pt Bis for correct answer**	**Item #**	**Estimate**	**Error**	**INFIT MNSQ**	**ZSTD (t)**	**Pt Bis for correct answer**
Item 01	-1.15	0.11	0.99	-0.3	0.27	Item 21	0.22	0.11	1.07	1.6	0.11
Item 02	0.69	0.12	0.94	-1.0	0.35	Item 22	1.34	0.14	0.98	-0.2	0.25
						Item 23	-0.63	0.11	0.99	-0.2	0.27
						Item 24	-0.22	0.11	0.95	-1.7	0.37
Item 05	-0.5	0.11	1.01	0.3	0.24						
Item 06	0.65	0.12	1.03	0.5	0.18	Item 26	1.06	0.13	1.00	0.1	0.19
Item 07	-0.7	0.11	0.95	-1.4	0.31	Item 27	-0.19	0.11	1.05	1.5	0.16
Item 08	-0.85	0.11	1.0	0.0	0.26						
Item 09	0.22	0.11	0.92	-2.0	0.4	Item 29	1.33	0.14	1.06	0.7	0.07
Item 10	0.28	0.11	1.03	0.7	0.21	Item 30	0.79	0.12	1.01	0.2	0.21
Item 11	0.71	0.12	1.01	0.2	0.22	Item 31	-0.74	0.11	0.99	-0.3	0.27
Item 12	-1.41	0.12	0.98	-0.3	0.26	Item 32	-0.13	0.11	1.03	0.8	0.20
Item 13	-1.01	0.11	0.96	-1.0	0.30	Item 33	0.41	0.12	0.93	-1.4	0.37
Item 14	-0.57	0.11	0.94	-1.9	0.37	Item 34	-0.09	0.11	1.00	0.1	0.27
Item 15	0.48	0.12	0.97	-0.6	0.31	Item 35	-0.25	0.11	0.96	-1.3	0.33
Item 16	-0.02	0.11	1.02	0.7	0.20	Item 36	-0.18	0.11	1.00	-0.1	0.27
Item 17	-1.07	0.11	0.99	-0.2	0.26	Item 37	0.29	0.12	1.03	0.6	0.20
Item 18	0.32	0.12	0.95	-1.0	0.33	Item 38	0.51	0.12	1.08	1.5	0.10
Item 19	-0.52	0.11	1.07	2.0	0.10	Item 39	0.33	0.12	1.02	0.4	0.20
Item 20	0.15	0.12	0.98	-0.5	0.29	Item 40	0.45*	0.12	1.11	2.0	0.07

Figure 5.1 below and Figure 5.2, (p.129) appear as examples of two items (Questions 04 and 22) representing the other items with their category probability curves and ICCs, indicating certain behavioural pattern for each item. Item 04 shows an under fit ICC with poor discrimination, and item 22 shows the majority of the students choosing an option that is not the correct key.

```
================================================================================
Item 4
------
item:4 (ITEM04)
Cases for this item      346   Item-Rest Cor.  0.04   Item-Total Cor.  0.12
Item Threshold(s):      0.01   Weighted MNSQ   1.11
Item Delta(s):          0.01
--------------------------------------------------------------------------------
 Label    Score     Count   % of tot  Pt Bis       t  (p)     PV1Avg:1 PV1 SD:1
--------------------------------------------------------------------------------
   A       0.00       103     29.77    0.05     0.95(.341)  -0.38     0.57
   B       1.00       143     41.33    0.04     0.66(.508)  -0.29     0.59
   C       0.00        81     23.41   -0.07    -1.23(.220)  -0.46     0.66
   D       0.00        18      5.20   -0.02    -0.43(.669)  -0.42     0.56
   Q       0.00         1      0.29   -0.15    -2.76(.006)  -2.09     0.00
--------------------------------------------------------------------------------
```

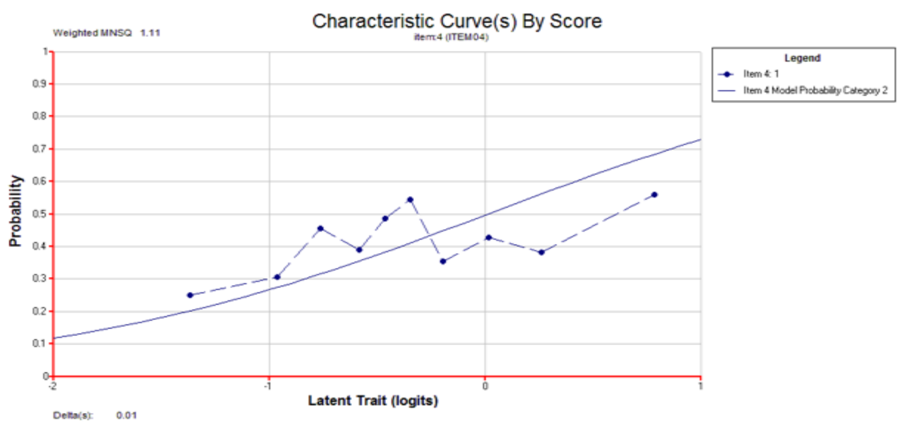


Figure 5. 1 ICC for item 4

One of the principles of classical test theory (CTT) is to measure students' performance using the raw score (true score), with discrimination varying depending on the sample of respondents that took the specific test. However, consistent with the purpose of this study, discrimination is discussed within the context of RMM; that the latent or unobserved variable can predict the performance of the respondents in the test without depending on the sample.

The discrimination index is the correlation between a respondent's score on an item and his/her total score on the mathematics test (Wu & Adams, 2007; Wu et al., 2016; Zainuri et al., 2016). In Figure 5.1, item 04's discrimination index (Item-Rest Cor) is 0.04. The item reflects a low level of independence and a very low correlation is expected between the score on the item and the total score of the test (Wu & Adams, 2007; Wu et al., 2016). This discrimination value of 0.04 is extremely low and is below the discrimination index value of 0.2 (Wu et al., 2016; Zainuri et al., 2016). It is therefore apparent that the discrimination for this item was poor and that it discriminated less between respondents according to their autonomy level despite the minimal positive relation reflected on the discrimination index value (Baghaei & Amrahi, 2011; Wu & Adams, 2007). As the item is nearing a discrimination value of 0, this suggests there is no relationship between the item and the total score.

Furthermore, the ICC of the items provided useful information about the behaviour of each item (Wu & Adams, 2007; Wu et al., 2016). The ICC for Figure 5.1 p.129 demonstrates that the respondents located at -0.4 logits performed above the 50% probability. This means the respondents' chances of responding to item 4 is greater at -0.4 logits, compared to the other ability levels. Figure 5.1 also shows that the ICC's level of autonomy increases along the horizontal axis, but that the observed probability of response category does not increase, while the theoretical probability of category increases with the level of ability. More importantly, the observed curves are relatively flatter, indicating that there is little relationship between the response categories and the level of autonomy (Wu & Adams, 2007; Wu et al., 2016). The item fit statistics discussed earlier in this chapter indicate the extent to which the item fit the RMM. As shown in Figure 5.1, the weighted MNSQ 1.11 demonstrates a fit index indicating that item 04 is less discriminating than the model predicts, because the INFIT MNSQ value is greater than 1. It is also evident from the ICC that the observed curve is flatter than the theoretical curve. Item 04 has less power in separating respondents on the latent variable scale because of the less discriminating power and, as such, this item is a candidate for deletion.

As shown in Figure 5.2 (item 22 below), more respondents (149) chose the distractor B, which is the wrong answer. Distractors C and D were answered by 42 and 92 respondents, respectively. However, 58 respondents with high ability correctly selected A as the correct answer. It is evident from the characteristics curve by category that the distractors for this item are less fairly distributed than other items. The weighted MNSQ of 0.97 demonstrates a good fit index (Wu & Adams, 2007). This INFIT MNSQ value indicates that item 22 was discriminating better than the model predicts, because the value is less than 1. It is also evident from the ICC that the observed curve is steeper than the theoretical curve. Item 22 is a good item that is able to separate respondents on the latent variable scale because of its discriminating power.

The Rasch model does not specify an absolute value for discrimination parameters (Wu & Adams, 2007; Wu et al., 2016). Further, it does not say anything about whether an item is good or bad in terms of its discrimination power. This is because fit statistics test whether items have equal discrimination, and an item-showing misfit in a given test may show good fit in another test (Wu et al., 2016). Even if the items are fitting the RMM, this does not mean that the test is a good test.

```
Item 22
-------
item:22 (ITEM22)
Cases for this item      342   Item-Rest Cor.   0.25   Item-Total Cor.   0.31
Item Threshold(s):      1.35   Weighted MNSQ    0.97
Item Delta(s):          1.35
```

Label	Score	Count	% of tot	Pt Bis	t (p)	PV1Avg:1	PV1 SD:1
A	1.00	58	16.96	0.25	4.85(.000)	0.01	0.68
B	0.00	149	43.57	-0.07	-1.23(.219)	-0.39	0.59
C	0.00	42	12.28	-0.08	-1.57(.118)	-0.55	0.56
D	0.00	92	26.90	-0.06	-1.10(.270)	-0.46	0.50
Q	0.00	1	0.29	-0.15	-2.84(.005)	-2.09	0.00

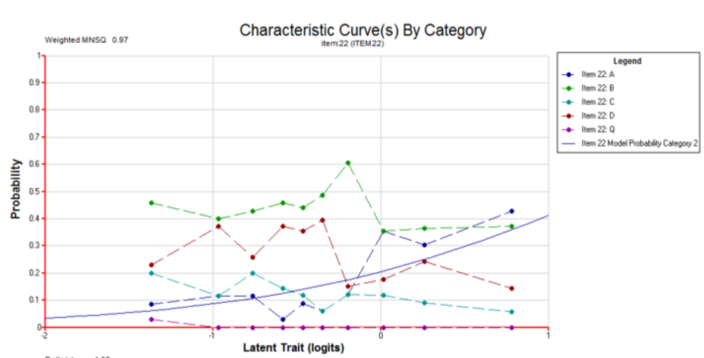


Figure 5. 2 Item 22 analysis and ICC

5.4.3 Differential item functioning

In this section, differential item functioning (DIF) analysis is conducted for the final 36 items that fitted the RMM, in order to assess the fairness and equity of the test items in terms of the different gender groups (male and female). As discussed in Chapter 3, where DIF is present in an item, this suggests that the item functions differently in different contexts; in other words, that two groups of people with the same ability have different probabilities of success in responding to an item (Hagquist & Andrich, 2015, 2017; Wu et al., 2016). For example, in this study, both male and female Grade 10 students with similar average mathematics abilities were given a test, in which an item was administered with the context of building a house. From this, it is observed that males in this group performed considerably better than females on this item, even though girls and boys performed similarly on other items. This is because the males were more familiar with the context of the question than the females, and so males found the item easier than females did. This example demonstrates an item that exhibits DIF for the two gender groups. This study therefore checked DIF both on the test level and the item level for clarity on the fairness and equality of the test items on gender.

5.4.3.1 Overall test level DIF for gender

This section analyses and compares the overall mathematics test performance for Grade 10 male and female students in Port Moresby. As shown in Table 5.4, p.131, the INFIT MNSQ values of the overall test for both males and females is within the acceptable value used in this study of 0.70 to 1.30. Additionally, Table 5.4 also shows the item difficulty parameter estimates for each of the 36 items; the negative sign for the 17 items on the parameter estimates indicates that they were easier for male students, while the 19 items parameter estimates with positive signs indicates that those items were more difficult for females (Wu & Adams, 2007; Wu et al., 2016). There was not much difference in difficulty levels between male and female respondents.

Table 5. 4 Test level gender difference statistics

Item No	Estimate	Error	INFIT MNSQ	Item No	Estimate	Error	INFIT MNSQ
Item 01	-1.15	0.12	1.0	Item 22	1.38	0.15	0.98
Item 02	0.72	0.13	0.94	Item 23	-0.63	0.11	0.99
Item 05	-0.49	0.11	1.03	Item 24	-0.21	0.11	0.93
Item 06	0.68	0.13	1.04	Item 25	-0.52	0.11	0.92
Item 07	-0.70	0.11	0.97	Item 26	1.10	0.14	1.01
Item 08	-0.83	0.11	1.01	Item 27	-0.16	0.11	1.05
Item 09	0.24	0.12	0.91	Item 28	-0.14	0.11	0.92
Item 10	0.31	0.12	1.04	Item 29	1.41	0.15	1.08
Item 11	0.74	0.13	1.02	Item 30	0.82	0.13	1.01
Item 12	-1.41	0.12	1.00	Item 31	-0.76	0.12	1.01
Item 13	-1.01	0.12	0.97	Item 32	-0.13	0.12	1.06
Item 14	-0.56	0.11	0.94	Item 33	0.44	0.12	0.92
Item 15	0.50	0.12	0.96	Item 34	-0.07	0.12	1.00
Item 16	0.00	0.12	1.05	Item 35	-0.24	0.12	0.98
Item 17	-1.08	0.12	1.00	Item 36	-0.19	0.12	1.02
Item 18	0.35	0.12	0.94	Item 37	0.34	0.13	1.04
Item 20	0.17	0.11	0.98	Item 39	0.35	0.12	1.03
Item 21	0.25	0.12	1.09	Item 40	0.48	0.13	1.12

Table 5. 5 Test gender differences in ability estimates

	Sex	Estimate	Error	INFIT MNSQ	ZSTD (t)
1	Female	0.042	0.042	1.03	0.3
2	Male	-0.042*	0.042	1.07	0.7

Chi-square test of parameter equality = 0.99, df = 1

The results of the mean estimates of male and female respondents for the overall test are shown above in Table 5.4. This table shows estimates for gender differences in ability estimates at test level. From these results, it is apparent that the test is easier for male students than for female students, indicated by the parameter estimate of -0.042 for males. The actual parameter estimate for male students is one (0.042/0.042), equal to its standard error estimate, and so the difference between the male and female means is obviously insignificant. However, if the actual parameter estimate for male students is two or three times larger than its standard error estimate, then the difference between the male and female means is obviously significant (Wu & Adams, 2007; Wu et al., 2016). This difference is associated with the chi-square value of 0.99 on one degree of freedom as shown in Table 5.5. Therefore, it can be concluded that the male students' mean performance is the same as the female students. As seen in Table 5.5, the difference between the actual parameter estimate of male and female students shows that males scored 0.084 logits lower than female students. Hence, a difference of 0.084 logits, which is much smaller than 0.5 logits, implies that the average performance levels of male and female students are not substantially different. The overall test INFIT MNSQ and t-values are within the range of 0.70 to 1.30 and -2 to 2, respectively. Specifically, females and males have INFIT MNSQ values of 1.03 and 1.07 with t-values of 0.3 and 0.7, respectively. Together, these statistical findings suggest that there is no DIF in the overall mathematics test for Grade 10 students. In the next section of this chapter, individual items on the test are checked to ascertain whether there exist DIF items for male and female students at the item level.

5.4.3.2 Item level DIF for gender

Following analysis for DIF at the overall test level, individual items from the mathematics test were analysed for DIF between male and female participants. This step was carried out due to the insignificant difference found for gender in the overall test analysis. At the item level of analysis, on the other hand it is expected that DIF would be detected on the items for gender. Item level analysis was conducted through three methods of identifying DIF in RMM: the item threshold approach, item fit approach, and group plot methods (Hungi, 2005; Scheuneman & Bleistein, 1999). These methods were used as they underpin the assumptions of item response theory (IRT), discussed in Chapter 3, which makes it useful to investigate DIF. The estimated parameters of the item response function (the probability that persons with lower ability have less of a chance to give the correct answer, while persons with higher ability are very likely to answer correctly) are unchanged for different samples drawn from the same population (Scheuneman & Bleistein, 1999). “Therefore, if parameters are estimated separately for two groups, the resulting item response functions of an item which is functioning equivalently for those groups should be the same” (Scheuneman & Bleistein, 1999, p. 229). In other words, the probability of a correct response for respondents at a given ability level is the same for males and females, since a true ability scale is used rather than observed test scores. This, therefore, allowed the researcher to use the three methods for RMM instead of other methods utilised by other researchers in classical test theory (CTT).

5.4.3.3 Item threshold approach

In Table 5.6 (p.135), negative values of difference in item threshold and difference in standardised item threshold, signify that the item is relatively easier for female students than for male students, while positive values for males imply the opposite (Abd-El-Fattah et al., 2014; Le, 2006). These differences are apparent in the Grade 10 test (see Table 5.7 p.139) with 17 items having negative values of item threshold difference and standardised item threshold difference. This implies that these items apparently favors female students over male students. The other 19 items have a positive value of difference in item threshold and difference in standardised item threshold, indicating that they are relatively more difficult for male respondents than for the female respondents.

Furthermore, the difference of the item threshold for an item d_1-d_2 <-0.5 and d_1-d_2>0.5 logits (Abd-El-Fattah et al., 2014; Le, 2006) indicates an item with DIF. This difference value is approximately equal to an extra one year of school to learn a distinct content area. When the difference in the item threshold value of an item between male and female respondents is below or above the predefined range of d_1-d_2 <-0.5 and d_1-d_2>0.5 logit range, such a difference may cause a significant concern regarding to item DIF(Abd-El-Fattah et al., 2014; Le, 2006).

In this study, as shown in Table 5.6, items 27, 29, 31 and 34 are within the acceptable item threshold logit value range of d_1-d_2 <-0.5 and d_1-d_2>0.5 (Hungi, 2005; Le, 2006). It is observed that items 09, 17,32,36 and 39 (see Table 5.5) are more difficult for male respondents than for female respondents, with differences of item threshold values less than -0.25 logits (Hungi, 2005). According to Hungi (2005), this indicates that male respondents require more time to learn some Grade 10 mathematics content compared to female respondents. On the other hand, items 08 and 37 are significantly more difficult for female students than male students with the differences of item threshold values are greater than 0.25 logits. In other words, female respondents need more time to learn this Grade 10 content compared to the male respondents (Hungi, 2005). Added to this, items 27, 29 and 31 posed a substantial amount of DIF and were deleted, with absolute difference values above and below + 0.5 logits to -0.5 logits threshold, as recommend by Hungi (2005). This would indicate that items 27 and 29 favored male respondents, and item 31 favored female respondents. Both genders require more time to learn the specific content area about these DIF items.

Table 5. 6 DIF results for the 36 mathematics test items by item threshold

Item No	Female		Male			
	Estimate (d1)	Error (e1)	Estimate (d2)	Error (e2)	d1-d2	st (d1-d2)
Item 01	-0.11	0.12	0.107	0.12	-0.28	-1.26
Item 02	0.05	0.13	-0.05	0.12	0.09	0.51
Item 05	0.07	0.11	-0.07	0.11	0.14	0.89
Item 06	0.05	0.13	-0.05	0.12	0.10	0.55
Item 07	-0.08	0.11	0.08	0.11	-0.16	-1.02
Item 08	0.19	0.11	-0.19	0.11	0.39	2.00
Item 09	-0.13	0.12	0.13	0.11	-0.27	-1.62
Item 10	0.11	0.12	-0.11	0.11	0.22	1.32
Item 11	-0.05	0.13	0.06	0.12	-0.11	-0.66
Item 12	0.07	0.12	-0.07	0.12	0.15	0.78
Item 13	0.11	0.12	-0.11	0.11	0.23	1.39
Item 14	0.10	0.11	-0.10	0.11	0.202	1.28
Item 15	0.05	0.12	-0.03	0.12	0.09	0.54
Item 16	-0.11	0.12	0.11	0.11	-0.23	-1.39
Item 17	-0.20	0.12	0.20	0.11	-0.41	-2.00
Item 18	0.01	0.12	-0.01	0.11	0.03	0.16
Item 20	-0.09	0.11	0.09	0.11	-0.18	-1.12
Item 21	0.01	0.12	-0.02	0.11	0.03	0.19
Item 22	-0.04	0.15	0.04	0.14	-0.09	-0.41
Item 23	-0.09	0.11	0.09	0.11	-0.19	-1.19
Item 24	-0.09	0.11	0.09	0.11	-0.19	-1.20
Item 25	-0.08	0.11	0.08	0.11	-0.16	-1.05
Item 26	0.03	0.14	-0.03	0.13	0.06	0.31
Item 27	0.32	0.11	-0.32	0.11	0.64*	4.03
Item 28	0.02	0.11	-0.02	0.11	0.03	0.22
Item 29	0.29	0.15	-0.29	0.15	0.58*	2.68
Item 30	0.07	0.13	-0.07	0.13	0.14	0.76
Item 31	-0.24	0.12	0.25	0.11	-0.49*	-2.92
Item 32	-0.13	0.12	0.13	0.11	-0.27	-1.63
Item 33	-0.00	0.12	0.01	0.12	-0.00	-0.05
Item 34	0.05	0.12	-0.05	0.11	0.09	0.56
Item 35	0.07	0.12	-0.07	0.11	0.15	0.89
Item 36	-0.16	0.12	0.16	0.11	-0.33	-1.94
Item 37	0.16	0.13	-0.16	0.12	0.32	1.79
Item 39	-0.15	0.12	0.151	0.12	-0.30	-1.74
Item 40	-0.03	0.13	0.03	0.12	-0.07	-0.38

5.4.3.4 Item fit approach

INFIT MNSQ values between the ranges of 0.70 to 1.30 are used to detect DIF in items for male and female respondents. The items with INFIT MNSQ values outside this acceptable range for males and females are assumed to be misfitting items, or items with DIF. However, it is evident from Table 5.7, (p. 139) that the final 36 items appearing in the Grade 10 mathematics test recorded INFIT MNSQ values that are within the predetermined range (0.70 to 1.30) for all respondents, both male and female. Therefore, based on INFIT MNSQ criteria, it is clear that gender DIF is not a significant problem in the 36 Grade 10 mathematics test items. That said, 16 items are easier for females and twenty items are more difficult for males; this, however, does not mean that there is DIF between male and female students. Instead, it indicates that Grade 10 male students in Port Moresby performed better compared to their female counterparts on the mathematics test.

Table 5. 7 Test item fit statistics for gender (male and female)

	Female			Male		
Item No	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ
Item 01	-0.11	0.12	0.99	0.11	0.12	1.01
Item 02	0.05	0.13	1.00	-0.05	0.13	0.88
Item 05	0.07	0.11	1.06	-0.07	0.11	1.01
Item 06	0.05	0.13	1.02	-0.05	0.13	1.06
Item 07	-0.08	0.11	0.96	0.08	0.11	0.98
Item 08	0.19	0.11	1.04	-0.19	0.11	0.98
Item 09	-0.13	0.12	0.88	0.13	0.12	0.95
Item 10	0.11	0.12	1.05	-0.11	0.12	1.02
Item 11	-0.06	0.13	1.08	0.06	0.13	0.96
Item 12	0.07	0.12	0.99	-0.07	0.12	1.01
Item 13	0.11	0.12	0.91	-0.11	0.12	1.03
Item 14	0.10	0.11	0.92	-0.10	0.11	0.96
Item 15	0.05	0.12	0.95	-0.05	0.12	0.97
Item 16	-0.12	0.12	1.13	0.12	0.12	0.98
Item 17	-0.20	0.12	0.98	0.2	0.12	1.02
Item 18	0.01	0.12	0.92	-0.01	0.12	0.96
Item 20	-0.09	0.11	1.04	0.09	0.11	0.93
Item 21	0.02	0.12	1.07	-0.02	0.12	1.10
Item 22	-0.04	0.15	0.95	0.04	0.15	1.01
Item 23	-0.1	0.11	1.04	0.1	0.11	0.95
Item 24	-0.1	0.11	0.92	0.1	0.11	0.95
Item 25	-0.08	0.11	0.94	0.08	0.11	0.91
Item 26	0.03	0.14	1.01	-0.03	0.14	1.01
Item 27	0.32	0.11	1.03	-0.32	0.11	1.08
Item 28	0.02	0.11	0.9	-0.02	0.11	0.94
Item 29	0.29	0.15	1.06	-0.29	0.15	1.09
Item 30	0.07	0.13	1.03	-0.07	0.13	0.99
Item 31	-0.25	0.12	0.99	0.25	0.12	1.02
Item 32	-0.14	0.12	1.06	0.14	0.12	1.06
Item 33	0.00	0.12	0.98	0.00	0.12	0.88
Item 34	0.05	0.12	1.01	-0.05	0.12	0.99
Item 35	0.07	0.12	1.01	-0.07	0.12	0.95
Item 36	-0.16	0.12	1.00	0.16	0.12	1.03
Item 37	0.16	0.13	1.11	-0.16	0.13	0.99
Item 39	-0.15	0.12	1.05	0.15	0.12	1.02

5.4.3.5 Group plot methods

The third approach used in RMM to identify DIF in this study is the group plot method. As with the above-discussed analysis, the group plot method is used to detect DIF due to gender. Figure 5.3 shows an example of a non-suspect item (item 06) while Figures 5.4,5.5 and 5.6 show the item characteristics curve of items 27,29 and 31, which are identified to be suspects of DIF in the previous section (item threshold approach). Figure 5.5 shows a comparison between male (light green curve) and female (dark blue curve) average scores on item 06 at each autonomy level, and indicates non-suspect DIF between female and male respondents on that item.

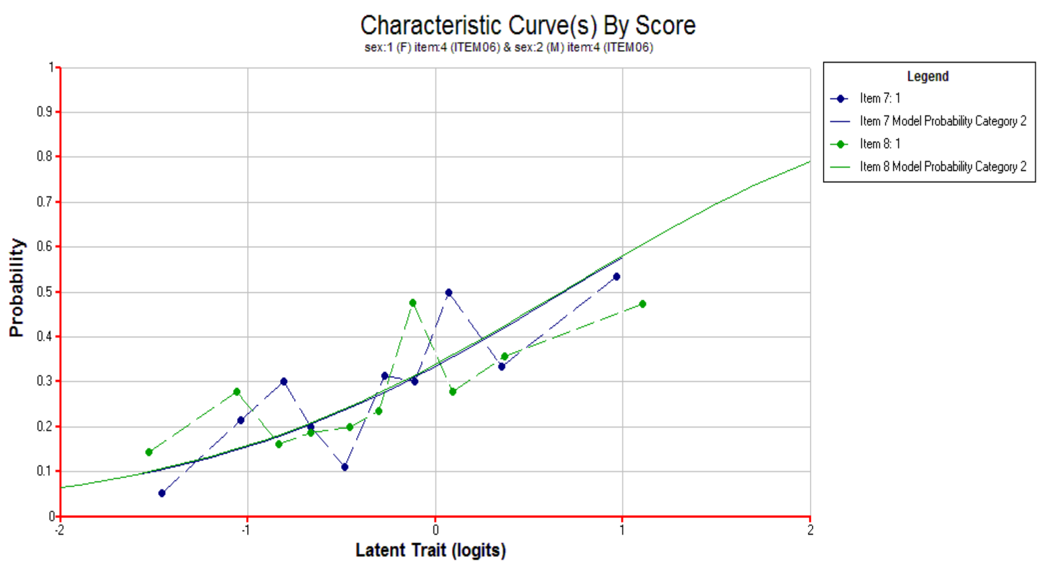


Figure 5. 3 ICC for Item 06 showing non-bias between male and female respondents

It can be seen from Figure 5.4 (item 27) and Figure 5.5 (item 29) that the ICCs for males (light green curve) are clearly higher than those for females (dark blue curve), which means that the males stand a greater chance than females of getting these items correct at the same ability level. These two items are suspected of DIF, with absolute difference values above +0.5 logits. This indicates that these two items are more in favour of males than females. On the other hand, the ICC shown in Figure 5.6 (item 31) is mostly higher for females than for males. This means that the item is biased in favour of the female respondents.

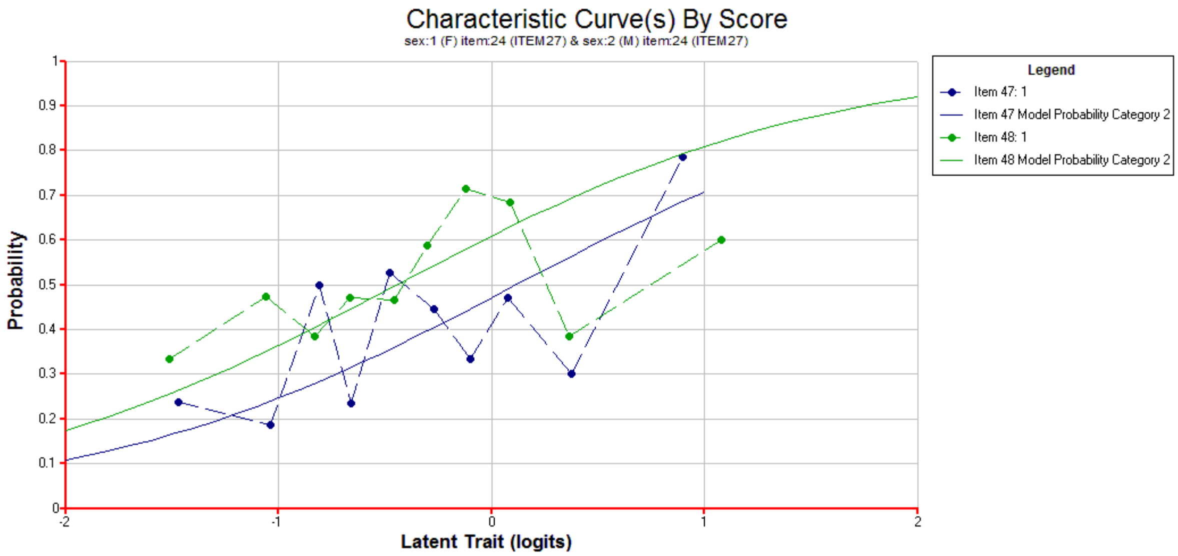


Figure 5. 4 ICC for items 27 biased in favour of male respondents

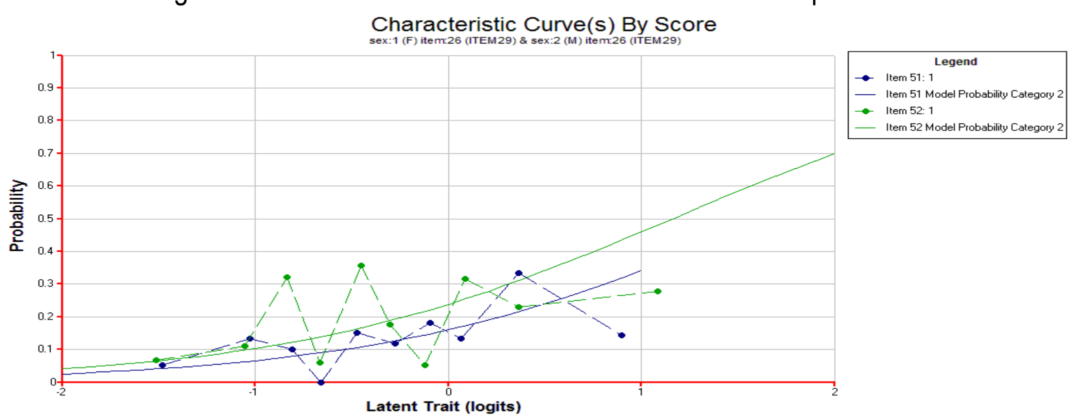


Figure 5. 5 ICC for items 29 biased in favour of male respondents

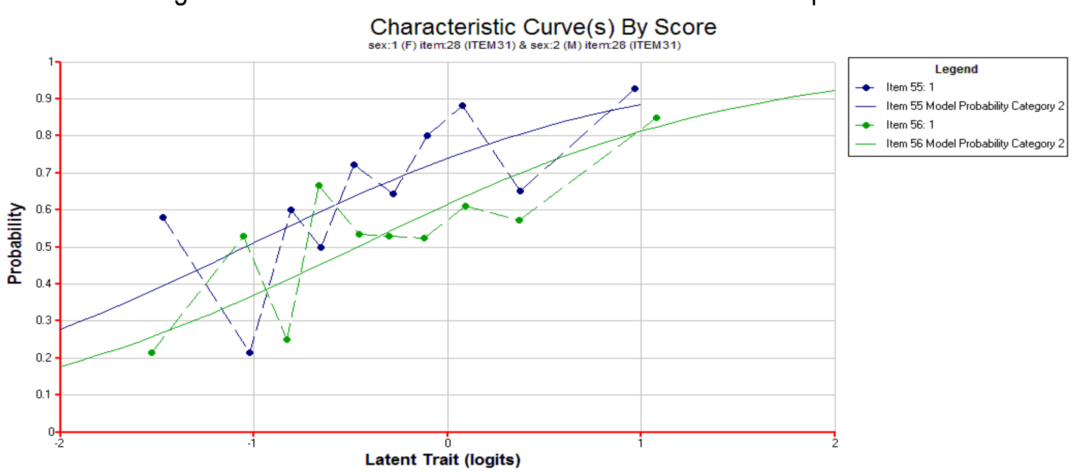


Figure 5. 6 ICC for items 31 biased in favour of female respondents

The DIF in Figure 5.4 (item 27), Figure 5.5 (item 29) and Figure 5.6 (item 31) shows that, though respondents have the same location (as evidenced by the analysis of the items), they scored differently depending on their gender. More importantly, the aforementioned figures show that, given the same total score, there is a difference by gender. According to Andrich and Styles (2004), this total score causes the items to show an opposite effect to some degree; a few items show a small amount of DIF, and a few items show the opposite to what is shown in Figures 5.4, 5.5 and 5.6. Therefore, if these items are to be studied further in terms of their relationships with the content of teaching, further investigations along these lines would need to be carried out.

5.4.4 Summary of the findings of the three methods

Overall DIF analysis provided information associated with equity and fairness of item functioning and respondent ability difference, thereby providing equity and fairness for disadvantaged groups. This analysis was carried out through three different methods (INFIT MNQS, item threshold and group plot approach) that identified DIF items with gender. These methods were utilised because they are not sample dependent. The item fit approach examined the INFIT mean square statistics of different sub-groups within the acceptable range of 0.70 to 1.3. This approach could not identify any item as DIF for two significant reasons; (1) the INFIT mean square statistics of the 36 items for the Grade 10 sample already existed within the acceptable range; and (2) the INFIT mean square statistics are for examining all items' fit within the Rasch theoretical curve. The item threshold approach and group plot approach provide similar results. Between these two, it is observed that the item threshold approach was the best method to provide the details of item thresholds and ability for male and female cohorts. The group plot approach presented graphs for item function through the range of ability for males and females.

The empirical findings obtained through the different methods, the quality of item in regard to fairness, and the ability gap between males and females, are all associated with DIF. The item threshold approach provided more specific findings; namely, that there were some items that favoured males and some that favoured females. Items 27, 29 and 31 were consistently found to have DIF between female and male students, both with item threshold and group plot approach. Items 27 and 29 were biased in favour of males, and item 31 is biased in favour of females. Therefore, this result indicates that there was a significant difference between the performances of male and female students in the Grade 10 test.

Though the overall test statistics indicate no significant difference in DIF for gender, however, the individual item analysis indicated that items 27, 29 and 31 have DIF within the RMM approaches. These three items were removed from the test for three reasons: a) the three items were not within the set statistical criteria for DIF, b) with a sample size of 355, the effect of DIF on the item results would be visible; and c), there were 36 items that were tested for DIF, and the results of these could have been affected considerably by the three items with DIF. This is because the number of items was small and the DIF influence in favour and against males and females was not evenly distributed, with items 27 and 29 favouring males and item 31 favouring females (Wu & Adams, 2007; Wu et al., 2016).

The three items with DIF detected featured content related to trigonometric functions under the topics Pythagoras theorem and trigonometric application, respectively. The observed gender differences on these topics may have been influenced by culture rather than content competence, because males are more dominant in the application aspect of mathematics in PNG, compared to females (Leder et al., 1996; Sukthankar, 1995). These three items (27, 29 and 31) show that there was difference between the male and female students in understanding and solving trigonometric functions relating to the PNG culture.

5.4.5 Final instrument

A final analysis was conducted through instrument calibration. As discussed earlier, four items showed DIF because they did not confirm to the RMM rules. The individual fit statistics in Table 5.3, (p.128) show that none of the items displayed insignificant misfit, with the INFIT MNSQ ranging from 0.92 to 1.11, *t*-value from -2 to 2 and biserial correlation (rpb) value from 0.07 to 0.4. To confirm overall statistical fit, the 36 items are again examined. As shown in Table 5.3, the final 36 items are discriminating between higher and lower abilities at a lower to certain degree. For instance, items 02, 07, 13, 14, 15,18,24,33 and 35 are very good items because the discrimination power is within the acceptable range of 3.0 and 3.9. These items are discriminating well and indicate that generally, the examinees who responded to the item correctly also did well on the test, while the examinees who responded to the item incorrectly tended to do poorly on the overall test (Adedoyin & Mokobi, 2013; Quaigrain & Arhin, 2017; Wu et al., 2016). On the other hand, items 06, 19, 21, 26, 27, 29, 38 and 40 had r_{ph} values from 0.07 to 0.19 below the acceptable value. Although, these items are discriminating between high and low ability levels, they did so at a poor level.

Detailed tables and figures of the Item Characteristics Curves (ICCs) for each item are provided in Figure 5.7 (p.143) as an example. As it can be observed from this figure, the items discriminated well with high and low performing respondents. The interpretation of the item correlation statistics, which are the correlation between the reposes for an item and the participants' total scores on the test items, is the same as that of item-total correlation (Johnsen, 2017; Wu et al., 2016). This implies that all participants had similar probabilities of answering the items correctly, regardless of their total test score (Johnsen, 2017; Plouffe, Paunonen, & Saklofske, 2017; Wu et al., 2016).

Question 09 is one example of an item (Figure 5.7) whose category probability curve, ICC and biserial point improved after removal of misfitting items. This item has an INFIT MNSQ value less than 1 and excellent discrimination, with a biserial point of 0.43 (Wu & Adams, 2007). This is evident in the characteristics curve by category, with an INFIT MNSQ of 0.91 and the correct answer's (B) observed curve steeper than the expected curve, resulting in high discrimination (Wu & Adams, 2007; Wu et al., 2016). Further, the graph in Figure 5.7 shows the modelled curve for item 09 and the matching empirical curve. This graph demonstrates the characteristics of the incorrect responses (the distractors). In particular, it shows the proportion of students in sequence of the ability grouping that responded with each of the possible responses. As shown in Figure 5.7's characteristics curve by category, answer B (which is the correct answer), was chosen by 127 respondents, with the other distractors evenly chosen and distributed among the respondents, with the three distractors having negative values, and the correct answer having a positive value.

```
------------------------------------------------------------------------------
Item 7
------
item:7 (ITEM09)
Cases for this item      346   Item-Rest Cor.  0.43    Item-Total Cor.   0.50
Item Threshold(s):      0.24   Weighted MNSQ   0.91
Item Delta(s):          0.24
------------------------------------------------------------------------------
 Label     Score      Count    % of tot   Pt Bis      t  (p)    PV1Avg:1 PV1 SD:1
------------------------------------------------------------------------------
   A        0.00        73      21.10    -0.18    -3.35(.001) -0.58      0.58
   B        1.00       127      36.71     0.43     8.75(.000)  0.05      0.67
   C        0.00        51      14.74    -0.16    -3.09(.002) -0.64      0.53
   D        0.00        94      27.17    -0.15    -2.85(.005) -0.53      0.60
   Q        0.00         1       0.29    -0.14    -2.65(.008) -2.27      0.00
==============================================================================
```

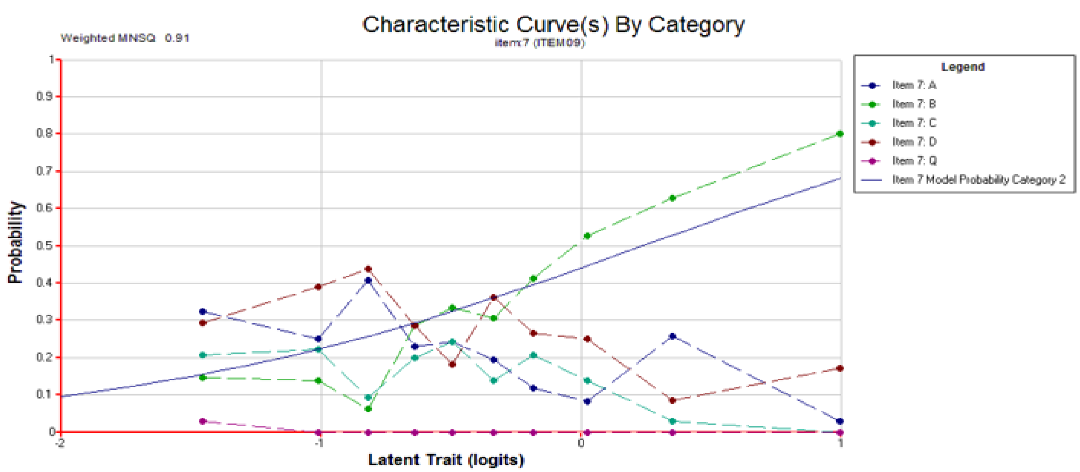


Figure 5. 7 Item 09 showing improved characteristics curve by category

5.4.6 Person-item map

The person-item map in Figure 5.8 (p.145) displays the location of person abilities and item difficulties, respectively, along the same latent dimension in Rasch logit units (Boone et al., 2013; Stelmack, Szlyk, Stelmack, & Babcock-Parziale, 2004). The distribution of person parameter with his/her ability level is located on the scale on the left of the map, while item difficulty distribution appears on the right section of the map (Boone & Scantlebury, 2006; Boone et al., 2013). The person who could perform best is located on the top of the map and the least able person is found at the base of the map (Boone & Scantlebury, 2006; Boone, Townsend, & Staver, 2011). The mean of the item difficulties on Figure 5.8 is adopted by default as the 0 point to start the mapping process (Bond et al., 2007; Wu et al., 2016). As shown in Figure 5.8, item 16 is calculated as being at the mean of the item difficulty estimates close to 0 logits and is located at the 0 point on the item-person map. The rest of the items are spread above and below 0 to represent their difficulties relative to item 16 that is located at 0. For instance, Figure 5.8 shows that a person with an ability estimate of 0 logits has a 50% probability of succeeding (or failing) on item 16 (Bond et al., 2007; Boone & Scantlebury, 2006). That same person have a greater than 50% chance of succeeding on items less difficult than item 16 (items such as 14,19, 26, 30,34 etc) and less than a 50% probability of succeeding on items that are more difficult than item 16 (i.e items 2,10,18, 20 ,20,24 etc).

Furthermore, Figure 5.8 also shows that for items 5, 20 and 24, the difficulty is highest and is located above logit of 1. These items are difficult, and so only a few respondents, all of whom had higher abilities responded correctly. Due to these items' difficulty level, respondents' probability of answering correctly is less than 50% (Boone & Scantlebury, 2006; Stelmack et al., 2004; Wei, Liu, & Jia, 2014). On the other hand, items 1, 12 and 17 are easy because they are located below logit -1 and respondents with both low and high ability level answered the items correctly.

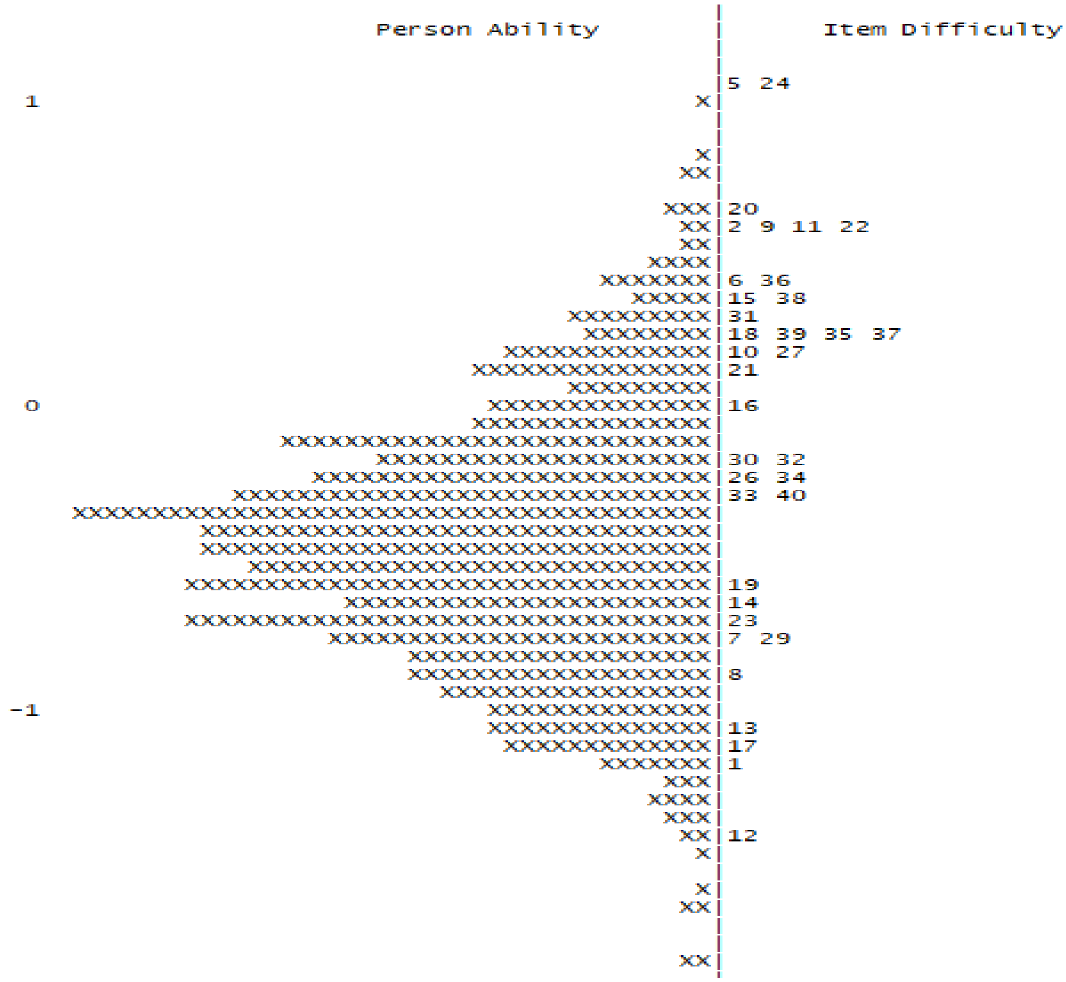


Figure 5. 8 Person-item map of Grade 10 test items

5.5 Transformation to W-scores

As discussed in Chapter 3, the WLE values are transformed into W-scores to eliminate negative value, and the need for decimal points (Benson et al., 2016; Linacre, 2009; Woodcock, 1999). Microsoft Excel is used for this transformation, using the formula W=9.1024 x logits + c, where c is a constant term, selected to eliminate the negative values (Benson et al., 2016; Woodcock, 1999). As mentioned in Chapter 3, the constant value used in the study is 500 in consistent with PISA studies. The W-score set the sign of the item difficulty and the person ability on the scales so that low values imply either low item difficulty or low person difficulty. High W-scores imply either high item difficulty or high person ability (Linacre, 2009;

Woodcock, 1999; Wu et al., 2016). These scores are used in the subsequent chapters on structural equation modelling and hierarchical linear modelling.

5.6 Summary

In this study, overall fit statistics were considered because these represent an indication of how well the data fits the Rasch model. Therefore, the overall test and the individual items are analysed using the INFIT MNSQ, ICCs and category probability curves. It was evident that by removing the misfitting items, the items produced better fitting ICCs, category probability curves and better fit statistics in most cases. Hence, the analysis reveals that the test followed principles of good test design because the data fit the model well after removing four faulty items. Further, analysis in this chapter has shown that the person-item location distribution showed a reasonable spread of items and respondents, and the test targeted well at the sample. Results of the analysis suggest that items based on complex procedures and those items involving problem solving were beyond the competencies of respondents. Analysis of the DIF in relation to gender was necessary, to examine the fairness and equity of the test in identifying biased items. The investigation of DIF showed a significant difference on the item level analysis and detected three items with DIF; two items biased towards males, and one towards females. The Rasch model's item threshold and group plot approaches were used to identify the DIF items. The overall item analysis was important for test and examination question improvement for students. Similar procedures and techniques employed in this chapter are used to validate and verify the Grade 12 mathematics items in the next chapter of this thesis.

Chapter 6: Grade 12 mathematics test validation

6.1 Introduction

This chapter follows on from Chapter 5, using the same chapter structure and analysis techniques to evaluate the mathematics test instrument administered to the Grade 12 students in Port Moresby. The test is analysed to identify its utility through the Rasch measurement model (RMM) to ensure that the items are reliable, valid and objective. This chapter begins by explaining the development of the test items through the table of specification (TOS) as discussed in Chapter 5. The next section explains the background of RMM that underpins the concept of item difficulty and person ability in responding to the test items. Then test item analysis procedure and results are discussed using RMM, and the chapter ends with a summary, highlighting the key ideas discussed.

6.2 Development of the mathematics test items

The development process for the Grade 12 test items employed in this chapter is similar to that described in Chapter 5. However, in this chapter the researcher combined both advanced and general mathematics units/topics together to form a single uniform test (see Chapter 1 for details), to be administered to students enrolled in both advanced and general mathematics. Since there were two different groups of students, similar units/topics studied for both mathematics were selected, and the TOS shown in Table 6.1 was designed accordingly. In order to maintain equity and fairness between the two groups of respondents, other units/topics that are different from each cohort were not included in the test. One of the reasons for having one test for both mathematics cohorts was to do with the practicalities involved with collecting data from schools and data analysis. For instance, if both groups of students received a distinct test, the scheduling of class (time) would be more difficult. Further, some secondary schools had very few students enrolled in the advanced mathematics course.

The TOS obtained from the Measurement Service Division (MSD) (see Appendix F) has cognitive domains distributed according to the Grade 12 advanced and general mathematics national examinations. In these examinations, low-level questions comprised 30%, medium-level questions consisted 50% and high-level questions made up 20% of the items (NDOE, 2017). The TOS for this chapter also adopts the same distribution of percentage as the TOS from MSD for the cognitive domains, with the same proportionality of items in the content areas that assess different mathematics skills (see Table 6.1). No questions/items specifically for advanced or general mathematics were used, but instead items were selected based on the similar topics covered for both types of mathematics course. Similar procedures in Chapter 5 were applied to estimate the percentage of content of the test items covered in the TOS, which was derived from Grade 11 and 12 units for both advanced and general mathematics. These techniques were employed to ensure that the test items in this study cover approximately a similar content area for alignment and uniformity purposes. For instance, basic algebra is covered in both advanced and general mathematics so an item that reflected this topic was chosen by the researcher. As discussed in Chapter 3, the selection of the item types was made possible through the researcher's experience teaching both advanced and general mathematics in secondary schools in PNG.

Table 6. 1 Combined Grade 12 advance and general mathematics test Table of Specification (TOS)

Content Area		Cognitive Domain				Percentage	
Unit (s)	Topic (s)	Low 30 %	Medium 50%	High 20 %	Score	Score %	MSD Content %
11.1 Number and Application	Basic Numeracy	31			8	20%	20%
	Ratio and proportion	4	11				
	Basic Algebra		2,37,38,29				
	Measurement			39			
11.3 Probability and Statistics/11.3 Managing Data	Exploring data	9	19,33		10	25%	22%
	Managing data	8	17,18				
	Probability	28	35	36			
	Data collection and presentation	32					
11.4 Geometry	Lines, angles, triangles and regular polygons	10	26		4	10%	12%
	Geometry construction		22,27				
12.1 Patterns and Algebra/12.4 Algebra and Graphs	Equations	5, 16	3,40	1	12	30%	26%
	Algebraic expressions		6,7	15			
	Graphs	12,13, 14					
	Sequences and series		30				
12.2 Trigonometry and Vectors/11.5 Trigonometry	Trigonometry		20,21	23, 34	4	10%	12%
12.2 Money	Using simple algebraic manipulation of financial formulae			24,25	2	5%	8%
	Total	12	20	8	40		
	Researcher Set Goal %	20%	50%	30%		100%	100%

6.3 Rasch measurement model

The item fit statistics in the Rasch measurement model (RMM) (discussed in Chapter 5) are utilised to evaluate individual student's responses to the test. Using ConQuest (version 4.0), these item fit statistics are used to assist and detect discrepancies between the RMM prescriptions and data, in order to verify the usability and portability of test analysis (Bond et al., 2007; Wu et al., 2016). Accordingly, the results of this analysis are determined through the test statistics of the 40 mathematics test items, which are then analysed and presented under key points to investigate their psychometric properties, and to locate items that are mis-fitting to the RMM.

6.3.1 Overall statistics

Analysis of the test was carried out using responses from 379 Grade 12 students from the surveyed 16 secondary schools in Port Moresby. The 40 item multiple-choice test items were subjected to fit the RMM, as discussed in Chapter 5, to determine whether all items fitted the model. As such, an initial run of the test data and the results of each item and person were obtained through ConQuest. The results of the data analysis were used to checked for validity of the items. In this chapter, four processes are used to perform the data analysis; 1. Running ConQuest, 2. Checking model fit, 3. Revising the model and 4. Re-running ConQuest until the data fit the model. From the initial Rasch analysis, the summary fit statistics (Table 6.2 p.153) such as estimates difficulty, INFIT MNSQ, standardized *t*-value and discrimination power are generated.

From the analysis, person and item parameters are estimated for the differences between their actual and expected responses. It is apparent that these fit statistics continue to highlight the assumption of unidimensionality in RMM, which was discussed in Chapter 3. In order to confirm this notion, the responses and function of items have to fit the theoretical characteristic curve of persons and items (Bond et al., 2007;

Wu & Adams, 2007). The distribution of the observed probabilities of responses are presumed to be normal for the purpose of the fit test (Bond & Fox, 2015; Bond et al., 2007; Wu & Adams, 2007).

The INFIT MNSQ results of the 40 items are within the acceptable value range of 0.70 to 1.30 (Bond & Fox, 2015; Wright, 1996; Wu & Adams, 2007) that was discussed in Chapter 3. However, the *t*-values for five items (04, 06, 13, 16 and 35) are not within the acceptable range of -2 to 2 (see Table 6.2). This is reflected in the analysis results appearing in Table 6.2, where the INFIT MNSQ ranged from 0.93 to 1.09 and the t-value from -2.3 to 3.3. The *t*-values and point biserial correlation with the ICCs from other statistics fit that were discussed in Chapter 5 are used as criteria for the deletion of items, as they represent how the data perfectly fits the RMM (Linacre, 2002; Wu et al., 2016). The items that did not conform to the RMM rule were removed one by one through calibration, until 35 items remained that fit the RMM.

In terms of reliability, the item separation index is used to estimate the internal consistency of the test. The separation index appearing in Table 6.2 is very good, with a value 0.98 (Wright & Stone, 1999), demonstrating that the estimation of a person's ability is consistent across the test. In this case, a figure of 0.98 indicates that the test-taker had been reliably assessed by the test (Wright & Stone, 1999). Table 6.2 shows that the item estimate difficulty ranged from 1.56 to -1.33 logits, which is associated with standard error that ranges from 0.14 to 0.11.

Table 6. 2 Results of the initial ConQuest analysis of the Grade 12 test items

Item No	Estimate	Error	INFIT MNSQ	ZSTD (t)	Pt Bis for correct answer	Item No	Estimate	Error	INFIT MNSQ	ZSTD (t)	Pt Bis for correct answer
Item 01	-1.15	0.11	0.98	-0.4	0.23	Item 21	-0.25	0.107	0.98	-0.8	0.24
Item 02	0.94	0.13	1.02	0.4	0.12	Item 22	-1.33	0.12	0.96	-0.7	0.28
Item 03	-0.36	0.11	1.07	-2.0	0.31	Item 23	-0.02	0.11	1.00	0.2	0.18
Item 04	-0.42	0.11	0.96	2.8	0.02	Item 24	0.59	0.12	1.09	1.8	-0.04
Item 05	-0.30	0.10	1.02	-1.8	0.31	Item 25	0.68	0.12	1.05	0.9	0.06
Item 06	-0.37	0.10	1.02	3.3	0.01	Item 26	0.61	0.12	0.92	-1.8	0.41
Item 07	0.16	0.11	1.04	-1.4	0.32	Item 27	1.28	0.13	1.00	0.0	0.16
Item 08	1.57	0.13	0.98	0.2	0.08	Item 28	0.78	0.12	0.96	-0.6	0.25
Item 09	-0.67	0.11	1.05	0.8	0.13	Item 29	0.61	0.11	1.08	1.6	-0.03
Item 10	0.28	0.11	0.94	1.0	0.12	Item 30	-0.49	0.11	0.96	-1.4	0.31
Item 11	-1.40	0.12	1.01	-0.4	0.24	Item 31	0.54	0.11	0.95	-1.1	0.31
Item 12	0.90	0.12	0.93	0.8	0.05	Item 32	0.49	0.12	1.04	0.9	0.08
Item 13	-0.40	0.10	0.94	-2.3	0.35	Item 33	0.82	0.12	1.06	1.0	0.00
Item 14	0.38	0.10	1.01	0.4	0.17	Item 34	0.72	0.12	1.02	0.3	0.13
Item 15	0.6	0.11	0.93	-1.4	0.34	Item 35	-0.56	0.11	1.09	3.2	-0.01
Item 16	-0.05	0.10	0.94	-2.3	0.35	Item 36	0.55	0.11	1.04	0.9	0.08
Item 17	-1.28	0.12	1.04	0.8	0.11	Item 37	0.32	0.11	0.94	-1.6	0.35
Item 18	-1.40	0.12	1.01	0.2	0.13	Item 38	-0.94	0.11	0.95	-1.4	0.32
Item 19	-0.74	0.11	1.01	0.4	0.18	Item 39	0.33	0.11	0.99	-0.1	0.21
Item 20	-0.16	0.11	0.95	-1.8	0.32	Item 40	-0.11*	0.11	0.96	-1.7	0.33

Separation Reliability = 0.981 Chi-square test of parameter equality = 1595.04, df = 34, Sig Level = 0.000

Table 6.3 (p.153) indicates that items 04, 06, 24, 29 and 35 are to be omitted, because of lack of fit to the model. These items are measuring something other than the intended content; that is, they can be considered test content irrelevant (Bond et al., 2007; Wu & Adams, 2007).

The biserial correlation (discrimination index) for the correct answers for each item ranges from -0.04 to 0.41 (Table 6.2). This indicates that most of the items are discriminating among the higher and lower achieving students. However, four items (24, 29, 34 and 35) are not discriminating at all because of their negative point biserial correction (Wu et al., 2016), which demonstrates that those respondents who were doing poorly on the overall test got those items correct. After rechecking and confirming the answer key for confirmation and the criteria for *t*-values as mentioned earlier, the researcher decided to remove items 04, 06, 24, 29 and 35 due to their violation of the RMM criteria. As shown in Table 6.3, this calibration process resulted in improved fit statistics for the misfit items, and at the same time the statistical fit for other items improved and some regressed slightly, but there was no longer any significant misfit.

Table 6. 3 Results of the final ConQuest analysis of the Grade 12 test items

Item #	Estimate	Error	INFIT MNSQ	ZSTD (t)	Pt Bis for correct answer	Item #	Estimate	Error	INFIT MNSQ	ZSTD (t)	Pt Bis for correct answer
Item 01	-1.21	0.12	0.98	-0.4	0.25	Item 21	-0.29	0.11	0.99	-0.4	0.26
Item 02	0.92	0.13	1.05	0.7	0.12	Item 22	-1.39	0.12	0.96	-0.6	0.29
Item 03	-0.4	0.11	0.96	-1.3	0.31	Item 23	-0.05	0.11	1.03	1.1	0.18
Item 04						Item 24					
Item 05	-0.35	0.11	0.96	-1.3	0.33	Item 25	0.66	0.12	1.09	1.5	0.04
Item 06						Item 26	0.59	0.12	0.91	-1.9	0.41
Item 07	0.13	0.11	0.96	-1.1	0.32	Item 27	1.28	0.14	1	0.1	0.18
Item 08	1.57	0.15	1.04	0.4	0.08	Item 28	0.77	0.14	0.98	-0.3	0.25
Item 09	-0.72	0.11	1.05	1.5	0.13	Item 29					
Item 10	0.26	0.11	1.06	1.6	0.12	Item 30	-0.54	0.121	0.96	-1.3	0.33
Item 11	-1.46	0.12	1.00	0.0	0.23	Item 31	0.52	0.11	0.95	-1	0.33
Item 12	0.88	0.13	1.08	1.2	0.08	Item 32	0.47	0.12	1.08	1.7	0.06
Item 13	-0.44	0.11	0.95	-1.7	0.36	Item 33	0.81	0.12	1.1	1.6	0.02
Item 14	-0.42	0.11	1.04	1.5	0.16	Item 34	0.69	0.12	1.03	0.4	0.16
Item 15	0.58	0.12	0.93	-1.4	0.37	Item 35					
Item 16	-0.09	0.11	0.94	-1.9	0.37	Item 36	1.57	0.15	1.02	0.2	0.14
Item 17	-1.35	0.12	1.05	1.0	0.12	Item 37	0.29	0.11	0.95	-1.2	0.33
Item 18	-1.47	0.12	1.03	0.5	0.15	Item 38	-0.99	0.11	0.97	-0.8	0.31
Item 19	-0.79	0.11	1.02	0.7	0.19	Item 39	0.30	0.11	1.01	0.2	0.22
Item 20	-0.20	0.11	0.96	-1.3	0.33	Item 40	0.14*	0.11	0.95	-1.5	0.42

6.3.2 Individual statistics

The 40 test items were checked for item fit statistics, ICC, and the functioning of the categories for each single item. These statistical diagnostic procedures were carried out through graphical analysis that provided the item characteristic curve (ICC) that is shown in Figure 6.1, which is used to check item function. Based on these diagnostic procedures, items that did not abide by RMM rules were identified and removed one by one, as discussed in Chapter 5.

The negative biserial correlation indicated that respondents who did poorly on the overall test did well on those items. Subsequently, the ICC and category probabilities for each item were checked and the data calibrated, resulting in an improved statistical fit (though some items regressed slightly). Figures 6.1 and 6.2 (items 38 and 32; see below) for the initial run are used as examples, to explain ICC and category probability curves indicating different behavioural pattern for the items (Wu & Adams, 2007; Wu et al., 2016). Item 38 shows an under fit ICC with very good discrimination, and item 32 shows most students choosing the correct answer with distractors well spread. Both of these items were retained following the statistical criteria.

As seen in Figure 6.1, item 38 has a discrimination index (Item-Rest Cor) of 0.32, with a good correlation between the score on the item and the total score of the test. This index further indicates that the discrimination is good and discriminated between respondents according to their ability (D'Sa & Visbal-Dionaldo, 2017; Wu & Adams, 2007). In Figure 6.2, the ICC provides useful information about the behaviour of an item. More importantly, the observed curves are relatively flatter, indicating that there is a relationship between the response categories and the level of autonomy (Wu & Adams, 2007). The item's weighted INFIT MNSQ value of 0.95 indicates that it is more highly discriminating than the model predicts, because the INFIT MNSQ value is less than 1.

As shown in Figure 6.2 (item 32), more respondents (115) chose option A as the correct answer. Distractors B, C and D were answered by 98, 79 and 69 respondents respectively. This shows that the distractors are distracting fairly among the options. Additionally, it is evident from the characteristics curve by category that the distractors are distributed fairly. The weighted MNSQ 1.09 demonstrates a fit index with an under fit item that is not discriminating as well as the model predicted, due to its INFIT MNSQ value of greater than 1 (D'Sa & Visbal-Dionaldo, 2017; Linacre, 2002).

```
Item 38
-------
item:38 (ITEM38)
Cases for this item      358   Item-Rest Cor.   0.32   Item-Total Cor.   0.41
Item Threshold(s):     -0.94   Weighted MNSQ    0.95
Item Delta(s):         -0.94
------------------------------------------------------------------------------
 Label     Score     Count    % of tot  Pt Bis      t   (p)    PV1Avg:1 PV1 SD:1
------------------------------------------------------------------------------
   A        1.00      231      64.53     0.32     6.44(.000) -0.22      0.54
   B        0.00       52      14.53    -0.17    -3.19(.002) -0.50      0.36
   C        0.00       49      13.69    -0.17    -3.30(.001) -0.52      0.52
   D        0.00       25       6.98    -0.11    -2.05(.041) -0.56      0.47
   Q        0.00        1       0.28    -0.17    -3.34(.001) -1.38      0.00
```

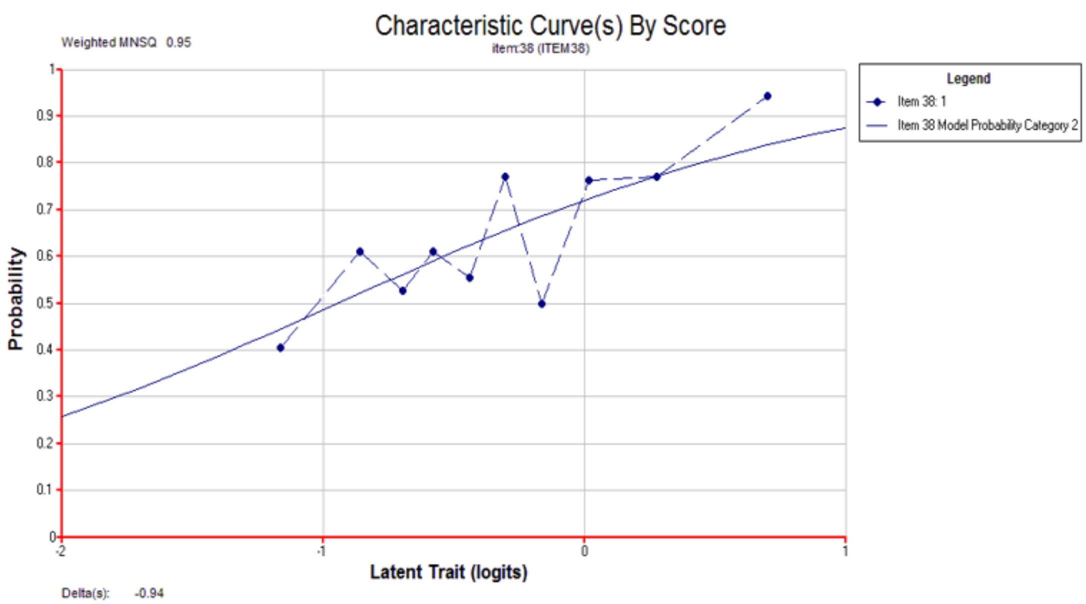


Figure 6. 1 Item characteristics curve for item 38

```
Item 32
-------
item:32 (ITEM32)
Cases for this item      362   Item-Rest Cor.   0.08    Item-Total Cor.   0.17
Item Threshold(s):      0.50   Weighted MNSQ    1.04
Item Delta(s):          0.49
------------------------------------------------------------------------------
 Label     Score      Count    % of tot  Pt Bis     t   (p)    PV1Avg:1 PV1 SD:1
------------------------------------------------------------------------------
   A        1.00       115       31.77     0.08     1.55(.122)  -0.21      0.52
   B        0.00        98       27.07     0.07     1.30(.194)  -0.31      0.55
   C        0.00        79       21.82    -0.09    -1.72(.087)  -0.42      0.51
   D        0.00        69       19.06    -0.06    -1.07(.287)  -0.46      0.47
   Q        0.00         1        0.28    -0.17    -3.33(.001)  -1.38      0.00
```

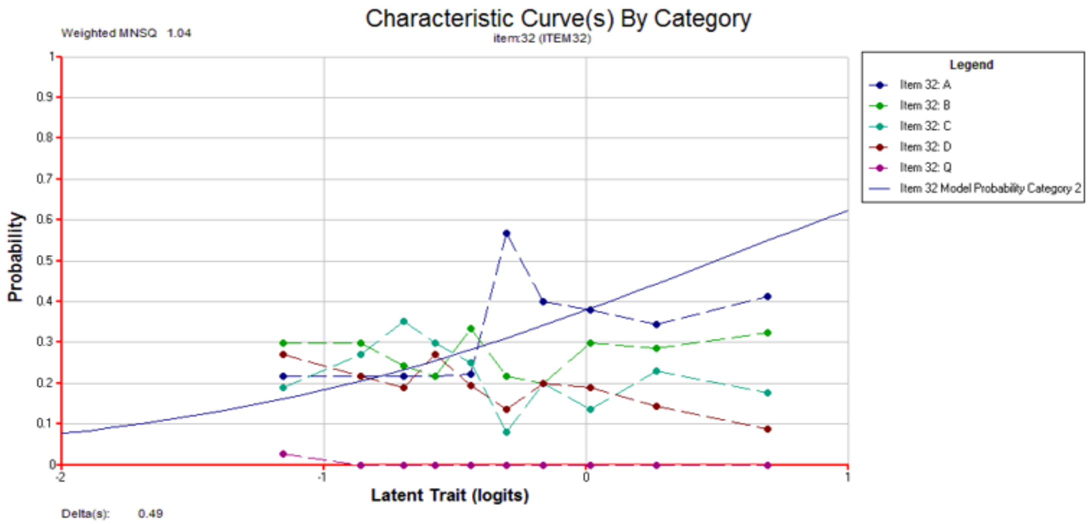


Figure 6. 2 Item characteristics curve by category for item 32

6.3.3 Differential item functioning

After analysing the utility of the items, the final 35 items are further analysed for DIF, to confirm whether the items favoured male or female students. The similar concept and procedure used for DIF in Chapter 5 is applied in this section. In this study, if the performance of respondents of different groups (males and females) with similar ability is different on a particular item, then that item is considered to be exhibiting DIF. The performance of the male and female students is controlled by the total score of the Grade 12 mathematics test items that were analysed. Similar methods to identify DIF in RMM that were discussed in Chapter 5 (item fit approach, item threshold approach and group plot methods) are employed again in this chapter.

6.3.3.1 Overall DIF level test for gender

This section analyses and compares the overall mathematics test performance for the participating Grade 12 male and female students in Port Moresby. Similar procedures and techniques of DIF analysis discussed in Chapter 5 is carried out in this chapter, to gauge the performance of males and females on the test, and so establish the existence of any bias in the test towards a particular gender group. Table 6.3 shows that the INFIT MNSQ values of the overall test for both male and female students are within the acceptable range of 0.77 to 1.30 that is used in this study (Hungi, 2005). Table 6.4 also shows the item difficulty parameter estimates for each of the 35 items. The negative sign for 17 items on the parameter estimates indicates that those items are easier for female students. On the contrary, the 18 positive items on the parameter estimates indicate that those items are more difficult for males. Overall, it can be observed that there is not much difference in the difficulty level between male and female respondents/students.

Table 6. 4 Test level gender difference statistics

Item No	Estimate	Error	INFIT MNSQ	Item No	Estimate	Error	INFIT MNSQ
Item 01	-1.21	0.12	1.01	Item 21	-0.27	0.11	1.00
Item 02	0.92	0.13	1.06	Item 22	-1.41	0.12	0.98
Item 03	-0.40	0.11	0.96	Item 23	-0.04	0.11	1.03
Item 05	0.35	0.11	0.96	Item 25	0.69	0.13	1.09
Item 07	0.13	0.11	0.96	Item 26	0.59	0.12	0.92
Item 08	1.61	0.15	1.05	Item 27	1.29	0.14	1.01
Item 09	-0.72	0.11	1.05	Item 28	0.76	0.12	0.98
Item 10	0.26	0.11	1.06	Item 30	-0.55	0.11	0.97
Item 11	-1.49	0.12	1.00	Item 31	0.54	0.12	0.95
Item 12	0.89	0.13	1.08	Item 32	0.47	0.12	1.08
Item 13	-0.44	0.11	0.95	Item 33	0.8	0.13	1.09
Item 14	-0.41	0.11	1.04	Item 34	0.7	0.13	1.03
Item 15	0.57	0.12	0.92	Item 36	1.57	0.15	1.01
Item 16	-0.09	0.11	0.94	Item 37	0.29	0.11	0.95
Item 17	-1.37	0.12	1.06	Item 38	-1	0.11	0.97
Item 18	-1.49	0.12	1.03	Item 39	0.3	0.12	0.99
Item 19	-0.8	0.11	1.02	Item 40	-0.153	0.11	0.95
Item 20	-0.2	0.11	0.96	Item 40	0.48	0.13	1.12

Table 6.4 (below) shows the mean estimates of male and female respondents for the overall test. From this table, it can be observed that the test is easier for female students than for male students. This is indicated by the parameter estimate of -0.005 for females. As discussed in Chapter 5, the actual parameter estimate for male students is 0.151 times (0.0056/0.037) larger than its standard error estimate. Therefore, the standard difference between the male and female student means is obviously insignificant. The chi-square value of 0.01 on one degree of freedom is consistent with this finding in Table 6.5 The fact that parameter estimates are less than the standard error, and that the chi-square p-value is small indicate that this difference is statistically insignificant. The difference between the actual parameter estimate of female and male students shows that female students scored 0.01 logits lower than male students (see Table 6.4). Therefore, the difference of 0.01 logits, which is much smaller than 0.5 logits, implies that the average performance levels of male and female students are not substantially different. The statistical findings in Table 6.4 show that there is no DIF in the overall mathematics test for Grade 12 students in Port Moresby. The next section in this chapter evaluates individual items of the test to ascertain whether there are DIF items for male and female students at that level.

Table 6. 5 Gender facet estimates for 35 mathematics items for Grade 12

	Sex	Estimate	Error	INFIT MNSQ	ZSTD (t)
1	Female	-0.005	0.037	1.00	0.0
2	Male	0.005*	0.037	1.05	0.5

Chi-square test of parameter equality = 0.01, df = 1

6.3.3.2 Item-level DIF for gender

After test-level analysis, the individual items in the mathematics test are further analysed for DIF between males and females. This approach is carried out due to the insignificant difference found for gender in the test level analysis. At the item-level of analysis, however, the researcher believes that DIF would be detected for gender. Item level analysis is conducted through three methods of identifying DIF in RMM, with their rationale discussed in Chapter 5.

6.3.3.3 Item fit approach

INFIT MNSQ values between the range of 0.70 to 1.30 (discussed in Chapter 5), are used to detect DIF in items between male and female respondents. The items with INFIT MNSQ values outside this acceptable range for males and females are assumed to be mis fitting items, or items with DIF. However, it is evident from Table 6.6 that the final 35 items for the Grade 12 mathematics test recorded INFIT MNSQ values within the predetermined range (0.70 to 1.30) for all respondents. Therefore, based on this INFIT MNSQ criterion, it is clear that gender DIF is not a problem in the 35 mathematics test items for Grade 12 respondents. That said, 21 items are easier for females and 14 items are harder for males. This, however, does not mean that there is DIF between male and female students; rather it indicates that Grade 12 female students in Port Moresby did poorly on those items compared male students on the mathematics test.

Table 6. 6 Test item fit statistics for gender (male and female).

	Female			Male		
Item No	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ
Item 01	-0.01	0.12	0.99	0.01	0.11	1.02
Item 02	-0.01	0.12	1.06	0.01	0.12	1.05
Item 03	0.06	0.11	0.95	-0.06	0.10	0.97
Item 05	-0.01	0.11	0.98	0.01	0.10	0.94
Item 07	-0.09	0.11	0.98	0.09	0.11	0.95
Item 08	0.27	0.15	1.01	-0.27	0.15	1.08
Item 09	-0.01	0.11	1.09	0.01	0.10	1.03
Item 10	0.05	0.11	1.04	-0.04	0.11	1.08
Item 11	-0.22	0.12	1.04	0.21	0.12	0.97
Item 12	0.05	0.13	1.04	-0.04	0.12	1.11
Item 13	0.07	0.11	0.95	-0.07	0.10	0.96
Item 14	0.19	0.11	1.04	-0.19	0.10	1.04
Item 15	-0.09	0.12	0.96	0.09	0.12	0.89
Item 16	0.03	0.11	0.95	-0.02	0.11	0.93
Item 17	-0.18	0.12	1.04	0.18	0.12	1.08
Item 18	-0.19	0.12	1.08	0.19	0.12	0.99
Item 19	-0.18	0.11	1.02	0.17	0.11	1.03
Item 20	-0.01	0.11	1.00	0.01	0.11	0.92
Item 21	0.28	0.11	1.01	-0.28	0.11	0.99
Item 22	-0.14	0.12	1.01	0.14	0.12	0.95
Item 23	0.18	0.11	1.03	-0.17	0.11	1.02
Item 25	0.28	0.13	1.13	-0.28	0.13	1.06
Item 26	0.06	0.11	0.94	-0.05	0.12	0.90
Item 27	0.07	0.14	1.01	-0.06	0.14	1.00
Item 28	-0.01	0.12	0.98	0.01	0.12	0.98
Item 30	-0.13	0.11	0.99	0.13	0.11	0.95
Item 31	0.15	0.12	0.98	-0.15	0.12	0.91
Item 32	-0.01	0.12	1.07	0.01	0.11	1.09
Item 33	-0.07	0.13	1.07	0.06	0.12	1.11
Item 34	0.08	0.13	1.04	-0.08	0.12	1.03
Item 36	-0.00	0.15	1.00	0.00	0.15	1.02
Item 37	-0.10	0.11	0.95	0.10	0.11	0.95
Item 38	-0.07	0.11	0.95	0.07	0.11	0.99
Item 39	-0.09	0.12	0.93	0.09	0.11	1.06
Item 40	-0.18	0.11	0.97	0.17	0.11	0.93

6.3.3.4 Item threshold approach

Table 6.7 below indicates that 14 items (positive values) are more difficult for males, and 21 items (negative values) are easier for females with a difference in item threshold (Abd-El-Fattah et al., 2014; Le, 2006). There are fewer items that the male students responded to correctly. The statistical criteria of d_1-d_2 >-0.50 and d_1-d_2 < 0.05 logits discussed in Chapter 5 is also employed to identify items with DIF in this chapter. As shown in Table 6.7, three items are outside the acceptable logit value of the criteria set of d_1-d_2 >-0.50 and d_1-d_2 < 0.05 and 32 items are within the acceptable range. Table 6.7 shows that items 08, 21 and 25 have a substantial amount of DIF, with absolute difference values above the 0.5 logits that is recommend by Hungi (2005). This indicates that these three items are in favour of the male respondents over the female respondents

Table 6. 7 Result of DIF for the 35 mathematics test items by item threshold approach

Item No	Female		Male		d1-d2	st (d1-d2)
	Estimate (d1)	Error (e1)	Estimate (d2)	Error (e2)		
Item 01	-0.01	0.11	0.01	0.11	-0.02	-0.13
Item 02	-0.01	0.12	0.01	0.12	-0.01	-0.08
Item 03	0.06	0.10	-0.06	0.10	0.12	0.79
Item 05	-0.01	0.10	0.01	0.10	-0.01	-0.07
Item 07	-0.09	0.11	0.09	0.11	-0.18	-1.19
Item 08	0.27	0.15	-0.27	0.15	0.54*	2.50
Item 09	-0.01	0.10	0.01	0.10	-0.01	-0.10
Item 10	0.04	0.11	-0.04	0.11	0.09	0.58
Item 11	-0.21	0.12	0.21	0.12	-0.43	-2.50
Item 12	0.04	0.12	-0.04	0.12	0.09	0.50
Item 13	0.07	0.10	-0.07	0.10	0.14	0.95
Item 14	0.19	0.10	-0.19	0.10	0.39	2.57
Item 15	-0.09	0.11	0.09	0.11	-0.18	-1.13
Item 16	0.02	0.10	-0.02	0.10	0.05	0.32
Item 17	-0.18	0.11	0.18	0.11	-0.36	-2.18
Item 18	-0.19	0.12	0.19	0.12	-0.39	-2.30
Item 19	-0.17	0.11	0.17	0.11	-0.35	-2.26
Item 20	-0.01	0.11	0.01	0.10	-0.01	-0.10
Item 21	0.28	0.11	-0.28	0.10	0.56*	3.65
Item 22	-0.14	0.12	0.14	0.12	-0.28	-1.63
Item 23	0.17	0.11	-0.17	0.11	0.35	2.25
Item 25	0.28	0.12	-0.28	0.12	0.56*	3.17
Item 26	0.05	0.11	-0.05	0.11	0.11	0.67
Item 27	0.06	0.14	-0.06	0.14	0.13	0.68
Item 28	-0.01	0.12	0.01	0.12	-0.01	-0.08
Item 30	-0.13	0.11	0.13	0.10	-0.26	-1.73
Item 31	0.15	0.11	-0.15	0.11	0.30	1.84
Item 32	-0.01	0.11	0.01	0.11	-0.03	-0.18
Item 33	-0.06	0.12	0.06	0.12	-0.13	-0.73
Item 34	0.0	0.12	-0.08	0.12	0.16	0.90
Item 36	-0.00	0.15	0.00	0.15	-0.01	-0.02
Item 37	-0.10	0.11	0.10	0.11	-0.20	-1.28
Item 38	-0.07	0.11	0.07	0.11	-0.14	-0.88
Item 39	-0.09	0.11	0.09	0.11	-0.19	-1.20
Item 40	-0.17	0.11	0.17	0.11	-0.35	-2.26
Item 40	-0.03	0.12	0.03	0.12	-0.06	-0.38

6.3.3.5 Group plot method

The group plot method is another means through which to detect DIF due to gender. Figures 6.3 shows an example of a non-suspect item (item 5), while Figures 6.4,6.5 and 6.6 show the item characteristics curve of the items 8, 21 and 25, which are identified as DIF items. It can be seen from Figures 6.4, 6.5 and 6.6 that the ICCs for males (light green curve) are clearly higher than those for female (dark blue curve). This means that the male students stood a greater chance than female students of getting these items correct at the same ability level. These items have DIF with threshold difference values above + 0.5 logits, which demonstrates that they are in favour of males, and that females require more time to learn specific content areas when compared to their male counterparts.

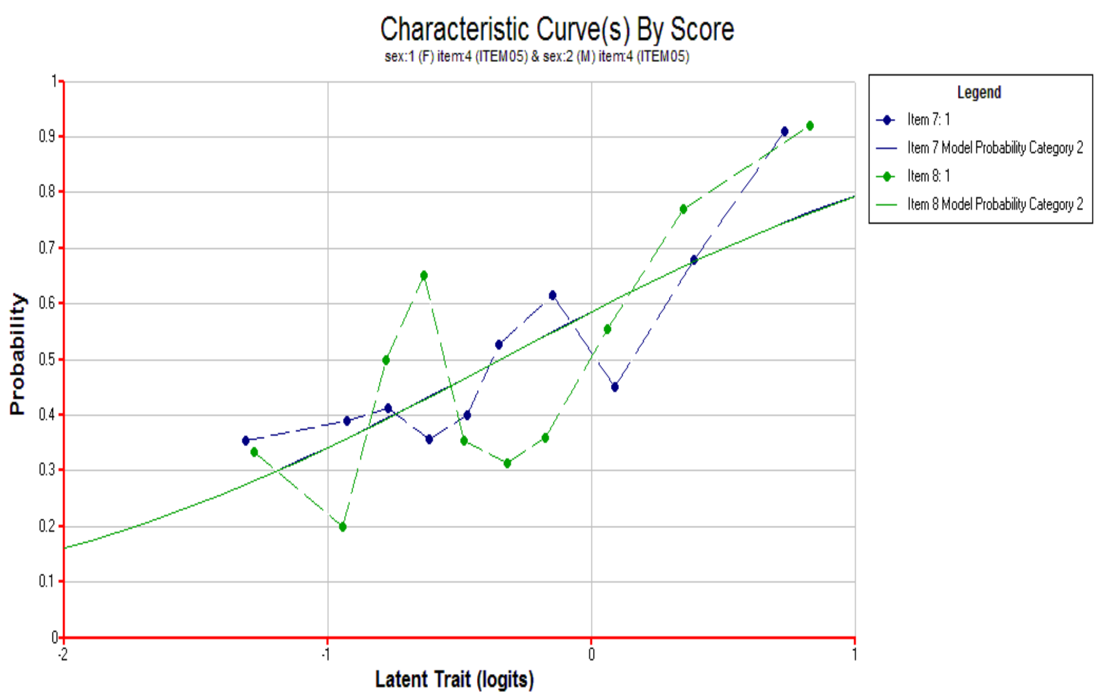


Figure 6. 3 ICC by score for gender DIF of item 05

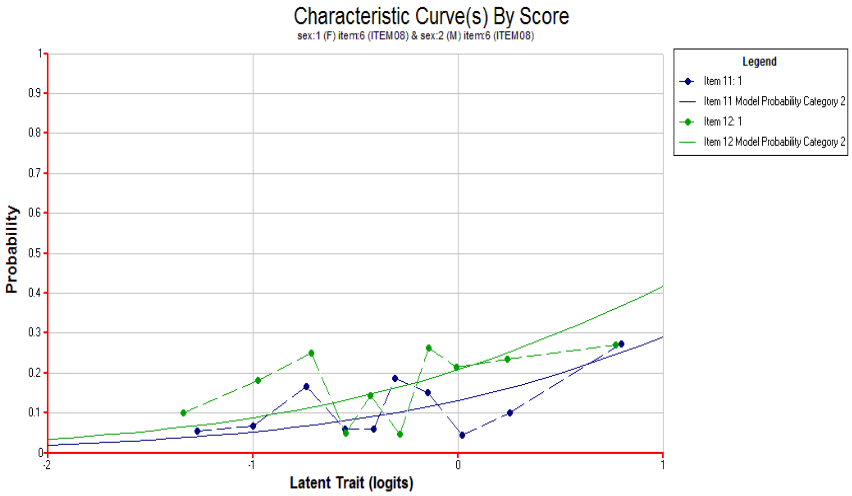


Figure 6. 4 ICC curve by score for gender DIF of item 08

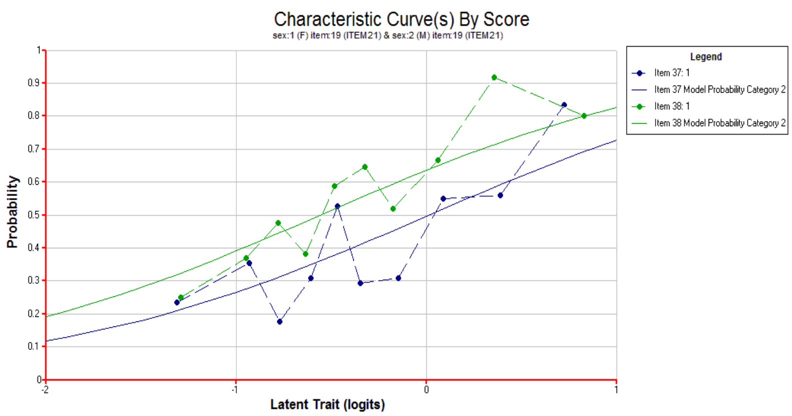


Figure 6. 5 ICC by score for gender DIF of item 21

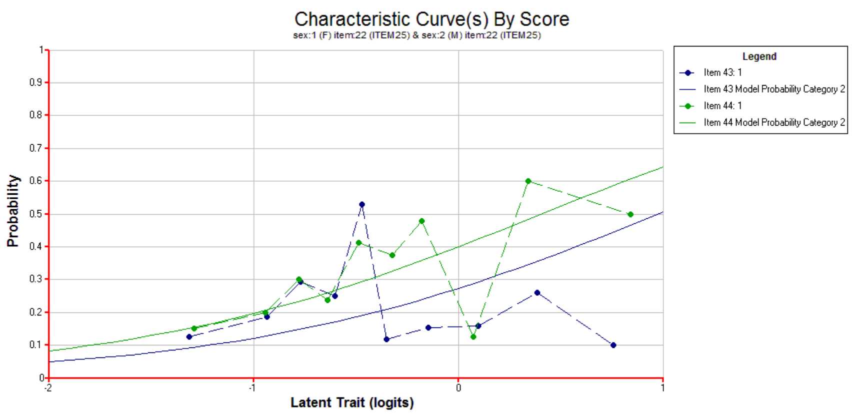


Figure 6. 6 ICC curve by score for gender DIF of item 25

6.3.4 Summary of the findings of the three methods

Overall findings from the three methods of DIF analysis provided information associated with equity and fairness of item functioning and respondent ability difference as discussed in Chapter 5. These methods are utilised because they are not sample dependent.

The item fit approach examined the INFIT mean square statistics of different sub-groups within the acceptable range of 0.7 to 1.3. This approach cannot suggest any item as DIF for two significant reasons, discussed in Chapter 5. The group plot approach presented graphs for item function through the range of person ability for males and females. The item threshold approach provides more specific findings. Some items were found to favour females, and some favoured males. Items 08, 21 and 25 were consistently found to have DIF, both with the item threshold and the group plot approaches, between the genders and so were deleted. The gender DIF in this study suggests a trend: with similar ability level, the three items favoured males over female students.

The overall test statistics indicated no significant difference in DIF for gender, but the individual item analysis through threshold and group plot approaches indicated that items 8, 21 and 25 have DIF. These three items showing gender differences in their ability level were removed due to the three main reasons regarding DIF that were highlighted in Chapter 5 (section 5.4.4). The content of these three items is related to statistics, trigonometric functions and money. Gender differences in responding to these topics may be influenced by content competence because, from experience teaching in PNG, males learn the content more than females. This may be associated with the constant exposure to mathematics by male students, and their competitive nature to achieve better grades than females (Matang, 2002; Saxe & Esmonde, 2012).

6.3.5 Final instrument

A final analysis is performed through calibration of the instruments after deleting the five items that did not conform to the RMM rules. The final 35 individual fit statistics (see Table 6.3 p.153) reveal that none of the items displayed an insignificant misfit, with the INFIT MNSQ ranging from -1.9 to 1.7, *t*-values from -2 to 2, and biserial correlation (r_{pb}) value from 0.02 to 0.41. This means that the 35 items are within the fit statistics ranges discussed in Chapter 3. The estimate is from -1.466 to 1.569, with an error ranging from 0.107 to 0.152. To conclude the overall statistical fit test, the 35 items that fitted RMM are looked at again. The items in each range of discrimination power are discriminating between the higher and lower ability respondents. This is shown in Table 6.3 with the biserial correlation (rpb) values.

6.3.6 Person-item map

The person-item map in Figure 6.7 is similar to the discussions of Chapter 5. As shown in Figure 6.7, item 23 is calculated as being at the mean of the item difficulty estimates, that is close to 0 logits and is located at the 0 point on the item-person map. The rest of the items are spread above and below 0 to represent their difficulties relative to items 23 that is located at 0. This means that respondents have a greater than 50% chance of succeeding on items less difficult than item 23 (items such as items 5,14,19, 22, 30 etc) and less than a 50% probability of succeeding on items that are more difficult than item 23 (i.e items 7,10, 27,33,37, etc).

Figure 6.7 shows that majority of the students' abilities are below the average (0 logit), and item difficulty fairly distributed along the logit scale. From the map, it can be seen that; items 08 and 27 are the most difficult in the test, and are located above a logit of 1. These items were difficult, and so only a few respondents, all of whom had higher abilities responded correctly. Due to the items' difficulty level, the respondents' probabilities of answering correctly is less than 50% (Boone & Scantlebury, 2006; Stelmack et al., 2004; Wei et al., 2014). These two items use complex processes, and the solutions were not straight forward, and require connection across other mathematics skills and concepts. On the other hand, items 1, 11, 17, 18 and 22 are the easiest located below a logit of -1, and so examinees with low and high ability answered them correctly. In other words, this indicates that more that 50% of the respondents could answer these items correctly.

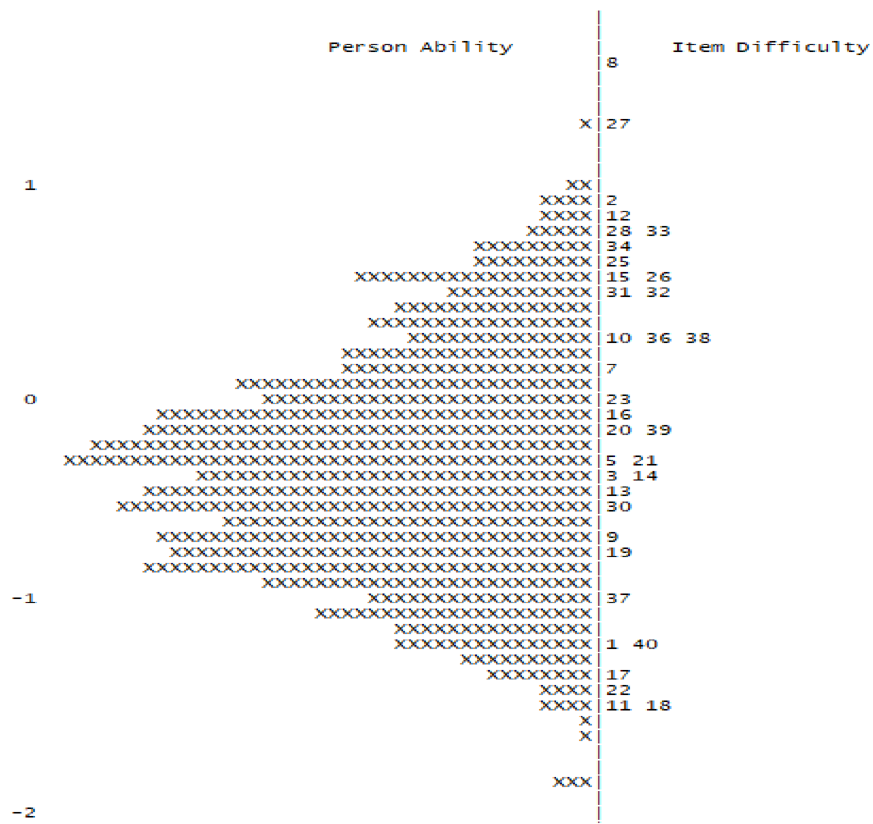


Figure 6. 7 Person-item map for Grade 12 items

6.4 Transformation to W-scores

As discussed in Chapter 3, the WLE values are transformed into W-scores to eliminate negative values, and the need for decimal points (Benson et al., 2016; Linacre, 2009; Woodcock, 1999). Microsoft Excel is used for this transformation, using the formula W=9.1024 x logits + c, where c is a constant term, selected to eliminate the negative values (Benson et al., 2016; Woodcock, 1999). As mentioned in Chapter 3, the constant value used in the study is 500 in consistent with PISA studies. The W- scores set the sign of the item difficulty and the person ability scales so that low values imply either low item difficulty or low person difficulty. High W-scores imply either high item difficulty or high person ability (Linacre, 2009; Woodcock, 1999; Wu et al., 2016)

6.5 Summary

In this chapter, overall fit statistics were analysed to investigate how well the data fit the RMM. The overall test and the individual items were analysed using statistical fits, category probability curves and ICCs, with, no significant response dependencies detected among them. This analysis indicates that there were three test items with DIF favouring male respondents. Generally, though, analysis revealed that the test followed principals of good test design, due to the data fitting the model after calibration. This is evident through INFIT MNSQ and *t*-values that are within the set criteria. Further, the items discriminated well between the higher and lower achieving students. It was found that two items are easy, and four items are difficult for students, as reflected on the person-item map. Despite this, however, the person-item location distribution shows a reasonable spread of items and persons, and the test targeted well the sample under investigation.

It has been shown that RMM can be used to contribute to improving the scoring of an assessment instrument, and that the methodology can be used in different assessment settings. This is because the person ability and difficulty level are matched at the same continuum analysis and inferences can be made accordingly. After this analysis of the 12 mathematics items, the subsequent chapters begin with the validation of the survey items on student-level constructs, such as attitude and motivation. Attitude is the first student-level construct analysed and discussed in the next chapter of this study.

Chapter 7: Attitude scale validation

7.1 Introduction

This chapter examines attitude scale at the individual level for the Grade 10 and 12 mathematics students surveyed in Port Moresby. Attitude in the context of this study is defined and discussed through the sub-scales *like, value*, and *self-belief* in mathematics. This approach is consistent with international studies such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA), where the items originated.

The chapter begins with a brief discussion of attitude sub-scales to set the context for discussion, and presents an introduction of the questionnaires/items that represent the scale. The next section is divided into two parts and discusses the data analysis techniques of this chapter. The first part discusses the confirmatory factor analysis (CFA) (discussed in Chapter 3) of the measurement model used to examine the construct validity of the scale (attitude); i.e., the relationship between the observed variables (instruments) and the construct attitude. This approach is employed to examine and uncover the underlying structure (ability to estimate relationships) among the instruments, with the construct adjusting for measurement error (Beccaria, Beccaria, & McCosker, 2018; Chamberlin, Moore, & Parks, 2017), and also to check and confirm the dimensionality of the scales used in Rasch model. The goodness of fit indices reported in Chapter 3 are investigated for the different models to ascertain whether the set of data fit a particular model. This technique is necessary for the establishment of valid and reliable constructs that provide accurate estimates for subsequent analysis of the student and teacher- level models, and is also a necessary step before Rasch measurement. The second part of the analysis examines the usability and portability of each item for the confirmed structures with CFA at the item-level analysis, using the multidimensional Rasch model discussed in Chapters 3. One of the main reasons for using the Rasch model is for scoring purposes in the subsequent analysis. The combination of these two techniques (the Rasch model & CFA) is important for two reasons: so that a) the results can be meaningfully interpreted to answer research question 1 highlighted in Chapter 1; and b) the results from this analysis can be used to make comparisons with the findings reported in the literature on students' attitudes towards mathematics outlined in Chapter 2. This chapter concludes with a summary that reiterates the key points of the preceding discussion.

7.2 Contextualisation of attitude scale

Improving students' attitudes towards learning is a major aim for many countries (Eivers & Chubb, 2017), and an abundance of educational research has documented the relationship between student attitudes and their academic achievements. The TIMSS and PISA studies have shown strong positive relationships within countries between students' attitudes toward mathematics and their mathematics achievements (Mullis et al., 2012; OECD., 2005). From these two studies, it is apparent that students with more positive attitudes toward mathematics achieve higher average mathematics results (Cheung, Mak, & Sit, 2018; Kadijević, 2006), and students with negative attitudes achieve lower mathematics results (Cheung et al., 2018). This trend underlies the importance of the general attitude students have towards mathematics, and the implications of this on their abilities, which is reflected in their mathematics results. Though PNG does not participate in the TIMSS and PISA studies, the attitude scale may have an impact on the mathematics results of PNG students, to some extent.

In this study, an attitude scale was developed with reference to the TIMSS and PISA assessments, where the items originated. According to TIMSS and PISA, 'attitude' is defined as students' enjoyment of mathematics and their tendency to value the importance of mathematics through intrinsic motivation and belief in their own abilities to perform better (Martin, Mullis, & Hooper, 2017; PISA, 2016).The two international assessments measure students' attitudes towards mathematics through sub-scales such as students *like of* mathematics, *value* felt towards mathematics, and their *self-belief* when engaging in mathematics (Martin & Kelly, 1996; Martin et al., 2017). These subscales items are selected from the TIMSS and PISA studies (refer to Table 7.1.p 176) as these two studies items are relevant and compatible to each other. Each of these subscales will now be defined, in order to better contextualise their meaning in this study.

The *like* subscale was developed to measure students' interest in and liking of mathematics (TIMSS, 2015). According to a longitudinal study by TIMSS (2015), students liking mathematics enables them to perform better, with the opposite being true for students who dislike mathematics. Therefore, one aim for educators is to encourage students who dislike mathematics to develop a better attitude towards the subject. This can be achieved by creating and promoting a positive atmosphere for learning, and avoiding framing mathematics as a hard and boring subject (Kennedy et al., 2014; Wilson & Mack, 2014). The degree to which students like mathematics is associated with their intrinsic motivation (which will be discussed further in Chapter 7) to learn the subject, which in turn promotes interest and enjoyment. The TIMSS 2015 study highlighted that intrinsic motivation energises students to like learning and doing mathematics, and ultimately to find the subject interesting and enjoyable (Mullis et al., 2012; Stephens et al., 2016). A strong relationship has been found between students' liking of mathematics and their achievements. For instance, in almost all countries in the 2011 TIMSS, students who indicated that they liked mathematics had higher averages in mathematics achievements than those who expressed that they did not like mathematics (Mullis et al., 2012; Thomson et al., 2012). In turn, those students who did not like mathematics had the lowest average mathematics achievement (Mullis et al., 2012; Thomson et al., 2012). This dislike of learning and doing mathematics may partially be a result of the framing of mathematics education (Mullis et al., 2012; PISA, 2013, 2016). This indicates that students who like and value mathematics perform better at it, and are more likely to enjoy learning mathematics (PISA, 2016; Prendergast & O'Donoghue, 2014) than those who dislike it.

The students' *value* subscale addresses students' attitudes regarding the importance and usefulness of mathematics for attainment value (Wigfield & Eccles, 2000). Importantly, the results of the TIMSS and PISA assessments have consistently shown a strong relationship between students valuing mathematics and their achievements (Mullis et al., 2012; PISA, 2013; Thomson et al., 2012). In the same way, the value subscale as a predictor of mathematics performance has an ability to predict future achievements (Eklöf, 2007a; Gaspard et al., 2015). A student could be confident in their ability to learn different aspects of mathematics, but if they do not value it as a subject, they may lack the motivation to continue studying it (PISA, 2013; Ray & Margaret, 2003). Studies have shown that students who reported that they valued mathematics typically had higher achievements than students who did not value the subject (Gaspard et al., 2015; Harris et al., 2015; Thomson et al., 2012).

The result of the student *self-belief* subscale from the TIMSS and PISA assessment cycles have shown a strong relationship between students' academic self-confidence and their mathematics achievements. Self-belief give students confidence that they can successfully perform the given academic tasks at designated levels (OECD., 2005; Reed et al., 2015), and the opposite is true for students who do not have self-belief. Thus, it is necessary for educators to assist students without self-belief in mathematics to have confidence in their own mathematics abilities (PISA, 2013). This confidence leads to higher levels of self-belief, which in turn leads to better performance in mathematics, regardless of students' abilities (Bandura, 1997; Schunk and Pajares, 2009 cited in TIMSS 2015). In other words, to accomplish mathematical tasks successfully, students need to believe in their mathematics abilities (OECD., 2005; Reed et al., 2015). Studies have revealed that self-belief acts to guide to assure students that they can do well and succeed (Pajares &

Schunk, 2001; Pitsia et al., 2017). A strong self-belief encourages students to engage positively with activities and show persistence, effort, and attentiveness (PISA, 2013; Reed et al., 2015). Subsequently, PISA (2013) and Mullis et al. (2012) have shown that mathematics self-belief tends to improve mathematics performance, with corresponding reductions in students' anxiety towards mathematics.

7.3 Attitude scale details in this study

In this study, attitude is measured through the three subscales (*like, value, and self-belief*) discussed above in the context of the TIMSS and PISA studies. This study also adopted and modified the student attitude scale items from TIMSS 2015 and PISA 2012, for the purpose of surveying the Grade 10 and 12 students in Port Moresby schools. As shown in Table 7.1, these two international studies provided 20 items (10 items each) for this study's attitude scale. The attitude scale consists of 18 positively-worded items and 2 negatively-worded items. These instruments use a four-point Likert-type scale: strongly agree, agree, disagree and strongly disagree (Penfield, Myers, & Wolfe, 2008; Thomas, Schmidt, Erbacher, & Bergeman, 2016), intended to encourage respondents to decide on responses without neutral and middle options that can psychometrically affect the attitude scale (Chyung, Roberts, Swanson, & Hankinson, 2017; Weems & Onwuegbuzie, 2001). The 20-scale items focus on the three subscales (*like, value and self- belief*) discussed above. In the questionnaire, these items were labelled from numbers 27 to 46. Hence, for data analysis process, the attitude scale items were labelled Attd27-Attd46, as shown in Table 7.1. The subscale *like* is represented by Attd27-Attd32, subscale *value* is represented by Attd33-Attd38 and subscale *self-belief* is represented by Attd39-Attd46. Item responses were coded 1,2,3 and 4 corresponding to the categories "strongly disagree"," disagree"," agree" and "strongly agree". Missing or omitted items responses were coded "9", which is an arbitrary value assigned for recognition with the statistical software as a non-response (Blackwell, Honaker, & King, 2017). The two negatively-coded items were reverse scored to maintain scoring consistency (Crenshaw, Christensen, Baucom, Epstein, & Baucom, 2017). These 20 items of the attitude scale were subjected to confirmatory factor analysis (CFA) to confirm whether each item has a relationship with the scale.

Table 7. 1 Summary of items in the attitude scale used in the study from TIMSS 2015 and PISA 2012

Item code	Item Text	Item adopted from/Item No	Subscale
Attd27	I like studying for my mathematics class outside of school.	TIMMS (20d)	Like
Attd28	I am studying mathematics because I like to learn new things	TIMSS (20g)	Like
Attd29	I look forward to my mathematics lessons.	PISA (31c)	Like
Attd30	I enjoy thinking about the world in terms of mathematics relationships	TIMSS (20l)	Like
Attd31	I do mathematics because I enjoy it.	PISA (31d)	Like
Attd32	I enjoy figuring out challenging mathematics.	TIMSS (20h)	Like
Attd33	Learning mathematics is important because it stimulates my thinking.	TIMSS (21a)	Value
Attd34	Learning mathematics is a worthwhile exercise	PISA (31e)	Value
Attd35	I am curious when I am learning mathematics.	PISA (36h)	Value
Attd36	When I do mathematics problems, it completely gets my attention.	TIMSS (20a)	Value
Attd37	I get a sense of satisfaction when I solve mathematics problems.	TIMSS (20b)	Value
Attd38	Mathematics is an important subject for me because I can use in my daily life	PISA (31g)	Value
Attd39	Sometimes the mathematics course material is too hard.	PISA (35d)	Value
Attd40	I feel confident in mathematics class.	PISA (31c)	Self-belief
Attd41	I want to study mathematics regularly.	PISA (36a)	Self-belief
Attd42	Making an effort in mathematics is worth it because it will help me in the work that I want to do later on.	TIMSS (21i)	Self-belief
Attd43	I want to do well in mathematics.	TIMSS (21b)	Self-belief
Attd44	I want to try harder in mathematics.	TIMSS (21e)	Self-belief
Attd45	I do badly in mathematics whether or not I study for my examinations	PISA (34f)	Self-belief
Attd46	I am very good at solving mathematics problems.	PISA (35a)	Self-belief

7.4 CFA of the measurement model for the attitude scale

The 20 items of the attitude scale were subjected to CFA as mentioned earlier to determine the attitude scale at the macro level, using the Mplus 7.0 software (Muthen & Muthen,2012). This technique is employed due to its ability to estimate the relationships (uncover the underlying structure) among variables adjusting for measurement error (Brown, 2014; Chamberlin et al., 2017; Han & Carpenter, 2014). The single-factor model/measurement model is first tested for the model fit to the data. This model is employed to understand the structure of the attitude scale. In checking the structure and relationship of the item with the construct a minimum factor loading or regression weight value of 0.3 is employed in this study for all factor analysis, consistent with other studies (Hailaya, 2014; Kline, 2014). Factor loadings of observed variables greater than the minimum value of 0.3 are considered meaningful indicators of a latent variable (Han & Carpenter, 2014; Ng, Wang, & Liu, 2017). A value below 0.30 for observed variable factor loading may suggest an indication of a factor other than the one it intends to reflect (Golay, Thonon, Nguyen, Caroline, & Favrod, 2018; Hailaya, 2014).

7.4.1 Structure analysis of the hypothesised measurement model

Theoretically, the attitude scale has multilevel dimensions as discussed in Chapter 2. However, in the actual data analysis, one factor structure CFA model is constructed first to test for fit statistics. This technique is employed to confirm whether all items were measuring one single factor, attitude scale (Han & Carpenter, 2014; Taasoobshirazi & Wang, 2016). As such, in this model the 20 items labelled "Attd" are loaded directly onto the latent variable 'attitude' with their residuals or error terms as indicated in Figure 7.1

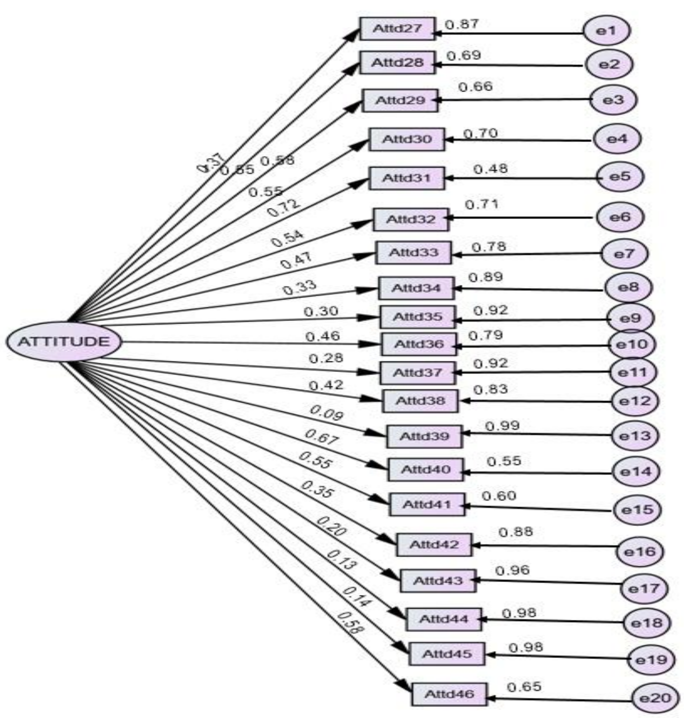


Figure 7. 1 Structure of one-factor model for attitude

7.4.2 CFA of the hypothesised measurement model

The attitude items (Attd27-Attd46) are examined using the factor loadings/regression weights to check whether the items demonstrate the latent variable they represented. Table 7.2 shows the factor loadings for the items under the attitude scale. This table indicates that 14 items (Attd27-Attd34, Attd36, Attd38, Attd40, Attd41, Attd42 and Attd46) have factor loadings above the adopted threshold and 6 items (Attd35, Attd37, Attd39, and Attd43-Attd45) exhibited unacceptable factor loading. It is evident from these results that the six items are not representing the attitude construct structure. Thus, the hypothesised single-factor model may not be appropriate for the attitude scale.

Table 7. 2 Factor loading of Attitude scale

Factor	Item(s)	Loading (se)*
	Attd27	0.36 (0.03)
	Attd28	0.55 (0.02)
	Attd29	0.58 (0.02)
	Attd30	0.55 (0.02)
	Attd31	0.71 (0.02)
	Attd32	0.53 (0.03)
	Attd33	0.47 (0.03)
	Attd34	0.33 (0.03)
	Attd35	0.26 (0.03)
Attitude	Attd36	0.46 (0.03)
	Attd37	0.28 (0.03)
	Attd38	0.41 (0.03)
	Attd39	0.09 (0.04)
	Attd40	0.67 (0.02)
	Attd41	0.55 (0.02)
	Attd42	0.35 (0.03)
	Attd43	0.20 (0.03)
	Attd44	0.13 (0.04)
	Attd45	0.14 (0.03)
	Attd46	0.58 (0.02)

*n=729

7.4.3 Model fit

The estimates of the parameters that are used to analyse the goodness of the model fit to the data are examined using the results of ratio of chi-square to its degree of freedom (x^2 / df), root mean square error of approximation (RMSEA), standardised root mean square residual (SRMR), comparative fit index (CFI) and Tucker-Lewis Index (TFI). These indicators of fit were used to evaluate how well the hypothesised model fit the data, and the criteria of these indicators were discussed in detail in Chapter 3.

Table 7.3 shows the results of the fit indices in testing the goodness of the model fit. All these fit indices demonstrate a poor fit indicating that attitude's single-factor structure does not fit the data. This result is consistent with the results of the low factor loadings of six items mentioned earlier. Therefore, alternative models are examined to test and analyse the factor loading, and model fit.

Table 7. 3 Results of fit indices for the single factor attitude structure

Fit Index	Obtained values	Remarks
X^2/df	4.98	Poor fit
RMSEA	0.07	Poor fit
SRMR	0.06	Poor fit
CFI	0.75	Poor fit
TLI	0.73	Poor fit

7.4.4 CFA analysis of the alternative models

Confirmatory factor analysis is used not only in hypothesis testing but also in comparing alternative models (Ng et al., 2017). The purpose of testing alternative models is to determine whether they are more consistent with the data compared to a single factor model (Byrne, 1998; Hooper et al., 2008; Ng et al., 2017). As such, alternative models, including the single factor model, orthogonal factor model, correlated factor model and hierarchical model discussed in Chapter 3 are examined and modification of these models are sought to obtain an expected model fit. The examination of the different alternative models fit indices shown in Table 7.4 indicate that the orthogonal factors model has poor model fit, with some items' factor loadings less than 0.3 even after deleting the five items with low factor loadings. This finding implies that the items have no relationship with the constructs. Instead, the hierarchal model fitted the data with the same five items (Attd39, Attd40 and Attd43-Attd45) removed. This process results in acceptable fit indices with acceptable factor-loading scores above 0.3, as shown in Table 7.4 and Table 7.5, respectively in each dimension. This finding indicates that the items are measuring different latent traits, as they belong to different scales. Therefore, the discussions below focus on the process of obtaining and describing the hierarchal model that will be used in the subsequent analysis.

Table 7. 4 Different model fit comparison for attitude towards mathematics

No	Model	X^2	*df*	X^2/df	*RMSEA*	*SRMR*	*CFI*	*TLI*
1	Single factor model 2*	227.58	73	3.12	0.05	0.04	0.91	0.89
2	Three-orthogonal factors model 1	1393.80	170	8.19	0.09	0.15	0.56	0.51
3	Three-orthogonal factors model 2*	808.09	90	8.97	0.1	0.16	0.64	0.58
4	Three-correlated factors model 1	745.73	167	4.47	0.06	0.06	0.79	0.77
5	Three-correlated factors model 2*	275.53	101	2.72	0.04	0.04	0.92	0.90
6	Hierarchical model 1	741.97	166	4.46	0.06	0.06	0.79	0.76
7	*Hierarchical model 2**	*228.59*	*86*	*2.65*	*0.04*	*0.04*	*0.93*	*0.91*

*Note. * items Mtvn39,Mtvn40,Mtvn43,Mtvn44 and Mtvn45 deleted*

The hierarchical model's factor loadings in Figure 7.2, (p.181) demonstrates that the items have a relationship with the constructs with three distinct clusters, or groups, that define theoretically the multilevel dimensions of the latent attitudinal traits discussed earlier in the literature review section. These traits are students' *like*, *value* and *self-belief* in mathematics. This model is further tested for its ability to provide the alternative model hypothesised by the attitude scale. In the hierarchical model *like* had six items (Attd27, Attd28, Attd29, Attd30, Attd31 and Attd32), *value* six items (Attd33, Attd34, Attd35, Attd36, Attd37 and Attd38) and *self-belief* five items (Attd41, Attd42 and Attd46).

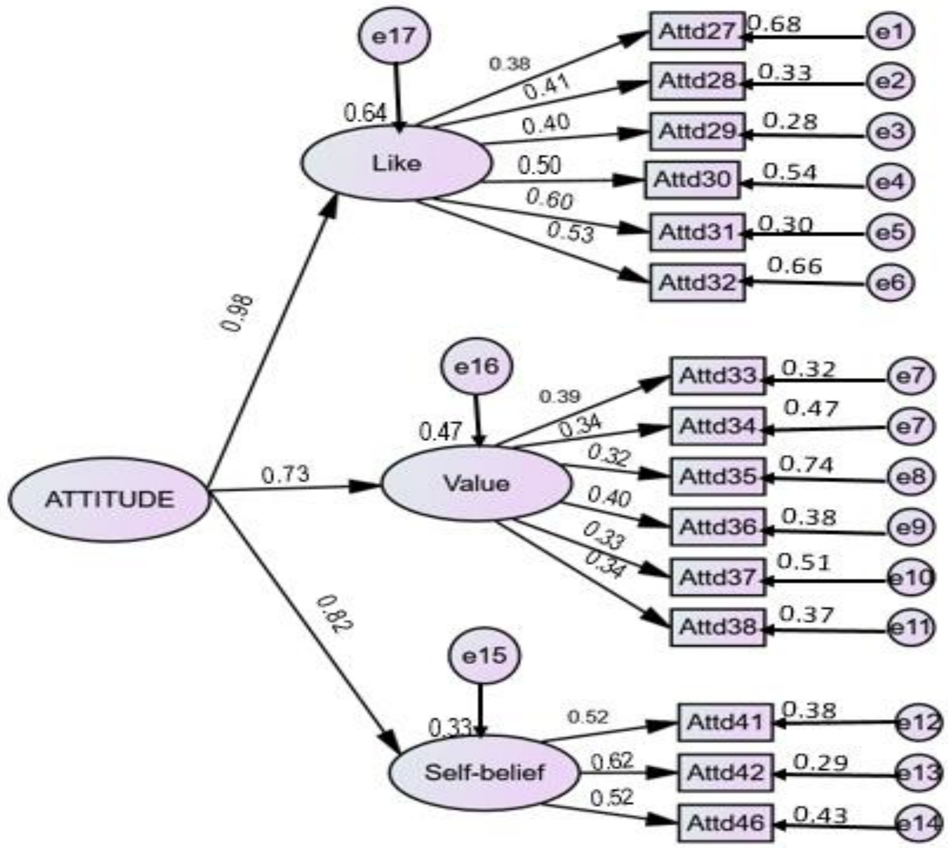


Figure 7. 2 Structure of the hierarchical model

7.5.5 CFA for the attitude alternative measurement model

Next, the attitude items are checked for overall model fit for the hierarchical model. Table 7.5 shows the results of the factor loading of the items representing the subscales *like, value*, and *self-belief*. Attd39 from the subscale *value,* and Attd40, Attd43, Attd44 and Attd45 from subscale *self-belief,* have low factor loadings (below 0.3). Thus, these five items with low factor loading are deleted from the CFA, with the remaining 15 items shown in Table 7.5 loading well to the three subscales with factor loadings above 0.3.

Table 7. 5 Results of the factor loading of the hierarchical model

Second-order factor	First-order factor	Loading	Item (s)	Loading (se)*
Attitude	Like	0.98	Attd27	0.34(0.03)
			Attd28	0.41(0.02)
			Attd29	0.40(0.02)
			Attd30	0.50(0.03)
			Attd31	0.54(0.02)
			Attd32	0.54(0.03)
	Value	0.73	Attd33	0.39(0.03)
			Attd34	0.34(0.03)
			Attd35	0.32(0.04)
			Attd36	0.40(0.03)
			Attd37	0.33(0.04)
			Attd38	0.34(0.03)
	Self-Belief	0.82	Attd41	0.52(0.03)
			Attd42	0.62(0.03)
			Attd46	0.52(0.03)

*n=729

7.4.6 Model fit

In evaluating the hypothesised model of the attitude scale, the overall model fit to the data is first examined using the results of ratio of chi-square to its degree of freedom (x^2 / df), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI) and Tucker-Lewis Index (TFI) (refer to Chapter 3 for a detailed discussion). Table 7.6 shows that five of the fit indices (RMSEA, SRMR, CFI, TLI and X^2/df) demonstrate acceptable fit. These five acceptable fit indices indicate that the hierarchical model structure fitted the data, with the variables in each subscale having acceptable factor loading. Therefore, it can be concluded from the CFA results that the hierarchical model appears appropriate for the attitude scale, and so is considered for subsequent analysis.

Table 7. 6 Result of the model fit

Fit Index	Obtained values	Remarks
X^2/df	228.92/87=2.6	Acceptable fit
RMSEA	0.04	Acceptable fit
SRMR	0.04	Acceptable fit
CFI	0.93	Acceptable fit
TLI	0.91	Acceptable fit

7.4.7 Summary of confirmatory factor analysis

The establishment of valid and reliable variables in this study is fundamental to provide accurate estimates for the subsequent analysis of the student-level model as well as hierarchal linear model. Therefore, CFA is employed to confirm the structure of the observed variables as well as the relationship between the observed variables and their underlying latent variable. After confirming their theoretical relationship/structure within the hierarchical model with CFA, the next section of the discussion focuses on checking the utility (usability and portability) of the 15 items at the item level through the multidimensional Rasch model. As the attitude construct is established through CFA, it is also important to check each item at the three dimensions (*like, value and self-belief*) through the multidimensional Rasch model. This is carried out by examining the person/item fit and the fit mean square to see how well an item or a person fits the ideal Rasch model.

7.5 Multidimensional Rasch analysis

The Multidimensional Rasch model (MRM) technique discussed in Chapter 3 is employed using ConQuest version 4.0 (Wu & Adams, 2007). This technique is employed based on the three dimensions (*like, value* and *self-belief*) for the construct attitude that are associated with the hierarchical factor model in CFA. The Rasch model is applied after removing the items that did not form a relationship with the construct. This model is used to ensure that the utility of each of the 15 items confirmed from the CFA is checked in validating the data, in order to make sound inferences. This analysis provides an assessment of the scale's internal construct validity (Bond & Fox, 2015; Yasin et al., 2015). Rasch analysis includes the tests for item testing of invariance, assessment of category ordering and the assessment of differential item functioning (Andrich & Marais, 2014; Guilera et al., 2010).

7.5.1 Item analysis with the partial credit model

The 15 of 20 items retained in the attitude scale following the CFA are subjected to item analysis using the partial credit model (PCM) discussed in Chapter 3. This step is carried out to test the multidimensional item response of the items measuring the subscale (*like, value* and *self-belief*) for the construct *attitude* and involves examining each item's fit statistics using the same statistical criteria and procedure. As such, Table 7.7 shows the statistical results of the items from the Port Moresby Grade 10 and 12 students' responses.

Table 7. 7 Model parameter estimates of attitude towards math (n=729)

			Weighted Fit						
Item(s)	Estimate	Error	MNSQ	CI	t	Pt Bis	Item Delta		
Attd27	0.42	0.04	1.22	(0.90,1.10)	4	0.46	-1.01	0.04	2.1
Attd28	-0.46	0.04	0.96	(0.90,1.10)	-0.7	0.55	-1.99	-0.78	1.55
Attd29	-0.66	0.04	0.96	(0.89,1.11)	-0.8	0.57	-1.75	-1.41	1.43
Attd30	0.63	0.04	0.96	(0.90,1.10)	-0.9	0.59	-1.25	0.53	2.38
Attd31	-0.03	0.04	0.86	(0.90,1.10)	-2.9	0.67	-1.75	-0.01	1.67
Attd32	0.09*	0.09	0.99	(0.90,1.10)	-0.2	0.60	-1.18	-0.09	1.52
Attd33	-0.35	0.04	0.95	(0.89,1.11)	-0.9	0.53	-1.34	-0.84	1.18
Attd34	0.38	0.04	1.05	(0.89,1.11)	0.8	0.42	-0.50	-0.44	2.02
Attd35	0.41	0.04	1.08	(0.90,1.10)	1.5	0.41	-0.98	0.14	1.97
Attd36	-0.07	0.04	0.98	(0.89,1.11)	-0.3	0.50	-1.12	-0.38	1.32
Attd37	-0.27	0.04	1.0	(0.86,1.14)	0.1	0.41	-0.59	-0.43	0.34
Attd38	0.12*	0.09	0.97	(0.89,1.11)	-0.6	0.51	-1.08	-0.6	1.33
Attd41	0.38	0.04	1.06	(0.90,1.10)	1.3	0.69	1.21	2.18	5.37
Attd42	-1.38	0.04	1.07	(0.88,1.12)	1.1	0.56	-2.83	-1.22	2.15
Attd46	0.99*	0.06	0.97	(0.90,1.10)	-0.6	0.61	-1.07	0.24	1.96

Separation Reliability = 0.995, Chi=square test of parameter equality=2142.93,df =12, Sig Level =0.000

It can be seen from the results in Table 7.7 that all items for the subscales fit the model, with the INFIT mean square value (MNSQ) criteria of 0.60 to 1.40 (Bond & Fox, 2015). As shown in Table 7.7 the item delta(s) indicate that response choices on the scale are in order. Moreover, the separation reliability is 0.99, showing minimal measurement error and high discrimination (Alagumalai et al., 2005; Ben, 2010). This further indicates that the items have more precise measurement and reliability (Hailaya, 2014; Wright & Stone, 1999). However, the results of Table 7.7 indicate that the *t*-values of Attd27 and Attd31 are not within the acceptable fit criteria of -2 to 2 and are therefore misfitting the model (Bond, Fox, & Christine, 2007). Since this is not a "high-stakes" test, a more lenient approach is taken with the *t*-values of these items (Bond & Fox, 2015). These two items (Attd27 and Attd31) are within the acceptable range of MNSQ value of 0.6 to 1.40, and the items are discriminating with high and low ability respondents. Therefore, these two items are not deleted because they may contain information or findings relevant to the study. Further, the item estimates, average expected values, biserial points and item thresholds are all in order.

After checking the utility of the 15 items, differential item functioning (DIF) for gender (shown in Table 7.8, below) is carried out, using similar criteria of INFIT MINSQ and item threshold (estimates) approaches discussed in Chapters 3, 5, and 6. As reported in Table 7.8, the INFIT MNSQ of the 15 items for both males and females are within the 0.60 to 1.40 criteria, and the differences between the threshold (estimates) of the genders are under 0.5 logits (Hungi, 2005), which is within the acceptable range. Thus, Table 7.8 indicates that there is no item detected with DIF for gender for both Grade 10 and 12 respondents.

Table 7. 8 INFIT MNSQ for gender and item threshold (estimates) of gender

	Male			Female			
Items No	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ	d1-d2
Attd27	0.04	0.03	0.6	-0.04	0.03	0.57	0.08
Attd28	-0.04	0.03	0.49	0.04	0.03	0.45	-0.08
Attd29	0.02	0.03	0.45	-0.02	0.03	0.42	0.04
Attd30	-0.06	0,025	0.49	0.06	0.03	0.46	-0.12
Attd31	-0.04	0.03	0.44	0.04	0.03	0.44	-0.08
Attd32	-0.06	0.03	0.52	0.06	0.03	0.59	-0.12
Attd33	-0.07	0.03	0.55	0.07	0.03	0.56	-0.14
Attd34	0.06	0.03	0.48	-0.06	0.03	0.57	0.12
Attd35	0.02	0.03	0.57	-0.02	0.03	0.57	0.04
Attd36	0.05	0.03	0.59	-0.05	0.03	0.58	0.10
Attd37	0.03	0.03	0.89	-0.03	0.03	0.84	0.06
Attd38	0.02	0.03	0.57	-0.02	0.03	0.58	0.04
Attd41	0.01	0.03	0.43	-0.01	0.03	0.46	0.02
Attd42	0.05	0.03	0.61	-0.05	0.03	0.62	0.10
Attd46	-0.04	0.1	0.47	0.04	0.1	0.47	-0.08

7. 5.2 Person-item map

The multidimensional Rasch analysis is shown in the person-item map on Figure 7.3, with each dimension or factor having its own column with estimates for respondents' attitudes. This map demonstrates a visual estimate of the latent trait in the sample, item difficulty on each dimension (*like, value* and *self-belief*).The items ranked at the top of the dimensions are more difficult to endorse than items at the bottom (Osteen, 2010). In Figure 7.3, the numbers on the far left are from -1 (weak attitude) to around +4 (positive attitude). The relative item difficulty is plotted on the right side of the scale, with personal ability estimates on the left side of the same scale on the person-map. The three dimensions items are fairly dispersed around the mean. However, the dimension *self-belief* shows that some respondents for this study found it very easy to endorse Attd42 for the attitude construct.

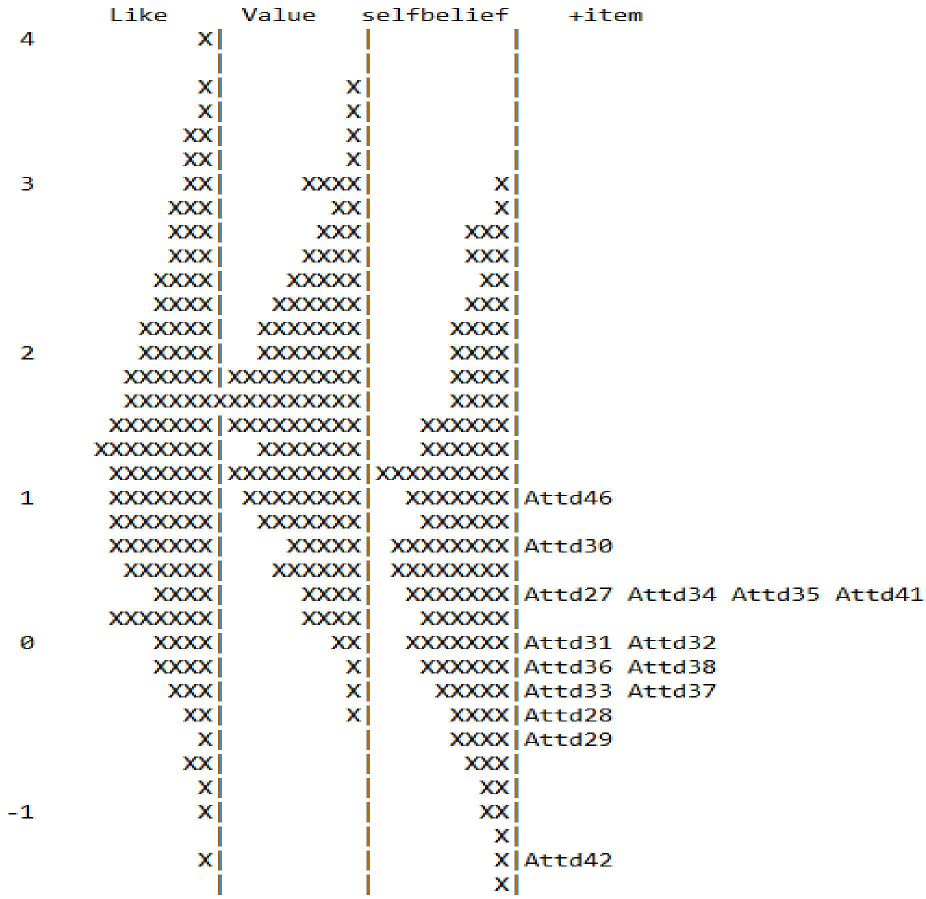


Figure 7. 3 Person-item map of attitude scale

Figure 7.3 shows that Attd31 and Attd32 are located at the mean (0 logits). The other items are spread above and below 0 to represent their difficulty relative to Attd31 and Attd32. Majority of the respondents have greater than 50% chance of endorsing items with difficulty level below their ability, and vice versa. . However, most of the respondents for the dimensions are located above the average score, with few respondents around and below the mean (Boone, Townsend, & Staver, 2010). This indicates that the items for the three dimensions were easy for the respondents to endorse. The distribution of item difficulty level is relatively low compared to the respondents' along the three dimensions.

7.6 Transformation to W-scores

As discussed in Chapter 3, the WLE values were transformed into W-scores to eliminate negative values, and the need for decimal points (Benson et al., 2016; Linacre, 2009; Woodcock, 1999). Microsoft Excel is used for this transformation using the formula W=9.1024 x logits + c, where c is a constant term, selected to eliminate negative values (Benson et al., 2016; Woodcock, 1999). As mentioned in Chapter 3, the constant value used in the study is 500 in consistent with PISA studies. The W- scores are then used in analysis in subsequent chapters, such as in structural equation modelling and hierarchical linear modelling.

7.7 Comparison of the different MRM and CFA results

The primary result from the CFA and Multidimensional Rasch Model (MRM) analysis is the establishment of the 'attitude' scale as a multidimensional measure. Both sets of analysis identified hierarchical model as having the best model fit compared to single-factor model. Tables 7.2 (p.178), 7.3 (p.179), 7.5 (p.182) and 7.6 (p.183) show the evaluation of the dimensionality of the items/constructs, and identify the traits that influence response pattern (Wright, 1996). The results in Table 7.7 (p.184) show the 15 items that were confirmed with the structure being analysed using the MRM at the three dimensions of *like, value* and *self-belief.* These sub-constructs items measure the attitude construct and are more informative at the item level. The MRM allowed analysis of the model factor structure and an estimation of the relationship between the factors. This model is found to be more informative for assessing individual item performances. Item analyses are obtained for attitude scale. Based on infit statistics, the 15 items fitted the Rasch model and the hierarchical model using CFA. Rasch analysis also provided estimates of the items' information function and the test information function (Osteen, 2010), making it possible to obtain the specific estimates of standard errors of measurement instead of relying on an averaged standard error of measurement obtained from the CFA (Osteen, 2010). The results reported in Table 7.8 (p.186) indicate that the items are not biased towards a particular gender group, as shown by the INFIT MNSQ and item threshold methods. Overall, these two methods address issues in measurement theory by applying the Rasch analysis to multidimensional latent variable measures (Osteen, 2010).

7.8 Model employed for the study

Comparison of the CFA results of the single factor model and other alternative models reveals that the hierarchical model for the attitude scale fit the data comparatively better. As such, from the multidimensional Rasch item analysis perspective, the items in the hierarchical model are viewed as forming three dimensions according to the variables' relationships with the subscales identified in CFA analysis. Thus, the items' utility at each of the three dimensions are checked through a multidimensional Rasch model, with responses varying in terms of the levels of agreement (Boman, Curtis, Furlong, & Smith, 2006; Grimbeek & Nisbet, 2006). The combination of Rasch analysis and CFA provides useful information that can be compared and matched accordingly. Hence, in this study, the hierarchical model with the scores at the three dimensions will be utilised in subsequent chapters.

7.9 Summary

This chapter focused on examining the relationship/structure of the attitude scale under each subscale through confirmatory factor analysis (CFA) and the Rasch measurement model (RMM). Specifically, the chapter checked the utility of the 15 items through multidimensional Rasch analysis at the individual item level of the subscales, after the 20 items have been examined with CFA. The results of both the CFA and MRM analysis reflected that the hierarchical model fitted the data, and the scores of each dimension will consequently be used in the subsequent analysis of this study. The individual score for each respondent (scale score) was generated by converting the weighted likelihood estimation generated by ConQuest 4.0 into W-scores (Warm, 1989), as discussed in Chapter 3. This conversion of the score is performed to categorise the range of abilities of respondents, which will be used in subsequent analysis (Martin, Mullis, Foy, & Arora, 2011). The next chapter employs similar techniques and procedures used in this chapter to analyse and discuss another student-level factor, the motivation scale.

Chapter 8: Motivation scale validation

8.1 Introduction

Another factor examined at the individual level by this study is how motivation towards mathematics influences Grade 10 and 12 students' mathematics achievements. Motivational factors such as *achievement goal, engagement* and *interest* have been found by other studies to encourage students' mathematics learning (Gasco & Villarroel, 2014; Hashim Ali, 2013; Pantziara & Philippou, 2015). In this chapter, these motivational factors (subscales) are discussed in the context of the TIMSS and PISA assessments, from which they have been derived. This chapter therefore contains a discussion of the quantitative analysis employed to establish relationships with the variables to confirm the utility of the instruments in measuring the motivation scale, similar to the analysis carried out in Chapter 7. This process allows the researcher to make meaningful interpretations to answer the research question in Chapter 1 (What are the school, teacher and student- level factors that are affecting Grade 10 and 12 students' mathematics results in Port Moresby, PNG?) and to make effective comparison with the related literature reviewed in Chapter 2.

The chapter begins by detailing the motivation scale used in the study, with a brief discussion of the items represented in the scale. It is then followed by confirmatory factor analysis (CFA) to confirm the relationship between the observed variables and constructs and, additionally, to check the dimensionality of Rasch measurement model as discussed in Chapter 7. The goodness of fit indices outlined in Chapter 3 are then investigated for the models, to ascertain whether the set of data fit a particular model according to the theoretical assumptions. The next section discusses the checking the utility of each item at the micro level/item-level using multidimensional Rasch modelling after confirming the structure and relationships of the items with the construct through CFA. Another reason for using Rasch analysis is for scoring purposes for the subsequent analysis. This chapter then concludes with a summary highlighting the key points discussed.

8.2 Contextualisation of motivation scale

Motivation is one of the most important ingredients of success in mathematics (OCDE, 2014; PISA, 2013). Greater motivation can drive individuals, especially those with less talent, to reach their goals (PISA, 2013; Tuomi, 2006). Motivated students are more likely to succeed than students who have talent alone, and they are more capable of setting goals for themselves to stay focused on their mathematics studies (Hopfenbeck & Kjaernsli, 2016; Ross, 2008). The motivation to achieve goals leads students to pursue work they perceive to be valuable and prompts them to compete with others (Hopfenbeck & Kjaernsli, 2016; Ross, 2008).

As with the attitude scale discussed in Chapter 7, the motivation scale items were derived from the TIMSS and PISA studies. These two studies define motivation as both intrinsic and extrinsic; a form of engagement that is regulated and sparked by interest, and by the desire to achieve a certain goal in mathematical tasks (Eklöf, 2007; Martin et al., 2017; PISA, 2016). In PISA and TIMSS, subscales such *as achievement goal, engagement,* and *interest* measure students' motivation towards mathematics learning, and these subscales are adopted by this study. The motivation scale items measuring the three subscales are selected from the TIMSS and PISA studies (refer to Table 8.1 p.195) due to their relevance and appropriateness in the context of the study. These items from these two studies are also compatible to each other. Though PNG does not participate in the two aforementioned international assessments, it is evident from reports of participating countries that the motivation scale can have an impact on the mathematics results of students.

Students with high *achievement goals* tend to do better at school (Gherasim et al., 2013; PISA, 2013), they set goals to achieve in mathematics that lead them to be more involved and engaged in their studies (Eklöf, 2007; Hopfenbeck & Kjaernsli, 2016). This, in turn provides students "higher autonomous and internalised achievement motivation with higher self-esteem, stronger cognitive awareness and greater efforts invested at mathematics" (OECD, 2017 p.94). Students who are highly motivated to achieve goals consciously value their own efforts in mathematics (OCDE, 2014) and are motivated to achieve good results. A study by the OECD (2013) highlights that motivated students are typically autonomous individuals who believe that they can learn in various positive ways to solve mathematical problems. Another OECD (2017) study proposes that this attitude stems from a student's sense of responsibility and obligation in their approach to learning mathematics. These two studies underline the significance of devoting time and effort towards achieving success through engagement in mathematics activities for the achievement of specific goals (Breakspear, 2014; Martin & Mullis, 2013; PISA, 2013).

The next subscale, *engagement,* involves students' participation in learning, and reflects behaviours of persistence, concentration, attention, asking questions and contributing to mathematics learning (Fredricks, Blumenfeld, & Paris, 2004). According to OECD studies, engagement in mathematics enables students to be involved in active learning as it encourages them to think on specific task (OECD, 2013; OECD, 2017). The role of engagement in mathematics learning is significant, and is based around student-centred activities that support conceptual competences (Mullis et al., 2012), and encourage active involvement in learning (Mullis et al., 2012; Wilson, 2011). Student-centred activities are advantageous when attempting to interest and engage students (OECD, 2013). Examples of such activities in mathematics are engaging students in experiments and projects that enable them to problematise the content and address problems using their own authority and relevant resources (Mullis et al., 2012; OCDE, 2014; PISA, 2013). This method assists students to understand the content of mathematics while maintaining an interest in their own learning (OECD, 2013; OECD, 2017).

Interest is a central component for students' motivation and engagement in learning (Lamb, Annetta, Meldrum, & Vallett, 2012; Nyman, 2017). The OECD (2004) reports that increasing students' interest in mathematics leads to higher test scores and student achievement in various contexts. A study by Cleary and Chen (2009) found a positive correlation between student interest, motivation, and better mathematics results. From these studies, it is clear that when students are interested in their studies, they are more motivated, connected, and engaged in mathematical tasks (Nyman, 2017; PISA, 2013). The OECD (2013) recommend that specific strategies applied in mathematics teaching and learning can capture students' interest. For instance, providing students with different materials and opportunities for learning can instil greater interest and engagement. This kind of strategy can assist students to conceptualise the mathematical knowledge needed to be successful in the subject (OCDE, 2014; PISA, 2013).

8.3 The motivation scale in this study

Motivation scale items from TIMSS 2015 and PISA 2012 were adopted and used in this study. These two international studies provided 20 items for the motivation scale (see Table 8.1). As this study has similar aims to TIMSS and PISA in terms of measuring students' motivation, the motivation scale items were adopted for Grade 10 and 12 students in Port Moresby. The scale consists of 19 positively-worded items and one negatively-worded item, using a four-point Likert-type scale: strongly agree, agree, disagree and strongly disagree (Penfield et al., 2008; Thomas et al., 2016). The motivation scale's 20 items were focused on three subscales, namely the above discussed *achievement goal, engagement* and *interest* scales. In the questionnaire booklet used in this study ('Examining factors that affect mathematics achievement for Grade 10 and 12 students' questionnaire') these items were labelled from 07 to 26. These motivation scale items were labelled Mtvn07-Mtvn46 for data analysis purposes, as shown in Table 8.1. The subscale *achievement goal* is represented by Mtvn07, Mtvn09-Mtvn13, subscale *engagement* is represented by Mtvn14-Mtvn19 and subscale *interest* is represented by Mtvn08, Mtvn20-Mtvn26. Item responses were coded 1, 2, 3 and 4, corresponding to the categories "strongly disagree"," disagree"," agree" and "strongly agree", respectively. The missing or omitted items responses were coded "9", which is an arbitrary value assigned for recognition with the statistical software as a non-response (Blackwell et al., 2017). The single negatively-coded item was reverse scored to keep scoring consistency (Crenshaw et al., 2017). These 20 observed variables of the motivation scale were subjected to confirmatory factor analysis (CFA) to confirm whether each variable has a relationship with the scale.

Table 8. 1 Summary of items in the motivation scale used in the study

Item code	Item Text	Item adopted from/Item No	Subscale
Mtvn09	Learning mathematics will help me get ahead in the world	TIMSS (21a)	Achievement goal
Mtvn10	It is important to do well in my mathematics class	TIMSS (21b)	Achievement goal
Mtvn11	Doing well in mathematics will help me get into university/colleges.	TIMSS (21e)	Achievement goal
Mtvn12	Learning mathematics will give me more job opportunities	TIMSS (21i)	Achievement goal
Mtvn13	Learning mathematics is worthwhile for me because it will improve my career prospects	TIMSS (21f)	Achievement goal
Mtvn07	I am prepared for my mathematics examinations	PISA (36c)	Achievement goal
Mtvn14	I keep studying until I understand mathematics material	PISA (36e)	Engagement
Mtvn15	I take part in mathematics competitions	PISA (38d)	Engagement
Mtvn16	I do mathematics more than 2 hours a day outside of school	PISA (38e)	Engagement
Mtvn18	I work hard on my mathematics homework	PISA (36b)	Engagement
Mtvn19	I keep my mathematics work well organized	PISA (36i)	Engagement
Mtvn08	Jobs that require mathematics skills seems interesting to me.	TIMSS (20j)	Interest
Mtvn17	I have my homework finished in time for mathematics class	PISA (36a)	Interest
Mtvn20	I talk about mathematics problems with my friends	PISA (38a)	Interest
Mtvn21	I listen and pay attention in mathematics class.	PISA (36f)	Interest
Mtvn22	I avoid distractions when I am studying mathematics	PISA (22h)	Interest
Mtvn23	Mathematics is one of my favourite subjects	TIMSS (20i)	Interest
Mtvn24	I am interested in the things I learn in mathematics	PISA (31f)	Interest
Mtvn25	The teacher did not get students interested in the material	PISA (35e)	Interest
Mtvn26	It is interesting to learn mathematics theory	TIMSS (20e)	Interest

8.4 CFA of the measurement model for the motivation scale

The motivation scale in this study applied to survey students in Port Moresby, is adapted to a different context than the items' original contexts in PISA and TIMSS assessment. The 20 items of the motivation scale are subjected to confirmatory factor analysis (CFA) (see Chapter 3), using Mplus 7.0 software (Muthen & Muthen, 2012), to determine the scale at the macro level. As discussed in Chapter 7, this method is employed due to its ability to estimate the relationships with the observed variables and the latent variable on their latent trait adjusting for measurement error (Brown, 2014; Chamberlin et al., 2017; Han & Carpenter, 2014). In testing the model fit to the data, the single factor model/measurement model was first analysed. The purpose of this is to better understand the motivation scale structure. As in Chapter 7, a minimum factor loading value of 0.3 was used consistent with other studies (Hailaya, 2014; Kline, 2014). A value lower than 0.30 for observed variable factor loading may indicate a factor other than the one it intends to reflect (Golay et al., 2018; Hailaya, 2014).

8.4.1 Structure analysis of the hypothesised measurement model

A similar procedure employed for the attitude scale in Chapter 7 is also used for motivation scale in this chapter. Hence, the motivation scale was measured through a number of observed variables in CFA (Han & Carpenter, 2014; Taasoobshirazi & Wang, 2016). In this model, the 20 items labelled "Mtvn" are loaded directly onto the latent variable *motivation,* as reflected in Figure 8.1, with the item residuals or errors in the model.

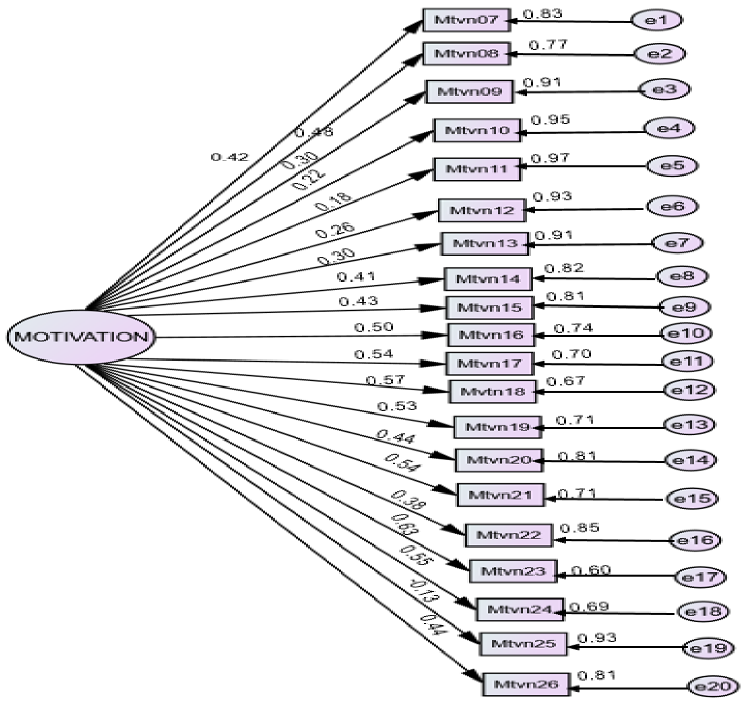


Figure 8. 1 Structure of one-factor model for motivation scale

8.4.2 CFA of the hypothesised measurement model

The factor loadings/regression weights are examined for the 20 motivation items to determine whether or not the items demonstrated the latent variable they represented as shown in Figure 8.1. The factor loading for the items under the motivation scale are shown in Table 8.2. This table indicates that five items (Mtvn10-Mtvn13, and Mtvn25) exhibit unacceptable factor loading below 0.3, while 15 items (Mtvn07-Mtvn09, Mtvn14-Mtvn24 and Mtvn26) have factor loadings above the adopted threshold. This means the five observed variables had a lower relationship with motivation, which is consistent with the poor results of the model fit in Table 8.3. Thus, the hypothesised single-factor model may not be appropriate for the motivation scale. Hence, it was necessary to modify the single-factor model to fit the data. This was carried out using alternative models, as discussed in Chapter 7.

Table 8. 2 Factor loading of motivation scale

Factor	Item	Loading (se)*
	Mtvn07	0.42 (0.03)
	Mtvn08	0.48 (0.03)
	Mtvn09	0.30 (0.03)
	Mtvn10	0.22 (0.03)
	Mtvn11	0.18 (0.04)
	Mtvn12	0.26 (0.03)
	Mtvn13	0.29 (0.03)
	Mtvn14	0.41 (0.03)
	Mtvn15	0.43 (0.03)
Motivation	Mtvn16	0.50 (0.03)
	Mtvn17	0.54 (0.03)
	Mtvn18	0.57 (0.02)
	Mtvn19	0.53 (0.03)
	Mtvn20	0.44 (0.03)
	Mtvn21	0.54 (0.03)
	Mtvn22	0.38 (0.03)
	Mtvn23	0.63 (0.02)
	Mtvn24	0.55 (0.03)
	Mtvn25	-0.13 (0.04)
	Mtvn26	0.44 (0.03)

*n=729

8.4.3 Model fit

Next, the results of ratio of chi-square to its degree of freedom (x^2/df), root mean square error of approximation (RMSEA), standardised root mean square residual (SRMR), comparative fit index (CFI) and Tucker-Lewis Index (TFI) are analysed to determine the goodness of the model fit to the data. However, in order to improve the model fit and fit indices, the observed variables of the motivation construct are correlated among themselves using the modification fit index results on the Mplus output. These indicators of fit index are used to evaluate and assess the hypothesised model for the motivation scale, with fit indices' criteria discussed in detail in Chapter 3. The results of this analysis (Table 8.3) show that all the fit indices (X^2/df, CFI, TLI RMSEA and SRMR) demonstrated poor fit, with values not within the acceptable threshold. As such, these results illustrate that, the motivation scale's single-factor structure did not fit the data. This is consistent with the results of the low factor loadings of the five items mentioned earlier. Thus, alternative models discussed in Chapter 7 are investigated and discussed in the next section to obtain a model that better fits the data to be used for subsequent analysis.

Table 8. 3 Results of fit indices for the single-factor motivation structure

Fit Index	Obtained values	Remarks
X^2/df	841.34/170=4.94	Poor fit
RMSEA	0.07	Poor fit
SRMR	0.06	Poor fit
CFI	0.73	Poor fit
TLI	0.71	Poor fit

8.4.4 Comparison of CFA of the different alternative models

As discussed in Chapter 7, confirmatory factor analysis (CFA) can be employed to compare alternative models and to test hypotheses (Ng et al., 2017); that is, to determine whether these alternative models are more consistent with the data than single-factor models (Byrne, 1998; Hooper et al., 2008; Ng et al., 2017). The same procedure in Chapter 7 is employed in this chapter to examine the fit indices for the alternative models such, as the orthogonal factor model, correlated factor model and hierarchical model. The results of the fit indices are demonstrated in Table 8.4, for comparison purposes among the different alternative models and their factor loadings.

Table 8. 4 Different model fit comparison for motivation towards mathematics

No	Model	X^2	df	X^2/df	RMSEA	SRMR	CFI	TLI
1	One factor model **	553.26	86	6.43	0.08	0.07	0.72	0.66
2	Orthogonal model	512.14	167	3.07	0.05	0.04	0.86	0.84
3	Three-correlated factors model 2*	337.47	142	2.39	0.04	0.04	0.92	0.91
4	Hierarchical model 1	512.14	167	4.51	0.05	0.04	0.86	0.84
5	Hierarchical model 2*	339.20	145	2.33	0.04	0.04	0.92	0.90

*Note. ** Item Mvtn25 deleted*

The fit indices results in Table 8.4 show that both orthogonal models, three-correlated model factor 1 (*X^2/df>3),* hierarchical model 1 have poor fit indices demonstrating a poor model. This is also consistent with the factor loading of some of the items in each model. Hierarchical model 2 fitted the data with the fit indices shown in Table 8.4. Similarly, hierarchical model 2 also demonstrated acceptable fit indices. However, hierarchical model 2 is more consistent with the theoretical framework of the study, as discussed in Chapter 2, and is used in this study. As shown Table 8.5, the hierarchical model 2 items at each subscale have relationships with the construct motivation, with all items displaying factor loading more than 0.3 and defining the three latent motivational traits. These factors/traits are students' *achievement goal, engagement* and *interest i*n mathematics. Together, these three traits formed the hierarchical model that the motivation scale hypothesised. As such, the trait *engagement* had six items (Mtvn07, Mtvn14-Mtvn16, Mtvn18 and Mtvn19), *interest* had eight items (Mtvn08, Mtvn17, Mtvn20-Mtvn24 and Mtvn26) and *achievement goal* had five items (Mtvn09-Mtvn13) that measured the motivation scale. However, Mtvn25 from the subscale *interest* is deleted as it displayed a factor loading below 0.3 and so is measuring something else other than motivation. This structure of the hierarchical model is presented below in Figure 8.2.

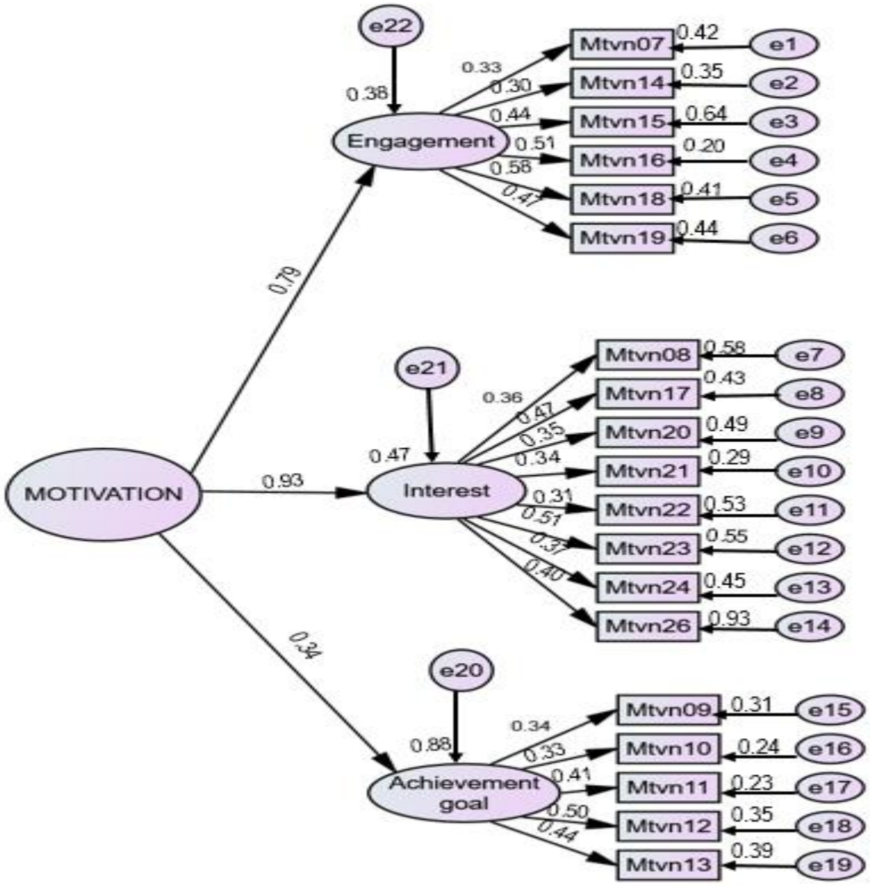


Figure 8. 2 Structure of the hierarchical model

8.4.5 CFA for the motivation hierarchical model

Items on the motivation scale are checked for the factor loadings for the hierarchical model to establish a relationship between the observed variable and the subscales. The results of the factor loading appearing in Table 8.5 clearly show all the items/variables represent each of the three subscales (*achievement goal, engagement* and i*nterest*), displaying factor loadings above 0.3. The items with the hierarchical model showed improvement over the single model with all factor loadings above 0.3. This result shows that the items have a relationship with the subscales in terms of the scale's structure.

Table 8. 5 Results of the factor loading of the model

Second-order factor	First-order factors	Loadings	Item (s)	Loading (se)*
Motivation	Engagement	0.79	Mtvn07	0.33 (0.03)
			Mtvn14	0.30 (0.03)
			Mtvn15	0.44 (0.04)
			Mtvn16	0.51 (0.04)
			Mtvn18	0.58 (0.03)
			Mtvn19	0.47 (0.03)
	Interest	0.93	Mtvn08	0.36 (0.03)
			Mtvn17	0.47 (0.03)
			Mtvn20	0.35 (0.03)
			Mtvn21	0.34 (0.02)
			Mtvn22	0.31 (0.03)
			Mtvn23	0.51 (0.03)
			Mtvn24	0.37 (0.03)
			Mtvn26	0.40 (0.03)
	Achievement goal	0.34	Mtvn09	0.34 (0.02)
			Mtvn10	0.33 (0.02)
			Mtvn11	0.47 (0.02)
			Mtvn12	0.50 (0.03)
			Mtvn13	0.44 (0.03)

*n=729

8.4.6 Model fit

As in Chapter 7, the results of ratio of chi-square to its degree of freedom (x^2 / df), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and Tucker-Lewis Index (TFI) are examined in evaluating the hypothesised model. This analysis is undertaken to gauge the overall model fit to the data. Table 8.6 shows that the five indices employed in this study demonstrated acceptable fit. The results show that the hierarchical model fits the data and even shows improvements in the factor loading compared to the factor loading of the single-factor model. Thus, it is concluded from these CFA results that the hierarchical model appears to be appropriate for the motivation scale, with better factor loadings and model fit indices.

Table 8. 6 Result of the model fit

Fit Index	Obtained values	Remarks
X^2/df	339.20/145=2.3	Acceptable fit
RMSEA	0.04	Acceptable fit
SRMR	0.04	Acceptable fit
CFI	0.92	Acceptable fit
TLI	0.90	Acceptable fit

As discussed in Chapter 7, it is important to establish the validity and reliability of the variables in this study to provide accurate estimates for the subsequent analysis of the student-level model. Nineteen out of the 20 items are confirmed with their structure, as well as the relationships between the observed variables and their underlying latent variable (motivation), through confirmatory factor analysis (CFA). This indicates that the motivation scale is made up of the subscales such as *engagement, interest* and *achievement goal*. The next section of the discussion focuses on checking the utility of these 19 items at the item level through a multidimensional Rasch model of the three dimensions. In other words, the Rasch analysis is used to verify that the scope of the structures of the scales confirmed in the CFA analysis fit the Rasch model.

8.5 Multidimensional Rasch analysis

Based on the hierarchical model of the motivation scale from the CFA analysis, a multidimensional items analysis is conducted. This was carried out to check the utility of the 19 items using the Rasch measurement model (RMM) with ConQuest version 4.0, as discussed in the previous chapters. The partial credit model (PCM) discussed in Chapter 3 is employed because partial successes on the item responses are expressed by partial credits, and are hierarchical (Eggert & Bögeholz, 2010; Huggins-Manley & Algina, 2015). This implies that a response that is given a higher partial credit is better than a response that is assigned a lower partial credit. This procedure is useful for surveys such as the motivation scale instrument, where items are not marked for correct or incorrect answers (Penfield et al., 2008). Survey results from the analysis, appearing in Table 8.7, demonstrate that data fitted the model well which indicates that PCM is a more parsimonious model (Wang & Wu, 2011).

8.5.1 Item analysis with the Partial Credit Model

The 19 items in the motivation scale analysed from the CFA are subjected to item analysis using multidimensional Rasch items analysis with PCM. This was carried out to test the unidimensionality of the 19 items measuring the construct (motivation). This involves examining each item's fit statistics using similar statistical criteria and procedures discussed in Chapter 7. Table 8.7 shows the results of the initial run of this analysis. From this table, it can be observed that all items fit the model with the INFIT MNSQ criteria of 0.6 to 1.40 (Bond et al., 2007). The separation reliability was 0.99 showing minimal measurement error and was discriminating between low and high-motivation respondents (Alagumalai et al., 2005; Ben, 2010). The items' estimates, delta(s), thresholds, average expected values and biserial points were ordered. However, the t- values of Mtvn08, Mtvn10, Mtvn12 and Mtvn22 are not within the acceptable fit criteria of -2 to 2, and as such are not within the criteria of the model (see Table 8.2). Moreover, as shown in Table 8.7, the Mtvn12 discrimination index is zero and the *t*-value is 8.0.This indicates that this item is not discriminating between the high and low motivation students, and so the researcher decided to delete the item due to violation of RMM requirement (Wu & Adams, 2007; Wu et al., 2016).

Table 8. 7 Model parameter estimates of motivation towards maths (n=729)

			Weighted Fit						
Item(s)	Estimate	Error	MNSQ	CI	t	Pt Bis	Item Delta		
Mtvn07	0.61	0.04	1.04	(0.89,1.11)	0.7	0.45	-1.09	0.15	2.79
Mtvn14	0.59	0.04	0.96	(0.90,1.10)	-0.7	0.55	-0.75	0.38	2.16
Mtvn15	-0.27	0.04	0.96	(0.89,1.11)	-0.8	0.40	-1.68	-0.46	1.32
Mtvn16	-0.95	0.04	0.97	(0.84,1.16)	-0.4	0.36	-2.87	-0.87	0.18
Mtvn18	-0.51	0.04	1.04	(0.86,1.14)	0.5	0.29	-1.10	-1.03	0.61
Mtvn19	0.53*	0.09	1.02	(0.90,1.10)	0.5	0.40	-1.03	0.25	2.36
Mtvn08	-0.64	0.05	1.17	(0.89,1.11)	2.9	0.38	-2.10	-1.28	1.46
Mtvn17	1.35	0.04	1.07	(0.90,1.10)	1.5	0.53	-0.47	1.78	2.75
Mtvn20	0.93	0.04	1.03	(0.90,1.10)	0.7	0.56	-0.84	1.36	2.27
Mtvn21	0.49	0.04	0.95	(0.90,1.10)	-1.0	0.57	-1.49	0.50	2.45
Mtvn22	-0.46	0.04	0.85	(0.90,1.10)	-3.1	0.62	-2.14	-0.50	1.26
Mtvn23	-0.11	0.04	0.91	(0.91,1.09)	-2.0	0.57	-2.56	0.31	1.91
Mtvn24	-0.35	0.04	1.08	(0.89,1.11)	1.4	0.48	-1.29	-0.91	1.16
Mtvn26	-1.21*	0.11	0.95	(0.89,1.11)	-0.9	0.54	-2.71	-1.62	0.71
Mtvn09	-0.33	0.04	0.98	(0.90,1.10)	-0.5	0.43	-1.93	-0.24	1.18
Mtvn10	0.08	0.03	0.88	(0.90,1.10)	-2.6	0.65	-0.67	0.15	0.77
Mtvn11	-0.67	0.04	0.88	(0.89,1.11)	-2.0	0.64	-1.59	-1.37	0.92
*Mtvn12***	*1.51*	*0.03*	*1.44*	*(0.90,1.10)*	*8.5*	*0.00*	*0.47*	*1.70*	*2.37*
Mtvn13	-0.59*	0.08	0.97	(0.89,1.11)	-0.5	0.49	-1.34	-1.34	0.92

** deleted item

Separation Reliability = 0.996 Chi-square test of parameter equality = 4163.39, df = 16, Sig Level = 0.000

After deleting Mtvn12, final multidimensional analysis was run. The results of the final run of analysis (shown in Table 8.8) indicate a significant improvement on the statistical fits of the items. The 18 items in Table 8.8 show that item difficulty ranges from -1.22 to 1.35. Item difficulty indicates the level of the latent trait at which the probability of the given response to the item was 0.50. The negative values indicate that the items are easier to responds to, and positive values indicate items that are harder to responds to. As shown in Table 8.8, the MNSQ values of the 18 items are within 0.6 to 1.40, and they are discriminating between high and low ability respondents. Besides this, separation reliability value of 0.99 implies less error and high discriminating power (Alagumalai et al., 2005; Hailaya, 2014). This indicates that the items still have high measurement capacity (Wright & Stone, 1999). Furthermore, the item delta(s) are in order, as demonstrated in Table 8.8, indicating that response choices on the scale are in order. However, three items (Mtvn08, Mtvn22 and Mtvn11) *t*-values are still above the criteria discussed in Chapters 3 and 7 despite significant changes in their values. Since this is not a "high-stakes" test, a more lenient approach is taken regarding the *t*-values of those these items. This approach is taken because these items may contain information or findings that are significant to the study (Bond, Fox, & Christine, 2007).

Table 8. 8 Model parameter estimates of motivation towards maths (n=729)

Item(s)	Estimate	Error	Weighted Fit MNSQ	CI	t	Pt Bis	Item Delta		
Mtvn07	0.62	0.04	1.05	(0.89,1.11)	0.8	0.45	-1.09	0.16	2.80
Mtvn14	0.60	0.04	1.00	(0.90,1.10)	0.0	0.53	-0.75	0.39	2.17
Mtvn15	-0.27	0.04	0.95	(0.89,1.11)	-0.8	0.41	-1.70	-0.45	1.33
Mtvn16	-0.98	0.05	0.92	(0.85,1.15)	-1.0	0.36	-2.17	-0.95	0.19
Mtvn18	-0.51	0.04	0.99	(0.86,1.14)	-0.1	0.30	-1.10	-1.05	0.62
Mtvn19	0.53*	0.09	1.03	(0.90,1.10)	0.6	0.41	-1.03	0.26	2.37
Mtvn08	-0.64	0.05	1.15	(0.89,1.11)	2.6	0.39	-2.12	-1.27	1.47
Mtvn17	1.35	0.04	1.11	(0.90,1.10)	2.1	0.52	-0.46	1.78	2.76
Mtvn20	0.93	0.04	1.00	(0.90,1.10)	0.1	0.57	-0.84	1.36	2.28
Mtvn21	0.49	0.04	0.94	(0.90,1.10)	-1.2	0.58	-1.49	0.50	2.45
Mtvn22	-0.46	0.04	0.89	(0.90,1.10)	-2.2	0.63	-2.15	-0.50	1,26
Mtvn23	-0.11	0.04	0.95	(0.91,1.09)	-1.0	0.58	-2.57	0.31	1.91
Mtvn24	-0.35	0.04	1.09	(0.89,1.11)	1.6	0.49	-1.29	-0.91	1.16
Mtvn26	-1.22*	0.11	0.95	(0.89,1.11)	-0.9	0.54	-2.76	-1.62	0.72
Mtvn09	-0.30	0.04	0.96	(0.91,1.09)	-0.9	0.44	-1.81	-0.21	1.12
Mtvn10	-0.62	0.04	0.89	(0.89,1.11)	-1.9	0.62	-1.43	-1.32	0.89
Mtvn11	1.45	0.04	1.28	(0.91,1.09)	3.6	0.02	0.49	1.63	2.23
Mtvn13	-0.53*	0.07	0.94	(0.89,1.11)	-1.0	0.48	-1.19	-1.17	0.88

Separation Reliability = 0.997 Chi-square test of parameter equality = 4163.39, df = 16, Sig Level = 0.000

After checking the utility of the motivation scale items, differential item functioning (DIF) is employed to ascertain whether the items are biased towards a particular gender. This is determined through the procedure and criteria of item fit and item fit threshold approaches discussed in Chapter 7.

Table 8. 9 DIF Item fit (INFIT MNSQ) results for gender and Item fit threshold results for gender

Items No	Female			Male			d1-d2
	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ	
Mtvn07	0.06	0.03	0.43	-0.06	0.03	0.43	0.12
Mtvn14	0.12	0.03	0.54	-0.12	0.03	0.53	0.24
Mtvn15	0.01	0.03	0.6	-0.01	0.03	0.61	0.02
Mtvn16	-0.19	0.04	0.77	0.19	0.04	0.74	-0.38
Mtvn18	-0.16	0.04	0.77	0.16	0.04	0.69	-0.32
Mtvn19	-0.01	0.03	0.51	0.01	0.03	0.57	-0.02
Mtvn08	-0.08	0.03	0.48	0.08	0.03	0.48	-0.16
Mtvn17	0.1	0.03	0.48	-0.1	0.03	0.53	0.20
Mtvn20	0.04	0.03	0.55	-0.04	0.03	0.53	0.08
Mtvn21	0.04	0.03	0.46	-0.04	0.03	0.45	0.08
Mtvn22	0.00	0.03	0.52	0.00	0.03	0.5	0.00
Mtvn23	0.07	0.03	0.45	-0.07	0.03	0.45	0.14
Mtvn24	-0.01	0.03	0.58	0.01	0.03	0.62	-0.02
Mtvn26	-0.01	0.03	0.5	0.01	0.03	0.49	-0.02
Mtvn09	-0.02	0.03	0.54	0.02	0.03	0.55	-0.04
Mtvn10	0.06	0.03	0.47	-0.06	0.03	0.5	0.12
Mtvn11	0.07	0.03	0.86	-0.07	0.03	0.84	0.14
Mtvn13	-0.02	0.12	0.48	0.02	0.12	0.57	-0.04

The INFIT MNSQ of the 19 items shown in Table 8.9 were within the range of 0.6 to 1.40, and the difference between the threshold (estimates) of the genders was under 0.5 logits (Hungi, 2005). These results show that the items are within the acceptable range, indicating that there is no item detected with DIF for gender for both Grade 10 and 12 respondents; though nine items are somewhat easier for female and 10 items are somewhat more difficult for males, as shown in Table 8.9.

8.5.2 Person-item map

The person-map in Figure 8.3. (p.206) (from left to right) is employed to assess the three dimensions (*engagement, interest* and *achievement goal*) of the motivation scale on which both items and respondents are calibrated on a logit scale. In Figure 8.3, the numbers on the far left are from -1 (low motivation) to around +3 (very high motivation). The relative item difficulty is plotted on the right side of the scale and personal ability estimates on the left side of the same scale on the person-map. The respondents at the top of the map shows higher motivation, while those at the bottom demonstrates lower motivation. Similarly, the items on the top are more difficult to endorse, while those at the bottom are relatively easier to endorse (Bond, Fox, & Lacey, 2007; Bond & Fox, 2015). The logit zero on the person-map is set at the average item difficulty and overall, the mean motivation of students was higher than the average difficulty.

The person-item map (Figure 8.3) of the three dimensions (*engagement, interest* and *achievement goal*) show that majority of the items are easy to endorse for most of the respondents. This indicates that majority of the respondents (above mean) are motivated to favourably respond to most of the motivation items. For instance, according to the person-item map of the three dimensions, difficult items such as Mtvn11 for the *achievement goal* dimension, and Mtvn17 for the *interest* dimension, were endorsed by motivated respondents. On the other hand, Mtvn16 for the *engagement* dimension and Mtvn26 for the dimensions *interest* are relatively easy to endorse. However, most of the items in the three dimensions below the average (o logits) are endorsed by more than 50% of the respondents, particularly for the *engagement* dimension.

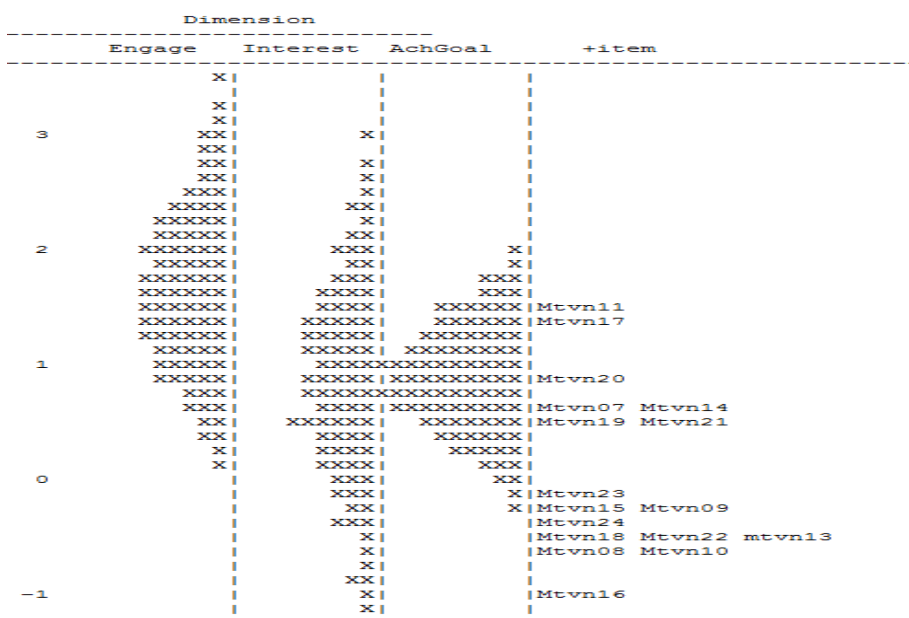


Figure 8. 3 Person-item map of motivation scale

8.6 Transformation to W-scores

As discussed in Chapter 3, the WLE values from the Rasch analysis were transformed into W-scores to eliminate negative values, and the need for decimal points (Benson et al., 2016; Linacre, 2009; Woodcock, 1999). Microsoft Excel is used for this transformation using the formula W=9.1024 x logits + c, where c is a constant term, selected to eliminate the negative values (Benson et al., 2016; Woodcock, 1999). The constant term used in this study is 500, which is consistent with PISA studies. The W- scores are used in subsequent chapters for structural equation modelling and hierarchical linear modelling.

8.7 Comparison of different RMM and CFA results

As in Chapter 7, the primary result of the CFA and multidimensional Rasch Model (MRM) analysis in this chapter was the establishment of the scale as a multidimensional measure. The results of the CFA analysis identified a three-factor model in which items loaded better with acceptable model fit compared to the single-factor model. Tables 8.2 (p.198) Tables 8.5 (p.202) and 8.6 (p.203) show the evaluation of the dimensionality of the items/constructs, and identify the traits that influence response patterns for the three dimensions (Wright, 1996). These tables reveal that CFA analysis is more informative at the subscale level with the hierarchical model. The results of Table 8.5 (p.202) show acceptable factor scores of the three-correlated-factor model for the subscales achievement *goal, engagement* and *interest*. These results indicate that the observed variables from the data possess a relationship with the constructs confirming the theory (Beccaria et al., 2018; Liao, Murphy, & Barrett-Lennard, 2018).

The results in Tables 8.7 (p.205), 8.8 (p.206) and 8.9 (p.207) show the evaluation of the dimensionality of the items/constructs, identifying the traits that influence the response pattern (Wright, 1996). This is demonstrated through the results of the statistical fits and the factor scores associated with errors. The results of the final 18 items appearing in Table 8.8 (p.206) show that the items are measuring the three dimensions *engagement, interest* and *achievement goal* for the motivation construct with statistical fits that meet the criteria of RMM. As in Chapter 7, the multidimensional Rasch model allowed for the model factor structure, and is more informative for assessing the performance of individual items. Further, item analysis was obtained for each subscale and for the whole attitude scale. Based on the infit statistics, the 18 items fitted the Rasch model and the hierarchical model using CFA. The Rasch analysis also provided estimates of the items' information function and the test information function (Osteen, 2010). This has allowed the researcher to obtain specific estimates of standard errors of measurement, instead of relying on an averaged standard error of measurement obtained from the CFA (Osteen, 2010).

8.8 Model employed for the study

The comparison of the confirmatory factor analysis (CFA) results from the single factor model and the hierarchical model for the motivation scale show that the hierarchical mode fit the data with a better structure. Further, from the Rasch item analysis perspective, the items in the hierarchical model are viewed as forming three dimensions, and the 18 items fitted the Rasch model, as the responses varied in terms of level of agreement (Grimbeek & Nisbet, 2006). This indicates that the combination of Rasch item analysis and CFA provides useful information for the hierarchical model, and it will consequently be employed in the subsequent chapters of this study.

8.9 Summary

This chapter focused on the context in which the motivation scale is used and the validation of the variables/items through confirmatory factor analysis (CFA) and multidimensional Rasch model techniques. CFA was carried out to establish the relationships between the observed variables and constructs (motivation) and multidimensional Rasch model analysis was employed to check the items in the three dimensions (*engagement, interest* and *achievement goal*) at the item level after CFA was carried out. The results of these two techniques clearly showed the factor structure of the variables for the hierarchical model that fit the data, and therefore the hierarchical model will be used in the subsequent analysis of this study. In the later chapters of this thesis, the individual score for each respondent (scale score) will be transformed by calculating the W-score from the WLE, as discussed in Chapter 3. The next chapter analyses and discusses the teacher-level scales' validation process using statistical techniques similar to those employed in Chapters 7 and 8.

Chapter 9: Validation of teacher-level scales

9.1 Introduction

Teacher-level factors such as quality of teaching, classroom environment and instructional resources play a significant role in predicting students' academic results in general (Gherasim et al., 2013; Green et al., 2018; Tuomi, 2006), and students' mathematics results specifically (Martin & Kelly, 1996; Mullis et al., 2012). This chapter follows on from Chapters 7 and 8 to evaluate the study's teacher-level scale items, which have been derived from existing scales in TIMSS and PISA, using the same chapter structure and statistical analysis techniques as the preceding chapters. This chapter concludes with a summary highlighting the key points and providing direction for subsequent chapters.

9.2 Contextualisation of teacher-level scales

This section contextualises and discusses briefly (for a detailed discussion, see Chapter 2) the different teacher-level scales used in this study: *quality of teaching, classroom environment* and *instructional resources,* which have been developed in accordance with the TIMSS and PISA assessments. In countries that have participated in these international studies, these teacher-level scales have been shown to have an impact on students' mathematics results (Eklöf, 2007; Martin et al., 2017; PISA, 2016). Athough, PNG has not participated in TIMSS and PISA, it is an assumption of this study that these scales would therefore also have an effect on the mathematics results of students in Port Moresby, PNG.

9.2.1 Quality of teaching

Quality of teaching is paramount for better student learning outcomes (Alton-Lee, 2003; Green et al., 2018). The concept encompasses teachers' content knowledge and their teaching pedagogy (OECD, 2005; PISA, 2016; Thomson, Wernert, Underwood, & Nicholas, 2008), and their ability to contribute meaningfully to student learning.

Teachers' knowledge of mathematics and its relationship to other content (both within and outside of mathematics) are essential elements for student learning (Martin & Mullis, 2013; Wilkins, 2008). As such, an effective mathematics teacher has a sound, coherent knowledge of mathematics that is appropriate to the student level they teach (OECD, 2011; Wilkins, 2008). High quality teachers present to students' an understanding of the broader mathematics curriculum, how mathematics is represented, communicated, and why mathematics is taught (Baumert, Blum, & Neubrand, 2004). These characteristics and teaching strategies allow them to effectively pass on knowledge to students.

Knowledge of teaching pedagogy includes an understanding of how students process, store, retain, and recall information, and of how teachers approach the activity of student instruction (Hiebert & Stigler, 2000). Further, it also includes knowledge of a variety of examples for each mathematical idea, specific instructional techniques, and instructional materials (Hiebert, 2003). These points emphasise that effective teachers of mathematics possess a rich knowledge of how students learn mathematics (Hiebert, 2003; Meyer & Benavot, 2013), as well as knowledge of students' mathematical development, such as through learning sequences, appropriate representations, models and language (Baumert et al., 2004; PISA, 2016).

In this study (as shown in Table 9.1), items from the *quality of teaching* scales in TIMSS 2015 and PISA 2012 were adopted and used. The *quality of teaching* scale instrument used in this study employs a five-point Likert-type scale: "almost never", "seldom", "sometimes", "often" and "very often" (DeLuca, Valiquette, Coombs, LaPointe-McEwan, & Luhanga, 2018; Krosnick, 2018). Since this scale was used for interviews as well, it was appropriate and necessary to include a middle or neutral position for comparison purposes with the themes developed (Creswell, 2008; Creswell & Clark, 2007). The subscale content knowledge is represented by QT06-QT12, QT17 and QT19, and teaching pedagogy by QT13-QT16, QT18 and QT20. Item responses were coded 1, 2, 3, 4 and 5 corresponding to the categories "almost never", "seldom", "sometimes", "often" and "very often".

Table 9. 1 Summary of items in the quality teaching scale used in the study

Item code	Item Text	Item originated from/Item No	Subscale
QT06	I Inspire students to learn mathematics	TIMSS 17a	Quality of teaching
QT07	I demonstrate to students a variety of problem-solving strategies.	TIMSS17b	Quality of teaching
QT08	I provide challenging tasks for the highest achieving students.	TIMSS17c	Quality of teaching
QT09	I adapt my teaching to engage students' interest.	TIMSS17d	Quality of teaching
QT10	I assist students' comprehension of mathematics.	TIMSS17f	Quality of teaching
QT11	I improve understanding of struggling students	TIMSS17g	Quality of teaching
QT12	I explain mathematics problem solving strategies step by step.	PISA	Quality of teaching
QT17	Students work on problems while I am occupied by other tasks.	TIMSS18d	Quality of teaching
QT19	I make mathematics relevant to students.	TIMSS17h	Quality of teaching
QT13	I assist in the development of students' higher-order thinking skills.	TIMSS17i	Quality of teaching
QT14	I explain new mathematics content appropriately.	TIMSS18a	Quality of teaching
QT15	I explain and demonstrate how to solve mathematics problems	TIMSS18b	Quality of teaching
QT16	I ensure students work on problems individually or with peers with my guidance	TIMSS18d	Quality of teaching
QT18	I solve mathematics like the examples in students textbook	PISA	Quality of teaching
QT20	I ask students to work on problems that there is no immediate obvious method of solution.	TIMSS17g	Quality of teaching

9.2.2 Classroom environment

An orderly and cooperative learning environment is required for effective student learning (Breakspear, 2014; Istance & Kools, 2013; Turner, 2017), and such successful environments are desired and valued by both students and teachers (Breakspear, 2014; Ray & Margaret, 2003). The learning environment encompasses the general atmosphere of the classroom including norms and values, as well as teacher-student relationships (Istance & Kools, 2013; OECD., 2005). A learning environment can be evaluated by several factors, including the degree of discipline among students (Istance & Kools, 2013), the quality of the relationship between students and their teachers, and how this affects teaching and learning in the classroom (Istance & Kools, 2013; Turner, 2017). Students learn more when they feel that their teachers take them seriously, which is established through strong and effective bonds between them (Istance & Kools, 2013; Istance & Paniagua, 2018). Through these positive relationships, "communal learning environments are created" (Istance & Kools, 2013; Turner, 2017), and adherence to norms conducive to learning are both promoted and strengthened (Birch and Ladd, 1998).

Table 9.2 shows the classroom environment scale instrument used in this study, which contains a four-point Likert-type scale: "strongly agree", "agree", "disagree", and "strongly disagree" (de Cássia Nakano & Primi, 2014; Huggins-Manley & Algina, 2015). The 13 items for the *classroom environment* instrument are designated from CE21-CE33. Item responses were coded 1, 2, 3 and 4 corresponding to the categories "strongly disagree"," disagree"," agree" and "strongly agree". Missing or omitted item responses were coded "9", an arbitrary value assigned for recognition with the statistical software as a non-response (Blackwell et al., 2017).

Table 9. 2 Summary of items in the classroom environment scale used in the study

Item code	Item Text	Item originated from/Item No	Subscale
CE21	I have a good rapport with students.		Classroom environment
CE22	I take personal interest in students' engagement in learning	TIMSS17a	Classroom environment
CE23	I help students with problems doing mathematics.	TIMSS17e	Classroom environment
CE24	I consider students feelings in learning.	PISA1a	Classroom environment
CE25	I decide where the students will sit.	PISA13b	Classroom environment
CE26	Student to teacher ratio is manageable	TIMSS11a	Classroom environment
CE27	There is classroom discussion.	TIMSS14d	Classroom environment
CE28	All students in the class do the same work.	TIMSS18e	Classroom environment
CE29	I teach without questions from the students	PISA8b	Classroom environment
CE30	I talk with each student in classroom	PISA12c	Classroom environment
CE31	There is adequate space for teaching in the classroom.	PISA12b	Classroom environment
CE32	Students ask the teacher question.	PISA12	Classroom environment
CE33	Students explain mathematics problems and graphs	TIMSS14b	Classroom environment

9.2.3 Instructional resources

The kinds of methods employed by educators in teaching mathematics at schools are, to a large extent, influenced by the kinds of resources and facilities available in the school (Martin & Mullis, 2013; Mullis & Chrostowski, 2004). According to Martin and Mullis (2013), teaching methods, in turn, influence the level and quality of student participation and performance in mathematics. In general, where resources and facilities such as teachers, textbooks, tools, equipment and teaching aids are adequate, teaching approaches tend to be student-centred (Martin & Mullis, 2013; Mullis & Chrostowski, 2004). Student-centred approaches are heavily dominated by students, and throughout the learning process they remain active participants, expected to discover things and learn (Martin & Mullis, 2013; Mullis et al., 2012). In this environment, the teacher, therefore, becomes a facilitator, and so should be adequately informed on the subject and possess strong communication skills to encourage student interest in the subject.

The *instructional resources* scale items are shown in Table 9.3, and are labelled from IR34–IR45. Item responses were coded 1, 2, 3 and 4, corresponding to the categories "strongly disagree"," disagree","

agree" and "strongly agree". As above, the missing or omitted item responses were coded "9" to indicate a non-response (Blackwell et al., 2017).

Table 9. 3 Summary of items in the curriculum scale used in the study

Item code	Item Text	Item originated from/Item No	Subscale
IR34	Instructional materials such as textbooks are adequate for students.	TIMSS 13A (a)	Instructional resources
IR35	There are adequate library resources relevant for mathematics lesson.	TIMSS 13B (c	Instructional resources
IR36	Students are provided calculators for mathematics lessons.	TIMSS 13B (d)	Instructional resources
IR37	Teachers have adequate instructional materials and supplies.	TIMSS 13A (b)	Instructional resources
IR38	Internet is available for research.	TIMSS 13B (b)	Instructional resources
IR39	The mathematics curriculum is regularly revised.	TIMSS 13B (a)	Instructional resources
IR40	Mathematics termly programs are evaluated at the end of each term.	TIMSS M6	Instructional resources
IR41	The mathematics curriculum is relevant and practical for the students learning.	TIMSS M8 (ii)	Instructional resources
IR42	Objectives of mathematics curriculum are measurable.	TIMSS M4	Instructional resources
IR43	All Grade 10 and 12 math topics are covered before the national exams.	TIMSS M8 (i)	Instructional resources
IR44	The mathematics curriculum is regularly revised.	TIMSS M2 B	Instructional resources
IR45	Mathematics teachers' guides are easy to follow in implementing the curriculum.	TIMSS M5	Instructional resources

9.3 CFA of the measurement models for the teacher-level scales

The items of the *quality of teaching, classroom environment* and *instructional resources* scales were further subjected to confirmatory factor analysis (CFA) to determine the relationship between the items/variables and the scales using Mplus 7.0 (Muthen & Muthen,2012). This technique was employed to examine and uncover the underlying structures among variables, adjusting for measurement error (Brown, 2014; Chamberlin et al., 2017; Han & Carpenter, 2014). The single-factor model was first tested for model fit with the data. As discussed in Chapters 7 and 8, the minimum factor loading value used in all factor analysis was greater than 0.30, consistent with other studies (Hailaya, 2014; Kline, 2014).

9.3.1 Structure analysis of the hypothesised measurement model

The CFA model was initially constructed using the one-factor structure. This means that the number of observed variables constituted the latent variables (*quality of teaching, classroom environment* and *instructional resources*) in CFA (Han & Carpenter, 2014; Taasoobshirazi & Wang, 2016). In this model the 15 items labelled “QT”, 13 items labelled “CE” and 12 items labelled “IR” are loaded directly onto the latent variables *quality of teaching, classroom environment* and *instructional resources*, respectively (see Figure 9.1). As shown in Figure 9.1, item errors/residuals in the model are indicated with numbers in the diagram.

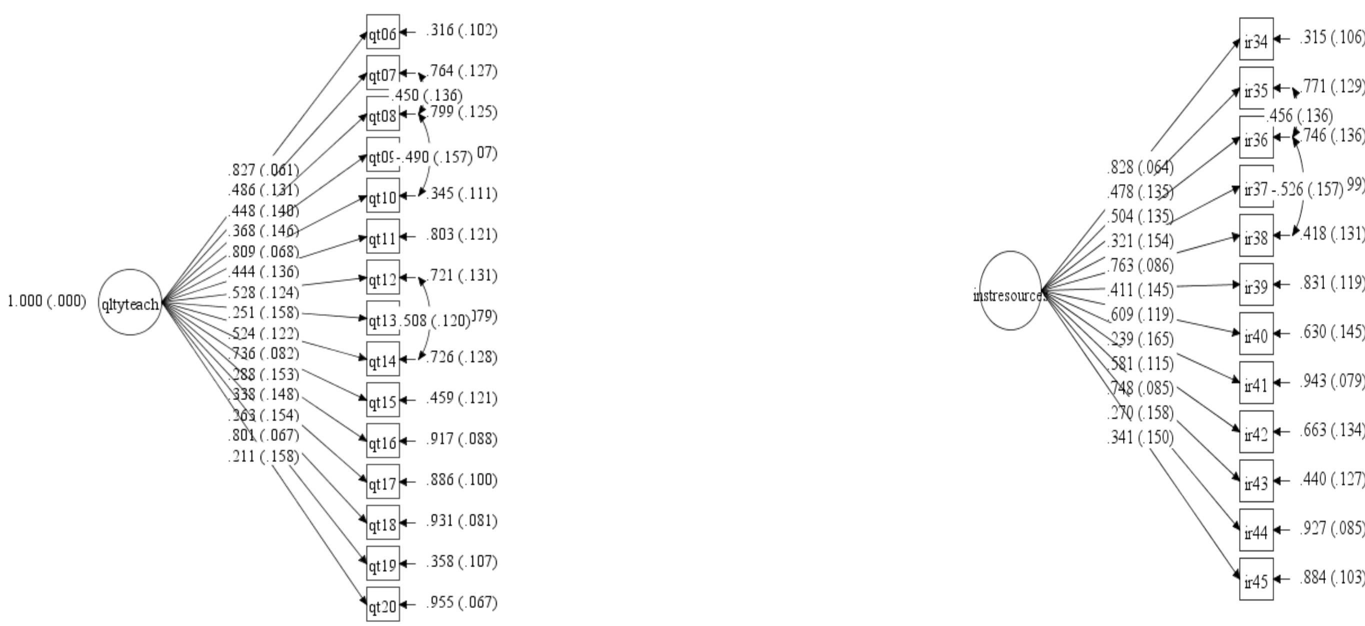


Figure 9. 1 Structure of one-factor model for teacher level factors

9.3.2 CFA of the hypothesised measurement model

The *quality of teaching* (QT06-QT20), *classroom environment* (CE21-CE33) and *instructional resources* (IR34-IR45) items are examined using the factor loadings to confirm whether or not the items demonstrated the latent variable they represented, as discussed in Chapters 7 and 8. Table 9.4 shows the factor loading for the items under the three scales. This table indicates that three items for *quality of teaching* (QT13, QT18, and QT20), three items for classroom environment (CE28, CE31 and CE33) and two items for *instructional resources* (IR41 and IR44) had factor loadings below the adopted value of 0.3, and therefore exhibited unacceptable factor loading. In other words, these items are not representing the scales' structures. Hence, the hypothesised single-factor model may not be appropriate for the *quality of teaching, classroom environment* and *instructional resources* scales.

Table 9. 4 Factor loadings of teacher level scales

Quality of teaching scale		
Factor	Items	Loading (se)*
	QT06	0.82 (0.06)
	QT07	0.48 (0.13)
	QT08	0.44 (0.14)
	QT09	0.36 (0.14)
	QT10	0.81 (0.06)
	QT11	0.44 (0.13)
QT	QT12	0.52 (0.12)
	QT13	0.25 (0.15)
	QT14	0.52 (0.12)
	QT15	0.73 (0.08)
	QT16	0.30 (0.15)
	QT17	0.33 (0.14)
	QT18	0.26 (0.15)
	QT19	0.80 (0.06)
	QT20	0.21 (0.15)

Classroom environment scale		
Factor	Items	Loading (se)*
	CE21	0.84 (0.06)
	CE22	0.47 (0.13)
	CE23	0.49 (0.13)
	CE24	0.31 (0.15)
	CE25	0.76 (0.08)
	CE26	0.40 (0.14)
CE	CE27	0.61 (0.11)
	CE28	0.22 (0.16)
	CE29	0.58 (0.11)
	CE30	0.74 (0.84)
	CE31	0.26 (0.15)
	CE32	0.34 (0.15)
	CE33	0.25 (0.15)

Instructional resource scale		
Factor	Items	Loading (se)*
	IR34	0.82 (0.06)
	IR35	0.47 (0.13)
	IR36	0.50 (0.13)
	IR37	0.32 (0.15)
	IR38	0.76 (0.08)
	IR39	0.41 (0.14)
IR	IR40	0.60 (0.11)
	IR41	0.23 (0.16)
	IR42	0.58 (0.11)
	IR43	0.74 (0.08)
	IR44	0.27 (0.15)
	IR45	0.34 (0.15)

9.3.3 Model fit

As with Chapters 7 and 8, the estimates of the parameters used to analyse the goodness of the model fit to the data is examined using the results of ratio of chi-square to its degree of freedom (x^2 / df), root mean square error of approximation (RMSEA), standardised root mean square residual (SRMR), comparative fit index (CFI) and Tucker-Lewis Index (TFI). These indicators of fit are used to evaluate the hypothesised models for *teaching of quality, classroom environment* and *instructional resources* scales. The criteria of these fit indices were discussed in detail in Chapter 3. The results of the scales fit indices are illustrated in Table 9.5.

Table 9. 5 Results of fit indices for the single teacher scales structures

Quality of teaching scale		
Fit Index	Values	Remarks
X^2/df	156.52/ 87=1.7	Acceptable fit
RMSEA	0.14	Poor fit
SRMR	0.13	Poor fit
CFI	0.7	Poor fit
TLI	0.64	Poor fit

Classroom environment scale		
Fit Index	Values	Remarks
X^2/df	140.9/6 3=2.2	Acceptable fit
RMSEA	0.17	Poor fit
SRMR	0.14	Poor fit
CFI	0.61	Poor fit
TLI	0.51	Poor fit

Instructional resource scale		
Fit Index	Values	Remarks
X^2/df	127.76/ 52=2.4	Acceptable fit
RMSEA	0.18	Poor fit
SRMR	0.148	Poor fit
CFI	0.61	Poor fit
TLI	0.51	Poor fit

Table 9.5 shows the results of the five fit indices employed in testing the goodness of the model fit. One of these fit indices, (X^2/df) demonstrated acceptable fit, while the other four (X^2, RMSEA, SRMR, CFI and TLI) demonstrated poor fit. These results of the fit indices indicate that the three teacher scales' single factor structures did not fit the data. This result is consistent with the results of the low factor loadings shown in Table 8.8. Despite these results, the three teacher-level scales were not tested for alternative models because they are regarded as a single-factor model in the literature review section of this study.

The next section of the discussion in this chapter focuses on checking the utility of the items at the item level after finding out the status of the relationship/structure of the items/variables for teacher-level constructs with confirmatory factor analysis (CFA). As with the previous two chapters, this process is carried out through the statistical technique Rasch measurement model (RMM) to better understand the scales' items unidimensionality though the structure has been already established through CFA.

9.4 Rasch analysis

The utility of the above items is checked using the Rasch measurement model (RMM) with ConQuest 4.0 as in Chapters 7 and 8. The partial credit model (PCM) discussed in Chapter 3 is employed because partial success on the item responses are expressed by partial credits and are hierarchical (Eggert & Bögeholz, 2010; Huggins-Manley & Algina, 2015). As mentioned in Chapters 7 and 8, this procedure is useful for analysis of surveys such as these teacher-level scales, where the items are not marked for correct or incorrect answers (Bertoli-Barsotti, 2005; Huggins-Manley & Algina, 2015). The survey results from the analysis in Table 9.8 demonstrate that the data fit the model, which indicates that PCM is a more parsimonious model (Wang & Wu, 2011; Wu et al., 2016).

9.4.1 Item analysis with the Partial credit model

The 15 items for the *quality of teaching*, 13 items for *classroom environment* and 12 items for *the instructional resource* scales were subjected to item analysis using the PCM. This technique was employed to test the unidimensionality of the items measuring the three constructs. This involves examining each item's fit statistics using similar statistical criteria and procedures to those discussed in Chapters 7 and 8.

The results appearing on Table 9.6 (p. 222),Table 9.7 (p. 223) demonstrate that all items of the scales *quality of teaching* (QT),*classroom environment* (CE) and *instructional resources* (IR) are within the INFIT MNSQ criteria of 0.6 to 1.40 (Bond et al., 2007) and fitted the RMM. The separation reliability for these scales are above 0.9, indicating minimal measurement error with high discriminating power (Alagumalai et al., 2005; Hailaya, 2014). This indicates that the items in each scale have high measurement capacity (Wright & Stone, 1999). Additionally, the *t*-values of all the items in the scales are within the acceptable fit criteria of -2 to 2 (see Table 9.8), again, demonstrating that the items are discriminating between low and high ability respondents (Bond et al., 2007; Linacre, 2002). The items' estimates, thresholds, average expected values and biserial points are all in order (Wetzel & Carstensen, 2014; Wu et al., 2016). This indicates that, at the item-level analysis, all items of the QT, CE and IR constructs are within the requirements of the RMM (Wu & Adams, 2007; Wu et al., 2016).

Table 9. 6 Statistical fit results for quality of teaching scale (n=41)

Item #	Estimate	Error	INFIT MNSQ	CI	ZSTD (t)	Pt Bis	Delta (s)			
QT06	-0.01	0.13	0.65	(0.59, 1.41)	-1.9	0.78	-0.01	-0.01	0.13	0.15
QT07	-0.35	0.14	0.90	(0.47, 1.53)	-0.3	0.63	-1.84	-1.72	-0.44	1.09
QT08	0.25	0.12	1.03	(0.44, 1.56)	0.2	0.54	-0.86	-1.36	0.59	2.10
QT09	0.33	0.14	0.98	(0.64, 1.36)	0.0	0.41	0.33	0.33	-0.53	1.20
QT10	-0.23	0.14	0.68	(0.61, 1.39)	-1.7	0.75	-0.23	-0.23	-1.20	0.73
QT11	1.03	0.13	0.96	(0.66, 1.34)	-0.2	0.56	1.03	1.03	1.27	1.79
QT12	-1.13	0.15	0.95	(0.39, 1.61)	-0.1	0.52	-1.16	-1.13	-1.08	-0.58
QT13	-0.03	0.13	1.12	(0.53, 1.47)	0.6	0.36	-0.03	-0.38	-0.92	1.21
QT14	-0.16	0.14	1.00	(0.62, 1.38)	0.0	0.54	-1.29	-0.16	-0.16	0.96
QT15	-0.63	0.14	0.79	(0.54, 1.46)	-0.9	0.70	-0.63	-0.63	-1.23	0.06
QT16	-0.20	0.14	1.05	(0.58, 1.42)	0.3	0.28	-0.58	-0.20	-0.20	0.17
QT17	0.46	0.11	1.38	(0.60, 1.40)	1.7	0.31	-0.21	-0.30	0.99	1.37
QT18	0.13	0.12	1.39	(0.60, 1.40)	1.8	0.34	-1.36	-0.36	0.77	1.50
QT19	-0.38	0.13	0.79	(0.53, 1.47)	-0.9	0.80	-1.11	-1.01	-0.8	0.75
QT20	0.93*	0.51	1.39	(0.58, 1.42)	1.7	0.33	- 1.12	-0.88	1.61	2.34

Separation Reliability = 0.929

Chi-square test of parameter equality = 175.91, df = 14, Sig Level = 0.000

Table 9. 7 Statistical fit results for classroom environment and instructional resource scales (n=41)

Summary statistical fit for classroom environment scale

Item #	Estimate	Error	INFIT MNSQ	CI	ZSTD (t)	Pt Bis	Delta (s)		
CE21	0.12	0.20	0.91	(0.75, 1.25)	-0.7	0.54	0.12	0.12	0.15
CE22	-0.50	0.17	1.50	(0.42, 1.58)	1.6	0.12	-0.58	0.28	2.29
CE23	-0.99	0.18	0.88	(0.22, 1.78)	-0.2	0.51	-1.31	-0.76	0.91
CE24	1.18	0.20	0.87	(0.78, 1.22)	-1.2	0.52	0.19	1.99	1.99
CE25	0.24	0.16	1.12	(0.59, 1.41)	0.6	0.44	-1.03	-0.10	1.87
CE26	1.02	0.13	1.00	(0.62, 1.38)	0.0	0.67	1.03	1.51	1.56
CE27	-0.65	0.18	0.66	(0.42, 1.58)	-1.2	0.78	-6.65	-3.65	1.05
CE28	-0.23	0.16	1.18	(0.60, 1.40)	0.9	0.4	-1.32	-0.15	0.78
CE29	-0.67	0.17	1.07	(0.53, 1.47)	0.4	0.46	-1.48	-1.02	-0.54
CE30	-0.36	0.17	0.91	(0.58, 1.42)	-0.4	0.53	-1.92	-0.56	1.39
CE31	0.45	0.13	0.78	(0.63, 1.37)	-1.2	0.77	-0.02	0.40	0.97
CE32	-0.59	0.17	1.06	(0.52, 1.48)	0.3	0.47	-1.52	-1.20	0.94
CE33	0.95*	0.60	1.02	(0.63, 1.37)	0.2	0.33	-0.96	-0.21	2.12

Separation Reliability = 0.936

Chi-square test of parameter equality = 184.15, df = 12, Sig Level = 0.000

Summary statistical fit for instructional resources scale

Item #	Estimate	Error	INFIT MNSQ	CI	ZSTD (t)	Pt Bis	Delta (s)		
IR34	0.02	0.144	1.11	(0.62, 1.38)	0.6	0.5	-1.26	0.01	1.30
IR35	0.442	0.141	0.9	(0.63, 1.37)	-0.5	0.67	-0.61	0.44	1.50
IR36	0.786	0.134	1.37	(0.62, 1.38)	1.8	0.41	0.06	0.09	1.40
IR37	-0.111	0.138	0.8	(0.62, 1.38)	-1	0.73	-0.83	-0.16	0.66
IR38	0.581	0.127	0.84	(0.60, 1.40)	-0.8	0.76	-1.18	-0.78	0.21
IR39	-0.18	0.139	1.01	(0.62, 1.38)	0.1	0.65	-1.12	-0.08	0.66
IR40	-1.049	0.152	1.02	(0.59, 1.41)	0.2	0.56	-2.53	-0.95	0.33
IR41	-0.85	0.158	1.11	(0.54, 1.46)	0.5	0.54	-2.53	-1.42	1.21
IR42	0.005	0.167	0.93	(0.58, 1.42)	-0.3	0.5	-1.55	-0.55	1.56
IR43	0.179	0.14	1.17	(0.63, 1.37)	0.9	0.6	-0.86	0.20	1.20
IR44	0.655	0.143	0.97	(0.63, 1.37)	-0.1	0.65	-0.41	0.25	2.12
IR45	-0.479*	0.479	1.11	(0.59, 1.41)	0.6	0.54	-1.44	-0.76	0.76

Separation Reliability = 0.940

Chi-square test of parameter equality = 166.73, df = 11, Sig Level = 0.00

After checking the utility of the scale items, differential item functioning (DIF) was then employed to ascertain whether the items are biased towards a gender. This is determined through the procedure and criteria of item fit and item fit threshold approaches that are discussed in Chapters 7 and 8.

Table 9.8 DIF for quality of teaching for gender

Items No	Male			Female			d1-d2
	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ	
QT06	0.07	0.12	0.89	-0.07	0.12	0.85	0.14
QT07	0.03	0.11	0.52	-0.03	0.11	0.73	0.06
QT08	-0.23	0.1	0.57	0.23	0.1	0.67	-0.46
QT09	-0.02	0.12	0.52	0.02	0.12	1.06	-0.04
QT10	0.27	0.12	0.41	-0.27	0.12	0.76	0.50
QT11	-0.06	0.11	0.47	0.06	0.11	0.92	-0.12
QT12	0.70	0.13	0.73	-0.70	0.13	0.98	0.14
QT13	-0.10	0.11	0.7	0.10	0.11	0.87	0.00
QT14	-0.18	0.12	0.67	0.18	0.12	1.03	-0.36
QT15	0.25	0.12	0.52	-0.25	0.12	0.98	0.50
QT16	0.43	0.12	0.89	-0.43	0.12	1.05	0.45
QT17	-0.22	0.09	0.81	0.22	0.09	0.96	-0.43
QT18	-0.41	0.1	0.61	0.41	0.1	0.52	-0.31
QT19	0.11	0.11	0.46	-0.11	0.11	0.82	0.23
QT20	-0.22	0.42	0.76	0.22	0.42	1.02	-0.44

The 15 items shown in Table 9.8 are within the INFIT MNSQ values of 0.6 to 1.40 (Bond et al., 2007), though items QT12, QT16, QT18 and QT20 are below -0.5 logits or above 0.5 logits (Hungi, 2005). This indicates that these items are detected to have DIF for gender and were removed, for both Grade 10 and 12 mathematics teachers. This means that there is significant difference between male and female ability levels for those items.

Table 9.9 DIF for instructional resources and classroom environment scales

Items No	Male			Female			d1-d2
	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ	
IR34	-0.06	0.12	0.51	0.06	0.12	0.67	-0.12
IR35	0.01	0.12	0.61	-0.01	0.12	0.60	0.02
IR36	-0.19	0.12	0.69	0.19	0.12	1.31	-0.38
IR37	0.07	0.12	0.78	-0.07	0.12	0.48	0.14
IR38	0.12	0.12	1.37	-0.12	0.12	0.88	0.24
IR39	0.05	0.12	0.72	-0.05	0.12	0.74	0.10
IR40	-0.06	0.13	0.63	0.06	0.13	0.63	0.00
IR41	0.18	0.13	0.75	-0.18	0.13	0.37	0.36
IR42	-0.15	0.13	0.5	0.15	0.13	0.58	-0.29
IR43	0.04	0.12	0.42	-0.04	0.12	0.96	0.07
IR44	-0.02	0.12	0.48	0.02	0.12	0.66	-0.04
IR45	0.01	0.4	0.52	-0.01	0.4	0.93	0.02
Items No	Male			Female			d1-d2
	Estimate (d1)	Error (e1)	INFIT MNSQ	Estimate (d2)	Error (e2)	INFIT MNSQ	
CE21	-0.31	0.81	0.86	0.31	0.18	0.92	-0.62
CE23	0.01	0.17	1.04	-0.01	0.17	1.02	0.02
CE24	-0.29	0.18	0.81	0.29	0.18	0.9	-0.58
CE25	-0.13	0.13	0.81	0.13	0.13	0.64	-0.26
CE26	-0.36	0.13	0.66	0.36	0.13	0.97	-0.72
CE27	0.61	0.16	0.79	-0.61	0.16	0.38	0.22
CE28	0.03	0.14	0.91	-0.03	0.14	1.20	0.06
CE29	0.05	0.15	1.00	-0.05	0.15	0.63	0.09
CE30	-0.06	0.14	0.45	0.06	0.14	0.53	-0.13
CE31	-0.06	0.13	0.85	0.06	0.13	0.52	-0.12
CE32	0.14	0.15	0.57	-0.14	0.15	0.76	0.28
CE33	0.20	0.5	0.69	-0.20	0.5	0.47	0.40

Table 9.9 shows the DIF results of the *instructional resources* and *classroom environment* scales. On the former scale, there are no DIF items detected for gender within Grade 10 and 12 teacher respondents because the 12 items' INFIT MNSQ are within the acceptable range of 0.6 to 1.40, and the difference between the threshold (estimates) for gender are under 0.5 logits (Hungi, 2005). However, within the

classroom environment scale, DIF is detected in the difference between the estimates of gender, in items (CE21, CE24, CE26, CE27 and CE33). These items showed difference of estimates (threshold values) greater than 0.5 logits and less than -0.5 logits (Hungi, 2005). In other words, items CE21, CE24, CE26 are biased towards male respondents while items CE27 and CE33 are biased towards female respondents, and so were removed.

9.4.2 Person-item map

Figure 9.3 shows a person-item map of the three teacher level scales. The mean of the item difficulties is set at 0 as a starting point for the mapping process (Bond et al., 2007; Wu et al., 2016). As shown in Figure 9.3, *quality-teaching* scale (item QT06), *instructional resources* scale (item IR42) and *classroom environment* scale (item CE21) have the mean of the item difficulty estimates of 0 logits on the item-person map. The rest of the items of the three scales are spread above and below 0 to represent their difficulties relative to items QT06, IR42 and CE21 that are located at 0. The majority of the respondents of *quality teaching* and *classroom environment* scales have greater that 50% chance of endorsing items with difficulty level below their ability, and vice versa. The items for *instructional resources* scales are spread out evenly between -1 and 1 logits, and again most of the respondents' abilities are above the average (0 logits) indicating that most of the respondents endorsed majority of the items. For instance, item QT12 was the easiest item to be endorsed for *quality teaching* scale, and CE24 for *classroom environment* scale seemed to be the most difficult item to endorse (Stelmack et al., 2004). As shown n Figure 9.3, most of the respondents favorably responded to the items, however, few respondents (appearing at the bottom of the map) endorsed the lower items such as items CE23 and IR40, which is the easiest item to be endorsed for the scales (Mursidi & Soeharto, 2017).

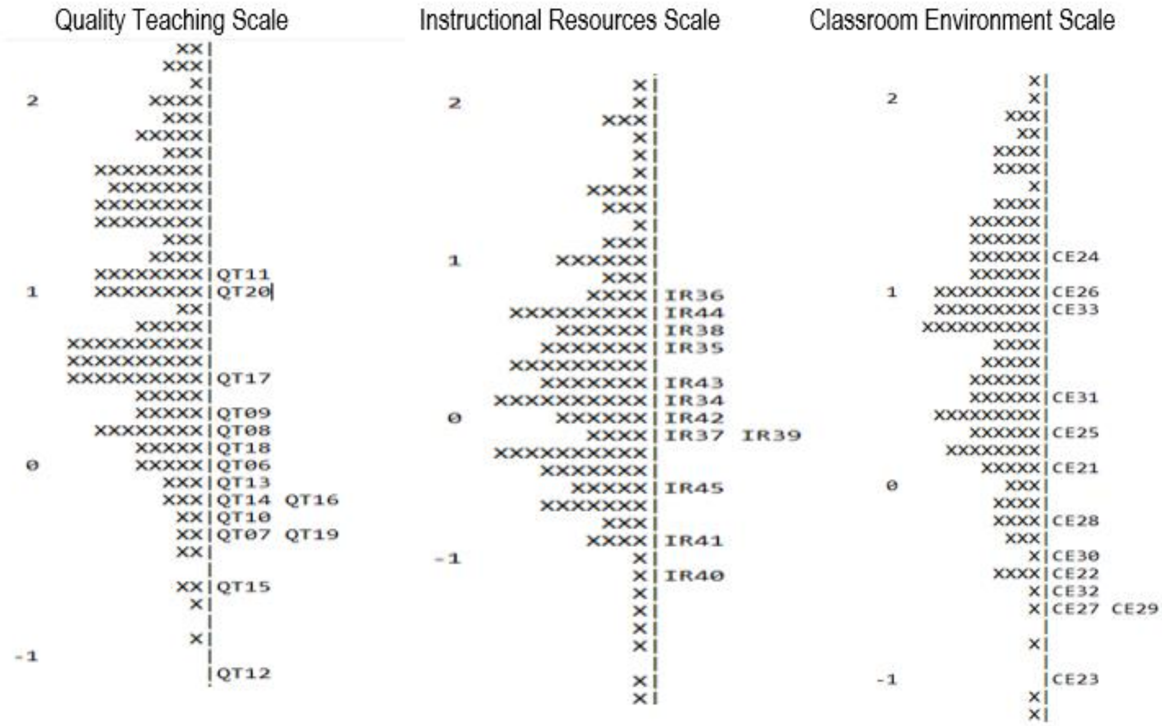


Figure 9. 2 Person-item map of teacher-level scales

9.5 Transformation to W-scores

As discussed in Chapter 3, the WLE values from the Rasch analysis for the scales and tests were transformed into W-scores to eliminate negative values, and the need for decimal points (Benson et al., 2016; Linacre, 2009; Woodcock, 1999). Microsoft Excel was used for this transformation, using the formula W=9.1024 x logits + c, where c is a constant term, selected to eliminate the negative values (Benson et al., 2016; Woodcock, 1999). The constant term used in this study is 500, consistent with PISA studies. The W-scores are used in the subsequent chapters' structural equation modelling and hierarchical linear modelling.

9.6 Comparison of the different RMM and CFA results

The results of factor analysis appearing in Table 9.4 (p. 217) shows the factor scores used to test whether items for each scale represented a single latent trait. However, through most items are found to represent the construct, factor scores of less than 0.3 for some scale items indicate that these items did not represent the construct. Further, the fit indices fit in Table 9.5 (p.218) indicates that the model did not fit the data.

Similarly, Tables 9.6 (p. 222), 9.7 (p. 223), 9.8 (p. 224) and 9.9 (p. 225) show the evaluation of the dimensionality of the items/constructs and identifying traits that influence the response pattern (Wright, 1996). Results shown on Tables 9.6 (p. 222) and 9.7 (p. 223) indicate that all the items of the *quality of teaching*, *classroom environment* and *instructional resources* scales fit the RMM, and are measuring each construct at the item level. Despite this, however, Tables 9.8 (p. 224) and 9.9 (p. 225) demonstrate that six items in the quality teaching and five items for the *classroom environment* scales are detected with DIF. Moreover, Tables 9.6 (p. 222) and Table 9.7 (p. 223) reflect the logit score with an interval scale, and Table 9.4 (p. 219) reflects the single-factor score with an ordinal scale of teacher-level factors. The former two tables show the direction in Rasch analysis with items ranging from hard to easy, while the latter table shows factor analysis, which is non-directional (Wright, 1996, p. 15).

9.7 Model employed for the study

The comparison of the confirmatory factor analysis (CFA) results from the single-factor models for the three teacher scales (*quality of teaching, classroom environment* and *instructional resources*) did not fit the data. However, from the Rasch item analysis perspective, the items for the single-factor model are viewed as forming a single scale, and the items' utility was checked with responses (Grimbeek & Nisbet, 2006), with the results of the items fitting the Rasch model. This means that the single-factor models for each of the teacher-level scales are represented by the items analysed through RMM and will be used for subsequent chapters (Baghaei & Amrahi, 2011; Wei et al., 2014).

9.8 Summary

The chapter focused on a structural-level analysis of the teacher-level scale items, with the items for each scale tested to ascertain the relationships between the variables/items and constructs. The single-factor structures were analysed using confirmatory factor analysis (CFA). After confirming the structure/relationship of the observed variables with each of the latent constructs, the utility of the teacher-level scales (*quality of teaching, classroom environment* and *instructional resources*) items were checked with the Rasch measurement model (RMM) at the micro/item level for the 41 Grade 10 and 12 mathematics teacher respondents. The results of the CFA clearly demonstrate that the single-factor model of each of the constructs using CFA did not fit the data for all the teacher-level scales. However, each of the items for each of the single level factors for the teacher-level constructs fitted the RMM model after removing the DIF items. As such, the WLE scores were converted to W-scores for the mathematics tests (refer to Chapters 5 and 6) and the student-level scales (refer to Chapters 7 and 8) together with these teacher-level scales (Chapter 9). These W-scores from the tests and scales will be employed in the subsequent chapters' structural equation modelling and hierarchical linear modelling. The next chapter examines and discusses structural equation modelling to ascertain possible relationships at the student-level and teacher-level variables/scales that might have an effect on the students' mathematics results.

Chapter 10: Structural Equation Modelling: Student and teacher-level models

10.1 Introduction

In this chapter, the relationships among the constructs/factors and the demographic variables (e.g., age, gender, educational level) are examined for the teacher and student-levels, respectively. The scores of the constructs and mathematics tests obtained the through validation and verification procedures in Chapters 5 to 9 are used to answer the general questions highlighted below.

1. What are the student and teacher-level factors that are affecting Grade 10 and 12 mathematics students' results in Port Moresby, PNG?
2. What are the relationships among student and teacher-level factors that are affecting Grade 10 and 12 mathematics students' results?

These general questions lead to three specific questions under the following headings:

Student-level

- What are the relationships among motivation, attitude, gender, parents' education level and parents' occupation that influence students' mathematics results?

Teacher-level

- What are the relationships among classroom environment, teaching quality, instructional resources, teachers' teaching experience and education level that influence students' mathematics results?
- What is the influence of school type (Catholic, public and private) on students' mathematics results?

The research questions in this study seek to examine the constructs/factors (*attitude, motivation, quality teaching, classroom environment and instructional resources*) and the demographic variables (*gender, parents' education level and employment*) that may have significant effects on students' mathematics results/achievements. These factors were divided into student, teacher and school levels. However, in the analysis the teacher and school-level data are combined into one level, as teacher-level data. This collapsing approach was taken to include all the 41 teachers' data rather than losing some data, because the data collected for some schools involved fewer than three teachers per school. At the student level, three models, namely the Combined, Grade 10, and Grade 12 models were analysed separately from each other. The rationale for analysing the Combined model was to examine the patterns that were common to both Grade 10 and 12 students. Additionally, the variable *Grade* was used in this model to control for the differences between Grade 10 and 12 students. Furthermore, the reason for using separate models for Grades 10 and 12, respectively, was to also examine in more detail the different aspects emerging from these two groups. The teacher model was further analysed separately from the three student models. Combined, the four models (three at student-level and one at teacher-level) were employed to investigate and evaluate the relationships among the factors through structural equation model (see Chapter 3 for details).

This chapter discusses the process employed in student and teacher-level analysis in order to answer all research questions (RQs) except RQ 3 ("What are the attributes that are deemed important by teachers regarding their mathematics teaching approaches and content knowledge?"), which is answered through qualitative responses (interviews with teachers) in Chapter 12. The different factors highlighted in the theoretical framework introduced in Chapter 2 are examined in this chapter through structural equation model analysis, to establish the possible relationships between factors that affect students' mathematics results. There are also casual links that determines some probability of causality. This chapter begins with a general description of the Mplus 7.0 software, followed by a brief discussion of Structural Equation Modelling (SEM), which has already been described in detail in Chapter 3. After that, this chapter describes the concepts and steps involved in building the model and testing the statistical assumptions proposed by this study. The chapter continues with a presentation of the results of analysis and discussion, using SEM to answer this study's research questions. The chapter ends by highlighting the main points of the results and discussions and outlining the techniques that will be employed in subsequent chapters.

10.2 Structural equation modelling

As discussed in Chapter 3, SEM is "a comprehensive statistical approach for testing hypotheses about relations among observed variables and constructs" (Holye, 1995, p.1 cited in Hailaya, 2014). SEM is known by other terms such as covariance structure analysis, covariance modelling, or analysis of covariance structures (Kline, 2015; Martynova et al., 2017; Ramlall, 2016). In this study, the empirical data are used in SEM to quantify and validate the theoretical model proposed in Chapter 2. This permits the researcher to examine and determine whether a hypothesised relationship exists (or does not exist) among the constructs and demographic variables at the student and teacher levels (Hooper et al., 2008; Ramlall, 2016). In this study, if the hypothesised relationships at these two levels are supported by the data, then the proposed theoretical model is accepted, and more complex models can be tested further. However, if the hypothesised relationships are not reflected in the data, the model can be modified and retested, following Ramlall's (2016) recommendation. Alternatively, the model can be rejected, and a new theoretical model developed and evaluated.

As discussed in Chapter 3, SEM is classified into two models, namely the measurement model and the structural model. The measurement model is concerned with the relationship between the construct and observed variables. For instance, this relates to the confirmatory factor analysis (CFA) that was discussed in detail in Chapter 3. On the other hand, the structural model deals with the relationship among the constructs confirmed through CFA. Ramlall (2016) states that the relationship or association may have direct or indirect effects among the constructs and demographic variables. As such, SEM examines the structure or relationship between the constructs (Kline, 2015; Schumacker & Lomax, 2004) in this study.

10.3 Mplus version 7.0

This study used the Mplus version 7.0, a statistical modelling program that provides researchers with a flexible tool to analyse data (Muthén & Muthén, 1998). Mplus estimates regression, path analysis, exploratory and confirmatory factor analysis (EFA and CFA), structural equation modelling (SEM) and discrete-and continuous analysis models (Byrne, 2013; Muthén & Muthén, 1998; Wang & Wang, 2012). Mplus offers researchers a wide choice of models, estimators, and algorithms in a program that has an easy-to-use interface (Muthén & Muthén, 1998; Wang & Wang, 2012) as well as graphical displays of data and analysis results (Muthén & Muthén, 1998). Furthermore, Mplus allows the analysis of single-level, multilevel data, data that comes from different populations with either observed or unobserved heterogeneity (Muthén & Muthén, 1998), and data that contain missing values (Byrne, 2013; Muthén & Muthén, 1998). These advantages of Mplus, particularly in SEM analysis, have prompted the researcher to use the software in this study.

Mplus version 7.0 is used by exporting data from a Microsoft Excel file to SPSS and then converting the SPSS file to a dat file to be used for analysis. This conversion to a dat. file allows the researcher to form and specify the path analysis by developing a syntax that can be accepted by Mplus. After that, the syntax is used in the software to run analysis of the data. The software then generates the unstandardised and standardised estimates with other statistical results for the specified model.

10.4 Models and presentation in quantitative research

In quantitative research, a model is used to represent the phenomena under investigation. The model may be illustrated in pictorial/graphical form, or in a series of structural equations to represent the relationship under study. The theory under study is made clearer through these representations. According to Ramlall (2016) and Kline (2015), models in graphical form use the following shapes to represent the symbols most often used in SEM diagrams. They are:

a) Ellipse symbols representing latent variables

b) Rectangle symbols representing observed variables

c) Straight single-headed arrows representing the impact of one variable on another variable (relationship) and,

d) Doubled-headed arrows representing covariance or correlation between pairs and variables

Specification, estimation, and evaluation are three general steps in building a model in any quantitative study (Ben, 2010; Byrne, 1998; Martynova et al., 2017). Specification refers to the careful examination of a given phenomenon of interest, and estimation permits the researcher to use equations or graphical forms to test a hypothesised model (Diamantopoulos et al., 2000); for instance, building a path diagram of relationships to demonstrate a theoretical model (Cangur & Ercan, 2015; Diamantopoulos et al., 2000; Ramlall, 2016). Model evaluation is employed to ascertain whether the model is consistent with the data and, consequently, if it can be concluded that the model fits the data well (Ben, 2010; Diamantopoulos et al., 2000). Consequently, this means that path analysis can be defined as a form of structural equation model, which require prior specification of inter-variable relations before being employed in the data analysis (Cangur & Ercan, 2015; Martynova et al., 2017).

10.4.1 Model trimming criteria

The *p-value* is the goodness of fit for the proposed model for students and teachers in this study, and was determined to examine the significance level of the variables. "The *p-value* is the probability of obtaining an effect equal to or more extreme than the one observed considering the null hypothesis is true" (Biau, Jolles, & Porcher, 2010, p. 886). This effect can be a difference in a measurement between two groups or any measure of association between two variables. 'P' stands for probability, and the *p*-value measures how likely it is that any observed difference between groups is due to chance (Biau et al., 2010; Dahiru, 2008).This study considers *p*-values of 0.05 to be significant, in line with international studies such as PISA and TIMSS on which this study is patterned. In other words, variables that have *p-values* less than 0.05 are considered to have a statistically significant effect with a 5% chance or probability of the event having occurred by chance alone; while variables with *p-values* greater than 0.05, are considered to show a non-significant effect on other variables, and therefore demonstrate no evidence of an effect from the data (Butler & Jones, 2018; Ramlall, 2016). At the structural level of the models, variables with $p>0.05$ were considered insignificant paths and were removed from the models. Moreover, at the measurement level, variables with factor loading more than 0.3 (discussed in Chapter 6) were accepted (Kline, 2015; Ramlall, 2016). These are the criteria employed for the teachers and three student-level models. The next section describes the variables and the hypothesised model as used in the student-level analysis.

10.5 Constructs and variables used in the student-level analysis

The student-level model indicates that there are three constructs and 14 demographic variables involved in the hypothesised model. These constructs and variables may influence the Grade 10 and 12 students' mathematics achievement.

The Weighted Likelihood Estimate (WLE) obtained from the constructs in Chapters 5 to 9 through Rasch analysis were converted to W-scores, as discussed in Chapter 3 and are used in the analysis instead of the original item responses scores, to simplify the model. This is due to the W-score's advantages over raw scores as (discussed in Chapter 3). At the student level, *attitude* (ATTD) and *motivation* (MTVN) scales were analysed for their impact on Grade 10 and 12 students' mathematics results. The *attitude* (ATTD) scale is divided into three subscales such as *like, value* and *self-belief,* as shown in Table 10.1. These subscales of attitude are consistent with the literature, as well as the TIMSS and PISA studies. TIMSS reports highlight that, when students like mathematics, they value the importance of it as a subject, and start having self-belief in their own mathematics ability (Martin & Mullis, 2013; Mullis et al., 2012). Similarly, the *motivation* (MTVN) scale is also divided into three subscales: *achievement goal, interest* and *engagement*, in accordance with other studies and PISA studies. PISA reports state that when students have high achievement goals in mathematics, they are interested and engaged in mathematics activities (PISA, 2016; Ray & Margaret, 2003). This step of dividing the scales was used to operationalise the two scales in the context of this study, as discussed in Chapters 7 and 8, and also to order the subscales for better classification. Other manifest variables such as students' grades, gender, and their parents' employment and education levels are included in the model as background variables that might have impact on the scales (attitude, motivation and mathematics achievement).

At the student-level, a dummy variable is used to distinguish the various treatment groups (Hayes, Montoya, & Rockwood, 2017; Henseler, Ringle, & Sarstedt, 2015). For instance, 0 and 1 dummy variables are used, where a variable is given a 0 if it is in the control group or a 1 if it is in the treated group. The three dummy (control) variables are public, Catholic, and private schools. This approach was applied to control for the effect of the significant differences in the variables and scales in the model. The public schools are confirmed as baseline variable. The latent and manifest variables used in the student-level model are shown in Table 10.1. From this Chapter onwards, the terms *factor, scale, construct and latent variable* are used interchangeably.

Table 10. 1 The variables used in the student-level model

Latent Variables/Scales	Subscales	Manifest variables	Description	Coding
		SCHOOL TYPE 1	Public	1=public, 0=non-public
		SCHOOL TYPE 2	Catholic	1=Catholic, 0=non-Catholic
		SCHOOL TYPE 3	Private	1=Private,0=non-private
		GRADE	Students year level	0=Grade 10 1=Grade 12
		GENDER	Students' gender	0=Female 1=Male
		FATHER/MOTHER EMPLOYMENT	Parents' employment	0=Informal 1=Formal
		FATHER/MOTHER EDU LEVEL	Parents' education level	0=Certificate 1=Diploma 2=Bachelors 3=Postgraduate
	Like	Attd27-Attd32	Like mathematics	
ATTITUDE	Value	Attd33-Attd39	Value mathematics	WLE/W-scores
				WLE/W-scores
	Self-belief	Attd40-Attd46	Self-belief in mathematics	WLE/W-scores
	Achievement Goal	Mtvn09-Mtvn13	Achieve in mathematics	WLE/W-scores
MOTIVATION	Engagement	Mtvn14-Mtvn19	Engage in mathematics	WLE/W-scores
	Interest	Mtvn20-Mtvn26	Interest in mathematics	WLE/W-scores
MATHS RESULTS		Q1-Q40	Test Achievement	WLE/W-scores

10.5.1 The hypothesised student-level model

Student factors were analysed for three groups of student participants: Grades 10 and 12 were combined to constitute one group, Grade 10 students only composed another group, and Grade 12 students only made up the third group. One of the reasons for analysing these three models was (as mentioned above) to examine both the common patterns for the combined model as well as the different aspects emerging from the separate models. This allows the researcher to obtain both an overall and specific picture of the directional relations among the factors at the combined student level and the two individual grade levels. The three models involved in the analysis of the relationships among student-level constructs and the demographic variables are mentioned in Table 10.1, above. In total, this analysis involved 354 and 375 participants from Grade 10 and 12, respectively, as well as an analysis of the 729 combined student respondents.

The hypothesised student-level model shown in Figure 10.1 was derived from the theoretical framework discussed in Chapter 2. The theoretical framework identified attitude (ATTD) and motivation (MTVN) as the two main factors. Attitude towards learning mathematics is defined in this theoretical framework in terms of students' liking, valuing of and self-belief in mathematics, as discussed in Chapter 7. Similarly, motivation towards learning mathematics is associated with students' achievement goals, interest and engagement, as discussed in Chapter 8. The student demographic variables such as gender, grade, parents' employment and education level, are also included in the model as observed variables as they affect achievements in mathematics at the student level. These demographic variables are hypothesised as exogenous (independent) variables as they are not influenced by other variables and scales in the student-level model. On the other hand, the scales such as attitude and motivation at the individual level are viewed as endogenous (dependent) variables, as they may interact with one another at the student level and are product variable (endogenous) due to the mediating effect or influence they have in the model. The proposed student-level theoretical framework in Figure 10.1 shows that the demographic variables mentioned earlier have an impact on the scales motivation and attitude, as well as on mathematics achievement both directly or indirectly through other mediating variables.

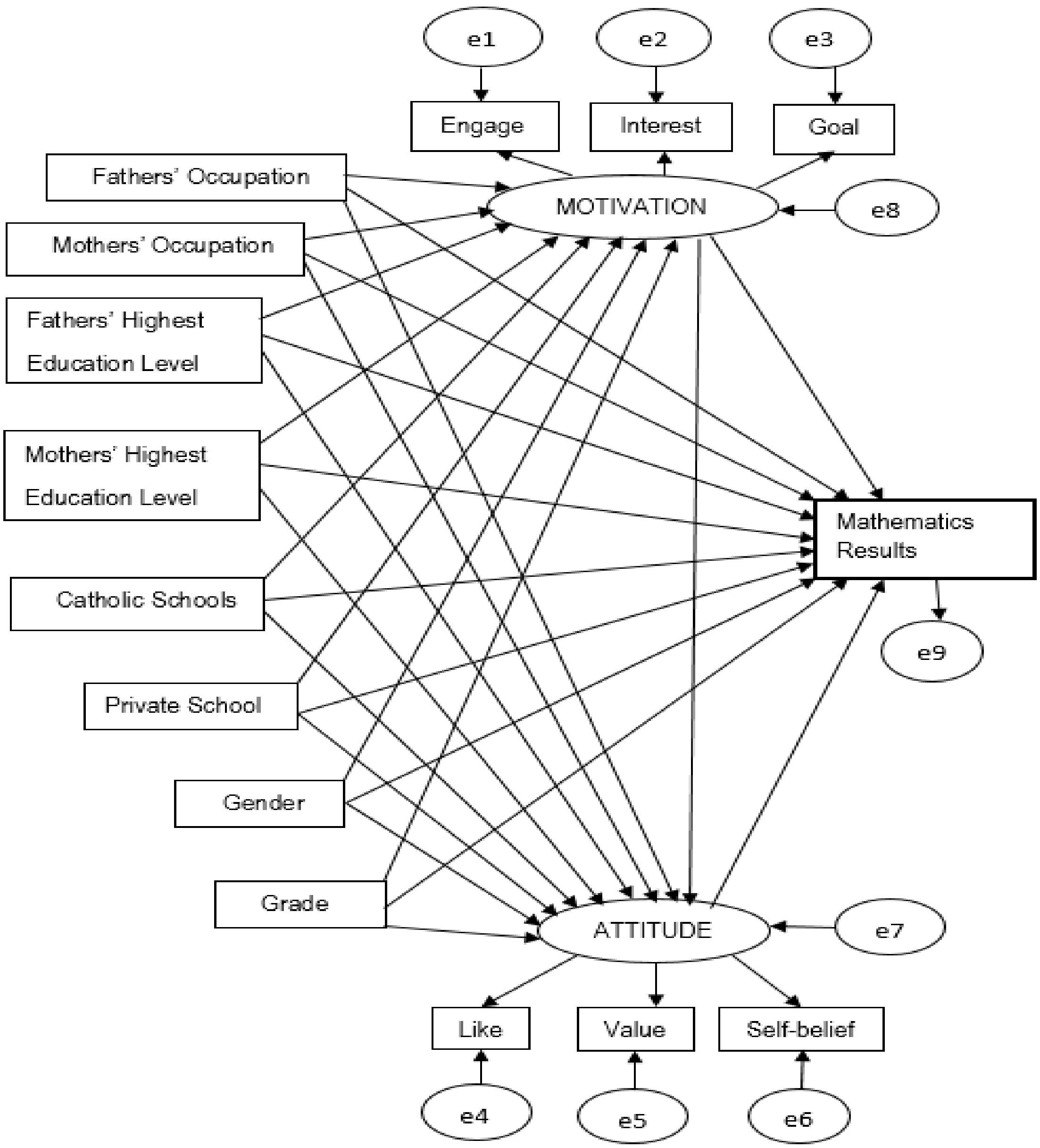


Figure 10. 1 The hypothesised student-level model

10.5.2 Student-level models results

The relationships between student-level factors discussed earlier are examined through the student-level model shown above in Figure 10.1. Prior to examination of the relationships among the variables this study employed multicollinearity and multivariate normality tests. The multicollinearity test was conducted to examine if the correlation is high between the measured pair of variables. It was revealed in the path analysis that the constructs ‘motivation’ and ‘attitude’ have a high correlation of 0.89. However, the multivariate distribution was normal and did not violate the accuracy of statistical test of the model (Khine, Fraser, Afari, Oo, & Kyaw, 2018). Khine et al. (2018) stress that this kind of result may bias and create problems in SEM analysis, as one of the variables/scales could be excluded. The scales with the correlated hierarchical structure are represented by one latent variable for attitude and motivation with three dimensions, which have three W-score indicators. As such, the student-level models are evaluated by the measurement model first, followed by the structural model; beginning with the results of the combined student model, and followed by the Grade 10 and 12 models, respectively.

10.6 Measurement model results for the combined student-level scales

Mplus 7.0 was used to assess the relationship between the subscales and scales. Assessment and evaluation based on several statistical indices such as unstandardised estimates (b), standardised estimates (β), standard errors (se) and the critical ratio (C.R) (the ratio associated with the probability of a sample; i.e. deviation from the mean to the standard deviation). The unstandardised estimates indicate the strength of the relationship between the observed variables and the scales (latent variables), whereas standardised estimates compare the extent of the relationship among the variables (Kline, 2015; Martynova et al., 2017). Further, the standard errors indicate the variability of the estimates. The unstandardised estimates are divided by the standard errors to obtain the C.R. In this section, factor loadings of 0.3 are an acceptable fit for the subscales and scales, as discussed in Chapters 7, 8 and 9. Table 10.2 shows the results of the measurement model of the student-level model with the two scales (attitude and motivation).

Table 10. 2 Measurement model results for combined student level scales

	B	β	S.E.	Est./S.E.	*p*
ATTD BY					
Value	1.00	0.64	0.03	21.33	0.00
Like	1.54	0.82	0.01	48.92	0.00
Self-Belief	1.41	0.73	0.02	34.89	0.00
MTVN BY					
Engagement	1.00	0.66	0.02	27.27	0.00
Interest	1.19	0.77	0.02	40.39	0.00
A/Goal	0.59	0.50	0.03	16.39	0.00

The measurement model of the *attitude* (ATTD) scale is indicated by three subscales (*value, like,* and *self-belief*). These three subscales and their loadings are values (β=0.64), *like* (β=0.82), and *self-belief* (β=0 .73). Similarly, *motivation* (MTVN) as a scale in the measurement model is also reflected by three subscales, that is, *engagement, interest,* and *achievement goal*. These variables have loadings of β=0.66, β=0.77 and β=0.50, respectively. The variables show a significant relationship with the two scales (attitude and motivation), at the same time also showing strong indicators of the scales.

10.6.1 Combined student-level model fit indices

The results of the fit indices of the final student-level model are shown in Table 10.3. The fit indices were examined using the results of ratio of chi-square to its degree of freedom (x^2/df), root mean square error of approximation (RMSEA), standardised root mean square residual (SRMR), comparative fit index (CFI), and the Tucker-Lewis Index (TFI), with their criteria discussed in Chapter 3. As shown in Table 10.3, the x^2/df, RMSEA and SRMR values are 2.8, 0.05, and 0.03, respectively. These indices indicate a good model fit for the student-level model. Moreover, the CFI (0.96) and TFI (0.94), provide acceptable values above

0.90 (i.e., closer to 1), denoting that the final combined student-level model is a better fitting model. The next section discusses the results of the analysis of the combined student-level structural model.

Table 10. 3 The results of the fit indices of the final combined student-level model

Fit Index	Obtained values	Remarks
X^2/df	99.33/35=2.8	Acceptable fit
RMSEA	0.05	Acceptable fit
SRMR	0.03	Acceptable fit
CFI	0.96	Acceptable fit
TLI	0.94	Acceptable fit

10.6.2 Combined Grade 10 and 12 student-level structural model results

The final combined Grade 10 and 12 student-level path diagrams, after deleting the non-significant variables, are shown in Figure 10.2 and Table 10.4 (p.242). The deleted demographic variables from the hypothesised combined student-level model are *gender, Catholic school, father's higher education level, and mothers' and fathers' occupations*, as these variables were not within the set statistical criteria.

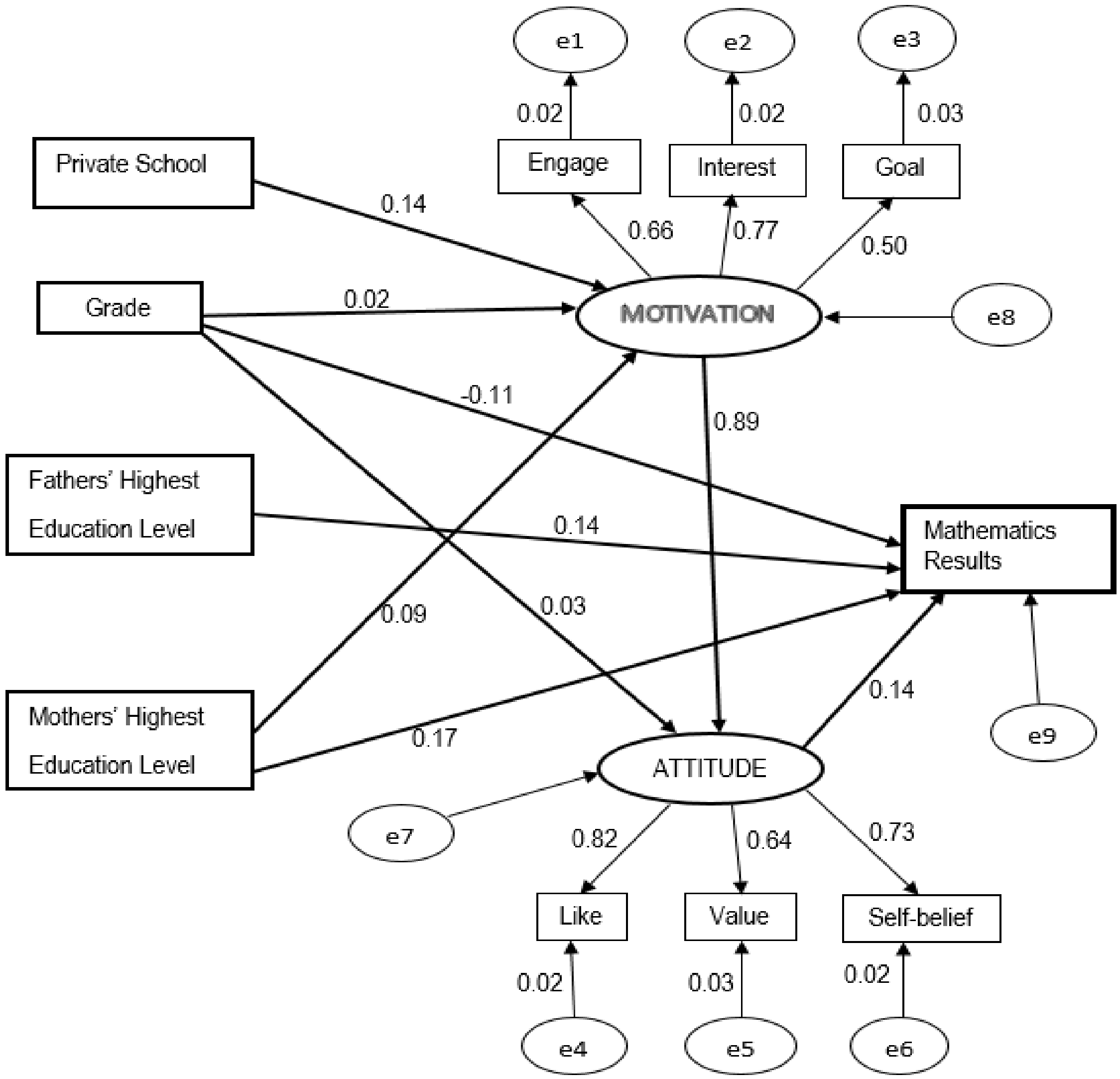


Figure 10. 2 The final combined student-level model (n=729)

Table 10.4 presents the outcomes with their predictors with indices such as standardised coefficients (β) for indirect, direct and total effect with significance level $p<0.05$. The standardised coefficients (β) are used to compare the results among the scales or variables, as the units for each variable or scale is similar. This section will now present and discuss in detail the direct and indirect results of the standardised coefficients (β) shown in Table 10.3.

Table 10. 4 Grade 10 and 12 students' estimated path models of the outcome of mathematics results

Dependent Variable	Independent	Combined Grade 10 and 12 students Direct (β)	Indirect (β)	Total (β)	*p*
Attitude					
	Motivation	0.89	-	0.89	0.00
	Grade	0.03	-	0.03	0.39
Motivation					
	Private schools	0.15	-	0.15	0.00
	Grade	0.02	-	0.02	0.58
	MHEDL	0.09	-	0.09	0.03
Maths Results					
	Motivation	-	0.14	0.14	0.00
	Attitude	0.14	-	0.14	0.00
	Grade	-0.11	0.007	-0.10	0.00
	FHEDL	0.14	-	0.14	0.00
	MHEDL	0.17	0.01	0.18	0.00
	Private schools	-	0.02	0.02	0.00

As shown in Figure 10.2 and Table 10.4, the *Grade* variable in this combined model is included to control any differences in the motivation, attitude and performance of the Grade 10 and 12 students. Therefore, it is kept in the model even though it is showing a non-significant result. Moreover, as shown in Figure 10.2 and Table 10.4, one variable has a direct effect or influence on the ATTD (attitude) of students towards mathematics: *motivation* (β=0.89). The scale *motivation* has a positive path coefficient and a statistically significant ($p<0.05$) influence on students' attitude towards mathematics. This implies that the higher the motivation of students is, the more positive their attitude towards mathematics. Similar to the findings for ATTD, private schools (β=0.15) and MHEDL (β=0.09) have direct effects students' motivation (MTVN) towards mathematics. *Private schools* and *mothers' highest education level* (MHEDL) have a positive path coefficient and significantly ($p<0.05$) effect on students' motivation towards mathematics, respectively. This suggests that students at *private schools* are more likely to be more motivated in mathematics compared to students at public and Catholic schools. Similarly, students whose mothers have a higher education levels are also likely to have higher levels of motivation in mathematics. This may be because more educated mothers are better equipped to assist students with their mathematics exercises at home.

For the mathematics results, after controlling for the differences between Grades 10 and 12, three variables have significant ($p<0.05$) direct effects on the students' mathematics results. These variables include ATTD ($\beta=0.14$), FHEDL ($\beta= 0.14$) and MHEDL ($\beta= 0.17$). The positive path coefficient of attitude (ATTD) on mathematics results indicates that the students with more positive attitudes are more likely to perform better in mathematics. Another important effect is the effect of *fathers' highest education level* (FHEDL) on mathematics results. This direct effect suggests that students whose fathers have higher education levels may perform better in mathematics. Similarly, the effect of mother's highest education level (MHEDL) on students' mathematics results has a positive path coefficient. This also suggests that students whose mothers have higher education levels may also do better in mathematics. However, the effect of mother's highest education level is slightly stronger than the fathers'. Overall, these results show that students' mathematics performance depends on their parents' highest educational level.

Additionally, *motivation* ($\beta=0.14$) has an indirect effect on mathematics results mediated by the attitude scale. This result indicates that motivation significantly ($p<0.05$) affects students' mathematics results through attitude. This signifies that Grade 10 and 12 students who are highly motivated with positive attitudes to learn mathematics obtain better mathematics results. Similarly, *mothers' highest education level* (MHEDL) can be observed from Figure 10.2 (p.243) as also having a significant indirect effect on students' mathematics results, mediated by motivation and attitude. As such, MHEDL ($\beta=0.01$) mediated by MTVN to ATTD shows a positive coefficient on mathematics results. This suggests that higher education levels of mothers drives students to have higher motivation, and consequently, promotes positive attitudes towards learning mathematics that could possibly have impact on their achievements in the subject.

Moreover, it can be gleaned from Figure 10.2 and Table 10.4 that *private schools* (β=0.02) appears to have a significant indirect positive impact on students' mathematics achievement. This finding is mediated from MTVN to ATTD to private schools to have an influence on students' mathematics results. This result shows that students at private schools have higher levels of motivation, which leads to more positive attitudes and ultimately, to higher performance level.

10.7 Grade 10 student level measurement model results

Similar statistical indices used in the previous analysis of combined student-levels are employed to evaluate the Grade 10 student-level model. Table 10.5 shows the results of the measurement model of the Grade 10 student with the two scales (attitude and motivation).

Table 10. 5 Measurement model results for Grade 10 student-level scales

	b	B	S.E.	Est./S.E.	*p*
ATTD BY					
Value	1.00	0.62	0.04	16.38	0.00
Like	1.54	0.81	0.02	33.30	0.00
Self-Belief	1.45	0.74	0.03	25.20	0.00
MTVN BY					
Engagement	1.00	0.64	0.04	17.46	0.00
Interest	1.23	0.79	0.03	30.63	0.00
Goal	0.62	0.51	0.04	11.48	0.00

The measurement model of the *attitude* (ATTD) scale is indicated by *value* (β=0.62), *like* (β=0.81) and *self-belief* (β=0.74). Similarly, the *motivation* (MTVN) scale is demonstrated by, *engagement* (β=0.64*), interest* (β =0.79) and *achievement goal* (β=0.51). These subscales of ATTD and MTVN indicate a significant relationship and better indictors of the scales, with factor loading results above the criteria of 0.3.

10.7.1 Grade 10 student level model fit indices

The results of the fit indices of the Grade 10 student-level model are shown in Table 10.6, with similar fit indices to the combined grades model discussed above. The results shown in Table 10.6 of indices X^2/df (2.47), RMSEA (0.06) and SRMR (0.06) indicate a good model fit for the Grade 10 student-level model, as these indices value are close to 0. Moreover, the CFI (0.94) and TFI (0.93), provide acceptable values above 0.90 (i.e., closer to 1), indicating that the Grade 10 student-level model is a better fitting model. The next section discusses the results of the analysis of the Grade 10 student-level structural model.

Table 10. 6 Grade 10 student-level fit indices

Fit Index	Obtained values	Remarks
X^2/df	76.59 /31=2.47	Acceptable fit
RMSEA	0.06	Acceptable fit
SRMR	0.06	Acceptable fit
CFI	0.94	Acceptable fit
TLI	0.93	Acceptable fit

10.7.2 Grade 10 student level structural model results

The final Grade 10 student-level path diagram with its variables is shown in Figure 10.3.

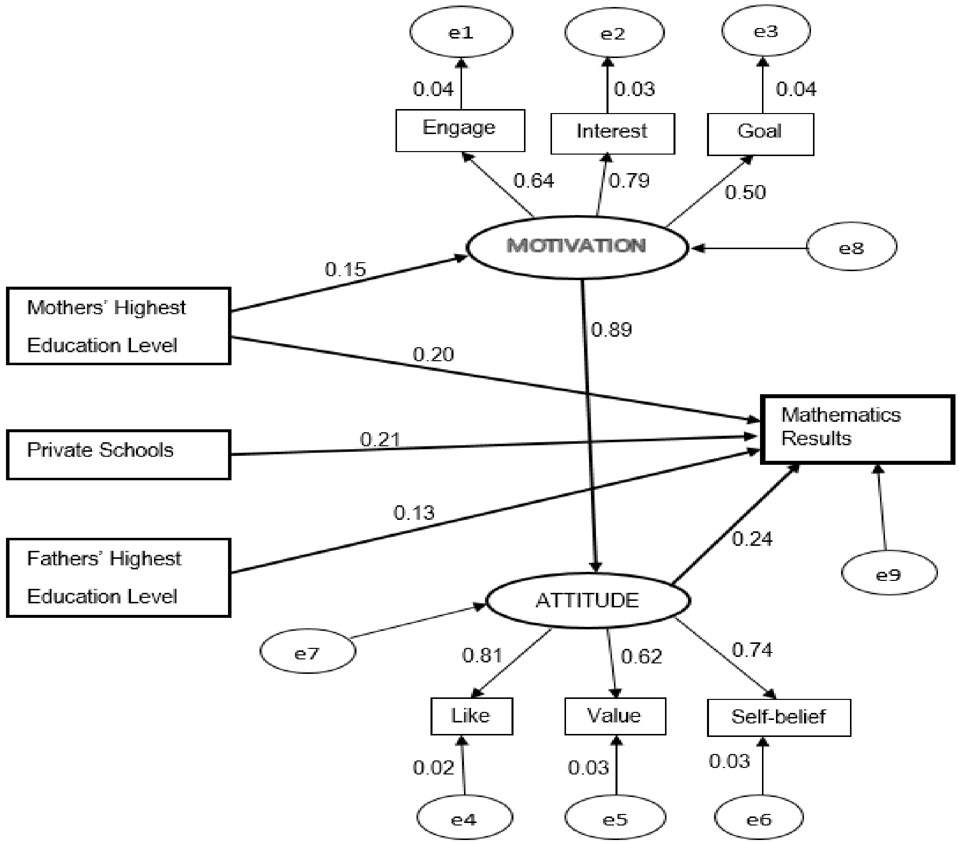


Figure 10. 3 The final Grade 10 student level model (n=354)

Table 10. 7 Estimated path models of the outcome of mathematics results for Grade 10 students

Dependent	**Variable**		**Grade 10 students**		
	Independent	**Direct (β)**	**Indirect (β)**	**Total (β)**	***p***
Attitude					
	Motivation	0.89	-	0.89	0.00
Motivation					
	MHEDL	0.15	-	0.15	0.00
Maths Results					
	Motivation	-	0.24	0.24	0.00
	Attitude	0.24	-	0.24	0.02
	FHEDL	0.13	-	0.13	0.00
	MHEDL	0.20	0.04	0.24	0.02
	Private schools	0.21	-	0.21	0.00

As shown in Figure 10.3 and Table 10.7, *motivation* (β=0.89) has a significant ($p<0.05$) direct effect on Grade 10 students' attitude towards mathematics. This positive coefficient suggests that the higher students' motivation, the more positive their attitude towards mathematics. Similarly, the *MHEDL* (β=0.15) has a significant ($p<0.05$) direct effect on the motivation of students towards mathematics. The positive path coefficient of MHEDL on MTVN shows that, the higher the students' mothers' highest education level, the higher their motivation level.

Furthermore, as demonstrated in Table 10.7, *ATTD, FHEDL, MHEDL* and *private schools* are found to have significant ($p<0.05$) effects on students' mathematics achievement. *ATTD* (β=0.24), has a positive coefficient with a significant influence on the Grade 10 students' mathematics results. This shows that students are likely to obtain better mathematics results when they possess more positive attitudes towards mathematics. On the same note, for *FHEDL* (β= 0.13) and *MHED* (β= 0.20), there are also significant ($p<0.05$) effects on mathematics results. These results show that the effect of mothers' highest education level on students' mathematics results is stronger that of fathers' highest education level. Despite that, both parents' highest education levels play a significant role in influencing students' mathematics results. This may be because when both parents are more educated, they are more likely to be able to teach their children at home. Furthermore, *private schools* (β= 0.21) have a significant ($p<0.05$) effect on the students' results with a positive path coefficient. This indicates that Grade 10 students at private schools obtain better mathematics results compared to the public and Catholic schools in Port Moresby.

Besides the direct effect the variables have on mathematics results, it is also clearly shown in Figure 10.3 and Table 10.7 that the *motivation* (β=0.24) has an indirect significant ($p<0.05$) effect on the Grade 10 mathematics results mediated by the attitude scale. This suggests that Grade 10 students with higher levels of motivation tend to have more positive attitudes, which in turn, leads to higher level of performance in mathematics. Moreover, *MHEDL* has an indirect effect on students' mathematics achievement. MHEDL (β=0.036) has a positive coefficient, indicating a significant ($p<0.05$) effect on mathematics achievement mediated through motivation and attitude towards mathematics. This positive relationship suggests that a higher mothers' education level leads students to have higher motivation and more positive attitudes that enable them to obtain better results.

10.8 Grade 12 student-level measurement model results

Similar statistical indices and criteria employed in the previous two student-level models are utilised in this section to assess and evaluate the Grade 12 student-level model. Table 10.8 shows the results of the measurement model of the Grade 12 student-level model with the two scales (attitude and motivation).

Table 10. 8 Measurement model results for Grade 12 student level scales

	b	β	S.E.	Est./S.E.	*p*
ATTD BY					
Value	1.00	0.66	0.03	19.53	0.00
Like	1.54	0.83	0.02	34.99	0.00
Self-Belief	1.38	0.72	0.03	24.32	0.00
MTVN BY					
Engagement	1.00	0.69	0.03	21.28	0.00
Interest	1.16	0.75	0.03	27.06	0.00
Goal	0.57	0.50	0.04	11.71	0.00

As mentioned earlier, the measurement model of the ATTD scale is indicated by v*alue* (β=0.66), like (β=0.83) and *self-belief* (β=0.72). Similarly, the MTVN scale is demonstrated by engagement (β=0.69), interest (β =0.75) and achievement goal (β=0.50). The three subscales of ATTD and MTVN show a significant relationship with each construct, with factor loading results above the criteria of 0.3, showing strong indicators.

10.8.1 Grade 12 student level model fit indices

The results of the fit indices of the Grade 12 student-level model are shown in Table 10.9 with similar fit indices found to those of the Grade 10 students' model mentioned earlier. The results show that the RMSEA (0.05) and SRMR (0.03) indicate values closer to zero. Moreover, CFI (0.96) and TFI (0.94) provide acceptable values above 0.90 (i.e., closer to 1). The X^2/df (2.1) also shows an acceptable fit. All these fit indices results suggest that the Grade 12 student level model demonstrated in Figure 10.4 is a better fitting model.

Table 10. 9 Grade 12 student level fit indices

Fit Index	Obtained values	Remarks
X^2/df	42.87/26=1.65	Acceptable fit
RMSEA	0.04	Acceptable fit
SRMR	0.03	Acceptable fit
CFI	0.98	Acceptable fit
TLI	0.97	Acceptable fit

The results of the combined and Grade 10 student models confirmed the proposed theoretical model introduced in Chapter 2, though the Grade 12 model was found to have an insignificant effect on the attitude scale influencing mathematics results. However, the model was made to fit better by trimming (deleting variables that were insignificant). So far, this chapter has discussed the direct and indirect path coefficients of each of the student model variables towards mathematics achievement, in order to define the relationships between the two constructs (ATTD and MVTN) and other demographic/background variables.

10.8.2 Grade 12 student-level model results of the structural model

The final Grade 12 student-level path diagram with its variables is shown in Figure 10.4

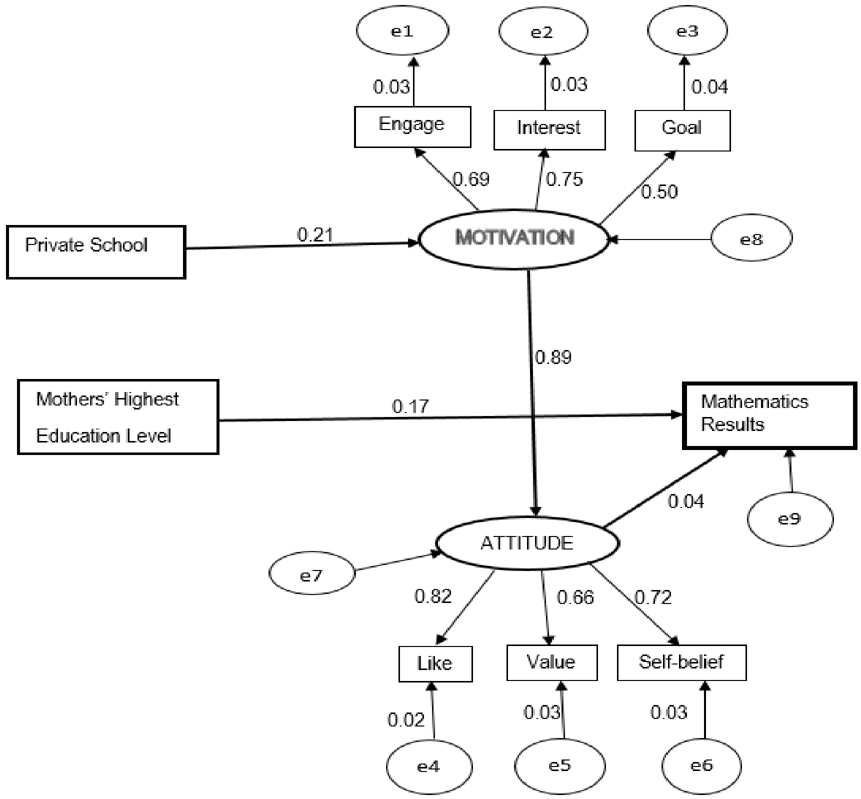


Figure 10. 4 The final Grade 12 student level model (n=375)

Table 10. 10 Estimated path models of the outcome of mathematics results for Grade 12 students

Dependent Variable		Grade 10 students			
	Independent	**Direct (β)**	**Indirect (β)**	**Total (β)**	***P***
Attitude					
	Motivation	0.89	-	0.89	0.00
Motivation					
	Private schools	0.21	-	0.21	0.00
Maths Results					
	Motivation	-	0.04	0.04	0.00
	Attitude	0.04	-	0.04	0.00
	MHEDL	0.17	-	0.17	0.00
	Private School	-	0.01	0.01	0.00

As shown in Table 10.10, *Motivation* ($\beta=0.89$) has a significant direct effect ($p<0.05$) on the attitude of students toward mathematics, with a positive path coefficient. This shows that a higher motivation enables Grade 12 students to have more positive attitudes about their ability to learn mathematics. Similarly, *private schools* ($\beta=0.21$) have a significant ($p<0.05$) direct effect on motivation. This result suggests that students at private schools are more motivated than those studying at the public and Catholic schools.

Moreover, as reflected in Figure 10.4 and Table 10.10, there are two variables that have a direct effect on students' mathematics results. These variables are *MHEDL* ($\beta=0.17$) and *attitude* ($\beta=0.04$). The variable mothers' higher education level has a positive path coefficient and is statistically significant in influencing Grade 12 students' results. This indicates that students who have their mothers with higher education level are likely to obtain better results in mathematics. In addition, attitude ($\beta=0.04$) has a significant ($p<0.05$) direct impact on students' mathematics results, with a positive path coefficient. This suggests that Grade 12 students with more positive attitude towards mathematics are likely to obtain better results. The finding is similar to other studies reviewed in Chapter 2.

Furthermore, *motivation* ($\beta=0.04$) has a positive indirect effect on mathematics results mediated by the attitude scale. The motivation scale shows a significant ($p<0.05$) effect on students' mathematics results through attitude. This indicates that Grade 12 students are more likely to have positive attitude when they are motivated and consequently are likely to obtain better results. In Figure 10.4, it can be seen that *private schools* ($\beta=0.21$) have a positive indirect effect on students' mathematics achievement, mediated by motivation to attitude scales. Private schools in Port Moresby have a significant ($p<0.05$) influence on the Grade 12 students' mathematics results compared to public and Catholic schools through their motivation to attitude. This suggests that the students at private schools are more motivated and are more likely to have more positive attitudes and as a result perform better in mathematics than public schools. The next section of this chapter discusses the results of the teacher-level model, following similar procedures and techniques employed in the previous three student-level models.

10.9 Scale and variables used at the teacher level analysis

The teacher-level model indicates that there are four latent (unobserved) variables and eight observed variables involved in the hypothesised model. These different types of constructs and variables may influence Grade 10 and 12 students' mathematics achievement at the teacher level. Similar to the student level models, a dummy variable was also used to distinguish the treatment group (Hayes et al., 2017; Henseler et al., 2015) for the teacher-level model. These three dummy variables are the government, Catholic and private schools where the teachers were working. The government schools were confirmed as the baseline variable, and are not represented in the model. This technique, as mentioned above in relation to the student-level models, may show significant differences for the variables and scales in the model.

Similar W-scores mentioned earlier at the student-level are also used at the teacher-level model analysis. At the teacher level, the W-Scores for the *quality teaching* (QT), *classroom environment* (CE), and *instructional resources* (RE) scales were analysed for their impact on Grade 10 and 12 students' mathematics results. Other observed variables that might also have impact on students' mathematics results are also used, including the grades teachers teach, their gender, highest education level and the major and minor subjects they studied at university. The latent and manifest variables used in the teacher-level model are shown in Table 10.11.

Table 10. 11 The variables used in the teacher-level model

Latent Variables/Scales	Manifest variables	Description	Coding
	SCHOOL TYPE 1	Government	1=public, 0 =non-public
	SCHOOL TYPE 2	Catholic	1=Catholic, 0 =non-Catholic
	SCHOOL TYPE 3	Private	1=private,0=non-private
	GRADE	Year level teachers teach	0=Grade 10
			1=Grade 12
	GENDER	Teachers' gender	0=Female
			1=Male
	Subject major	Major specialisation at uni	0=Maths
			1=Others
	Subject minor	Minor specialisation at uni	0=Maths
			1=Others
	Years of teaching	Experience	0=1-10 years
			1=11-20 years
			2=21 -30 years
			3= 30 years over
	Teachers' education	Teachers' education level	0=Certificate
			1=Diploma
			2=Bachelors
			4=Postgraduate
Quality of teaching	QT	Teachers' content and strategies of teaching	WLE/W-scores
Classroom environment	CE	Learning atmosphere	WLE/W-scores
Instructional resources	IR	Resources used for teaching	WLE/W-scores
Maths results	Q1-Q40	Test achievement	WLE/W-scores

10.9.1 The hypothesised teacher-level model

The hypothesised teacher-level model shown in Figure 10.5 (p.254) is derived from the theoretical framework discussed in Chapter 3. The theoretical framework identified that QT, CE and IR are affecting Grade 10 and 12 students' mathematics results. As highlighted in the literature review section (Chapter 2), these constructs, combined with other demographic variables such as teaching experience, level of education, specialisation of subjects and school type influence teaching methods at the teacher-level, and consequently, have an impact on students' mathematics results. These demographic variables are hypothesised as exogenous (independent) variables as they are not influenced by other variables and scales in the teacher-level model. The constructs QT, CE and IR at the teacher level are viewed as endogenous (dependent) variables, as they may interact with one another at the teacher-level and with the student-level model. As mentioned earlier in relation to the student-level model, scales or variables in the process are product variables and are viewed as endogenous, due to the mediating effect or influence they have on the teacher-level model. The proposed theoretical framework in Figure 10.5 shows that the observed variables and two scales (CE and IR) influence the scale QT, therefore having an indirect impact on students' mathematics results.

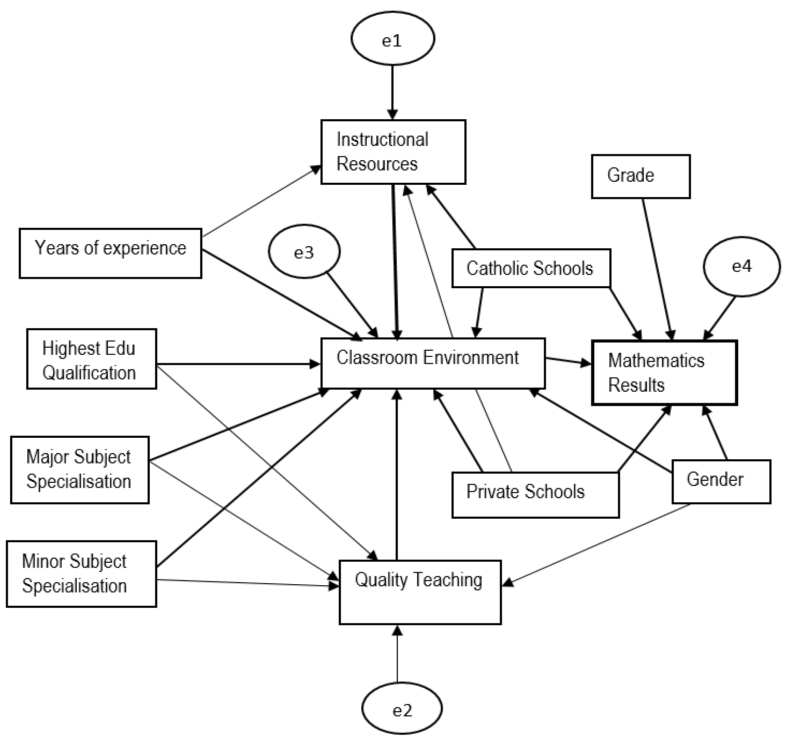


Figure 10. 5 The proposed teacher level model (n=41)

10.9.2 The final teacher-level model

After deleting the variables not conforming to the statistical criteria, the final teacher-level path is shown in Figure 10.6. The deleted demographic variables from the hypothesised teacher-level model are: grade, gender, highest education level and minor areas of specialisation, as they were not within the set statistical criteria.

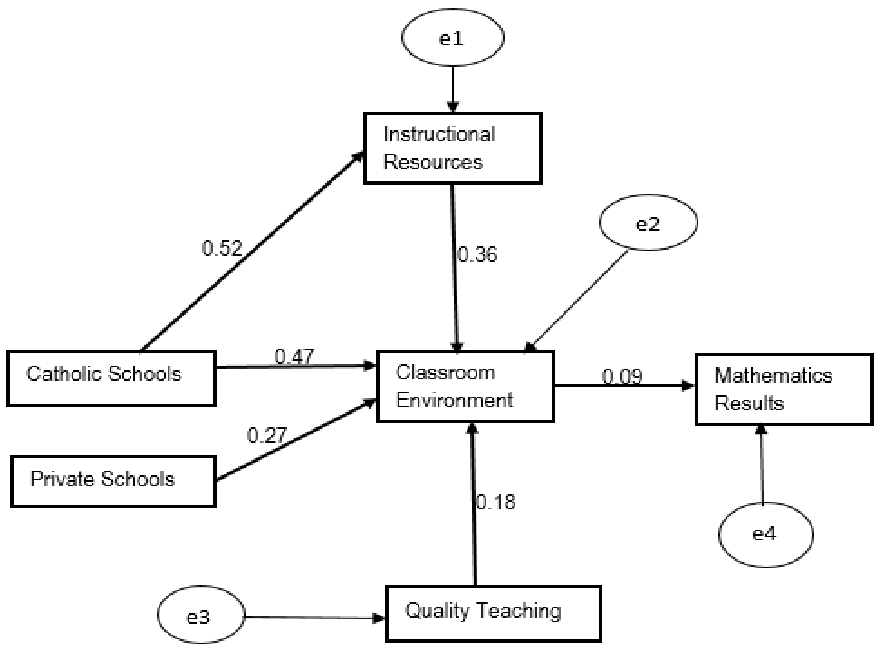


Figure 10. 6 The final teacher level model (n=41)

10.9.3 Measurement model results for teacher-level scales

As discussed in Chapter 9, the scales at the teacher-level (*quality teaching* (QT), *classroom environment* (CE) and *instructional resources* (IR)) had single-factor structures and were represented by a single WLE score (converted to W-scores in this chapter). Therefore, the three constructs were treated as observed variables for the path model in terms of analysis. This means that all the variables at the teacher level reflected in Figure 10.6 involve only observed variables with no latent components. As a result, there was no discussion about the measurement model reflected in this section.

10.9.4 Teacher-level model fit indices

Table 10.13 shows the results of the fit indices of the final teacher-level model. Similar fit indices employed in the student-level model are used in this section. As shown in Table 10.12, the RMSEA value of 0.06 and SRMR values of 0.03 reflected a good model fit, as they are close to 0. Moreover, the CFI (0.97) and TLI (0.94), also provide acceptable values above 0.90 (i.e., closer to 1), indicating that the final teacher level-model is a parsimonious model.

Table 10. 12 The results of the fit indices of the final teacher level model

Fit Index	Obtained values	Remarks
X^2/df	7.09/6=1.18	Acceptable fit
RMSEA	0.06	Acceptable fit
SRMR	0.03	Acceptable fit
CFI	0.97	Acceptable fit
TLI	0.94	Acceptable fit

10.9.5 Results of teacher-level structural model

The Mplus 7.0 structural model results at the teacher level are presented in Table 10.13. The next section discusses in detail the standardised estimates for the direct and indirect effect of the variables on mathematics results.

Table 10. 13 Teacher level variables with structural model results

Dependent Variable	Independent	Teachers Direct (β)	Indirect (β)	Total (β)	*p*
IR					
	Catholic	0.52	-	0.52	0.00
CE					
	IR	0.36	-	0.36	0.00
	QT	0.18	-	0.18	0.04
	Catholic	0.47	0.06	0.06	0.00
	Private	0.25	-	0.25	0.04
Maths Results					
	QT	-	0.02	0.02	0.03
	IR	-	0.03	0.03	0.04
	CE	0.09	-	0.09	0.56

Figure 10.6 and Table 10.13 show that *Catholic schools* (β=0.52) have a significant ($p<0.05$) direct effect on instructional resources (IR). The path coefficient of Catholic schools on instructional resources indicates that Catholic schools provide more adequate instructional resources for teaching compared to public and private schools. Similarly, *IR* (β=0.36), *QT* (β=0.18), *Catholic* (β=0.47) and *private* (β=0.25) have direct effect on classroom environment experience (CE). *Instructional resources* have a significant ($p<0.05$) effect on the classroom learning environment, implying that teachers' use of different sources to more creatively impart knowledge produces more successful learning spaces. Moreover, *Catholic and private schools* have a significant impact on the teaching environment, suggesting that teachers at these schools encourage a conducive environment for both teaching and learning in classrooms, compared to the public schools (baseline schools). Furthermore, the *quality of teaching* (β=0.18) of teachers has a significant ($p<0.05$) effect on the classroom environment, with a positive path coefficient. This suggests that higher-quality teaching produces a better learning atmosphere for teachers and students to work collaboratively in the classroom. Moreover, *Catholic schools* (β=0.06) exerted a positive indirect effect that is mediated by instructional resources and classroom environment on mathematics results. This suggests that Catholic

schools also offer better instructional resources to support teachers' instructional practices, compared to both private and public schools, thereby promoting better classroom learning environments.

As indicated in Table 10.13, *classroom environment* (β=0.09) has a direct impact on the students' mathematics results. However, this effect is non-significant ($p>0.05$) even though the path coefficient is positive. One of the main reasons for this result is the small sample size ($n=41$) of teachers used in the study. Finally, the constructs *IR* (β=0.02) and *QT* (β=0.03) have indirect positive path coefficient effects on students' mathematics results. Both of these constructs have significant impact ($p>0.05$) on the students' mathematics results.

10.10 Summary

In this chapter, two separate models were analysed: the student and teacher-level models. Within these, at the student-level, three separate models were analysed: a combined model for both Grade 10 and 12 students, and separate Grade 10 and Grade 12 models. Both the teacher and different student-level models were analysed using Mplus 7.0.

The chapter examined the effects of student-level constructs (*motivation and attitude*) and teacher level variables on mathematics achievements. On the basis of these findings, it can be stated that there is a strong correlation between *motivation* and *attitude* that significantly ($p<0.05$) affects mathematics results for the combined, Grade 10 and Grade 12 student-level models in the theoretically expected direction. The demographic variables *private schools, FHEDL and MHEDL* also had a significant ($p<0.05$) effect on the mathematics results. These results show strong support for the hypothesised relationship in Chapter 2 and Figure 10.2, (p.241) as *motivation* and *attitude* with the aforementioned demographic variables have an effect on students' mathematics results.

The common findings in the three-student-level models (combined, Grade 10 and 12) is that *attitude, mothers' highest education level* and *private schools* had a significant effects on the students' mathematics results. However, the difference between the models is that, though *fathers' highest education level* had a significant influence on the combined and Grade 10 models, it had a non-significant effect on the Grade 12 model.

The SEM results of the three models indicate that the mother's highest education level (MHEDUL), private schools and attitude (ATTD) are the similar/common factors that have a significant effect on students' mathematics results. As it is evident in the results, MHEDUL is one of the factors that influences the children to perform better in school. For instance, mothers with higher education check their children's homework to succeed at school. Furthermore, the private schools influence students' results due to the positive effect they have on students' attitude. The private schools have conducive learning environment and better teachers in PNG that promotes positive attitude for students to perform well in mathematics. The difference among the three models is that fathers' highest education level (FHEDUL) has significant effect on the combined and Grade 10 models but not for Grade 12 model. These results of the two models indicate that fathers' highest education level influence students to perform well in mathematics in Port Moresby. This could be due to the educated fathers standing in families and society to influence young children.

Moreover, at the teacher-level analysis, variables such as *instructional resources, quality teaching* and *Catholic schools* had a significant ($p<0.05$) relationship with the classroom environment. This result reflects the hypothesised relationship in Chapter 2 and section 10.5.1 that *instructional resources, quality teaching* and *Catholic schools* are correlated with *classroom environment*. In spite of this result, however, the *classroom environment* had a non-significant ($p>0.05$) effect on the students' mathematics results. The other finding from the analysis is that the subscales/factors and the two constructs (*attitude and motivation*) had a significant relationship in the measurement model for the three different student models.

The constructs and variables analysed in this chapter make up two levels (student and teacher-levels) because the data collected in some schools were less than two teachers per school. These data will be further analysed in the next chapter, using Hierarchical Linear Modelling (HLM). This step is taken to examine how these variables interact with each other at the two different levels, and how their factors directly influence the Port Moresby Grade 10 and 12 students' mathematics results.

Chapter 11: Hierarchical Linear Modelling

1.1 Introduction

In Chapter 10, SEM was employed separately for both the student and teacher-level data using Mplus 7.0. The data collected were multilevel or hierarchical, and so could introduce bias when aggregated or disaggregated (as discussed in Chapter 3) to form a single level of data. This is because the data obtained from this study possess a multi-level or hierarchical structure, which is consistent with other educational studies. Therefore, to overcome the issue of encountering bias from the aggregation of multi-level data, the hierarchical linear modelling (HLM) technique is employed to analyse the data in this chapter. This technique is used to examine the direct effects of different predictors at the student and teacher levels on the mathematics results, and simultaneously investigates the cross-level interaction effects that occur between student-level (level-1) variables and teacher-level (level-2) variables that may have a significant effect on the results. In this chapter, three different HLM analyses are employed, paralleling the three path analyses for student and their teacher results in Chapter 10. The SEM results in Chapter 10 shows that the attitude scale has a significant effect on the combined, Grade 10 and Grade 12 students' results. Motivation scale also has significant indirect influence on combined, Grade 10 and Grade 12 students' mathematics results. Given these results, it is necessary to examine the predictors and interaction effects for the three student groups and their teachers' data at the hierarchical level, in order to obtain more meaningful insights into the results.

This chapter begins with a brief discussion of the HLM technique and the HLM 6.08 software used in the analysis of data in this chapter. The chapter then outlines the variables used at the student and teacher-levels. After that, the combined, Grade 10 and the Grade 12 students' null models with their final models' direct and cross- interaction effects on mathematics results are discussed separately. Finally, the chapter concludes with a summary highlighting the chief findings of the analysis.

11.2 Overview of hierarchical linear modelling (HLM)

As discussed in Chapter 3, HLM provides both direct effects from various levels and the interaction effects between variables at different levels (Ramlall, 2016). The hierarchy shown in the multilevel data is made up of units grouped at different levels (Geng, 2014; Hox, Moerbeek, & Van de Schoot, 2017; Raudenbush & Bryk, 2002). For example, in the three-level model concerning students, teachers and schools, individual students can be assigned as level-1, teachers as level-2, and schools as level-3 units. This means that students (who occupy the first hierarchical level) are clustered or nested within teachers (who are located in the second level), and both students and teachers are clustered within schools, (which are located at the third level of the hierarchy). This concept of hierarchy is applied in this context of HLM.

HLM is more efficient than other analytical techniques in estimating of the variance among variables at different levels (Goldstein, 2011; Hox, 2013; Hox et al., 2017). In this study, HLM serves three analytical purposes: it (a) improves estimation of individual effects; (b) models cross-level effects; and (c) partitions variance-covariance components across levels in order to apply significance tests more appropriately (Bryk & Raudenbush, 2002).

11.3 Application of HLM 6.08

There are several software products available for multilevel data analysis. One of these products is HLM 6.08, which is capable of multilevel analysis such as hierarchical linear modelling with complex calculations of data (Hox et al., 2017). These calculations are carried out in estimating the model coefficients and predicting the random effects associated with sampling variables at each level (Raudenbush, Bryk, Cheong, & Congdon Jr, 2004). In this study, HLM 6.08's is used to estimate the effects of the student level variables on the outcome variable, and also estimates the teacher-level variables on the coefficients from the student-level analysis. The development of HLM's 6.08 capabilities and functionalities has progressed through the years, and it provides a wide choice of estimation options including an ability to manage three- and four-level models (Raudenbush et al., 2004). The software reads data in a particular format from a source. Its importing capabilities are enhanced by reading data not only from a plain text (ASCII) format but also from data saved in SPSS and other statistical software (Ben, 2010; Hailaya, 2014). In this study, HLM 6.08 is used to undertake the student and teacher- level analysis to estimate the effect of level-1 variables (student-level) on the outcome variables, and to estimate the effects of the level-2 variables (teacher-level) on the coefficient of level-1 variables, and on the response variables in level-1.

The analysis of the multilevel model using HML 6.08 involves three steps (Garson, 2013). The first step is to import the data set into HLM 6.08 to create a multivariate data matrix (MDM) file. Running the analysis using the MDM file is the second step. Finally, the third step is evaluating model fit with relevant estimates and residual files. The MDM file is constructed from raw data saved in SPSS (Garson, 2013). The teacher and student-level data (2 levels) are required for the input, meaning that the student file was prepared for level-1 and the teacher file was prepared for level-2. These two data files are linked by the level-2 ID variable (the Teacher ID). The MDM file produced is then used as input in all subsequent analysis. The MDM file can be viewed as a "system file" in the standard computing packages that contains both the summarised data and the names of all the variables (Ben, 2010; Garson, 2013).

11.4 Variables in the two-level model

The sets of variables at the student and teacher levels required identification before conducting HLM analysis. As mentioned in Chapter 3, the variables/items for the scales were subjected to Rasch analysis (except for the demographic variables). This technique was carried out to obtain the Weighted Likelihood Estimate (WLE), which later converted to W-scores, consistent with the PISA studies. These W-scores for the scales that were used in Chapter 10 are also used in the HLM analysis of this chapter. The teacher-level scales (*teaching quality, classroom environment* and *instructional resources*) have a single-factor structure and are denoted by one set of W-scores as discussed in Chapter 3. However, at the student-level, the scales *attitude* and *motivation* have three subscales with three W-scores that indicate a hierarchical structure, which are simplified into a one factor structure. The factor scores of the W-scores of the three subscales under each scale (*attitude* and *motivation*) are calculated in SPSS to be employed in HLM analysis.

Table 11.1 lists the teacher-level (level-2 or macro-level) and student-level (level-1 or micro-level) variables that are examined in this study. In this chapter, variables are listed in upper case with reference to variables used in the HLM analysis that require reporting.

Table 11. 1 List of variables in the three-level model

Variable Name	Variable label	Description
Teacher-level variables		
SCHOOL 1	Type 1	Catholic
SCHOOL 2	Type 2	Private
SCHOOL 3	Type 3	Public
GRADE	Class	Year level teachers teach
GENDER	Male/female	Teachers' gender
MAJOR	Major subject	Major specialisation at university
MINOR	Minor subject	Minor specialisation at university
YOTE	Years of teaching experience	Experience teaching mathematics
TEL	Teachers education level	Teachers education Level
TQ	Quality of teaching	Teachers content knowledge and teaching strategies
CE	Classroom environment	Classroom learning atmosphere
IR	Instructional resources	Resources/materials for teaching
Student-level variables		
CLASS TEACHING	Grade	Students' year level
SEX	Male/female	Students' gender
MOCC	Mothers occupation	Job mothers do for living daily
FOCC	Fathers occupation	Job fathers do for living daily
MHEDUL	Mothers Highest Edu Level	Mothers highest education level attained
FHEDUL	Fathers Highest Edu Level	Fathers highest education level attained
ATTD	Attitude	Attitude towards mathematics
MTVN	Motivation	Motivation towards mathematics

The scales analysed under HLM are in W-scores, (i.e., the standardised form with the mean of 500: see Chapter 3), but not the demographic variables, and it is important to define the location of the variables for better interpretation of the results. In HLM, information about all of the variables from the two different levels is centred for clarity purposes (Thien, Darmawan, & Ong, 2015; Woltman et al., 2012). The two common ways of centring variables are group-mean centring and grand-mean centring (Luke, 2004; Stephen & Anthony, 2002; Thien et al., 2015). In group mean centring, the mean of the group is subtracted from the corresponding individual scores, whereas grand mean centring is when the overall mean is subtracted from all values of the variable (Hox et al., 2017; Thien et al., 2015). In this study, a grand-mean centring approach is employed for level-1 continuous variables (such as *attitude* and *motivation*) and level-2 continuous variables (*quality teaching, classroom environment*, and *instructional resources*). This approach assumes that group means are uncorrelated with predictors and interaction estimates, and tests are unaffected, to improve computation and for easy interpretation of the results of the “main effects” with fixed mean at 0 when cross-level interactions are tested (Hofmann & Gavin, 1998; Hox et al., 2017). All the categorical variables (demographic variables) for both levels are entered as un-centred. Thus, it is important to be familiar with these two different approaches of centring for proper and meaningful interpretation of results and numerical stability in estimating HLM.

11.5 The two-level model of mathematics achievement

As mentioned above, two-level HLM is carried out to examine the relationships between level-1 and level-2 predictors and outcome variables (mathematics results) in the model. In this study, the terms ‘student-level’, ‘level-1’ and ‘micro-level’ refer to the same data and, are used interchangeably, and the same is true for the terms ‘teacher-level’, ‘level-2’ and ‘macro-level’. The two-level model of the mathematics results is shown in Figure 11.1. The level-1 variables allow the examination of the independent effects of level-1 and level-2 predictors, which facilitates better interpretation of the intercept as the expected value of the outcome variable, and other multilevel analysis results (Hox et al., 2017; Woltman et al., 2012).

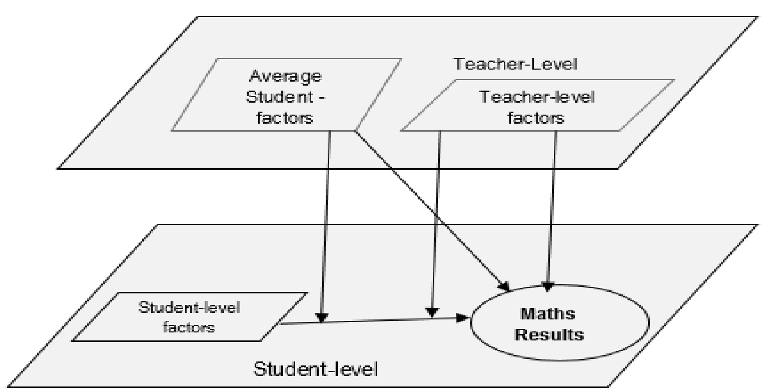


Figure 11. 1 Two-level model of mathematics results

11.5.1 The hypothesised model and analysis framework

The hypothesised two-level (2L) model of mathematics results derived from the theoretical framework introduced in Chapter 2 is shown in Figure 11.1 (p.275). The reason that the 2L/HLM model serves as this study's theoretical framework instead of the 3L/HLM model (as outlined in Chapter 2) is that even though school types appear as one of the categorical variables, six out of the 16 schools have less than three teacher participants, making it impossible for HLM 6.08 to estimate the coefficient. This means the data associated with the six schools would be removed completely during analysis with HLM 6.08, resulting in the loss of valuable information.

The 2L/HLM theoretical framework proposed has ten variables at the teacher level and eight variables at the student level. The variables at teacher level include: *teacher quality* (TQ), *classroom environment* (CE), *instructional resources* (IR), *private* and *Catholic schools, gender, years of teachings experience* (YOTE), *highest education level* (HEDL), *major specialisation subject* and *minor specialisation subject* at universities. The student level factors include: *attitude, motivation, grade, gender, mothers' occupation* (MOCC), *father's occupation* (FOCC), *fathers' highest education level* (FHEDUL) and *mothers' highest education level* (MHEDUL).

According to the theoretical framework, the factors at each level affect Grade 10 and 12 students' mathematics results. Although the 2L/HLM model cannot examine indirect effects, it can investigate direct and interactive effects for 2L/HLM variables using HLM 6.08. In other words, the variables at each stage directly influence variables at each level and also provide interactive effects for students' mathematics results at the two levels.

The level-1 (student-level) data is aggregated into level-2 (teacher-level) data. This process forms the level-2 compositional factors/variables, as shown in Figure 11.1 (p.274). These variables at level-2 have effects on level-1 variables and consequently directly affect students' mathematics results. For instance, this study investigates the relationship between an individual student's mathematics results and the student's attitude together with their parents' average highest education level. The effect of the *attitude* variable from level-1 on the student's mathematics result is named 'individual' because it refers to individual attributes of the student (this student has a particular attitude). The effect of the *parents' average highest education level* variable from level-2 (see Figure 11.1) on the student's performance is a compositional effect (Pokropek, 2015). This means that parents' highest education level is an attribute of each of the students in a particular classroom that combines as one factor to have an effect on the learning outcomes (Darmawan & Keeves, 2009). On the other hand, contextual factors/variables are the actual factors that are associated with the teachers' contexts at the classroom level (see Figure 11.1) that predict or influence the students' mathematics results. In other words, the contextual variables at the teacher-level contribute by specifying the contexts in which phenomena are expected to occur in the classroom to influence students' mathematics results (Tortorella, Marodin, Miorando, & Seidel, 2015).

Figure 11.2 shows the direct effect of student and teacher-level variables on mathematics results. Theoretically, the arrows from each of the student-level variables should interact with teacher level variables to have an impact on mathematics results. However, for simplicity and to avoid complication with the arrows pointing from each of the teacher-level variables toward each of the student-level effects representing the interaction effects, only the direct effects from both levels are presented in Figure 11.2.

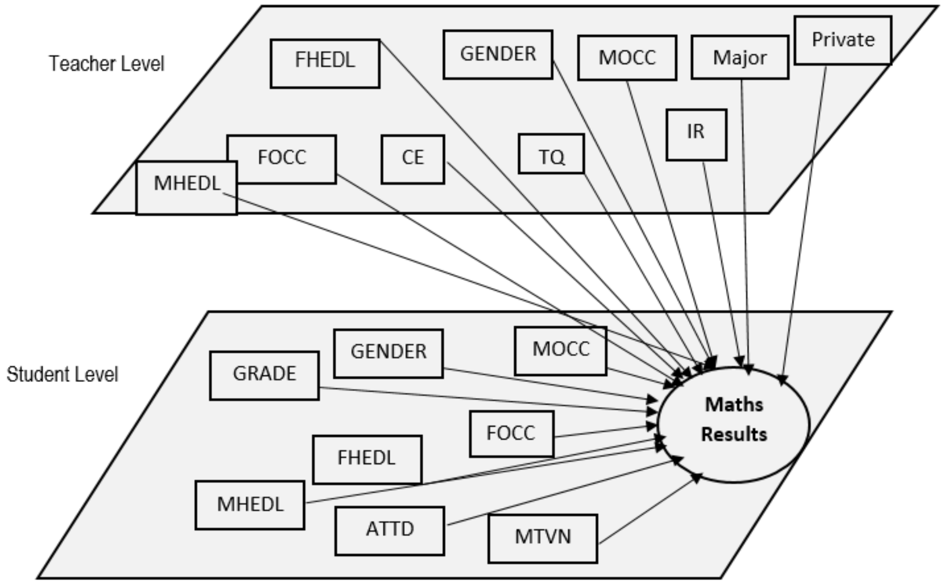


Figure 11. 2 The hypothesised two-level model of mathematics results

11.6 Model building and evaluation

A bottom-up approach was taken in building the multilevel model employed in this study (Hox et al., 2017). This approach begins with simple model and proceeds by adding variables one at a time. A variable that has a non-significant effect on mathematics results is then removed to obtain a parsimonious model. A bottom-up strategy is more productive than the top-down approach and allows the researcher to identify the best predictors, as well as tending to avoid multicollinearity problems (Hox, 2013; Hox et al., 2017; Stephen & Anthony, 2002).

Moreover, the evaluation and assessment for HLM in this study is based on statistical analysis such as the chi-square test (x^2), intraclass correlation coefficient (ICC), reliability of intercepts and slopes, and deviance statistics (Bryk & Raudenbush, 2002). The x^2 is a significance test and indicates the existence of variance in the dependent variable that is attributable to level-2 factors. The ICC is another statistical value that examines the null model to determine whether HLM is needed, and it is calculated by dividing the variance of level-2 units by the sum of variance of level-1 and level-2 units (Woltman et al., 2012). The multilevel model is needed when a high ICC indicates that there is significant variance in the outcome variable. As such, an ICC greater than 10% of the total variance in the outcome warrants the use of HLM (Ben, 2010; Hailaya, 2014). Similarly, an ICC threshold of 0.25 and above is taken as impetus for the multilevel model analysis to proceed (Go, 2005).

Further, the reliability of intercepts and slopes is an important part of the output. "Reliability here is viewed in terms of the degree of variability present between groups compared to total variability (i.e., between-group variance plus error variance)" (Hailaya, 2014, p. 302). The reliability threshold is 0.05, and any reliability below this threshold causes a variety of numerical difficulties. Therefore, the parameter that has been estimated should be fixed (Darmawan & Keeves, 2009). Deviance is the final statistical output that is employed to measure the fit between the data and the model and can be utilised to compare models. A significant reduction in the value of deviance indicates a better fit of the model (Hox et al., 2017; Stephen & Anthony, 2002).

In a regression equation, the error term (residuals) is sometimes denoted by "e" or "R". The residual expresses the part of the equation of dependent variable "Y" and residuals represent the unexplained variance in each level of a multilevel model.

11.7 Null model

The analysis of the two-level HLM model is carried out by first running the null model, or fully unconditional model, to obtain the estimates of the amount of variance available to be explained in the model. The estimates do not involve any predictor at the student and teacher levels. Thus, the null model is similar to one-way analysis of variance (ANOVA) with random effects (Darmawan & Keeves, 2002; Thien et al., 2016), and represents how much variation there is in the outcome (mathematics results) that is allocated across the student and teacher levels. In other words, the null model also allows the partitioning of the variance of the outcome variable (mathematics results) at the two levels (Darmawan & Keeves, 2002; Thien et al., 2016). The purpose of running the null model is to obtain the empirical or statistical evidence required to decide whether HLM is necessary and to estimate other coefficients (Darmawan & Keeves, 2002).

According to Darmawan and Keeves (2002), the null model is given in the following equation.

$$\text{Level-1 Model:} Y_{ij} = \beta_{0j} + r_{ij} \qquad [11.1]$$

where:

Y_{ij} is the mathematics result according to the student *i* under the teacher *j*

β_{oj} is the intercept for teacher *j* (the mean outcome (mathematics result) for the j^{th} teacher),

r_{ij} *is* a random error (unique random effect of student *i* under teaher j)

The indices i and j denote level-1 units (students) and level-2 units (teachers), respectively, where there are *i*=1,2,…N, students with J teachers; and *j*=1,2,....N teachers (Ben, 2010).

In Equation 11.1, mathematics results for student *i* under teacher *j* are considered to be similar to the teacher mean plus a random error. This implies that the fully unconditional model assumes no difference in the mathematics results between students with the students at level-1. It is assumed that each level-1 error, r_{ij} *is* normally distributed with a mean of zero and a constant level-1 variance, σ^2 (Bry & Raudenbush, 1992 cited in Darmawan & Keeves, 2002 p.50).

$$\text{Level-2 Model: } \beta_{0j} = \gamma_{00} + u_{oj} \qquad [11.2]$$

where:

β_{0j} is the intercept for the teachers

γ_{00} is the mathematics results all teachers (the grand mean outcome in the population),

u_{oj} is the unique random effect associated with teacher *j (i.e, the deviation of teacher j's mean from grand mean).*

In Equation 11.2, average mathematics result scores for teacher *j* are considered equivalent to the mean across all teachers plus a random error (Darmawan & Keeves, 2002). In other words, the fully unconditional model assumes no differences in the mathematics results between teachers at level-2. This is under the assumption that the effect associated with teacher j, u_{oj} has a normal distribution with the mean of zero and variance τ_β (Bryk & Raudenbush, 1992 cited in Darmawan & Keeves, 2002, p.51).

Running the null model provides a point estimate and confidence interval for the grand mean, Y_{oo} (Darmawan & Keeves, 2002). The σ^2 parameter represents the student-level variability, while the τ_β parameter represents the teacher-level variability (Bryk & Raudenbush, 1992; Bryk & Raudenbush, 2002). According to Bryk and Raudenbush (2002), the null model allows the estimation of the proportions of variation within and among teachers, respectively, by the following mathematical expressions:

$\sigma^2 / (\sigma^2 + \tau_\beta)$ the proportion of variance within teachers [11.3]

$\tau_\beta / (\sigma^2 + \tau_\beta)$ the proportion of variance among teachers [11.4]

Moreover, the reliability represents the degree to which the teacher-level units can be discriminated between using the least square estimates of β_{oj} (Bryk & Raudenbush, 2002 cited in Hailaya, 2014, p.305). The variance at each level is by σ^2 for the level-1 model, and τ_β for level-2 model. Bryk and Raudenbush (2002) further state that the reliability estimate for the student sample mean for each teacher group can be calculated using the following equation

$$\text{Reliability } (\beta_{oj}) = \tau_\beta / (\tau_\beta + \sigma^2 / n_{jk}) \quad [11.5]$$

The average of the reliabilities across teachers or groups of teachers may be viewed as measures of reliability of the teacher means as reflected by the Equation 11.5 (Thien et al., 2016).

11.8 Combined student sample results

Table 11.2 (p. 274) shows the HLM unconditional model results for the combined student data. These results indicate that the between-teacher variance exhibits statistical significance (u0j= 12.50, x^2 *(40)* = 308.05, p<0.00). In other words, the mathematics results of the combined students' data varied across the teacher group. The ICC 0.287 (12.50/ (12.50+30.95) =0.287 or 29%) indicates that the combined data has a hierarchical structure, which consequently allows multi-level analysis to proceed. This result shows that

29% of the variability in student mathematics results is attributable to teachers, while the remaining 71% is attributable to student characteristics. This is an indication that most of the variance comes from students rather than their teachers.

The result of the reliability shown in Table 11.2 indicates the extent to which the mean of the dependent variable can be discriminated among level-2 variables (Ben, 2010; Hailaya, 2014b). The reliability of 0.85 of the null model indicates that the teachers' sample mean can estimate well the student outcome (mathematics results). This is because the "reliability is a good indicator of how well the teachers sample mean is in this study to estimate the unknown parameter β_{oj}" (Ma et al.2008 cited in Haiyala, 2014 p.306). These results of the null model show the presence of variance at the teacher and student levels and confirm the suitability of multilevel analysis.

Table 11. 2 Fully unconditional model results for 2L/HLM for the combined student model

Final estimation of fixed effects						
Fixed Effect		Coefficient	Standard Error	T-ratio	DF	Approx. P-value
For INTRCPT1, β_{0j}						
INTRCPT2, β_{00}		495.79	0.59	840.23	40	0.00
Final estimation of variance components						
Random Effect	Reliability	Standard Deviation	Variance Component	DF	Chi-square	P-value
INTRCPT1, u_{0j}	0.85	3.54	12.50	40	308.05	0.00
level-1, rij		5.56	30.95			
Statistics for current covariance components model						
Deviance	4651.81					
Number of estimated parameters		2				

11.8.1 Analysis steps for the HLM student models

After running the null model, the variables are entered into the equation using the bottom-up approach explained earlier. First, the level-1 variables are entered one at a time into the equation according to the magnitude and statistical significance of the path coefficient. This step is necessary to examine how much variance the level-1 predictors (Bry & Raudenbush, 1992) explain. The results are then examined, and variables that have coefficients that are found to be non-significant are deleted, with the next probable variable entered into the equation. This procedure is repeated until the significant variables for level -1 is achieved and finalised. The next step is entering the teacher level variables into the equation. This step requires exploratory analysis to check the possible variables for level-2 to be included in the model. As such predictors that are found to have t-ratios more than 2 are retained from the model. This means the teacher-level variables that are retained are entered one by one, starting with the variable with highest t-ratio. This step is repeated until a final model is obtained with an equation showing only significant effects (predictors) at both level-1 and level-2, accordingly. For consistency purposes, student-level variables that had a significant effect on mathematics results in Chapter 10 are used as a starting point for level-1 variables in this multi-level analysis, as shown in Table 11.4 HLM's results.

11.8.2 The final combined students' model

The final combined model statistical fit for the final model (in terms of coefficients, reliability and significance level) is presented in Table 11.3 below.

Table 11. 3 The final model results for two levels model for combined students' mathematics results

Combined final estimation of fixed effects						
Fixed Effect		Coefficient	Standard Error	T-ratio	Approx D.F	P-value
For INTRCPT1, β_{0j}						
INTRCPT2, γ_{00}		492.47	2.00	245.96	38	0.00
Maths Major Spec, γ_{01}		1.87	0.85	2.21	38	0.03
Ave-FatherOCC, γ_{02}		6.63	3.05	2.19	38	0.03
For GRADE slope, β_{1j}						
INTRCPT2, γ_{10}		-2.08	1.01	-2.05	722	0.04
For MHEDUL, β_{2j}						
INTRCPT2, γ_{20}		0.81	0.27	3.02	40	0.00
For ATTD slope, β_{3j}						
INTRCPT2, γ_{30}		1.09	0.27	3.93	39	0.00*
AVE_MHEDUL, γ_{31}		1.81	0.91	1.99	39	0.03*
Final estimation of variance components						
Random effect Deviation	Reliability Component	Standard Deviation	Variance Components	df	Chi-square	P-value
INTRCPT1, U0j	0.71	2.68	7.18	37	159.13	0.00
MHEDUL slope, U2j	0.03	0.39	0.16	39	34.44	>.50
ATTD slope, U3j	0.29	0.96	0.94	38	51.72	0.06
level-1, Roj		5.35	28.62			
Statistics for current covariance component model						
Deviance				4582.72		
Number of estimated parameters				7		

**Cross-level interaction effect (P<0.05)*

The final model's variables shown in Table 11.3 is specified by the following equation.

Level-1 model

$$Y_{ij}=\beta_0 + \beta_{1j}(\text{Grade}) + \beta_{2j}(\text{MHEDUL}) + \beta_{3j}(\text{ATTD}) + r_{ij} \quad [11.6]$$

Level-2 model

$$\beta_{0j}=\gamma_{00}+\gamma_{01}(\text{Mathsmajor}) + \gamma_{02}(\text{Ave-FOCC}) + u_{0j} \quad [11.7a]$$
$$\beta_{1j}= \gamma_{10}+ u_{1j} \quad [11.7b]$$
$$\beta_{2j}= \gamma_{20} + u_{2j} \quad [11.7c]$$
$$\beta_{3j}= \gamma_{30}+ \gamma_{31}(\text{Ave-MHEDUL}) +u_{3j} \quad [11.7d]$$

The final model is represented by the equation resulting from substituting Equations 11.7a to 11.7d into Equation 11.6.

$$\gamma_{ij} = \gamma_{00} + \gamma_{01}(\text{Maths major}) + \gamma_{02}(\text{Ave-FatherOCC}) + \gamma_{10}(\text{Grade}) + \gamma_{20}(\text{MHEDL}) + \gamma_{30}(\text{ATTD})$$

$$+ \gamma_{31}(\text{Ave-MHEDUL})(\text{ATTD}) + u_{0j} + u_{1j}(\text{Grade}) + u_{2j}(\text{MHEDL} + u_{3j}(\text{ATTD}) + r_{ij} \quad [11.8]$$

The equation above indicates that the combined students' mathematics results are defined as the function of the overall intercept (γ_{00})*, five direct or main effects, one cross-level interaction effects and a random error* [u_{0j} +u_{1j}(Grade) + u_{2j}(MHEDUL) + u_{3j}(ATTD) + r_{ij}]. The five main effects include the direct effect from student level factors such as *mothers highest education level (MEDUL), attitude (ATTD)* of the students towards mathematics and *Grade.* The *Grade* is added to compare variables after the results were controlled for the grade difference. The teacher- level factors include teachers with mathematics as a *major area of specialisation (Math major)* and the average number of students in the class with their *fathers' occupation (Ave-FOCC).* Further, the cross-level interaction effects include students' *attitude (ATTD)* towards learning mathematics at level-1, and the composition of average students in the class with their *mothers' highest education level (Ave-MHEDUL)* at level-2.

Figure 11.3 reflects the results presented in Table 11.3 and equation 11.8, illustrating the direct and cross-interaction effects of variables from levels-1 and 2 on the combined students' mathematics results.

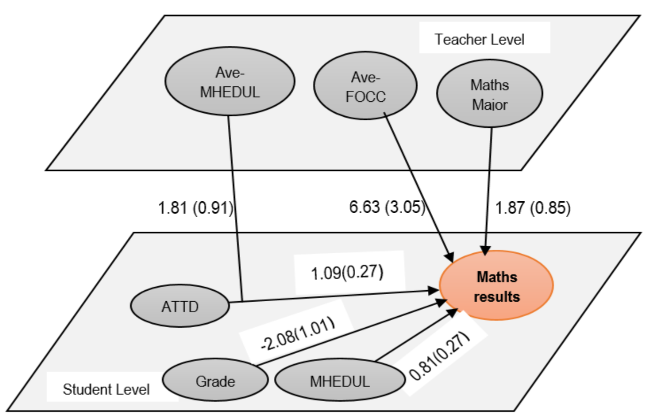


Figure 11. 3 Two-level model of students' mathematics results for the final combined model

11.8.3 The direct effect of level -1 and level-2 predictors on mathematics results

The results in both Table 11.3 (p.276) and Figure 11.3 (p.277) show that level-1 variable *MHEDUL* (0.81) has a significant ($p<0.05$) direct effect on the combined students' mathematics results. This indicates that *mother's highest education level (MHEDUL)* has a positive influence on students' mathematics results. This result suggests that students whose mothers have obtained a higher education level are more likely to perform better in mathematics. Moreover, since the results of the Grade 10 and 12 students' tests were not equated, the variable *grade* was added to allow the estimation of the effects of other variables after the results were controlled for the grade difference. Similarly, *ATTD* (1.09) has a significant ($p<0.05$) positive direct effect on the students' mathematics results. This suggests that students with more positive attitudes towards mathematics obtain better results in the subject. These 2L/HLM level-1 variables result from the combined students' data are consistent with the combined students' SEM/path analysis for the single-level combined student results in Chapter 10.

Furthermore, as shown in Table 11.3 (p. 276) and Figure 11.3 (p. 277), the variables *maths major* (1.87), *Ave-FOCC* (6.63) and *Ave-MHEDUL* (1.81) at level-2 (teacher-level) show significant ($p<0.05$) positive direct effects on mathematics results. This indicates that teachers with mathematics as their major specialisation have a positive effect on their students' mathematics results. This further suggests that some students have better mathematics results because their teachers have adequate mathematics content knowledge and teaching pedagogy. Similarly, students in the classroom whose fathers are employed show better mathematics results. Additionally, students in classes with higher average mothers' *education level (Ave-MHEDUL)* have better mathematics results.

11.8.4 The cross-interaction effects

Three variables are required for the analysis of cross-level interaction: the outcome variable, the predictor at level-1 (student-level), and the level-2 (teacher-level) predictor that has an influence on the level-1 predictor's effect on the outcome variable (mathematics result). The HLM results in Figure 11.3 (p. 277) show that Ave-*MHEDUL* from level-2 and *ATTD* from level-1 have a cross- level interaction. This interaction can be expressed in an equation form. As such, part of the equation for the final combined student model can be drawn to show this cross-level interaction effect by setting the remaining terms to zero. The equation is as follows:

$$\gamma_{ij} = \gamma_{00} + \gamma_{30}(\text{ATTD}) + \gamma_{31}\ (\text{Ave-MHEDUL})\ (\text{ATTD}) + r_{ij}$$

where $\gamma_{00} =$ 492.47 (average mathematics results), $\gamma_{30} =$ 1.09 and $\gamma_{31} =$ 1.81

Thus, the cross-level interaction equation is

$$\gamma_{ij} = 492.66 + 1.09(\text{ATTD}) + 2.02(\text{Ave-MothersHEDUL})\ (\text{ATTD}) + r_{ij}$$

This equation can be used to calculate teacher-level coordinates to obtain a graphical representation of the cross-level interaction effect. The conditions used to calculate the coordinates for teachers used by Ben (2010, p. 389) and Lye (2016b, p. 305) are adapted in this study. These conditions are:

a) One standard deviation above the average on ATTD and AVE-MHEDUL
b) One standard deviation above the average on ATTD and one standard deviation below the average on AVE-MHEDUL
c) One standard deviation below the average on ATTD and one standard deviation above the average on AVE-MHEDUL
d) One standard deviation below the average on ATTD and one standard deviation below the average on AVE-MHEDUL
e) Average on ATTD and one standard deviation above the average on AVE-MHEDUL

f) Average on ATTD and one standard deviation below the average on AVE-MHEDUL

The above guide is used to calculate the coordinates using the AVE-MHEDUL and ATTD with standard deviations of 0.33 and 0.41, respectively. These standard deviation results are obtained through SPSS analysis from level-2 data. The coordinates are:

i. High mothers' education level and positive attitude towards mathematics (AVE-MHEDUL=0.33, ATTD=0.41); γ (mathematics results) = 492.47 + 1.09(0.41) +1.81(0.33) (0.41) =493.16

ii. High mothers' education level and negative attitude towards mathematics (AVE-MHEDUL=0.33, ATTD=0.41); γ (mathematics results) = 492.47 + 1.09(-0.41) +1.81(0.33) (-0.41) =491.78

iii. Low mothers' education level and positive attitude towards mathematics (AVE-MHEDUL=0.33, ATTD=0.41); γ (mathematics results) = 492.47+ 1.09(0.41) + 1.81(-0.33) (0.41) =492.67

iv. Low mothers' education level and negative attitude towards mathematics (AVE-MHEDUL=0.33, ATTD=0.41); γ (mathematics results) = 492.47+ 1.09(-0.41) + 1.81(-0.33) (-0.41) =492.26

v. Average mothers' education level and positive attitude towards mathematics (AVE-MHEDUL=0, ATTD=0.41); γ (mathematics results) = 492.47 + 1.09(0.41) +1.81(0) (0.41) =492.91

vi. Average mothers' highest education level and negative attitude towards mathematics (AVE-MHEDUL=0, ATTD=0.41); γ (mathematics results) =492.47+ 1.09(-0.41) +1.81(0) (-0.41) =492.02

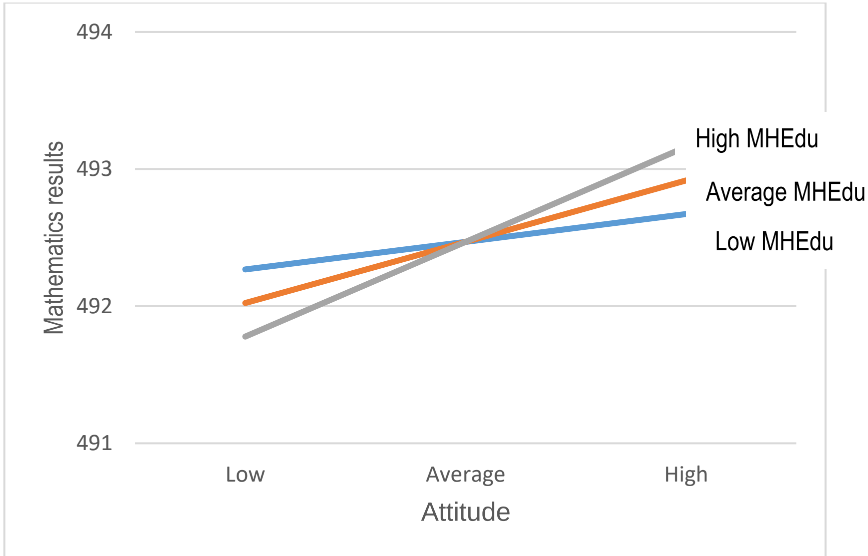


Figure 11. 4 Cross-level interaction effect of students in class on Ave-MHEDUL on the slope of combined students' attitudes towards mathematics.

The results in Table 11.3 indicate that the class average of *mothers' highest education level (Ave-MHEDUL)* interacts with students' *attitude* (ATTD) towards mathematics with an interaction effect coefficient of 1.90. This suggests that *average mothers' highest education level* of the classroom has a positive effect on the slope of students' *attitude* towards mathematics on the outcome variable, mathematics results.

Furthermore, Figure 11.4 shows that *attitude*, on average, has a positive effect on mathematics performance. In classes where the average level of students' mothers' education is high, the effect of attitude on mathematics achievements is stronger. In contrast, for students in classes where the average level of mother's education is low, the effect of attitude on mathematics achievements is weaker.

11.8.5 Variance explained for the two-level combined model

Table 11.4 provides important information on the variance component in the model. Examination of the variance component assesses the proportion of variance explained in the two levels (students and teachers). As such, Table 11.4 shows the estimation of variance components for the outcome variable (mathematics results) with most of the variance (71.2 %) attributable to the students at level-1, with 28.8% of the variance attributable to teachers at level-2 in the null model. Compared to the null model, the final model, which includes the level-1 and level-2 predictors for mathematics results, explains 7.5% of the variance at the student level and 42.6% at the teacher level. These results reflect the notion that the estimated variance component of the final model is reduced as a result of the inclusion of predictors at the two levels.

Table 11. 4 Estimation of variance components for the final two-level model for combined students

Estimation of variance components		
Model	Between students (n=729)	Between teachers (n=41)
Null model	30.95	12.50
Full model	28.62	7.18
Variance available at each level		
	Between students	30.95/(30.95 + 12.50)= 0.7123=71.2%
	Between teachers	12.50/ (30.95 + 12.50) =0.2877=28.8%
Proportion of variance explained by final model		
	Between students	(30.95-28.62)/30.95= 0.0752 =7.5%
	Between teachers	(12.50-7.18)/12.50= 0.4256=42.6%
Proportion of total available variance explained by final model		
(0.7123 x 0.0752) + (0.2877 x 0.4256) = **0.1760= 17.6%**		

In total, 17.6% of the variable available from both levels has been explained by the final model at the two levels. This indicates that the final model may lack some potential factors that are necessary to explain and predict the outcome variable (mathematics results). This finding in turn suggests that the final model can be improved through future research in PNG by adding more potential predictors into the model. To compare this finding, the next section of this chapter analyses separately the results of the Grade 10 students' model.

11.9 The Grade 10 students' model

Table 11.5 (below) shows the HLM unconditional Grade 10 model results. These results indicate that the between-teacher variance exhibits statistical significance (u0j= 20.41, x^2 *(14)* = 176.14, p<0.05). In other words, the mathematics results of the Grade 10 students varied across the teacher group. The ICC 0.3756(20.41/ (33.92+20.41) =0.3756 or 37.6%) indicates that the Grade 10 group has a hierarchical structure, which permits analysis of 2L/HLM. This result demonstrates that 37.6% of the variability in the student mathematics results is attributed to teachers, while the remaining 62.4% is contributed by student characteristics.

Table 11. 5 Fully unconditional model results for the 2L/HLM for Grade 10 student model

Final estimation of fixed effects						
Fixed Effect		Coefficient	Standard Error	T-ratio	DF	Approx. P-value
For INTRCPT1, β_{0j}						
INTRCPT2, β_{00}		497.21	1.16	425.13	14	0.00
Final estimation of variance components						
Random Effect	Reliability	Standard Deviation	Variance Component	DF	Chi-square	P-value
INTRCPT1, u_{0j}	0.92	4.51	20.41	14	176.14	0.00
level-1, rij		5.82	33.92			
Statistics for current covariance components model						
Deviance	2287.33					
Number of estimated parameters		2				

The reliability of 0.92 of the null model shows that the teacher sample mean can estimate well the student outcome (mathematics results). A similar procedure to the one employed in the previous section to analyse the combined student model is used here to evaluate the result of 2L/HLM for Grade 10 students. The student-level variables that had a significant effect on mathematics results in Chapter 10 for Grade 10 students are again used as a starting point for consistency purposes in this multi-level analysis, as shown in Table 11.6:

Table 11. 6 Final results for the two-level model for Grade 10 mathematics results

Grade 10 Final estimation of fixed effects					
Fixed Effect	Coefficient	Standard Error	T-ratio	Approx.D.F	P-value
For INTRCPT1, β_{0j}					
INTRCPT2, γ_{00}	482.99	3.23	149.54	13	0.00
Ave_FatherOCC, γ_{01}	18.42	4.68	3.94	13	0.00
For MotherHEDUL slope, β_{1j}					
INTRCPT2, γ_{10}	1.04	0.40	2.57	14	0.02
For Private slope, β_{2j}					
INTRCPT2, γ_{20}	4.44	1.47	3.01	348	0.00
For ATTD slope, β_{3j}					
INTRCPT2, γ_{30}	2.19	0.25	8.51	13	0.00*
Ave_MotherHEDL, γ_{31}	3.27	0.73	4.45	13	0.00*

Final estimation of variance components						
Random Effect Deviation	Reliability	Standard Deviation	Variance Components	df	Chi-square	P-value
INTRCPT1, U0j	0.82	3.00	8.41	13	63.27	0.00
MHEDUL slope,U1j	0.12	0.63	0.39	14	16.11	0.30
ATTD Slope, U3j	0.04	0.27	0.07	13	7.65	>.500
level-1, Rij		5.38	29.03			

Statistics for current covariance components model	
Deviance	2209.96
Number of estimated parameters	7

**Cross-level interaction effects (P<0.05)*

The results in Table 11.6 of the final 2L/HLM model for Grade 10 students can be specified by the equation below.

Level-1 model

$$Y_{ij}=\beta_0 + \beta_{1j}(MotherHEDUL) + \beta_{2j}(Private) + \beta_{3j}(ATTD) + r_{ij} \quad [11.9]$$

Level-2 model

$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{FatherOCC}) + u_{0j}$ [11.10a]

$\beta_{1j} = \gamma_{10} + u_{1j}$ [11.10b]

$\beta_{2j} = \gamma_{20} + u_{2j}$ [11.10c]

$\beta_{2j} = \gamma_{30} + \gamma_{31}(\text{Ave-MotherHEDUL}) + u_{3j}$ [11.10d]

The final model is represented by the equation resulting from substituting Equations 11.10a to 11.10d into Equation 11.9:

$$\gamma_{ij} = \gamma_{00} + \gamma_{01}(\text{FatherOCC}) + \gamma_{10}(\text{MotherHEDUL}) + \gamma_{20}(\text{Private}) + \gamma_{30}(\text{ATTD})$$
$$+ \gamma_{31}(\text{ATTD})(\text{Ave-MotherHEDUL}) + u_{0j} + u_{1j}(\text{MotherHEDUL}) + u_{2j}(\text{Private})$$
$$+ u_{3j}(\text{ATTD}) + r_{ij} \qquad [11.11]$$

The equation above indicates that Grade 10 students' mathematics results are defined as function of the overall intercept (*γ_{00}*), four direct or main effects, one cross-level interaction effects and a random error *[u_{0j}* $+u_{1j}(\text{MotherHEDUL}) + u_{2j}(\text{Private}) + u_{3j}(\text{ATTD}) + r_{ij}$].

Figure 11.5 reflects the results presented in Table 11.6 and Equation 11.11 above with a diagram that illustrates the direct and cross-interaction effects of variables from levels-1 and 2 on the Grade 10 students' mathematics results.

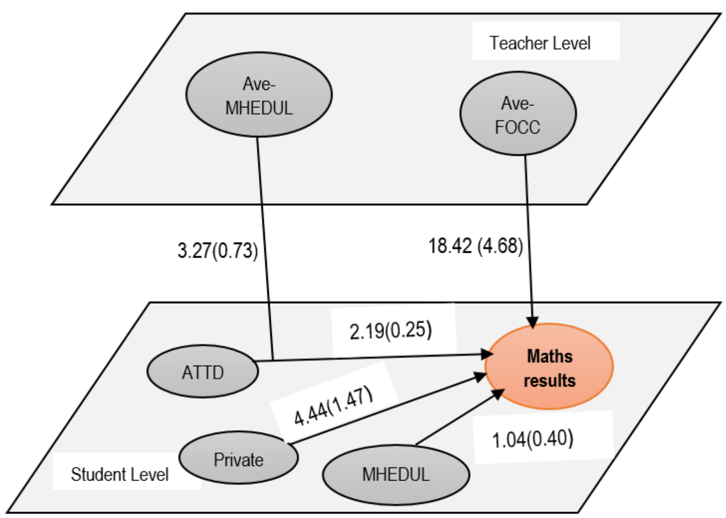


Figure 11. 5 Two-level model of students' mathematics results for Grade 10 sample

11.9.1 The direct effect of level -1 and level-2 predictors on mathematics results

Table 11. 6 shows that *private school* (4.44) at level-1 has a significant ($p<0.05$) direct effect on Grade 10 students' mathematics results, which is consistent with the SEM/path analysis findings in Chapter 10. This suggests that students enrolled at private schools perform better in mathematics compared to those enrolled at Catholic and public schools. This may be because private schools are more likely to provide a conducive learning environment than the public and Catholic schools as stated in the literature. Similarly, *MHEDUL* (1.04) has a significant ($p<0.05$) effect on Grade 10 students' mathematics results. This implies that mothers' highest education levels positively influence students' mathematics results. The higher the mothers' education level, the better the students perform.

Furthermore, *ATTD* (2.19) has a direct significant ($p<0.05$) effect on Grade 10 students' mathematics results. This indicates that Grade 10 students who have more positive attitudes towards mathematics perform better in the subject. These findings of the level-1 variables are consistent with the combined, Grades 10 and 12 models of the path analysis results in Chapter 10. Moreover, at level-2, *Ave-FOCC* (18.42) has a direct significant ($p<0.05$) effect on Grade 10 students' mathematics results. This suggests that students in classrooms that, on average, have fathers' who are employed perform better in mathematics. Similarly, *Ave-MHEDUL* (3.27) has a significant ($p<0.05$) effect on Grade 10 students' mathematics results. This indicates that students in a class with higher average mothers' highest education levels have better mathematics results.

11.9.2 Cross-interaction effects

As discussed earlier three variables are required for the cross-level interaction. These variables are from level-1, level-2 and the outcome variable. As shown in the HLM results in Figure 11.5 (p. 285), the variables

Ave-MHEDUL from level-2 (teacher-level) and *attitude* from level-1 (student level) have a cross-level interaction that has an effect on Grade 10 students' mathematics results.

As such, part of the equation for the final model of the Grade 10 students' data can be drawn to show this cross-level interaction effect by setting the remaining terms to zero. The equation is as follows:

$$\gamma_{ij} = \gamma_{00} + \gamma_{30}(ATTD) + \gamma_{31}(ATTD)(Ave\text{-}MHEDUL) + r_{ij}$$

where $\gamma_{00} = 482.99$ (average mathematics results), $\gamma_{30} = 2.19$ and $\gamma_{31} = 3.27$

Thus, the cross-level interaction equation is

$$\gamma_{ij} = 481.99 + 2.19(ATTD) + 3.27(Ave\text{-}MHEDUL)(ATTD) + r_{ij}$$

Similar conditions employed in the combined student model are used to calculate teacher-level coordinates to obtain a graphical representation of the cross-level interaction effects in the Grade 10 student model, for consistency purposes. These conditions are:

a. One standard deviation above the average on ATTD and AVE-MHEDUL
b. One standard deviation above the average on ATTD and one standard deviation below the average on AVE-MHEDUL
c. One standard deviation below the average on ATTD and one standard deviation above the average on AVE-MHEDUL
d. One standard deviation below the average on ATTD and one standard deviation below the average on AVE-MHEDUL
e. Average on ATTD and one standard deviation above the average on AVE-MHEDUL
f. Average on ATTD and one standard deviation below the average on AVE-MHEDUL

The above guide is used to calculate the coordinates using the AVE-MHEDUL and ATTD standard deviations of 0.36 and 0.34, respectively. These standard deviation results are obtained through SPSS analysis from level-2 data. The coordinates are:

I. High mothers' education level and positive attitude towards mathematics (AVE-MHEDUL=0.36, ATTD=0.34); γ (mathematics results) =482.99 + 2.19(0.34) +3.27(0.36) (0.34) =484.13
II. High mothers' education level and negative attitude towards mathematics (AVE-MHEDUL=0.36, ATTD=0.34); γ (mathematics results) = 482.99 + 2.19(-0.34) +3.27(0.36) (-0.34) =481.84
III. Low mothers' education level and positive attitude towards mathematics AVE-MHEDUL=0.36, ATTD=0.34); γ (mathematics results) = 482.99+ 2.19(0.34) + 3.27(-0.36) (0.34) =483.33
IV. Low mother's education level and negative attitude towards mathematics (AVE-MHEDUL=0.36, ATTD=0.34); γ (mathematics results) = 482.99 + 2.19(-0.34) + 3.27(-0.36) (-0.34) =482.64
V. Average high mothers' education level and positive attitude towards mathematics (AVE-MHEDUL=0.36, ATTD=0.34); γ (mathematics results) = 482.99+ 2.19(0.34) +3.27(0) (0.34) =483.73
VI. Average mothers' education level and negative attitude towards mathematics (AVE-MHEDUL=0.36, ATTD=0.34); γ (mathematics results) =482.99 + 2.19(-0.34) +3.27(0) (-0.34) =482.24

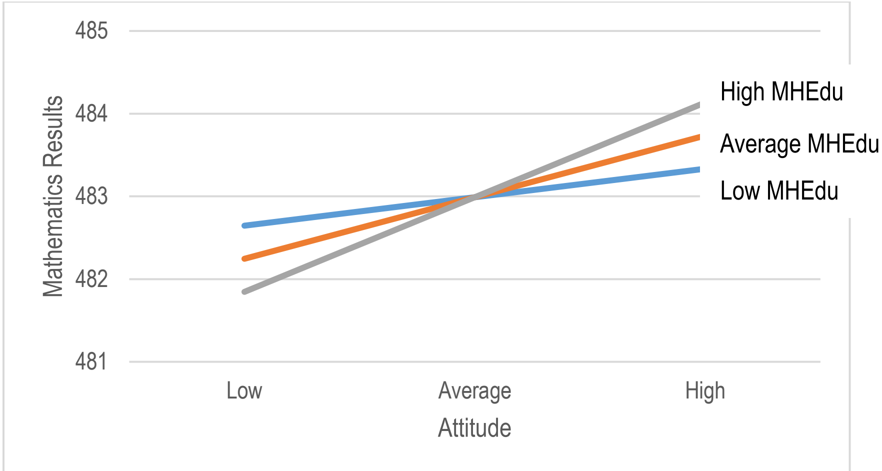


Figure 11. 6 Cross-level interaction effect of students in the class on Ave-MHEDUL on the slope of Grade 10 students' attitude towards mathematics

The results in Table 11.6 indicate that *average mothers' highest education levels* (Ave-MHEDUL) at level-2 interacts with students' *attitude* (ATTD) towards mathematics at level-1, with an interaction effect coefficient of 2.16. This means that the *average mother's highest education level* in a given classroom has a positive effect on the slope of students' attitudes towards mathematics on the outcome variable, mathematics results.

Moreover, Figure 11.6 shows that *attitude*, on average, has a positive effect on mathematics performance. In classes where the students' average level of mothers' education is high, the effect of attitude on mathematics achievement is stronger. In contrast, for students in classes where the average level of mother's education is low, the effect of attitude on maths achievement is weaker. These findings are similar to the combined student model.

11.9.3 Variance explained for the two-level Grade 10 model

As mentioned earlier, examination of the variance component assesses the proportion of variance explained at the two levels for the outcome variable (mathematics results). Table 11.7 (p. 290) shows that 62.4% of the variance is attributable to the students' responses at level-1, and 37.6% of the variance is attributable to teachers at level-2. Compared to the null model, the final model, which includes the level-1 and level-2 predictors for mathematics results, explains 14.4% of the variance at the student level and 58.8% at the teacher level. These results reflect the notion that the estimated variance component of the final model is reduced as a result of the inclusion of predictors at the two levels.

Table 11. 7 Estimation of variance components for the final two-level model for Grade 10 students

Estimation of variance components		
Model	Between students (n=729)	Between teachers (n=41)
Null model	33.92	20.41
Full model	29.03	8.41
Variance available at each level		
	Between students	33.92/(33.92 + 20.41)= 0.6243=62.4%
	Between teachers	20.41/(33.92 + 20.41)=0.3756=37.6%
Proportion of variance explained by final model		
	Between students	(33.92 - 29.03)/33.92= 0.1441 =14.4%
	Between teachers	(20.41–8.41)/20.41= 0.5879 =58.8%
Proportion of total available variance explained by final model		
(0.6243 x 0.1441) + (0.3756 x 0.5879) = **0.3107= 31.1%**		

In total, 31.1% of the initial variance has been explained by the final model at the two levels, taking into account the amount of variance explained by the final model at each level and the amount of available variance to be explained at that level. This result indicates that the final Grade 10 model involves factors that explain and predict well the students' mathematics results. The last section of this chapter discusses the results of the Grade 12 students' model using similar statistical techniques employed in the combined and Grade 10 students' models.

11.10 The Grade 12 students results

Table 11.8 (below) shows the HLM unconditional model results for the Grade 12 students' data. These results indicate that the between-teacher variance exhibits statistical significance (u0j= 6.55, x^2 *(25)* = 112.79, p<0.00). In other words, the mathematics results of the Grade 12 students varied across the teacher group. The ICC value of 0.1895 (6.55/ (6.55+28.01) =0.1895 or 18.9%) indicates that the Grade 12 group has a hierarchical structure and guarantees the analysis of 2L/HLM. This result shows that 18.9% of the variability in the students' mathematics results is attributable to teachers, while the remaining 81.1 % is contributed by student characteristics.

Table 11. 8 Null model results for the two-level model for Grade 12 mathematics results

Final estimation of fixed effects						
Fixed Effect		Coefficient	Standard Error	T-ratio	DF	Approx. P-value
For INTRCPT1, β_{0j}						
INTRCPT2, β_{00}		494.91	0.57	866.05	25	0.00
Final estimation of variance components						
Random Effect	Reliability	Standard Deviation	Variance Component	DF	Chi-square	P-value
INTRCPT1, u_{0j}	0.74	2.56	6.55	25	112.79	0.00
level-1, rij		5.29	28.01			
Statistics for current covariance components model						
Deviance	2349.06					
Number of estimated parameters		2				

The reliability of 0.74 of the null model indicates that the teachers' sample mean can estimate well the student outcome (mathematics results). These results of the null model show the presence of variance at the teacher and student-levels, and confirm the viability of multilevel analysis. A similar procedure to that used to analyse the combined student and Grade 10 models in the previous sections is now used to evaluate the results of 2l/HLM for the Grade 12 students. The final results of this analysis are presented below in Table 11.9.

Table 11. 9 The final model results for the two-level model for Grade 12 mathematics results

Final estimation of fixed effects						
Fixed Effect		Coefficient	Standard Error	T-ratio	Approx D.F	P-value
For INTRCPT1, β_{0j}						
INTRCPT2, γ_{00}		495.39	0.75	654.27	25	0.00
For Gender slope, β_{1j}						
INTRCPT2, γ_{10}		-2.32	0.59	-1.83	24	0.04
Ave_MHEDUL, γ_{21}		1.44	0.60	2.43	24	0.02
For FOCC slope, β_{2j}						
INTRCPT2, γ_{20}		1.05	0.56	1.86	25	0.04
Final estimation of variance components						
Random effect	Reliability	Standard	Variance	df	Chi-square	P-value
Deviation	Component	Deviation	Components			
INTRCPT1, U0j	0.73	3.36	5.64	17	78.38	0.00
Gender slope, U1j	0.31	1.88	3.53	16	22.96	0.11
FOCC slope, U2j	0.02	0.45	0.20	17	18.48	0.35
level-1, Rij		5.18	26.86			
Statistics for current covariance component model						
Deviance				2326.77		
Number of estimated parameters				7		

**Cross-level interaction effect (P<0.1)*

Based on the results in Tables 11.9, the final 2L/HLM model for Grade 12 students can be specified by the equation below.

Level-1 model:

$$Y_{ij}=\beta_0 + \beta_{1j}(\text{Gender}) + \beta_{2j}(\text{FatherOCC}) + r_{ij} \qquad [11.12]$$

Level-2 model

$$\beta_{0j}= \gamma_{00} + u_{0j} \qquad [11.13a]$$
$$\beta_{1j}= \gamma_{10}+ \gamma_{11} (\text{MHEDUL}) +u_{1j} \qquad [11.13b]$$
$$\beta_{2j}= \gamma_{20} + u_{2j} \qquad [11.13c]$$

The final model is represented by the equation resulting from substituting Equations 11.13a to 11.13c into Equation 11.12.

$\gamma_{ij} = \gamma_{00} + \gamma_{10}(Gender) + \gamma_{20}(FOCC) + \gamma_{21}(Ave\text{-}MHEDUL)(Gender) + u_{0j} + u_{1j}(Gender) + u_{2j}(FOCC) + r_{ij}$

[11.14]

The equation above indicates that the Grade 12 students' mathematics results are defined as the function of the overall intercept (*γ_{00}*), two direct or main effects, one cross-level interaction effect and a random error [$u_{0j} + u_{1j}(Gender) + u_{2j}(FOCC) + r_{ij}$].

11.10.1 The direct effect of level -1 and level-2 predictors on mathematics results

At level-1, Table 11.9 (p. 293) shows that *gender* (-2,32) has a significant ($p<0.05$) negative direct effect on mathematics results. This indicates that female Grade 12 students perform better than their male counterparts in mathematics. Moreover, *FatherOCC* (1.05) has a significant ($p<0.05$) positive effect on the Grade 12 students' mathematics results, suggesting that students whose fathers are employed perform better in mathematics, possibly because their fathers are more educated and are therefore more likely to teach mathematics concepts or assist with their homework. The results in Table 11.9 show that, at level-2, *Ave-MHEDUL* has a direct significant ($p<0.05$) effect on Grade 12 students' mathematics results. This indicates that students in the classroom who have on average more highly educated mothers perform better in mathematics.

Figure 11.7 reflects the results presented in Table 11.9 and Equation 11.14, illustrating the direct and cross-interaction effects of variables from levels-1 and 2 on the Grade 12 students' mathematics results.

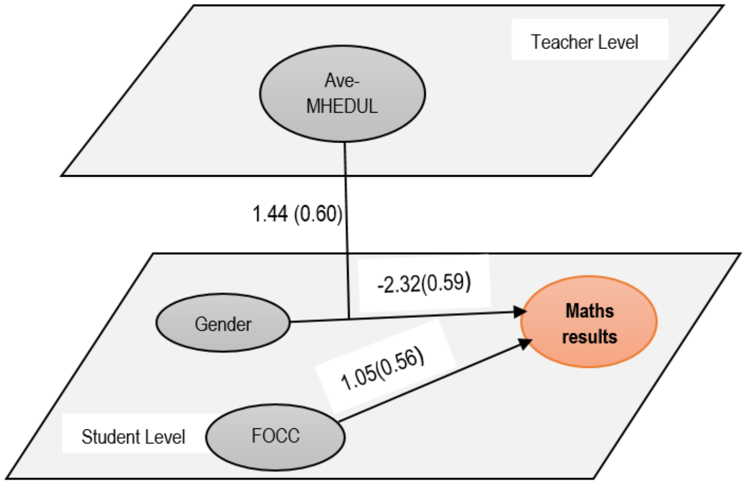


Figure 11. 7 Two-level model of students' mathematics results for the Grade 12 student model

11.10.2 Cross-interaction effects

As shown in the HLM results in Figure 11.7, the variable *Ave-MHEDUL* from level-2 and *gender* from level-1 have a cross level interaction. As such, part of the equation for the final model for the Grade 12 students' data can be drawn to show this cross-level interaction effect by setting the remaining terms to 0. The equation is as follows:

$$\gamma_{ij} = \gamma_{00} + + \gamma_{20}(\text{Gender}) + \gamma_{21}(\text{Ave-MHEDUL})(\text{Gender}) + r_{ij}$$

where γ_{00} =495.39 (average mathematics results), γ_{20} =-1.08 and γ_{21} =1.44

Thus, the cross-level interaction equation is

$$\gamma_{ij} = 495.39 - 1.38(\text{Gender}) + 1.44(\text{Ave-MHEDUL})(\text{Gender}) + r_{ij}$$

This equation can be used to calculate teacher-level coordinates to obtain a graphical representation of the cross-level interaction effect, similar to the previous two models. The conditions used to calculate the coordinates for teachers used by Ben (2010, p. 389) and Lye (2016, p. 305) are adapted in this study. These conditions are:

a. One standard deviation above the average on Gender and AVE-MHEDUL

b. One standard deviation above the average on Gender and one standard deviation below the average on AVE-MHEDUL
c. One standard deviation below the average on Gender and one standard deviation above the average on AVE-MHEDUL
d. One standard deviation below the average on Gender and one standard deviation below the average on AVE-MHEDUL
e. Average on Gender and one standard deviation above the average on AVE-MHEDUL
f. Average on Gender and one standard deviation below the average on AVE-MHEDUL

The above guide is used to calculate the coordinates using the AVE-MHEDUL and Gender standard deviations of (0.33, 0.49), respectively. These standard deviation results are obtained through SPSS analysis from level-2 data. The coordinates are:

I. High mothers' education level and high female students (AVE-MHEDUL=0.33, Gender=0.49); γ (mathematics results) =495.93 -1.08(-0.49) +1.44 (0.33) (-0.49) =495.70
II. High mothers' education level and low female students (AVE-MHEDUL=0.33, Gender=0.49); γ (mathematics results) = 495.93 – 1.08(-0.49) +1.44(-0.33) (-0.49) =496.15
III. Low mothers' education level and high male students (AVE-MHEDUL=0.33, Gender=0.49); γ (mathematics results) = 495.93 -1.08(0.49) + 1.44(0.33) (0.49) =494.07
IV. Low mothers' education level and low male students (AVE-MHEDUL=0.33, Gender=0.49); γ (mathematics results) = 495.93 -1.08(0.49) + 1.44(-0.33) (0.49) =496.62
V. Average high mothers' education level and gender (AVE-MHEDUL=0.33, Gender=0.49); γ (mathematics results) = 495.93 -1.08(0.49) +1.44(0) (0.49) =495.40
VI. Average mothers' education level and gender (AVE-MHEDUL=0.33, Gender=0.49); γ (mathematics results) =495.93 - 1.08(-0.49) +1.44(0) (-0.49) =496.45

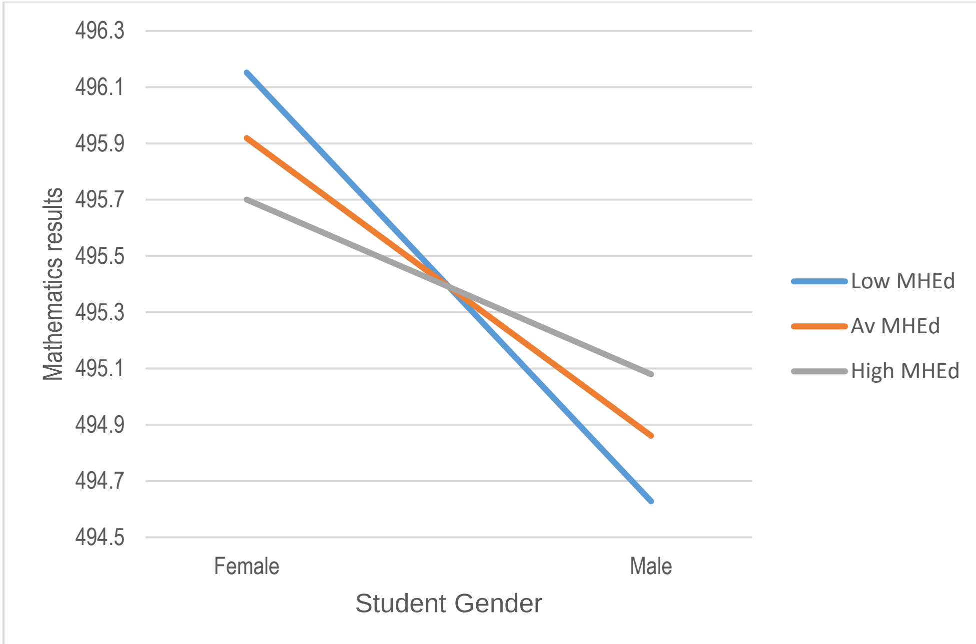


Figure 11. 8 Cross-level interaction effect of students in class on Ave-MHEDUL on the slope of Grade 12 students' gender

The results in Table 11.9 (p. 292) show that average *mothers' highest education level (Ave-MHEDUL)* from level-2 interacts with students' *gender*, with an interaction effect coefficient of -2.32. Moreover, Figure 11.8 shows that *gender*, on average, has a negative effect on mathematics performance. In other words, female students do better in mathematics. The difference is bigger in classrooms where, on average, mothers' highest education levels are low. In classrooms where mothers' highest education levels are high, the gender difference is smaller. These findings reflect the cultural aspects of PNG, where girls are more attached to their mothers compared to boys, and want to be like their mothers (Spark, 2010). Instead, boys are attached to the fathers and see them as their role models.

11.10.3 Variance explained for the two-level Grade 12 model

Table 11.10 shows that 81.0 % of the variance is attributable to the students' responses at level-1 and 18.9% of the variance is attributable to teachers at level-2. Compared to the null model, the final model, which includes the level-1 and level-2 predictors for mathematics results, explains 4.1% of the variance at the student level and 13.9% at the teacher level, respectively. These results reflect the notion that the estimated variance component of the final model is reduced as a result of the inclusion of predictors at the two levels.

Table 11. 10 Estimation of variance components for the final two-level model for Grade 12 students

Estimation of variance components		
Model	Between students (n=729)	Between teachers (n=41)
Null model	28.01	6.55
Full model	26.86	5.64
Variance at each level		
	Between students	28.01/(28.01+ 6.55)= 0.8104= 81.0%
	Between teachers	6.55/(28.01 + 6.55)=0.1895= 18.9%
Proportion of variance explained by final model		
	Between students	(28.01 – 26.86)/28.01= 0.041=4.1%
	Between teachers	(6.55 – 5.64)/6.55= 0.1389=13.9%
Proportion of total available variance explained by final model		
(0.8104 x 0.041) + (0.1895 x 0.1389) = **0.0595=5.9%**		

In total, 5.9% of the total variance available has been explained by the final Grade 12 model at the two levels. The final model therefore suggests that other factors are not covered to fully explain and predict the outcome variable (mathematics results), which indicates that the final Grade 12 model can be improved and addressed through relevant future research in PNG.

11.11 Summary

This chapter presented the two-level hierarchical linear modelling (HLM) analysis of the combined, Grade 10 and Grade 12 students' mathematics results. The HLM approach was employed to examine the direct effect and cross-interaction effects of the hierarchical structure on students' mathematics results. In order to carry out the HLM analysis, the W-scores discussed in Chapters 3 were used in the analysis of all constructs First, this chapter discussed the null model for each individual model, followed by separate discussion of each initial and final full model of the three groups of students, consistent with the presentation of SEM results in Chapter 10. These models were comprised of the student and teacher-level constructs and demographic variables.

The results of analysis of the final combined student model showed that at level-1, *mothers' highest education-level (FHEDUL) and attitude (ATTD)* had a positive direct significant ($p<0.05$) influence on mathematics results. These results are also consistent with the single level SEM/path analysis in Chapter 10. At level-2, *mathematics major* and students in the class on average fathers' who are employed (Ave-FOCC) show a positive direct significant ($p<0.05$) effect on students' mathematics results. The other finding is that students in class on average *mother's highest education level (Ave-MHEDUL),* a level-2 compositional variable, was found to be interacting with students' *attitude (ATTD)* at level 1, which has an influence on mathematics results. The average mother's *highest education level (Ave-MHEDUL)* of the classroom has an interaction effect with students' *attitude* (ATTD) towards mathematics results. This suggests that the *average mothers' highest education level* of the classroom has a positive effect on the slope of students' *attitude* towards mathematics on the outcome variable, mathematics results.

For the second model, of Grade 10 students' results demonstrated that at level-1, *mothers' highest education-level (FHEDUL)* and *attitude (ATTD)* have a significant ($p<0.05$) direct effect on Grade 10 students' mathematics results. These findings are consistent with the results of SEM and the combined student model discussed earlier. Similarly, *private schools* have a significant ($p<0.05$) direct effect on Grade 10 students' mathematics results, which is consistent with the SEM results in Chapter 10. Furthermore, at level-2, compositional variables such as students in class that on average have *fathers who are employed (Ave-FOCC)* and *mothers' highest education level (Ave-MHEDUL)* have a direct significant ($p<0.05$) effect on Grade 10 students' *mathematics results.* Another finding is that Grade 10 students' *average mothers'* *highest education level (Ave-MHEDUL)* at the level-2 variable was found to be interacting with students' *attitude (ATTD)* at level-1, which leads to an effect on mathematics results. The average mothers' *highest education level (Ave-MHEDUL)* at level-2 is interacting with the students' *attitude (ATTD) at level-1* to positively influence mathematics results. This finding is consistent with the results for combined student model.

The Grade 12 students' final model showed that *fathers' occupation (FOCC) at level-1 has a significant* ($p<0.05$) effect on their mathematics results. This finding implies that students with fathers who are employed are more likely to perform better in mathematics. Additionally, this result also demonstrated that HLM analysis is a more refined technique than path analysis, as the variance of the variables is partitioned to student and teacher levels, with path analysis being the explanatory stage of analysis to obtain an overview of the data.

Additionally, the model also indicates that average *mothers' highest education level* (Ave-MHEDUL) at the level-2 variable is found to be interacting with gender at level-1, which leads to an effect on the outcome variable, mathematics results. The average *mothers' highest education level (Ave-MHEDUL)* on the slope of *gender* on mathematics results *is* a negative one. This suggests that the average *mothers' highest education level* has an effect on the rise of the difference between males and females' mathematics results. In other words, the performance gap in mathematics between male and female students grows bigger as the average mothers' highest education level of the classroom becomes lower.

In addition, the results of the combined, Grade 10, and Grade 12 student model variance components indicate that the final full models explain 17.6%, 31.1% and 5.9% of the total variances of the three models, respectively. These findings indicate that the Grade 10 student model explained more than a quarter of the total variance. On the other hand, the combined students' model explained more than 16% of the total variance, while the Grade 12 student model explained slightly more than 5% of the total variance. Based on these results, the Grade 10 students' final model seems to cover constructs and variables that explain and predict the outcome variables (mathematics results), while the combined and Grade 12 student models cover less of the variables that predict mathematics results. These findings can be improved and addressed through relevant future research in PNG.

Overall, the findings discussed in this chapter show that the HLM results provide a useful and meaningful interpretation for the understanding of the effects of the predictors from the two levels (student and teacher-levels) on the outcome variable (mathematics results). However, to obtain a deeper understanding of the issues at the teacher-level that are affecting students' mathematics results, the next chapter turns to analyse qualitative (teacher interview) data.

Chapter 12: Teacher-level interview analysis

12.1 Introduction

In this chapter, the interview data of teachers are analysed to identify the common themes emerging from the teacher group. Interviews were only for the teacher-level to obtain perspectives on educators' teaching strategies and content knowledge that might affect their students' mathematics achievements. As discussed in Chapter 3, using a single data collection technique such as the survey questionnaires used in this study (a quantitative approach), can alone never adequately shed light on a phenomenon. Supplementing this method by using interviews (a qualitative approach) can therefore assist in providing a deeper understanding of the factors affecting mathematics results for Grade 10 and 12 students. These two research methods used in this study overlap, complement, and at times contradict each other, which has the overall effect of balancing out each method and providing a richer and better understanding of the results (Bamberger, 2012; Cresswell, 2008; Onwuegbuzie & Combs, 2011; Sandelowski, 2000). In this Chapter, the triangulation of quantitative (surveys) and qualitative (interviews) results are carried out.

As the literature indicates that students' mathematics achievements are affected by teacher-level factors, it is important to discuss the issues that are identified in the interviews with teachers. The results of SEM analysis in Chapter 10 indicated that *quality teaching* and *instructional resources* at teacher-level show significant result except *classroom environment* that indicated insignificant result. Therefore, interview analysis can capture information that might have been missed in teacher surveys, and can facilitate an in-depth exploration of the attributes and factors at the teacher-level that are affecting Grade 10 and 12 students' mathematics results. This chapter will answer the research question "What are the attributes that are deemed important by teachers regarding their mathematics teaching approaches and content knowledge?" In order to answer this question, several main themes are identified and discussed, including teaching methods, content knowledge, outcome of teaching methods, and encouragement towards mathematics, alongside associated sub-themes. In analysing the interviews, the main themes are first discussed, followed by discussion of participants' responses under each theme. After that, the discussion of the responses are presented, according to the main themes and sub-themes identified. This approach provides a more exploratory perspective on the statistical results presented in Chapter 11, to provide balanced findings at the teacher-level through triangulation.

12.2 Thematic analysis approach

Thematic analysis is the most common analysis approach used in qualitative research. This approach emphasises pinpointing, examining and recording patterns (themes) within data. Themes are patterns across data sets that become the categories for analysis, and are important in describing a phenomenon associated with a specific research question. In this approach, themes are used to capture the essence and spread of meaning; they unite data that might otherwise appear disparate, and correct meanings that occur in multiple and varied contexts. "Thematic analysis can be an essentialist or realist method, which reports experiences, meanings and the reality of participants" (Braun & Clarke, 2006, p. 81). Therefore, thematic analysis works both to reflect reality and to unpick or unravel the surface of 'reality'.

In this study, thematic analysis is performed following the six processes of coding phases outlined by Brinkmahnn and Kvale (2015) and Braun & Clarke, (2006) to create established and meaningful patterns. The first step is familiarisation with the data to sort out ideas through transcribing, reading and re-reading. Second, codes are generated in a systematic approach across the entire data set, in order to collate data that are relevant to each code. Third, themes are identified for coding, and to gather the data for each relevant and potential theme. Fourth, these themes are reviewed to ascertain that they work in relation to the coded extracts and the entire data set, to generate a thematic ‘map’ of the analysis. After that, the themes are defined and named to tell a clearer story of the data. Finally, a scholarly report of the analysis is produced that relates back to the research question and literature. These six steps to analyse the quantitative data using the thematic approach were organised and expedited through use of the NVivo 12 software.

12.3 NVivo 12

NVivo 12 is used to analyse the interview data in this chapter. NVivo is a data management tool (Hart & Achterman, 2017), that organises and assists in making sense of data during analysis (Hamrouni & Akkari, 2012; Hart & Achterman, 2017). NVivo organises, stores and retrieves data more efficiently than manual methods, saves time, and helps to rigorously back up findings with evidence (Hamrouni & Akkari, 2012). The data were imported from a text file and analysed with NVivo’s visualisation tools. The software allows the researcher to classify, sort and arrange information; examine relationships in the data; and combine analysis with linking, shaping, searching and modelling (Hamrouni & Akkari, 2012; Hart & Achterman, 2017). The researcher can test theories, identify trends and cross-examine information in a multitude of ways using the software’s search engine and query functions.

There were several steps involved in analysing the interview data with NVivo 12. First, the teachers’ transcribed interviews were loaded into NVivo. After that, themes and sub-themes from the scripts were identified and created. Next, the teacher transcripts were sorted into the appropriate themes and sub-themes and, a mind map was formulated through the NVivo software to create relationships among the themes. This process created a frequency of the number of teachers under each of the themes, making it easier to trace the teachers’ responses on the themes. The themes were identified from the NSW quality of teaching model discussed in the literature review of Chapter 2. For instance, in this study, four themes were identified, and the software assisted in putting together all the responses of the participants according to their respective themes, with the percentages of responses and the frequency of respondents listed under each theme. This organisational approach of the software requires less time and allows more flexibility than manual methods in comparing interview responses and analysis of data.

12.4 Emerging themes

This section highlights each of the main themes and their sub-themes that were identified in the data by frequency analysis of the teacher-level interviews. The four key themes that emerged from the data analysis are: (1) teaching methods (student and teacher centred methods), (2) content knowledge, (3) outcome of the teaching methods used in the classroom, and (4) encouragement towards mathematics (practice exercise and students' attitudes/approaches). All of these themes seem to have an influence on the mathematics outcomes of students. The teaching quality of teachers was measured in the survey questionnaires (discussed in Chapter 10) that scale *quality of teaching* (teaching methods and content knowledge). This scales' confirmatory factor analysis results (the quantitative method) in Chapter 9 indicated that the 15 instruments have a relationship with the scale, and the scores were used to predict the outcome of the students' mathematics results. However, the scale's SEM analysis results in Chapter 10 indicated that the quality of teaching (teaching methods and content knowledge) of teachers has a significant effect on the students' mathematics results. Therefore, these interview results related to the teaching methods and content knowledge of teachers will shed more light into how *quality of teaching* affect Grade 10 and 12 students' mathematics results. These two themes are in turn also interconnected with the other themes identified above which also have an impact on students' mathematics results.

12.5 Theme 1: Teaching methods

The kind of teaching approach employed by teachers can have an impact on students' mathematics learning in the classroom. This is evident in the mathematics teachers' responses, in which they report using different methods in delivering their mathematics lessons. It was clear from the responses that either traditional (teacher-centred) or student-centred teaching methods are the approaches most often used to teach mathematics at schools in Port Moresby. However, there are a few teachers that apply both teaching methods in the delivery of their lessons. These teaching strategies are now discussed separately to present a deeper understanding of how they affect students' mathematics results.

12.5.1 Traditional (teacher-centred) method

The interview analysis revealed that most teachers use traditional methods to deliver their lessons. In other words, the teachers are verbally explaining the mathematics concepts/ideas using a blackboard and/or textbooks, handouts, worksheets and charts. This suggests that these teachers are more dependent on writing notes from textbooks on the blackboard, verbally explaining the mathematics ideas, and giving handouts to assist students in exercises and activities. One of the teacher participants said:

> *"Ok.I use two basic methods. A) Use the normal method using black board. I stand at the front write down the topic introduce topic briefly and use examples on the board. The main method used.*
> *B) Issuing Textbooks and handout. They have the resource with them write topic, give the example there on the board, and tell them you have the example in the textbook or handout you have. I go through the example briefly with the students and write one exercise on the board and write questions in the textbook for them on the board to do it as activity for that period. They give them instruction only what to do because everything is in their textbook" [Teacher 4]*

Another teacher participant highlighted that charts and visual aids assist to verbally explain the main points with examples in the lessons, and also noted that they give exercises to the students derived from these resources. This implies that these teachers are not providing guidance to students individually, but instead are using a lecture approach similar to a higher education setting. The teacher participant highlighted that:

> *"When I go to class, I use chart or visual aid and stick them on the black board. Put up the main points and explain to the students. After discussing the main points, I explain and go through the examples. Then I tell the students to do the activities. It's like lecture type" [Teacher 10].*

Interestingly, another teacher participant mentioned the specific step-by-step strategies they employed in the classroom, and identified their own approach as a teacher-centred method. This method includes giving out handouts to students, with verbal explanations carried out on the blackboard. However, the effectiveness of the handout method was not explained in detail by the participant. The teacher participant said:

> *"Ok ah. Firstly, I provide notes in the form of handout, all the explanations esp. examples and exercise are on the handout. While they are looking through the handout I explain to them. When I am explaining I go through the same examples in the handout on the blackboard because this is mathematics and sometimes they might not understand what they are reading. So, I have to do it on board by writing them again, explain the maths problems step by step. For them to understand what they see on the paper. The normal and usual teaching method. I always provide handouts and teach accordingly" [Teacher 14].*

Another teacher participant remarked that the explanation of details with examples assists students to understand and better practice mathematics exercises. Within this method, more practice exercises are therefore encouraged for students to better understand mathematics ideas. The teacher participant stated that:

> *"Before I give the activities, I explain and I have to go in detail explaining the examples. They have to understand first before I give them exercises to do. First, I give the practice exercise and once they are done, I give them allocated time for these practice exercises. That is to see whether they understand the examples given" [Teacher 7].*

Similar to teacher participant 4, a great deal of dependency on textbooks and work sheets is evident in teacher participant 9's method of teaching. This teacher feels comfortable with providing summarized work sheets to students so that they can follow a set method employed in response to challenges, as there is shortage of textbooks at the school. This particular teacher highlighted that, with this method, lessons become more teacher-centred and students are not actively involved, as they get confused working by themselves. The teacher participant emphasised that:

> *"Just textbook alone. Teaching with textbook is the old passion ah but teachers need to be more prepared ah it's the preparation part that always make teaching of mathematics more interesting. It's the interest ah it how you prepare and present the lesson. So textbooks we may have few textbooks but this days teachers do not turn to produce work unlike before we prepare work and lessons are planned before you present. To have methods to present like worksheets. I believe that worksheets are better because it is actually what teacher takes out from the textbook. Now teachers tend to use textbooks more than producing work. It's like what is prepared in the textbook is what is taught. More like lecture type not more*

of student involvement. I use variety of methods, I use worksheet more to cover, worksheets makes it short cut the process of passing and sometimes there are more complicated words. The students do not understand when they are by themselves, so when we simplify into worksheets they are better presented". [Teacher 9].

Another teacher participant said:

"Like you mean Chalkboard. For that, we using especially with Chalkboard sometimes and textbook. But we do not have plenty textbooks. We just using handouts, duplicating handouts from textbooks. Sometimes we use charts to write our notes on the paper so that students can copy the notes, examples and exercises on charts. We use these methods". [Teacher 2].

One of the reasons teachers rely on using the chalkboard (blackboard) is because there are too few textbooks for each student in the classroom. This forces teachers to duplicate handouts from the only textbook they possess.

12.5.2 Student-centred method

Student-centred teaching is one of the most effective methods used in many classrooms by mathematics teachers. This method is increasingly favoured by teachers in PNG over the traditional, teacher-centred method detailed above. However, only four teachers interviewed for this study emphasised that they use student-centred methods such as group work, peer discussions and presentations. These teachers give group work with specific instructions and organise students according to their ability levels to assist each other in learning, while acting as a facilitator.

Teacher participant 6 focuses on group work in the classroom with a student within the group taking on leadership responsibilities, thereby allowing the students to take responsibly for their own learning and present their findings to the class. Besides these methods, this teacher also promotes students with different ability levels to work together in order for them to learn from each other. These two approaches allow the teacher to facilitate classroom learning while the students take control of their own learning. This teacher participant stated that:

"This method is to give students group work and select someone to be the leader. After the explanations are done, the group leader takes care of the groups. Later they do their presentation. Another method is

they work in pairs instead of a big group. I select students who are fast workers, those who can work with supervisor and I put them with someone who is very weak and slow. In that way the person who can be able to understand more helps or assists and I move around to assist in one way or the other. (Group work and demonstration are the methods)". [Teacher 6].

A similar approach employed by teacher participant 6 is evident in the interview of teacher 3, who told of how the teacher-centred approach is strategised in the delivery of the lessons in their classroom. First, concepts are explained, and then group/peer work is given for students to check their own work and solve problems together. The teacher participant said:

"First one is ah teacher centred kind of lesson that the teacher talks and explains ok. Teacher to students and another strategy is using groups, checking work in groups ha giving problems to each, allocating a problem to each groups and students work in groups. Working in pairs. Identifying students work in pairs solving problems together" [Teacher 3].

Furthermore, teacher 7 reported that, to facilitate group work, they drill students for their speed and accuracy skills in a given time frame, and later further explain details of mathematics problems, as PNG students typically assume that mathematics is a difficult subject. This indicates that students' attitudes towards mathematics can have an effect on their mathematics learning. However, in their classroom, this teacher offers a detailed, clear explanation and facilitation of group work designed to assist students to overcome this challenge. As such, this teacher participant highlighted that:

"Once they understand within the allocated with for instance 10 minutes. After that, I ask them to work in pairs or in groups in order for them to help each other. Sometimes maths is a difficult subject in PNG; there are many students who finds it very hard to understand maths so sometimes in my lesson I try to break the example down to detail to explain for students to understand the concept on how to solve the particular problem. I encourage them to work in groups or pairs so that they can understand from each other if they do not understand from me" [Teacher 7].

Teacher participant 19 pointed out that their teaching method involves:

Mostly to do with demonstrations and explanations. Get students to work in pairs and as well as in groups [Teacher 19].

For teacher 19, explanation and demonstration of ideas are employed first, followed by students being encouraged to work in pairs to explore and understand the concepts.

12.6 Using both teaching methods

Three out of the 19 teachers interviewed use both teaching methods (teacher and student-centred methods) to deliver their mathematics lessons. Interestingly, two of the participants highlighted that:

> *First one is ah teacher centred kind of lesson that the teacher talks and explains ok.Teacher to students and another strategy is using groups, checking work in groups ha giving problems to each, allocating a problem to each groups and students work in groups. Working in pairs. Identifying students work in pairs solving problems together. Ha the other one is teacher to student. One on one assistance but not in class outside of class [Teacher 3].*

> *I give a problem that may be on a paper or on the board. It may be 10 simple questions to do with multiplication, division, word problem or a graph. I give back to them and they must come up with an answer. Tell them to stop after 10 minutes and they have group discussion for 2 minutes. Before the slow kids had problems understanding, plenty work for me but I put them in groups for group talking and someone have the answer to help the others. I work around to hear what they are doing to and assist them [Teacher 18].*

The two teachers above have used both methods in their teaching, with the traditional method employed in the first part of the lesson to explain ideas and procedures, while the student-centred approach is used in the second part to emphasise the importance of solving mathematics problems in groups for better understanding. It is interesting to note how they have planned their lessons in a similar format. These two teachers seem to understand the importance of both methods in the students learning process, and have employed them meaningfully in their lessons.

However, there are also challenges associated with employing these teaching methods effectively. Two participants said:

> *Most basically like our population, current population before it use to be 1:35 students, currently this 21st century specially we have like 1: 65 to 70 and so forth .The previous methods of you know teaching ah some strategies that we used we have use in the past we don't apply them at the present [Teacher 1].*

> *Teaching methods are aids like handouts, textbooks we have limited textbook as I've said earlier and then access to internet and so forth that is a major problem in our school. Which most of the students and then the school we don't have internet access so forth that is a major problem in our school which most of the students then the school we do not have internet access in order for us to teach them those lessons or so forth that can really keep them up to the latest standards [Teacher 2].*

These participants indicated that the increase of class sizes and a lack of learning resources at schools have hindered and affected teaching methods in the classroom. The issue of limited learning resources has been discussed in the literature review and in the quantitative results in Chapter 9 shows that have instructional resources have a significant effect on the students' mathematics results. The teachers' responses about learning resources were also collected through the survey questionnaires. However, the SEM analysis results in Chapter 10 of instructional learning resources showed significant effect on the mathematics results. It is evident in the interview analysis that learning resources promote effective mathematics learning, and consequently improve students' mathematics results. In other words, adequate learning resources with a manageable student population in the classroom promote better learning outcomes.

The next step in analysis is to compare the three teaching methods used by the teachers described above in qualitative analysis through quantitative methods. This comparison is conducted by analysing the teachers who are using each of the respective methods. A simple statistical one-way analysis of variance (ANOVA) test was employed to predict which of the three teaching methods had an influence on the students' mathematics results. The procedure was employed to gauge a clear understanding of the methods that affect mathematics results. The outcomes of this analysis are displayed in the three tables below with their descriptions.

Table 12. 1 Descriptive statistics of maths results for the three different teaching methods

Maths Results					95% Confidence Interval for Mean			
Teaching Methods	N	Mean	SD	Std Error	Lower Bound	Upper Bound	Minimum	Maximum
Teacher Centred	209	494.91	6.12	0.42	494.77	495.74	458.79	514.85
Student Centred	47	500.06	7.75	1.13	497.79	502.33	484.75	519.62
Both Methods	47	500.94	5.48	0.82	499.27	502.60	491.61	515.45
Total	303	496.60	6.79	0.39	495.83	497.37	458.79	519.62

The descriptive statistics associated with the three teaching methods employed by teachers are reported in Table 12.1. It can be seen that the teacher-centred method is associated with results numerically below the mean level (M=494.91), and the student-centred method (M=500.06) and both methods (M=500.94). The total mean is 496.60, which is not too different to the mean of the teacher-centred method.

Table 12. 2 ANOVA analysis result of the overall teaching method.

Maths Results

	Sum of Squares	df	Mean Square	F	Sig
Between Groups	1988.02	2	99401	24.94	0.00
Within Groups	11836.11	297	39.85		
Total	13824.13	299			

Table 12.2 shows the output of the ANOVA analysis, indicating whether there is a statistically significant difference between our group means. The results shown in Table 12.2 indicate that the significance value is 0.000 [F (2,994) = 24.94, p = .000)], which is below 0.05 and, therefore, there is a statistically significant difference in the mean of the three teaching methods used to determine the students' mathematics results.

Table 12. 3 Post hoc tests for the each of the teaching methods for comparison

Dependent Variable: Maths Results

Tukey HSD

(I) Teaching Methods	(J) Teaching Methods	Mean Difference (I-J)	Std.Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Teacher Centred	Student Centred	-5.15	1.02	0.00	-7.55	-2.75
	Both Centred	-6.02	1.04	0.00	-8.49	-3.56
Student Centred	Teacher Centred	5.15	1.02	0.00	2.75	7.55
	Both Centred	-0.87	1.32	0.78	-3.99	2.23
Both Methods	Teacher Centred	6.02	1.04	0.00	3.56	8.49
	Student Centred	0.87	1.32	0.78	-2.25	3.99

It is apparent from the results that there are statistically significant differences between the groups as a whole but the differences between the three teaching methods have not yet been shown. Therefore, the post hoc test of one-way ANOVA shown in Table 12.3 illustrates the multiple comparisons, showing how the teaching methods differed from each other. There is no significance ($p>0.05$) difference between the student-centred teaching method and mixing both teaching methods. However, it is also evident from the mean results from Table 12.1 and post hoc test results from Table 12.3 that the student-centred teaching method and mixing both teaching methods are significantly ($p<0.05$) higher than teacher-centred teaching methods. These results suggest that employing student-centred teaching methods and mixing both teaching methods makes a significant difference compared to using only the teacher-centred teaching method.

12.6.1 Discussion of teaching methods

As expected, *teaching method* is one of the main themes that emerged across for the interviews. All participants reported that teaching methods have an impact on students' mathematics results; however, the findings of the interviews reveal that there is more impact for some participants than others due to the different teaching methods employed in the delivery of the mathematics lessons. Most of the participants employ traditional methods, with some participants using student-centred methods, and a few incorporating both methods.

It is important to note that teachers' actions and inactions may impact positively or negatively on students learning experiences in mathematics (Ampadu, 2012; Yasmin, Naseem, & Masso, 2019). This is because students' learning experiences are to a large extent controlled by their teachers, and teachers tell students which questions to solve and which methods to use. Most of the participants in this study believe that students learn and perform better in mathematics when teachers are at the centre of the teaching process (i.e., the teacher-centred approach). This means that teachers explain concepts and take full control of the learning session (Yasmin et al., 2019). However, the results of the post hoc test in Table 12.3 reveal that the student-centred teaching method, and mixing both teaching methods, have high significant ($p<0.05$) difference on students' mathematics results, compared to teacher-centred methods. In other words, teachers employing student-centred methods and mixed-methods in teaching at secondary schools in Port Moresby are more likely to influence Grade 10 and 12 students' mathematics results than the teacher-centred teaching method. The teacher participants who employed student-centred methods stressed that these methods enable students to be more responsible for their own learning, with more group discussions to assist each other's learning. Studies have also argued that student-centred methods promote discovery learning to understand and learn new ideas, as well as encouraging students to work cooperatively with peers when tackling mathematics problems, and ultimately assisting them to obtain better results (Emre-Akdogan & Yazgan-Sag, 2019; Lahdenperä, Postareff, & Rämö, 2019; Leong, Kaur, Lee, & Toh, 2019). This argument is supported by the ANOVA test results in Table 12.3; that teachers using student-centred methods and both teaching methods are likely to influence students' mathematics results. The teachers surveyed who incorporate both teacher and student-centred teaching approaches use the former to explain step-by-step process on the blackboard, and the latter to actively engage students' in-group work. This approach is similar to that found in a case study by Ampadu (2012) in Ghana regarding students' perceptions of teachers. The combination of both teaching methods seems to assist students to understand mathematics ideas and concepts, and they are likely to perform better in mathematics.

These findings clearly indicate that student-centred teaching methods and both methods make a significant difference on the mathematics results of the students in Port Moresby. This result also supports the researcher's experience teaching in secondary schools in Papua New Guinea. However, the teachers interviewed who adopted these two methods did not go into detail about how they understood and developed their students' skills, in order to improve their teaching practices and better communicate mathematics concepts in the classroom.

12.7 Theme 2: Content knowledge

Mathematics content knowledge is important for teachers to possess in order to deliver effective lessons to students. As discussed above, teachers variously use teacher-centred, student-centred, and mixed methods to deliver the content of mathematics topics, and it is believed that showing confidence with the content knowledge towards a particular topic is likely to have an impact on students' mathematics results. As mentioned earlier, the content knowledge of the teachers interviewed was measured with quantitative analysis to ascertain its impact on the students' mathematics achievements. Therefore, the results of teacher interviews will instead provide an in-depth exploration of how the content knowledge of mathematics teachers can affect students' mathematics results through their teaching.

The results of the interviews indicate that eight participants have adequate content knowledge and are confident in teaching all topics in the mathematics syllabus. Six other participants are confident in teaching both Geometry and Trigonometry. Two teachers are confident in teaching Series and Sequences, Vectors, Statistics and Money, and only one teacher is confident in teaching Calculus. Algebra is the only topic that all the 19 mathematics teachers are confident in teaching. These findings indicate that some of the teachers are less confident with their content knowledge when teaching some of the mathematics topics. Moreover, the results also reveal that there are six teacher participants who are less confident in teaching Calculus, and four participants who are less confident in teaching Vectors, respectively. Furthermore, three other participants are less confident in teaching Probability, with another two participants are lacking confidence in teaching Money. Geometry, Trigonometry, and Statistics are topics that one teacher each has less confidence in teaching. These results indicate that teachers' confidence in teaching mathematics topics can have influence on students' mathematics achievements.

Some participants gave their reasons for feeling less confidence in teaching these mathematics topics. One of the participants stated that:

> *"Trigonometry in Grade 12 especially, to do with quadrants and all that yeah and then going on to circular functions, the graphs because there a lot of details to be analysed. Information like the graphs to be analysed for the sine function, cos functions, the features of the graphs, you know what this letter is on the equation, what its meaning and all that, and they have to remember all those. Too many properties and inequalities to remember about certain equations make me you know to forget many things when I am teaching that topic and it makes me feel not confident" [Teacher 3].*

Another two teachers also shared similar sentiments:

> "Sometimes when I see that there is no time. *May be because the nature of the topic. So many things in there where we need to put together and understand. Find some ways to simplify for the students to understand" [Teacher 12].*

> *"It takes time and effort for teachers to sit down to read through different books to see the way it's been set up. Where you understand it you can be able to tackle it" [Teacher 16].*

The above three teachers' views indicate that teachers lack knowledge in the detailed aspects of topics such as Trigonometry, and as a result consume more time and effort attempting to consult different sources in preparing their lessons. Teachers also struggle to confidently deliver content knowledge that they are less familiar with, even despite extra time and effort put into the preparation of lessons.

These problems arise when teachers are less trained or are recruited directly from non-teacher training universities without the necessary content knowledge. Two teachers emphasised that:

> *Vectors and small part of calculus. That very area that I am not confident is because I did not pick at my studies at university [Teacher 4].*

> *I am a person I am not a trained teacher but I graduated from university and looking for a job and joined teaching. Because of my background as physics [Teacher 5].*

These two participants' views suggest that teacher-training institutions lack the capacity to impart much needed content knowledge to the student teachers in the necessary topics. Further, schools are employing teachers who are not trained to teach mathematics. This results in teachers who lack confidence in teaching some mathematics topics. As a result, students in turn might have different attitudes towards learning mathematics, which can affect their results. One teacher highlighted that:

> *"Sometimes maths is a difficult subject in PNG; there are many students who finds it very hard to understand maths so sometimes in my lesson I try to break the example down to detail to explain for students to understand the concept on how to solve the particular problem" [Teacher 7].*

This attitude aspect of students are also found in the quantitative results scetion of this study, showing that students' positive attitudes have a significant effect on their mathematics results. However, the results of the interviews also show that students can have negative attitudes towards mathematics (e.g. that is 'too hard') if their teachers struggle to prepare and deliver lessons.

Teachers' feelings of confidence/lack of confidence to teach mathematics content are compared below in Table 12.4. This comparison is conducted through independent samples t-test analyses.

Table 12. 4 Descriptive statistics of the teachers' confidence in teaching mathematics content

Group statistics

	Confident Level	N	Mean	Std.Deviation	Std.Error Mean
Maths Results	Confident	119	498.91	7.64	0.70
	Not confident	181	495.08	5.72	0.43

The descriptive statistics associated with the confidence level for teachers are reported in Table 12.4. From this table, it can be seen that teachers who are confident with mathematics content are associated with the high mean (*M*=498.91,*S.D*=7.636) compared to teachers who are not confident with mathematics content (*M*=495.08,*S.D*=5.72).

Table 12. 5 The output of independent sample t-test analysis

Independent Samples test

		Levene's Test for Equality of Variances		t-test Equality of Means					95% Confidence interval of the Difference	
		F	Sig.	*t*	df	Sig.(2-tailed)	Mean Difference	Std.Error Difference	Lower	Upper
Maths Results	Equal variances assumed	4.95	0.02	4.95	298	0.00	3.82	0.77	2.30	5.35
	Equal variances not assumed			4.66	203.06	0.00	3.82	0.82	2.21	5.44

Table 12.5 shows the output of the independent sample t-test analysis indicating whether there is a statistically significant difference between teachers with confidence in teaching mathematics content and teachers who are not confident. Both the results of the assumption of homogeneity of variances via Levene's F test, $F(298)=4.94, p=0.027$ and the independent *t*-test $t(298)=4.95, p=0.00$,95% *C.I* [2.303-5.345] are associated with statistically significant effects on the teachers' confidence levels. Thus, the teachers who are confident in teaching mathematics are associated with a statistically significant ($p<0.05$) effect with larger means than the teachers without confidence, in determining students' mathematics results.

Furthermore, professional development (PD) is crucial to overcome challenges associated with less confident teachers. PD can help with gaps in teachers' content knowledge, and provide methods to enable them to gain confidence in the effective teaching of mathematics topics. However, 16 out of the 19 teachers interviewed stated that they do not have access to PD for the topics they are less confident in teaching. These teachers stated that PD is not initiated at the department level, nor at their schools as a whole. Six of the teachers said:

> *"No. In that case no. Specially that particular topic ah. Any professional development or in-service there is nothing. But I am interested in taking up any of those" [Teacher 1].*
>
> *"I haven't any. The only way is to do is to do research to understand examples and deliver to the students" [Teacher 16].*
>
> *"No in-service. We lack teachers so we go in and teach as long as we achieve our objective" [Teacher 3].*
>
> *"So far we do not have any in-service. May be because of the head of department have not done a schedule" [Teacher 17].*
>
> *"Zero. There is no assistance in that area, no in-service in topics we are not confident to teach" [Teacher 3].*
>
> *"No in-service in the school and department" [Teacher 7].*

These responses indicate that professional development is less prioritised by most of the mathematics departmental heads at schools in Port Moresby. This could have multiple effects on the teaching and learning of teachers and students: teachers may remain less confident in teaching topics, and cannot deliver content knowledge to students effectively or with creativity. This implies that students are likely to get less out of the teachers, and that may consequently affect their mathematics results.

12.7.1 Discussion of content knowledge

It is clear from the findings and literature review that the impact of teachers' content knowledge of mathematics topics on their students' results can be quite significant. The interview data indicate that the impact seems to be greater when teachers are less confident in teaching some mathematics topics. This could be attributed to the fact that only eight of the teachers interviewed are competent to teach all of the mathematics topics, while 11 are confident in teaching just some of the topics. Besides that, most of the teachers still lack the skills and knowledge in topics such as Vectors and Calculus. The teachers' knowledge gaps in mathematics topics in turn can affect their students' mathematics results (Kelcey, Hill, & Chin, 2019), as these topics frequently appear in the Grade 12 national examination. A study by Kelcey, Heather, Hill and Chin (2019) indicates that there are significant correlations between teacher content knowledge, instruction, and student achievements. This argument links well with the ANOVA results of the different teaching methods discussed earlier. This is because using student-centred teaching method and mixing both teaching methods have a high significant influence on students' mathematics results and these methods can be used to effectively deliver content knowledge to students in the classroom. Teachers using these methods consistently, and who have adequate content knowledge in all mathematics topics can teach and guide students to obtain better mathematics results. This is because teachers having sufficient mathematics knowledge allows them to focus on teaching, readily provide alternative explanations, and incorporate additional resources into lessons (Hallman-Thrasher, Connor, & Sturgill, 2019). However, for teaching to be successful, it is also key that teachers possess content knowledge in multiple ways and have an ability to make this knowledge accessible to students. For this to happen, the evaluation of knowledge in terms of content and pedagogy are important as they are interrelated for teachers' success in their practice. This is because the knowledge needed to make specific content accessible to students, which forms a critical component of teacher knowledge, has been shown to have an impact on teaching effectiveness (Hallman-Thrasher et al., 2019). This involves teachers' knowledge of content, students, teaching and curriculum in a holistic approach. This knowledge also includes teacher understanding of the student learning process of a particular topic, for instance, the process of how students might solve a problem, and common conceptions and misconceptions about a topic. Further, the knowledge of how a teacher might teach a particular concept needs sharpening, for instance in areas such as representations, sequencing, pacing, and providing examples that are most appropriate for a given topic (Appova & Taylor, 2019; Hallman-Thrasher et al., 2019). The knowledge of content and curriculum involves knowing what instructional resources are available to teach a given topic and under what conditions they should be used (Appova & Taylor, 2019). It is true that content knowledge that is common to any mathematical profession is needed to teach and learn mathematics, for instance, knowing how to solve an algebraic equation. However, the components mentioned above are fundamental to fully utilise this specific content knowledge in mathematics. In other words, teachers need to develop not just a deeper knowledge of mathematics content but also an understanding of the mathematical process of inquiry and problem solving to enrich their teaching practices and to encourage the development of critical thinking skills in their students (Albert 2012). Teachers are required to know and understand not only mathematics, but also their students learning processes, teaching strategies, as well as how to challenge and support the classroom-learning environment. In addition, teachers need to know and use mathematics for teaching that combines mathematical and strategical knowledge, as they must continue to learn new or additional mathematics content and study how their students learn mathematics (Zhao & Zhao, 2016). These approaches allow teachers to see their knowledge in teaching mathematics in a wider perspective, and are likely to positively influence students' mathematics results.

However, many participants reported that professional development is not given priority by school heads to improve on those topics that they are not confident to deliver in the classroom. This reflects the mixed feelings reported in relation to the impact of teachers' relationships with school heads who do not organise PD opportunities, which may be affecting their students' mathematics results. Though some participants have an interest in PD, they are not given the chance to improve on their weaknesses in content knowledge areas. Studies have shown that PD equips teachers with content knowledge and skills they lack to teach students (Jacob, Hill, & Corey, 2017), as it upgrades and updates their teaching processes and allows them to deliver mathematics content on topics more confidently.

12.8. Theme 3: Outcomes of the teaching methods

Outcomes of the teaching methods in the delivery of content knowledge (topics) by teachers are important to determine and evaluate students' learning processes. This evaluation of teaching methods is observed through the students' positive responses towards lesson participation, approaches towards exercises, results of their assessment tasks (e.g., tests and assignments) in classrooms and the results of their standardised national examination. These elements of the students' learning processes indicate the effectiveness of the particular teaching method employed by their teachers. The effects of these teaching methods determine the outcome of the students' mathematics learning, and are also evident in the teachers' interview responses.

Of the teachers interviewed, 12 out of the 19 teachers employing traditional teaching methods (discussed earlier) gave responses on how these methods influence learning outcomes. Two of the participants stated that:

> *I give students exercises and when they get them correct I know they understand. Sometimes students have to explain correctly in a logical order and that means that they really understand the concept [Teacher 14].*
>
> *They learn very well I think. Students do more examples and want more practical examples to do the exercises. The method that I use is very good and I think students are learning more on that [Teacher 10].*

These participants' responses indicate that students' understanding of the mathematics concepts are determined by the competence level (number of correct answers) that students display when attempting practical exercises, and teachers' logical explanation of ideas.

The other three participates mentioned that student participation and interaction with teachers shows that students have sufficiently understood the mathematics content. These participants said:

> *I see the students' response favourable to the method that I use. There seems to be an interaction between a teacher and the student. When I ask questions, students are able to raise their hands and answer the question. In that where I know and realise that, the students are into the lesson that is been presented and are following. I also call the attention of students who do not really participate, who are doing other things. When I am teaching I look around and see that I direct questions to them. In that, way I make them pay attention to what is been discussed [Teacher 7].*
>
> *The learning is OK. It's not high standard of learning as expected. They need various strategies in order to fully understand the subject [Teacher 13].*
>
> *Students learning, the most method they are catching up in learning is through chalkboard doing calculation on the board and doing corrections, word problems you know this. The board is a good way of learning that is improvement [Teacher 2].*

It is evident in these participants' responses that participation in the lesson indicate that students are learning from the teaching method. However, to facilitate students' participation in lessons, teachers must expend effort to get all students' attention by asking questions. Furthermore, teachers' use of the blackboard for doing calculations, and correcting mathematics problems together with students, effectively improves student learning in mathematics. However, in order to better disseminate mathematics ideas to students, it is still necessary to also utilise other teaching strategies such as student-centred methods. The other six teachers using the student-centred methods also provided responses detailing their methods. Two of these participants highlighted that:

> *Ok other one when we give in groups, we find that learning is quite effective, because within their own groups, they kind of check on each other. And everyone has part to play in the group and they report to each other so nobody seems to be sleeping [Teacher 3].*
>
> *One way is they are helping each other during group work. They share ideas in groups esp. those bright ones are assisting the weak ones to understand what was taught in class and doing their presentation to does the same thing [Teacher 12].*

These two participants' responses denote that student-centred methods are effective, as students work in groups sharing ideas to assist each other with lessons. This suggests that high-achieving students could assist weaker students when group and peer learning is encouraged, with students checking each other's learning and sharing ideas through presentations. Teachers act to facilitate their students learning processes, and students take ownership of their own learning for desirable learning outcomes.

Another teacher further stated that:

> *What I find is in certain topics where I demonstrate but the other where I feel that they can be able to help with their peers I give the task to them and they enjoy working with the peers. This is because they understand their peers and friends [Teacher 6].*

This response supports the previous two participants' responses emphasis on students of working together in groups to learn from each other, indicating that students work and learn with their peers better than with their teachers.

12.8.1 Discussion of the outcomes of the three teaching methods

The interview results show that the three teaching methods (student-centred, teacher-centred and both methods mixed) are employed by teachers in Port Moresby. As shown in the ANOVA results earlier, student-centred methods and mix methods have a high significant ($p<0.05$) difference on the students' mathematics results, compared to the solely teacher-centred teaching method. One of the main reasons for this may be because of the shift of teachers' teaching methods from traditional (teacher-centred) methods to the student-centred and mixed teaching methods. Despite that, the traditional method is still prominent in schools in PNG. This suggests that students are generally dependent on teachers for their learning, and are not taking ownership of their own learning. As discussed above, teacher-centred methods treat students as passive learners, and encourage less creativity in learning. Studies have shown that teacher-centred methods have a negative influence on students' academic achievements (Basso, 2019; Olson & Stoehr, 2019). This phenomenon can be evidenced in the mathematics results of students who are taught by teachers using student-centred or mixed methods in Port Moresby, which demonstrate improvement over traditional methods.

Moreover, other studies have argued that student-centred and mixed methods have a positive influence on student academic outcomes (Saadati, Cerda, Giaconi, Reyes, & Felmer, 2019; Ulandari, Amry, & Saragih, 2019). This is because students are active and are more engaged in their own learning. With these methods, students are also able to construct their own knowledge through research and presentation in groups as mentioned by the participants in this study. Drawing from the researcher's experience as a teacher in Port Moresby, this may be tied to teachers' access to instructional resources, teachers' willingness to adapt, and many other strategies related to student-centred method and mixed methods. Despite this, however, the researcher believes that there remain many issues that are likely to continue to hinder teachers to effectively practice the non-traditional methods in delivering mathematics lessons in the classroom (PNG NDOE, 2006, 2009).

12.9 Theme 4: Encouragement

Teachers' encouragement of their students is important in improving their mathematics results. Encouragement from teachers can influence students' approaches to learning mathematics. The two important sub-themes that emerged from analysis are practice with more exercises and students' attitudes/approaches towards mathematics.

Working consistently on mathematics exercises can give students greater confidence in solving mathematics problems, and deepen their understanding of the subject. For instance, the more they practice and actively solve exercises, the better they will be in mathematics. This view was confirmed by three participants' responses. They highlighted that:

> *Maths is not something that you can study and memorise, but maths is something you must practically do. I encourage them to do more activities/exercise and do everything so that you can be used to [Teacher 10].*

> *Tell them to do a lot of exercise because they do more practice to understand the concept in mathematics [Teacher 14].*

> *You must do a lot of exercises [Teacher 4].*

The responses above clearly indicate that teachers encourage students to do more practice with mathematics exercises to learn and master mathematics knowledge and skills. Similarly, a teacher participant emphasised the importance of relating abstract mathematics knowledge to more practical applications in society. The participant stated that:

> *One of the things I tell them esp. those who do not show much interest in mathematics is whatever we learn in school we apply outside in real life. Put in practice what we learn in the classroom. Get yourselves involved and learn more to put in practice outside. Money or formula can be applied in real world esp. every way you go in shops or markets there is mathematics. Therefore, you must learn and really understand mathematics [Teacher 7].*

Students' attitudes/approaches towards mathematics is another sub-theme identified from the analysis. Of the teacher interviewed, 12 of the 19 encourage students to have positive attitudes or approaches towards mathematics. The sub-theme of students' attitudes towards mathematics was also discussed in the quantitative analysis of this study (Chapter 7). For instance, three of the participants emphasised that:

> *Cheer them up and tell them you can do it. Mathematics is not that difficult it's fun. I make sure they do questions I give and check their work [Teacher 16].*
>
> *Maths is interesting because it's all to do with numbers. It interesting in the way I teach mathematics to students [Teacher 19].*
>
> *I always tell them that it's like you learning 1 + 1 like you did in Primary school. So you do not learn something very hard. Know your basics well then you can do better and understand because that will be applied and carried on to Grade 9, 10, 11 and 12 [Teacher 12].*
>
> *I encourage students to be active and be energetic when learning mathematics [Teacher 2].*

These responses suggest that the teachers encourage their students to have positive attitudes towards mathematics, for instance, encouraging students by communicating that mathematics is not hard but is interesting and fun. These teachers support their students to understand basic mathematics skills in order to progress to problems at higher-grade levels that require higher skill levels. This encouragement can be prompted through active learning with greater energy. Similarly, some teachers encourage students to approach them (or others they are comfortable with) when they have difficulty understanding mathematics ideas/concepts. Two teacher participants pointed out that:

> *Encourage Feel free to come and see me when they need help or assist. In addition, I encourage them to help each other. Otherwise, they are not comfortable with me as their maths teacher I encourage them to see their friends, parents or seek assist from other mathematics teachers. My biggest encouragement is to open up, to love the subject and to do what they can be able to do. Any difficulty they can come to see me [Teacher 6].*
>
> *I always encourage students especially those with problems, those who find work difficult to come and seek help. My own experience in this school for the last five years I give help to anybody because I opened the door. Anyone with problem even they are not from my class I help as the head of department I make my business to do that. I am helping average students and lower achieving students but the bright ones do not come [Teacher 9].*

These responses show that teachers are willing to allocate time to assist students that have difficulty understanding mathematics concepts. This approach for students seeking help is important to unpack their uncertainties about mathematics problems, and encourages them to improve in their learning practices thereby subsequently improving their mathematics results.

12.9.1 Discussion of encouragement

From the interview results, it is evident that the teachers encourage their students in studying mathematics through having positive attitudes and taking efforts to make mathematics more practical. This outcome is consistent with this study's quantitative results in Chapters 10 and 11 with the literature review in Chapter 2. Several other studies have also shown a correlation between students' attitudes towards mathematics and their achievements in the subject. Teachers' encouragement contributes to positive student attitudes influencing their decisions and improving their results (Khun-Inkeeree, Omar-Fauzee, & Othman, 2016; Pepin & Roesken-Winter, 2015; Prendergast, Hongning, & Block, 2016). As discussed in the literature review in Chapter 2, mathematics can be presented as a fun and interesting subject, and when students perform successfully in mathematics, this affects their attitudes towards mathematics in positive ways, and this perception can continue throughout their schooling (Mullis et al., 2012; Stephens, Landeros, Perkins, & Tang, 2016). This is because negative attitudes towards mathematics can cause some students to lose self–belief in their abilities (Iqbal, Mirza, & Shams, 2017; Kiss, 2018; Soni & Kumari, 2017). In the interviews, the teachers' efforts to eradicate their students' negative feelings towards mathematics is evident, as is the belief that more positive encouragement is likely to influence students.

Practical exercises further motivate students to learn mathematics, by allowing them to understand the concepts and steps involved in solving a problem. It is evident from existing studies that students gain interest and are engaged when concepts discussed in the classroom are practically demonstrated in the real world. In the mathematics context, this is achieved through relating the basic principles of how mathematics concepts apply to the every-day. For instance, Kiemer, Gröschner, Pehmer, and Seidel

(2015) discuss the case of an engineer who went to teach mathematics in India. In their study, Kiemer et al. (2015) demonstrated how the teacher motivated their students by relating what had been learnt in the lesson to what was outside the classroom. This approach was initiated due to the students' interest in applying mathematics theory into practice. Overall, the interview data show that there seems little doubt that teachers can better motivate students to study mathematics by emphasizing the relevance of mathematics in the real world. For instance, Ali (2013) shows how teachers motivate students through solving mathematics problems relating to their everyday lives. This approach can generate excitement among the students to persevere in mathematics because of the practical skills and knowledge that underpin the mathematics lessons. Mathematics in this context promotes reasoning skills that are helpful in students' everyday lives. These kind of encouragements by teachers not only motivate students but also promote positive learning approaches that are likely to have an impact on the students' mathematics results.

12.10 Summary

This chapter presented the results and discussion from the interviews with the 19 Grade 10 and 12 teachers who participated in this study. The findings were presented in four sections that corresponded to the primary themes that emerged from the results: (1) teaching methods (student and teacher-centred methods), (2) content knowledge, (3) outcomes of the teaching methods used in the classroom, and (4) encouragement towards mathematics. Sub-categories within each theme assisted to support and provide insight into the broader themes. The aim of this chapter was to capture information that might have been missed in teacher survey to facilitate an in-depth exploration of the quality of teaching that affect students mathematics results at the teacher-level

It is clear from the findings of these interviews that the student-centred method and mixed methods have significant ($p<0.05$) effects, compared to teacher-centred method. As the interview results have shown, the teaching methods adopted by teachers can assist and promote students' learning. Furthermore, the confidence of teachers towards their own mathematics content knowledge has a positive influence on their students' mathematics results. This means teachers' knowledge of mathematics content, pedagogy, instructional resources and curricula are contributing elements in producing better results. However, as identified in the analysis, teachers also face challenges such as student population increases and a lack of learning resources in classrooms that may affect the practical delivery of lessons to effectively communicate content knowledge. On a positive note, some of the participants involved in this study acknowledged that they encourage students to have positive attitudes towards mathematics learning. They also suggested approaches to help students overcome their struggles towards mathematics when the subject becomes difficult, assisting them to believe in their own mathematical abilities in order to obtain better results.

These qualitative results complement the quantitative results discussed earlier, through the triangulation process applied in this chapter. For instance, the teachers' interview results on the three teaching methods were triangulated through the simple one-way ANOVA and content knowledge through independent t-test (quantitative approaches) analysis to obtain a balanced view. This interview results support the literature review and SEM results in Chapter 10 that *quality of teaching* at the teacher-level has a significant effect on students' mathematics results. In other, words teaching methods and content knowledge identified in this chapter that reflect quality of teaching in this study influence students mathematics results. The next chapter of this thesis provides a summary of the findings of the different chapters and discussions of this study, and presents the implications, and recommendations of these findings, as well as providing concluding remarks.

Chapter 13: Discussion and Conclusion

13.1 Introduction

Though many studies have been conducted into the possible causes of poor and declining student mathematics results, the majority of these have focused on Western countries. A limited number of studies have been conducted in the Pacific Islands context, and no study has been conducted specifically into factors affecting students' mathematics results in the Papua New Guinea (PNG) context, meaning that these factors have been to date unexplored.

Thus, this study's goal was to examine and identify the factors that are affecting Grade 10 and 12 students' mathematics results in Port Moresby, PNG. The study explored student-level factors (attitude, motivation) and teacher-level factors (quality of teaching, classroom environment, and instructional resources). In addition, the study also explored both student and teacher demographic variables such as gender, years of teaching experience, parents' highest education level, parents' employment and type of school. The interrelationship of the constructs at the student and teacher-levels with their demographic variables were also examined, to determine whether there was influence on the students' mathematics results. The study investigated three main questions:

1. What are the student and teacher-level factors that are affecting Grades 10 and 12 students' mathematics results in Port Moresby, PNG?
2. What are the relationships among student and teacher-level factors that are affecting Grades 10 and 12 students' mathematics results?
3. What are the attributes that are deemed important by teachers regarding their mathematics teaching approaches and content knowledge?

This study employed a quantitative-dominant mixed-research approach. The student-level constructs, demographic variables and Grade 10 and 12 tests were measured by the development of survey questionnaires and test items. These instruments were calibrated, validated, and verified to ascertain the measurement properties of variables/scales and mathematics tests. The students' achievements were measured through different test items taken from the topics they had covered in Grade 10 and 12, respectively, as discussed in Chapters 5 and 6. Rasch analysis was employed to validate, calibrate and verify test items for both grades. The scales included students' *attitude* (ATTD) and *motivation* (MTVN). Confirmatory factor analysis (CFA) and Rasch analysis from the Partial Credit Model (PCM), discussed in Chapters 7 and 8, were used to validate, calibrate and verify the scales.

Similarly, in Chapter 9 statistical analysis techniques were used to validate, calibrate, and verify the teacher-level scales such as *quality of teaching* (QT), *classroom environment* (CE) and *instructional resources* (IR). The demographic variables included were *gender, type of school,* and *years of teaching experience*. Additionally, qualitative data from interviews with teachers teaching Grade 10 and 12 students were collected and analysed with NVivo using a thematic approach to obtain a deeper understanding of the quality of teaching at the teacher-level.

In this concluding chapter, the findings of this study are discussed according to the main research questions advanced in Chapter 1, with reference to evidence from the previous studies reviewed in Chapter 2. After presenting the study's significant findings according to the factors identified at the student and teacher-levels, this chapter discusses the effect of the demographic variables on the students' mathematics results. Next, the chapter identifies the limitations and significance of the study, with recommendations for future research directions in the area. The chapter then concludes by reiterating the key findings of the research.

13.2 Summary of the findings

This section discusses the main findings of the study.

13.2.1 The effect of students' attitudes and motivation on their mathematics results

In this study, combined, Grade 10 and Grade 12 models showed that students' attitudes towards mathematics influences their mathematics results. The findings of the structural equation model (SEM) show that students' *attitude* had a significant ($p<0.05$) direct positive effect on mathematics results. This means that those students who possess more positive attitudes towards mathematics, as well as having clear perceptions of the subject relevance, value and difficulty, are likely to obtain better results. This result is supported by other studies, which have also shown a relationship between students' attitudes towards mathematics and their achievements in the subject (Chen et al., 2018; Kiwanuka et al., 2017; Yavuz Mumcu & Cansız Aktaş, 2020).

Moreover, the SEM results of this study also found that, all three models (combined group, Grade 10 and Grade 12) show students' *motivation* towards mathematics has a significant ($p<0.05$) indirect positive influence on their mathematics results. This suggests that more motivated students are more likely to have more positive attitudes, which leads them to perform better in mathematics. The SEM results also indicate that the students with higher motivation towards mathematics perform better in mathematics. This finding is similar to a study by Hidi and Renninger (2006, p. 111), who highlighted "the level of a person's interest has repeatedly been found to be a powerful influence on learning". Interest is thus an important motivational factor (Ufer et al., 2017) which leads to students being intrinsically motivated in mathematics.

Similar findings are also evident in the results of hierarchical linear model (HLM) analysis, which found that the combined and Grade 10 models show that students' *attitudes* had a significant ($p<0.05$) direct effect on their mathematics results. This suggests that students with more positive attitudes obtained better results in mathematics. As indicated in Chapter 2, several studies have also shown a correlation between students' attitudes towards mathematics and their achievements in the subject (Chen et al., 2018; Kiwanuka et al., 2017; Yavuz Mumcu & Cansız Aktaş, 2020). This relationship is highlighted in a study by Benken et al. (2015), which concludes that attitude towards the mathematics is one of the most significant factors influencing student participation in the subject. Further, these findings are also consistent with international studies such as TIMSS and PISA; that, the more positive students' attitudes, the better their academic outcomes (in this case, mathematics achievement). Students' attitude and motivation towards mathematics is significant. It is perhaps important for teachers to consider that Grade 10 and 12 students in PNG are experiencing changes in their learning daily , so it is important to help them sustain positive attitude and stronger motivation to study mathematics. This can be encouraged through professional development to improve students' motivation and attitude to in mathematics.

13.2.2 The effect of students' demographic variables of their mathematics results

The demographic characteristics of the study showed that students' mothers' highest education level influences their mathematics results. As evidenced by the findings from both SEM and HLM, at the student level, *mothers' highest education level* has a significant ($p<0.05$) positive effect on the students' mathematics achievements. This shows that students whose mothers have higher education levels are likely to perform better in mathematics. The finding from this study is consistent with those reported by Baliyan et al. (2012) and Tan et al. (2020), who simmilarly found that mothers' education level has an influence on their children's academic achievements. In addition, the SEM results of the combined and Grade 10 models show that *fathers'* highest education level has a direct significant ($p<0.05$) positive influence on their mathematics results. This suggests that students whose fathers have highest education levels obtain better mathematics results. Similarly, SEM results of the three different models (Combined, Grade 10 and 12) indicate that *private schools* have a positive ($p<0.05$) direct influence on students' mathematics results. This could mean that students in private schools have better resources, compared to public and Catholic schools. This result is consistent with studies by researchers such as Lefebvre et al., (2011), and Siddiqui and Gorard (2017), who found that private schools offer more conducive learning environments, featuring high-quality teachers with strong experience who are equipped with adequate content and teaching pedagogy knowledge.

Moreover, the HLM result of the Grade 12 model also showed that students' *gender* had a significant ($p<0.05$) positive direct influence on mathematics results. This result suggests that female students perform better than their male counter parts. This finding is inconsistent with studies by Zhu, Kaiser, & Cai (2018) and Reilly, Neumann, & Andrews (2019), who discovered the opposite result. However, researchers such as Goldman & Penner (2016), Stewart et al, (2017), and Uwineza, Rubagiza, Hakizimana, and Uwamahoro, (2018) found results that are similar and consistent with the findings of this study.

In addition, the HLM results of the Grade 12 model further indicated that *fathers' occupation* had a significant ($p<0.05$) positive direct influence on students' mathematics results. This finding suggests that Grade 12 students whose fathers are employed obtain better mathematics results. Similarly, the HLM results of the Grade 10 model show that private schools had a positive significant ($p<0.05$) direct effect on the students' mathematics results. This result is similar to the SEM result of the three models discussed earlier.

13.2.3 The influence of teacher-level factors on students' mathematics results

The SEM results of this study indicate that the constructs *quality of teaching* and *instructional resources* at the teacher-level had significant effect on the students' mathematics results. However, *classroom* *envornemnet* had non-significant effect and could be due to the small sample size (41) at the teacher level. The HLM results also show that all contextual factors of the three models had no influence on students' mathematics results. However, the HLM results of compositional factors such as *average mothers' highest education level (Ave-MHEDL)* for the combined, Grade 10 and 12 models have a direct effect on the students' mathematics results. This finding suggests that students in classrooms with a higher average mothers' education level perform better in mathematics. This result is similar to studies such as those by Baliyan et al. (2012), Tan et al. (2020) and Okumu et al. (2008), which indicated that there was a significant relationship between mothers' education level and students' academic results. In addition, the HLM results of the combined and Grade 10 models showed that students in class where on average fathers with occupation (*Ave-FOCC*) has direct effects on the students' mathematics results. This indicates that fathers who are employed are more likely to influence their children's mathematics results.

13.2.4 The effect of teacher demographic variables on students' mathematics results

The HLM results of combined and Grade 10 models of teachers' demographic characteristics demonstrate that teachers *with mathematics majors* had a significant ($p<0.05$) positive effect on their students' mathematics results. This implies that teachers who are specialised in mathematics as a major at the university level are more likely to positively influence their students' mathematics results. This finding is consistent with that by Odumosu and Fisayi (2018), who found that mathematics teachers with an in-depth knowledge of the subject have an influence on students' mathematics results and these teachers are more qualified to teach algebra at schools.

13.2.5 The interaction effect of student and teacher-level factors on mathematics results

The HLM findings of the combined, Grade 10, and Grade 12 models in this study reveals that *average mothers' highest education level (Ave-MHEDL)* at the teacher-level had a positive interaction effect with *attitude* at the student-level, and consequently influenced the mathematics results of the students. In classes where the average level of students' mothers' education is high, the effect of attitude on mathematics achievement becomes stronger. By contrast, for students in classes where the average level of mothers' education level is low, the effect of attitude on mathematics achievement is weaker.

Moreover, the HLM results for the Grade 12 model showed that *average mothers' highest education level (Ave-MHEDL)* at the teacher-level had an interaction effect with *gender* at the student-level, with a negative effect on mathematics performance. In other words, the female students perform better in mathematics. This difference was bigger in classrooms where, on average, mothers' higher education level was low. In classrooms where mothers' higher education was high, the difference was smaller. These findings reflect cultural aspects of PNG, where females are more attached to their mothers, and are more inclined than males to imitate their mothers in their learning journey (Edwards, 2015).

13.2.6 Qualitative interview results

Four main themes were identified from the analysis of the interview data: teaching methods, content knowledge, outcome of the teaching methods, and encouragement towards students learning mathematics. Three teaching methods (the teacher-centred method, student-centred method and both methods mixed) were identified from the first theme (teaching methods). Out of these three teaching methods, the student-centred method and mixing both teacher and student-centred methods have a significant difference on the Grade 10 and 12 students' mathematics results, compared to teachers solely using the teacher-centred method. As such, these two methods assist students to better learn and do well in mathematics, as they become the centre of the teaching process. This means that students take full control of the learning process, while teachers explain the main points through facilitation. These findings are consistent with those reported by Ampadu, (2012), Emre-Akdogan & Yazgan-Sag (2019), Lahdenperä, Postareff, & Rämö (2019), and Leong, Kaur, Lee, and Toh,(2019).

On the other hand, the results showed that teacher-centred methods has a non-significant effect on the students' results. This suggests that teachers using teacher-centred methods are unlikely to influence students' mathematics results. This result is inconsistent with Yasmin et al., (2019), who indicated that teacher-centred approaches influence students' academic results.

Moreover, teachers who are confident with content knowledge in teaching mathematics have a significant positive effect on their students' mathematics results. This could mean that teachers influence mathematics results by having confidence in the content knowledge of mathematics topics, pedagogies, and instructional resources. These results are consistent with Chapter 2's discussion on some of the attributes of quality of teaching adopted from the NSW Department of Education and Training. Furthermore, this study found that students who are continuously encouraged by their teachers to do more practice in mathematics had more positive attitudes towards mathematics. This result is similar to the quantitative results of this study (SEM and HLM); that students' positive attitudes towards mathematics had a significant influence on their mathematics results.

13.3 Significance of the Study

This study will significantly contribute to the discipline in three important ways:

13.3.1 Methodological significance

This research study involved questions that examine school, teacher and student-level factors, and the possible relationships that exist among them. The study's theoretical framework (introduced in Chapter 2) was developed in relation to information gathered from the relevant literature, to investigate these teacher and student-level factors, and the relationships between factors that may influence Grade 10 and 12 students' mathematics results. Within this theoretical framework, quantitative data for the constructs and demographic variables involved were collected and analysed to test these relationships. The survey instruments for scales with the Grade 10 and 12 mathematics test items were validated and calibrated to obtain reliable data for subsequent analysis. This validation of the construct and test items was carried out through CFA and Rasch Model, using Mplus 7.0 and ConQuest 4.0, respectively.

The use of the Rasch model to transform raw scores to Weighted Likelihood Estimates (WLE) scores made data interpretation more meaningful and valid, as WLE scores are interval scales of equal distance. Compared to other similar transformation techniques, this approach has the advantage of a minimised estimation bias (Woodcock, 1999). Further, the WLE scores were transformed into a W-score. The main advantage of the W- score is the elimination of negative and decimal values, which makes interpretation more convenient. Furthermore, item/person fit analysis based on the Rasch models enabled identification of respondents who misbehaved as a result of not understanding the items or rushing in answering the questions. The WLE scores provided more accurate estimates of the respondents' abilities. The W-score was used in the SEM and HLM analysis to predict the variables that had a significant effect on the students' mathematics results. As mentioned in Chapters 3 and 11, the data collected in this study were hierarchical in nature, and analysis involves data from both the teacher and student-levels.

13.3.2 Theoretical significance

Other studies have also examined factors affecting students' mathematics outcomes. It has been evident in the existing literature that school-level (instructional resources, type of school) teacher-level (teacher quality, classroom environment) and student-level (attitude, motivation, and gender) factors have influenced students' results. A number of mathematics education studies have also investigated the factors affecting mathematics teaching to provide evidence of teacher and student-level factors that could be developed to improve outcomes.

Despite continuous emphasis by the existing literature on the importance of these teacher and student-level factors, however, there are still some aspects that require further attention. First, existing research studies on these factors have to date been inadequate, as many studies have identified only a few factors, and have not examined many factors that have interrelationships with or impact on mathematics results. Besides that, most existing studies have been from developed countries such as Australia, the USA, and other Western countries, and such studies have not been widespread in Pacific Island nations such as Papua New Guinea, Fiji, Solomon Islands and Samoa. In other words, to date there has been a lack of research in PNG on factors affecting students' mathematics results, and so investigation of the problems associated with the declining and poor mathematics results of the country is important. This study was conducted after identifying these gaps in the literature.

Moreover, this study provided information on the factors associated with variations in Grade 10 and 12 students' mathematics results, and the insights offered here can be made part of the bases for improving student learning. However, while these factors were deemed important and related to the issue of declining mathematics results, the findings cannot be generalised to other contexts. As new findings, they are still subject to confirmation or refutation by other studies. Nevertheless, the results (factors affecting mathematics) can assist in providing information for the development and testing of new frameworks in the future. This study has assisted in highlighting relevant issues, and provides new information with empirical evidence on the factors affecting mathematics results in Port Moresby, PNG. This kind of empirical study is the first of its kind conducted for PNG, and possibly for any Pacific Island country, to investigate the factors at the student and teacher-level that are affecting Grade 10 and 12 students' mathematics results.

13.3.3 Practical significance

Results obtained from this study's quantitative and qualitative analysis showed that several factors had an influence on the Grade 10 and 12 students' mathematics results. These factors include students' mothers' highest education level, private schooling, fathers' occupation, whether teachers undertook a mathematics major at university, and students' motivations and attitudes towards mathematics. These findings have implications for education policy, curricula, development programs and educational research in the area of mathematics education.

Through these findings, this study provides evidence on the factors that affect Grade 10 and 12 students' mathematics results, which in turn provides an explanation for PNG's declining or poor mathematics results. Instead of blaming students for their poor results, or blaming teachers for their poor teaching, attention should be given to the factors identified by this study. The research model developed in this study took into account the two stages of teaching and learning mathematics, and the student and teacher factors that were included to examine factors contributing to improving Grade 10 and 12 students' mathematics results were investigated empirically. The findings of this study show that both teacher and student-level factors had a direct influence on the students' mathematics results, and that teacher and student-level factors also had cross-interaction effects with an influence on these results. Improvement of the teaching practice in PNG should be encouraged through adequate professional development (e.g., refresher courses on mathematics content knowledge and pedagogy) for teachers to develop confidence in delivering mathematics lessons. This should allow teachers to enrich themselves with the necessary skills to assist students to learn mathematics effectively.

The qualitative results of this study show that different teaching methods have different effects on students' performance in the classroom. Student-centred approaches and mixed teacher and student-centred approaches were shown to make a significant difference in learning outcomes, compared to teacher-centred teaching methods. These pedagogies establish effective teaching environment and create positive relationship between the students and the teachers, encouraging dialogue approaches to enable deeper and more meaningful leaning, and develop respect that recognise individual students' value, importance and contribution.

Therefore, more professional development related to these teaching pedagogies should be encouraged in all schools to upgrade teachers' professional capabilities. Supportive policies, coherent teacher education programs and further research are among the key areas that warrant review to fully utilise the three different methods of teaching mathematics.

13.4 Limitations and future research

This present study contributes significantly to mathematics education scholarship, especially to literature related to teaching and learning mathematics. However, there are limitations in some aspects of this study, which mirror the limitations of other research studies in the area. The first limitation is that the study focused on Grade 10 and 12 students in secondary schools only in the city of Port Moresby. As a result, this may restrict the generalisation of findings to other parts (towns and rural areas) of Papua New Guinea. Second, the teacher and student-level factors that are included in this study did not include all the factors at the different levels that might influence students' mathematics results. In other words, many other factors at the student and teacher-levels are not identified by this study that may also contribute to students' mathematics results. Third, to obtain stronger findings, a longitudinal study would be ideal for collecting data over a long period, especially for examining the interrelationships between the student and teacher-level factors that influence student results. However, this study took a cross-sectional approach due to time constraints and limited resources.

These limitations of this study suggest several avenues for future research.

First, a sample from the schools in all four regions (Highlands, Islands, Momase and Southern) of PNG needs to be included to report balanced findings. This would provide researchers with a broader understanding of the factors affecting mathematics results. Second, future research can be conducted to investigate other school, teacher and student-level factors that may influence Grade 10 and 12 students' mathematics results. Third, longitudinal studies should be conducted to examine the interrelationships among student, teacher and school-level factors and their effect on results. This is because longitudinal studies allow researchers to make observations of students and teachers over a period of time, to more easily trace changes in their behaviour. Fourth, this study utilised and modified instruments from international studies such as TIMSS and PISA, which are developed in other countries. Although these instruments were validated and found to have acceptable measurement properties, the development of new instruments that are more relevant specifically to the PNG student and teacher context is suggested for more meaningful outcomes. Additionally, interview questionnaires need to be revised to obtain more information about the tested variables at the teacher-level, and to provide a more in-depth interpretation of the qualitative findings for better triangulation. Finally, should similar studies be conducted in the future, actual observation or video study of teachers using teaching methods in the classroom is suggested to cross-check with interviews and quantitative variables, to obtain better interpretations of the relationships of the factors that affect mathematics results.

13.5 Recommendations

Though Papua New Guinea (PNG) is rich in mineral, agricultural, forestry and fisheries resources, development in terms of the human resources is still in its early stages, with educational challenges encountered within the education system, particularly in mathematics education. Therefore, in view of the results of the study, it is the opinion of this researcher that:

1. The government of Papua New Guinea (PNG) should invest more money in training secondary school mathematics teachers at training institutions and universities. This is because confident teachers with adequate content knowledge and effective teaching pedagogies have more influence on their students' mathematics outcomes.
2. Secondary school training institutions/universities in PNG should enrol students with strong analytic and mathematical skills in mathematics majors. These trained teachers should be the only ones teaching mathematics at secondary schools rather than teachers specialised in other subjects, as this study has shown that teachers with major in mathematics influences their students' results.
3. School principals/head teachers should plan and encourage mathematics teachers to have in-service or refresher courses on teaching pedagogy and content knowledge in order to gain confidence in teaching mathematics. This is due to the effect that confident teachers have on their students' results, as identified by this study.
4. The government should encourage and empower more women to undertake higher education because in the long run, it will benefit the country through the effects on their children's education. As it was evident in this study, students whose mothers have a higher education level are more likely to obtain better results.
5. The government through the National Department of Education (NDOE) should empower principals/head teachers of secondary schools to train teachers in how to maintain positive students' attitudes towards mathematics. Teachers should be encouraged to adopt best practices

that are employed in other countries, as this study found that students' attitudes towards mathematics have a positive significant effect on their outcomes.

6. Government and Catholic schools should learn best teaching and learning practices from private schools, and integrate these with their own practices accordingly. This recommendation derives from this study's identification of private schools' significant positive effect on student results.
7. The NDOE should investigate the gender issues that are affecting students' general education and results. This is because students' gender plays a significant role in PNG's education system due to the cultural views that general society has on male and female education.
8. The NDOE should empower principals/head teachers of secondary schools to assist and develop their teachers with strategies on how to motivate students towards mathematics. Teachers should be encouraged through professional development on how to consistently engage and motivate students' in mathematics as this study found that students' motivation towards mathematics has a significant positive effect on their outcomes.
9. The NDOE should encourage PNG teachers to use more teaching methods that are student-centred and both teacher and student-centred as this study clearly shows that these methods have significant effects on students' mathematics results.
10. The NDOE should have a mechanism in place to overcome complex teaching and learning issues that are interrelated from school, teacher and student levels. This is to avoid educational issues being seen at only one level of the hierarchy and drawing substantive conclusions from this limited perspective.
11. The PNG government should take part in well-recognised and established large-scale international studies such as the Trends in International Mathematics and Science study (TIMSS) and Programme for International Student Assessment (PISA). These two international studies are the

benchmarks used internationally, and PNG should be part of it in order to compare the progress in PNG with other countries around the world. This can be an opportunity for PNG to learn from other countries and gain from their experiences. This is because the factors identified in this study are similar to the factors highlighted by TIMSS and PISA studies.

12. Community programs that help parents encourage their children towards education (not just mathematics) should be initiated and put in place.
13. Teaching and learning are complex issues and the NDOE should encourage educational policies to be informed by rigorous empirical studies in the context of PNG.

13.6 Conclusion

The results obtained from this study have shown that factors such as mothers' highest education level, fathers' occupation, teachers' mathematics major, private schools, students' gender, attitude, motivation and student centred-teaching methods are found to have significant ($p<0.05$) positive effects on Grade 10 and 12 students' mathematics results. Among these significant factors identified in this study, mothers' highest education level ($\beta=3.75$) stands out as the most critical factor influencing students' mathematics results. In addition, multiple interrelationships have also been found among the student-level factors and the teacher-level variables.

Identification of the student and teacher-level factors influencing students' mathematics results in this study contributes to the knowledge and theoretical literature in the area of mathematics education. Moreover, this empirical study used rigorous statistical techniques such as Rasch analysis, SEM and HLM that provided more valid results and further confirms that the model is appropriate for analysing mathematics teaching and learning approaches in this context. Overall, the findings of this study offer a meaningful understanding of the student and teacher-level factors, and their interrelationships, that affect Grade 10 and 12 students' mathematics results in Port Moresby, PNG.

References

Abd-El-Fattah, S. M., Al-Sinani, Y., El Shourbagi, S., & Fakhroo, H. A. (2014). Using Rasch Analysis to Examine the Dimensionality Structure and Differential Item Functioning of the Arabic Version of the Perceived Physical Ability Scale for Children. *Australian Journal of Educational & Developmental Psychology, 14*, 29-44. Retrieved from https//www.newcastle.edu.au/journal/ajedp/

Abdulkadiroğlu, A., Pathak, P. A., & Walters, C. R. (2018). Free to choose: can school choice reduce student achievement? *American Economic Journal: Applied Economics, 10*(1), 175-206. doi:10.1257/app.20160634

Abou El-Komboz, B., Zeileis, A., & Strobl, C. (2014). Detecting Differential Item and Step Functioning with Rating Scale and Partial Credit Trees. *Department of Statistics*: *Technical Reports,* No.152, Department of Statistics University of Munich. doi: 10.5282/ubm/epub.17984

Abuya, B. A., Mutisya, M., & Ngware, M. (2015). Association between mothers' education and grade six children numeracy and literacy in Kenya. *Education 3-13, 43*(6), 653-665. doi: 10.1080/03004279.2013.855250

Adams, R., & Wu, M. (2010). *Multidimensional models*. Australian Council of Educational Research. Retrieved from https://www.acer.org/files/Conquest-Tutorial-7-MultidimensionalModels.pdf

Adams, R.J.& Wu, M.(Eds).(2002).*PISA 2000 Technical Report*.Paris,France:OECD Publishing.

Adedoyin, O., & Mokobi, T. (2013). Using IRT psychometric analysis in examining the quality of junior certificate mathematics multiple choice examination test items. *International Journal of Asian Social Science, 3*(4), 992-1011. Retrieved from http://www.aessweb.com/journal-detail.php?id=5007

Agirdag, O., Driessen, G., & Merry, M. S. (2017). The Catholic school advantage and common school effect examined: a comparison between Muslim immigrant and native pupils in Flanders. *School Effectiveness and School Improvement, 28*(1), 123-135. doi:10.1080/09243453.2016.1251469

Ahmed, W., & Mudrey, R. R. (2019). The role of motivational factors in predicting STEM career aspirations. *International Journal of School & Educational Psychology, 7*(3), 201-214. doi: 10.1080/21683603.2017.1401499

Aina, J. A., & Adekanye, E. A. (2013). Audio-visual resources availability and use for library services among colleges of education in Lagos State Nigeria. *International Journal of Library and Information Science, 5*(10), 417-425. doi: 10.5897/IJLIS2013.0384

Aji, C. A., & Khan, M. J. (2019). The Impact of Active Learning on Students' Academic Performance. *Open Journal of Social Sciences, 7*(3), 204-211. doi: 10.4236/jss.2019.73017

Akyuz, D., Dixon, J. K., & Stephan, M. (2013). Improving the quality of mathematics teaching with effective planning practices. *Teacher Development, 17*(1), 92-106. doi:10.1080/13664530.2012.753939

Alagumalai, S., & Curtis, D. D. (2005). Classical test theory.In S. Alagumalai, D.D Curtis, & N. Hungi (Eds.), *Applied Rasch measurement: A book of exemplars* (pp. 1-14). Netherlands: Springer.

Alagumalai, S., Curtis, D. D., & Hungi, N. (2005). Our experience and conclusions. In S. Alagumalai, D.D Curtis, & N. Nungi (Eds.), *Applied Rasch measurement: A book of exemplars* (pp 343-346). Netherlands: Springer.

Alton-Lee, A. (2003). *Quality teaching for diverse students in schooling: Best evidence synthesis*: Ministry of Education. Wellington, New Zealand:. Retrieved from https://www.minedu.gov.nz

Ambrosetti A., Capeness R., Kriewaldt J., Rorrison D. (2018). Educating Future Teachers: Insights, Conclusions and Challenges. In: Kriewaldt J., Ambrosetti A., Rorrison D., Capeness R. (eds) Educating Future Teachers: Innovative Perspectives in Professional Experience. Springer, Singapore. Retrived from https://doi-org.proxy.library.adelaide.edu.au/10.1007/978-981-10-5484-6_14

Ampadu, E. (2012). Students' Perceptions of their Teachers' Teaching of Mathematics: The Case of Ghana. *International Online Journal of Educational Sciences, 4*(2). Retrieved from https://librarysearch.adelaide.edu.au

Andersen, E. B. (1997). The rating scale model. *Handbook of modern item response theory* . Springer New York.

Anderson, P., & Morgan, G. (2008). Developing tests and questionnaires for a national assessment of educational achievement. *In Developing tests and questionnaires for a national assessment of educational achievement* (Vol. 2). World Bank. doi.10.1596/978-0-8213-7497-9

Andrich, D. (1978). Rating formulation for ordered response categories.Psychometrika,*43*(4),561-573.doi:10.1007/BF02293814

Andrich, D., & Hagquist, C. (2015). Real and artificial differential item functioning in polytomous items. *Educational and Psychological Measurement, 75*(2), 185-207. doi.10.1177/0013164414534258

Andrich, D., & Marais, I. (2014). Person proficiency estimates in the dichotomous Rasch model when random guessing is removed from difficulty estimates of multiple choice items. *Applied Psychological Measurement, 38*(6), 432-449. doi:10.1177/0146621614529646

Andrich, D., & Marais, I. (2019). The Idea of Measurement. *In A Course in Rasch Measurement Theory* (pp. 3–11). Springer Singapore. doi.10.1007/978-981-13-7496-8_1

Andrich, D., & Styles, I. (2004). *Final Report on the Psychometric Analysis of the Early Development Instrument EDI) Using the Rasch Model: A Technical Paper Commissioned for the Development of the Australian Early Development Instrument (AEDI)*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.495.5875&rep=rep1&type=pdf

Appova, A., & Taylor, C. E. (2019). Expert mathematics teacher educators' purposes and practices for providing prospective teachers with opportunities to develop pedagogical content knowledge in content courses. *Journal of Mathematics Teacher Education, 22*(2), 179-204. doi:10.1007/s10857-017-9385-z

Archibald, M. M. (2016). Investigator triangulation: A collaborative strategy with potential for mixed methods research. *Journal of Mixed Methods Research, 10*(3), 228-250. doi:10.1177/1558689815570092

Australian Institute and Leadership School. (2011). Australian profesional standards for teachers: Retrieved from https://www.aitsl.edu.au/docs/default-source/national-policy-framework/australian-professional-standards-for-teachers

AusAid report. (2018). Aid program performance report 2017-2018. Papua New Gunea. Retrieved from https://www.dfat.gov.au/sites/default/files/papua-new-guinea-appr-2017-18.

Aytekin, E., & Isiksal-Bostan, M. (2019). Middle school students' attitudes towards the use of technology in mathematics lessons: does gender make a difference? *International Journal of Mathematical Education in Science and Technology, 50*(5), 707-727.doi: 10.1080/0020739X.2018.1535097

Baghaei, P., & Amrahi, N. (2011). Validation of a Multiple Choice English Vocabulary Test with the Rasch Model. *Journal of Language Teaching & Research, 2*(5).doi: 10.4304/jltr.2.5.1052-1060

Baliyan, S., Rao, K. M., & Baliyan, P. (2012). Influence of parental education and income level on students' performance in senior secondary school mathematics in Botswana. *Global Research Journal on Mathematics and Science Education, 1*(2), 135-158.

Bamberger, M. (2012). Introduction to mixed methods in impact evaluation. *Impact Evaluation Notes, 3*(3), 1-38.

Bansilal, S. (2015). A Rasch analysis of a Grade 12 test written by mathematics teachers. *South African Journal of Science, 111*(5/6), 68-69. doi:10.17159/sajs.2015/20140098

Barkley, E. F. (2018). Terms of Engagement: Understanding and Promoting Student Engagement in Today's College Classroom. *In Deep Active Learning* (pp. 35–57). Springer Singapore. doi:10.1007/978-981-10-5660-4_3

Barreiro, P. L., & Albandoz, J. P. (2001). Population and sample. Sampling techniques. *Management mathematics for European schools, 1*(1), 1-18. Retrieved from http://www.mathematik.unikl.de/mamaeusch

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42*(5), 815-824. doi:10.1016/j.paid.2006.09.018

Barry, A. (2019). Gender Differences in Academic Achievement in Saudi Arabia: A Wake-Up Call to Educational Leaders. *International Journal of Education Policy and Leadership, 15*(15). doi:10.22230/ijepl.2019v15n15a890

Basso, A. (2019). Mathematics Anxiety in Pre-Service Elementary Teachers and its Impact on their Teaching. 1-6. Retrived from https://doi.org/10.31235/osf.io/536qg

Baumert, J., Blum, W., & Neubrand, M. (2004). Drawing the lessons from PISA 2000. *Zeitschrift für Erziehungswissenschaft, 7*(3), 143-157.

Beccaria, L., Beccaria, G., & McCosker, C. (2018). A confirmatory factor analysis of the student evidence-based practice questionnaire (S-EBPQ) in an Australian sample. *Nurse Education Today*, *62*, 69–73. doi:10.1016/j.nedt.2017.12.010

Belfield, C. R., & Levin, H. M. (2015). Modeling School Choice: A Comparison of Public, Private–Independent, Private–Religious, and Home-Schooled Students *Privatizing Educational Choice* (pp. 89-102): Routledge.doi: 10.4324/9781315632575-11

Ben, F. (2010). *Students' uptake of physics: a study of South Australian and Filipino physics students.* (Unpublished Doctoral thesis, University of Adelaide, Australia). Retrieved from https://librarysearch.adelaide.edu.au

Benken, B. M., Ramirez, J., Li, X., & Wetendorf, S. (2015). Developmental mathematics success: Impact of students' knowledge and attitudes. *Journal of developmental education, 38*(2), 14-31.

Benson, N. F., Beaujean, A. A., Donohue, A., & Ward, E. (2016). W Scores: Background and Derivation. *Journal of Psychoeducational Assessment*, doi.0734282916677433.

Berends, M., & Waddington, R. J. (2018). School choice in Indianapolis: Effects of charter, magnet, private, and traditional public schools. *Education Finance and Policy, 13*(2), 227-255. doi: 10.1162/edfp_a_00225

Bertoli-Barsotti, L. (2005). On the existence and uniqueness of JML estimates for the partial credit model. *Psychometrika, 70*(3), 517-531. doi:10.1007/s11336-001-0917-0

Biau, D. J., Jolles, B. M., & Porcher, R. (2010). P value and the theory of hypothesis testing: an explanation for new researchers. *Clinical Orthopaedics and Related Research, 468*(3), 885-892.doi 10.1007/s11999-009-1164-4.

Blackwell, M., Honaker, J., & King, G. (2017). A unified approach to measurement error and missing data: Details and extensions. *Sociological Methods & Research, 46*(3), 342-369.doi: 10.1177/0049124115589052

Boman, P., Curtis, D., Furlong, M. J., & Smith, D. C. (2006). Cross-validation and Rasch analyses of the Australian version of the Multidimensional School Anger inventory–Revised. *Journal of Psychoeducational Assessment, 24*(3), 225-242.doi: 10.1177/0734282906288472

Bond, T. G & Fox, C. M. (2007). Applying the Rasch model. *Fundamental measurement in the human sciences* (1st Ed.). New York: Taylor & Francis Group, LLC.

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*: (2nd ed.).Psychology Press.

Bond, T. G., & Fox, C. M. (2016). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York and London: Routledge.

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*: Routledge.

Boone, W. J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education, 90*(2), 253-269.doi: 10.1002/sce.20106

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Rasch Analysis in the Human Sciences. In *Rasch Analysis in the Human Sciences* . Springer Netherlands. doi.10.1007/978-94-007-6857-4

Boone, W. J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education, 90*(2), 253-269. doi:10.1002/sce.20106

Boone, W. J., Townsend, J. S., & Staver, J. (2010). Using Rasch theory to guide the practice of survey development and survey data analysis in science education and to inform science reform efforts: An exemplar utilizing STEBI self-efficacy data. *Science Education, 95*(2), 258-280. doi: 10.1002/sce.20413

Boone, W. J., Townsend, J. S., & Staver, J. (2010). Using Rasch theory to guide the practice of survey development and survey data analysis in science education and to inform science reform efforts: An exemplar utilizing STEBI self-efficacy data. *Science Education, 95*(2), 258-280. doi:10.1002/sce.20413

Brandenberger, C. C., Hagenauer, G., & Hascher, T. (2018). Promoting students' self-determined motivation in maths: results of a 1-year classroom intervention. *European Journal of Psychology of Education, 33*(2), 295-317.doi. I 10.1007/s10212-017-0336-y

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101. doi:0.1191/1478088706qp063oa

Breakspear, S. (2014). *How does PISA shape education policy making? Why how we measure learning determines what counts in education.* Paper presented at the Centre for Strategic Education Seminar Series Paper (pp.1-13).Melbourne, Australia. Retrieved from https:// http://www.cse.edu.au/

Brodersen, J., Meads, D., Kreiner, S., Thorsen, H., Doward, L., & McKenna, S. (2007). Methodological aspects of differential item functioning in the Rasch model. *Journal of Medical Economics, 10*(3), 309-324. doi: 10.3111/13696990701557048.

Brown, L. (2013).*Teaching Secondary Mathematics as if the Planet Matters* : Routledge. New York, USA.

Brown, L. I., & Kanyongo, G. Y. (2010). Gender Differences in Mathematics Performance in Trinidad and Tobago: Examining Affective Factors. *International Electronic Journal of Mathematics Education, 5*(3), 115-124.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research (Second Edition)*: Guilford Publications.Boston.USA.

Brownlee, J. M., Farrell, A., & Davis, J. (2012). Understanding learning and teaching in Papua New Guinea: Elementary teacher trainers engaged in cultural authorship in the context of national educational reforms. *Australian Journal of Teacher Education, 37*(2), 25-40. doi.10.14221/ajte.2012v37n2.4

Buchanan, J. (2017). How do the standards stand up? Applying quality teacher frameworks to the Australian Professional Standards *Teacher Education Policy and Practice* (pp. 115-128): Springer.Singapore.

Budhwani, H., & De, P. (2017). Human papillomavirus vaccine initiation in Asian Indians and Asian subpopulations: a case for examining disaggregated data in public health research. *Public health, 153*, 111-117.doi: 10.1016/j.puhe.2017.07.036

Butler, J., & Jones, P. (2018). Theoretical and empirical distributions of the p value. *Metron, 76*(1), 1-30. doi:10.1007/s40300-017-0130-2

Byrne, B. M. (1998). Structural equation modeling with LISREL, PRELIS and SIMPLIS:Basic concepts, applications,and programming.Mahmah,New Jersy:Lawrance Erlbaun Associates,Inc.

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*: Routledge.

Byrne, J., & Prendeville, P. (2019). Does a child's mathematical language improve when they engage in cooperative group work in mathematics? *Education 3*(13), 1-15.doi: 10.1080/03004279.2019.1636109

Cai, E. Y. L., & Liem, G. A. D. (2017). 'Why do I study and what do I want to achieve by studying?'Understanding the reasons and the aims of student engagement. *School Psychology International, 38*(2), 131-148. doi: 10.1177/0143034316686399

Cangur, S., & Ercan, I. (2015). Comparison of model fit indices used in structural equation modeling under multivariate normality. *Journal of Modern Applied Statistical Methods, 14*(1), 14. doi: 10.22237/jmasm/1430453580

Chamberlin, S. A., Moore, A. D., & Parks, K. (2017). Using confirmatory factor analysis to validate the Chamberlin affective instrument for mathematical problem solving with academically advanced students. *British Journal of Educational Psychology, 87*(3), 422-437. doi:10.1111/bjep.12157

Chao, T., Chen, J., Star, J. R., & Dede, C. (2016). Using Digital Resources for Motivation and Engagement in Learning Mathematics: Reflections from Teachers and Students. *Digital Experiences in Mathematics Education, 2*(3), 253-277.doi: 10.1007/s40751-016-0024-6

Charalambous, C. Y., Hill, H. C., Chin, M. J., & McGinn, D. (2019). Mathematical content knowledge and knowledge for teaching: exploring their distinguishability and contribution to student learning. *Journal of Mathematics Teacher Education*, *23*:579-613.doi:10.1007/s10857-019-09443-2

Chawla, A., & Saxena, S. (2016). A confirmatory factor analysis of knowledge management assessment instrument in Indian higher educational institutions. *International Journal of Quality & Reliability Management, 33*(7), 1019-1029. doi:10.1108/IJQRM-07-2014-009

Chen, L., Bae, S. R., Battista, C., Qin, S., Chen, T., Evans, T. M., & Menon, V. (2018). Positive attitude toward math supports early academic success: Behavioral evidence and neurocognitive mechanisms. *Psychological science, 29*(3), 390-402. doi:10.1177/0956797617735528

Cheng, L. (2017) *Exploring the Factors that Affect Secondary Student's Mathematics and Portuguese Performance in Portugal*. Masters dissertation Technological University Dublin doi:10.21427/D7P33K

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling, 9*(2), 233-255.doi: 10.1207/S15328007SEM0902_5

Cheung, K. C., Mak, S. K., & Sit, P. S. (2018). Resolving the attitude–achievement paradox based on anchoring vignettes: evidences from the PISA 2012 mathematics study. *Asia Pacific Education Review*(19), 389-399. doi: 10.1007/s12564-018-9526-9

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement, 56*(10), 15-23.doi: 10.1002/pfi.21727

Clarke, B., Clarke, D., & Sullivan, P. (1996). The mathematics teacher and curriculum development *International handbook of mathematics education* (pp. 1207-1233): Springer. Netherlands

Cleary, T. J., & Kitsantas, A. (2017). Motivation and self-regulated learning influences on middle school mathematics achievement. *School Psychology Review, 46*(1), 88-107.

Cleverley, J. (2007). Schooling in Papua New Guinea. *Going to school in Oceania*, Greenwood Publishing Group,Inc.USA.

Coleman, J. S. (2018). *Parents, their children, and schools*: Routledge.

Conroy, J. D. (2017). *Education, employment and migration in Papua New Guinea*: Development Studies Centre, The Australian National University. Canberra, ACT: Australia.

Cordier, R., Speyer, R., Schindler, A., Michou, E., Heijnen, B. J., Baijens, L., . . . Joosten, A. V. (2018). Using Rasch Analysis to Evaluate the Reliability and Validity of the Swallowing Quality of Life Questionnaire: An Item Response Theory Approach. *Dysphagia*, 33:441–456.doi: /10.1007/s00455-017-9873-4.

Craft, A. (2018). *International developments in assuring quality in higher education* : Routledge.

Crenshaw, A. O., Christensen, A., Baucom, D. H., Epstein, N. B., & Baucom, B. R. (2017). Revised scoring and improved reliability for the Communication Patterns Questionnaire. *Psychological assessment, 29*(7), 913-925.doi. 10.1037/pas0000385

Creswell, J. W. (2008). Educational research. *Planning, conducting, and evaluating quantitative and qualitative research*. New Jersey: Pearson Prentice Hall.

Creswell, J. W. (2014). *Research design internatonal student edition: Qualitative, quanatitative, and mixed methods approaches* (4th ed.): SAGE Publications, Inc.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*: London: Sage publications.

Crossley, M., & Sprague, T. (2014). Education for sustainable development: Implications for small island developing states (SIDS). *International Journal of Educational Development, 35*, 86-95. doi: 10.1016/j.ijedudev.2013.03.002

Curtis, D. D. (2005). Comparing classical and contemporary analyses and Rasch measurement Applied Rasch measurement: In S. Alagumalai, D. D. Curtis, & N. Hungi (Eds.), *Applied Rasch measurement: A book of exemplars* (pp. 179-195): Netherlands: Springer.

Curtis, D. D., & Boman, P. (2007). X-ray your data with Rasch. *International Education Journal: Comparative perspectives*, *8* (2) 249-259.

D'Sa, J., & Visbal-Dionaldo, M. (2017). *Analysis of Multiple Choice Questions: Item Difficulty, Discrimination Index and Distractor Efficiency 9* (3).109-114. doi: 10.5958/0974-9357.2017.00079.4

d'Ambrosio, U. (2016). A Response to the Changing Role of Mathematics in Society. *Critical mathematics education: Theory, praxis and reality*, (pp.23-34).Age Publishing Inc.USA.

Daher, W. (2020). Values in the mathematics classroom. *Educational Philosophy and Theory, 52*(3), 284-299. doi:10.1080/00131857.2019.1618276

Dahiru, T. (2008). P-value, a true test of statistical significance? A cautionary note. *Annals of Ibadan postgraduate medicine, 6*(1), 21-26. doi:10.4314/aipm.v6i1.64038

Darmawan, I. G. N. (2003). *Implementation of information technology in local government in Bali, Indonesia*: Flinders University. Adelaide, South Australia: Shannon Research Press.

Darmawan, I. G. N., & Keeves, J. P. (2002). Two-level model of information technology adoption in local government of Bali. *International Education Journal, 3*(1), 47-60. Retrieved from http://dspace.flinders.edu.au/dspac

Darmawan, I. G. N., & Keeves, J. P. (2006). Suppressor viraibles and multilevel mixture modeling. *International Education Journal, 7*(2), 160-173.

Darmawan, I. G. N., & Keeves, J. P. (2009). Using multiple analysis. In C. R . Aldous, I. G. N. Darmawan, & J. P. Keeves (Eds.), *Change over time in numeracy learing and litaracy in rural and remote schools: An investigation into the performance of students and schools in numearcy and litaracy in rural, regional and remote South Australia (SiMERR-SA)* (Vol.2, pp. 48-60): Shannon Research Press.

Davis, B. (2008). Complexity and education: Vital simultaneities. *Educational Philosophy and Theory, 40*(1), 50-65. doi:10.1111/j.1469-5812.2007.00402.x

Dawis, R. V. (2000). Scale construction and psychometric considerations. In H. E. A. Tinsley & S. D. Browns (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 65-94): Academic press. doi:10.1016/B978-012691360-6/50004-5

de Cássia Nakano, T., & Primi, R. (2014). Rasch-master's partial credit model in the assessment of children's creativity in drawings. *The Spanish journal of psychology, 17*(35), 1-16. doi:10.1017/sjp.2014.36

Deieso, D., & Fraser, B. J. (2019). Learning environment, attitudes and anxiety across the transition from primary to secondary school mathematics. *Learning Environments Research, 22*(1), 133-152. doi:10.1007/s10984-018-9261-5

Dekar, Y. (2016). *The Influence of Parents' Educational Background on the Academic Achievement of the Government High School Students in Thimphu.* (Unpublished Doctoral thesis), Mizan-Tepi University, Ethiopia

Delaney, J. M., & Devereux, P. J. (2019). Understanding gender differences in STEM: Evidence from college applications. *Economics of Education Review, 72*, 219-238. doi: 10.1016/j.econedurev.2019.06.002

DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: a large-scale survey. *Assessment in Education: Principles, Policy & Practice, 25*(4), 355-375. doi:10.1080/0969594X.2016.1244514

Diamantopoulos, A., Siguaw, J. A., & Siguaw, J. A. (2000). *Introducing LISREL: A guide for the uninitiated*: Sage.

Dowker, A., Cheriton, O., Horton, R., & Mark, W. (2019). Relationships between attitudes and performance in young children's mathematics. *Educational Studies in Mathematics, 100*(3), 211-230. doi:10.1007/s10649-019-9880-5

Drent, M., Meelissen, M. R., & van der Kleij, F. M. (2013). The contribution of TIMSS to the link between school and classroom factors and student achievement. *Journal of curriculum studies, 45*(2), 198-224. doi: 10.1080/00220272.2012.727872

Durksen, T. L., Way, J., Bobis, J., Anderson, J., Skilling, K., & Martin, A. J. (2017). Motivation and engagement in mathematics: a qualitative framework for teacher-student interactions. *Mathematics Education Research Journal, 29*(2), 163-181. doi: 10.1007/s13394-017-0199-1

Edwards, J. (2015). Gender and Education Assessment, Papua New Guinea: *A review of the literature on girls and education.* Report prepared for the Australian High Commission, Papua New Guinea and the Education Capacity Development Facility (managed by GRM International). Retrieved from https://www.dfat.gov.au/sites/default/files/png-gender-and-education-assessment-review-literature-girls-education.

Eggert, S., & Bögeholz, S. (2010). Students' use of decision-making strategies with regard to socioscientific issues: An application of the Rasch partial credit model. *Science Education, 94*(2), 230-258. doi:10.1002/sce.20358

Eivers, E., & Chubb, E. (2017). *Shaping Schools: What TIMSS tells us about education systems*: Educational Research Centre. ERC Research Series: Retrieved from https://d1wqtxts1xzle7.cloudfront.net/58108935/Shaping-Schools-Series-1

Eklöf, H. (2007a). Self-Concept and Valuing of Mathematics in TIMSS 2003: Scale structure and relation to performance in a Swedish setting. *Scandinavian Journal of Educational Research, 51*(3), 297-313. doi:10.1080/00313830701356141

Eklöf, H. (2007b). Test-taking motivation and mathematics performance in TIMSS 2003. *International Journal of Testing, 7*(3), 311-326.

Eldridge, K., Larry, L., Baird, J., & Kavanamur, D. (2018). A collaborative governance approach to improving tertiary education in Papua New Guinea. *Asia Pacific Journal of Education, 38*(1), 78-90. doi:10.1080/02188791.2018.1423949

Elliott, L., Bachman, H. J., & Henry, D. A. (2020). Why and how parents promote math learning with their young children: A mixed-methods investigation. *Parenting, 20*(2), 108-140. doi: 10.1080/15295192.2019.1694830

Emre-Akdogan, E., & Yazgan-Sag, G. (2019). Transformation of theoretical knowledge into instructional practice: A mathematics teacher's journey. *Issues in Educational Research, 29*(1), 55-69. Retrieved from http://hdl.handle.net/20.500.12485/359

Engelhard, G. J. (2013). *Invariant measurement. Using Rasch models in the social behavioral, and health sciences.* New York and London: Routledge.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences, 116*(12), 5472-5477.doi: 10.1073/pnas.1818430116

Escribano, R., Treviño, E., Nussbaum, M., Irribarra, D. T., & Carrasco, D. (2020). How much does the quality of teaching vary at under-performing schools? Evidence from classroom observations in Chile. *International Journal of Educational Development, 72*, 102125. doi: 10.1016/j.ijedudev.2019.102125

Evans, D. K., & Popova, A. (2015). *What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews*: The World Bank. doi:10.1596/1813-9450-7203

Fafunwa, A. B. (2018). *History of education in Nigeria*: Routledge.

Faize, F. A., & Dahar, M. A. (2011). Effect of mother's level of education on secondary grade science students in Pakistan. *Research Journal of International Studies, 19*, 13-19.

Fan, H., Xu, J., Cai, Z., He, J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year meta-analysis, 1986–2015. *Educational Research Review, 20*, 35-54. doi: 10.1016/j.edurev.2016.11.003

Farooq, M. S., & Shah, S. Z. U. (2008). Students' attitude towards mathemtaics. *Pakistan Economic and Social Review*, *46*(1), 75-83. Retrieved from http://www.jstor.org/stable/25825325

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9. doi:10.1016/j.learninstruc.2013.07.001

Fleming, D. J., Lavertu, S., & Crawford, W. (2018). High School Options and Post-Secondary Student Success: The Catholic School Advantage. *Journal of Catholic Education, 21*(2), 1-25.doi: 10.15365/joce.2102012018

Fleming, J., Gibson, R., Anderson, M., & Martin, A. J. (2020). Quality teaching frameworks and arts education: Seeking a compatible approach. *Journal of curriculum studies*, 1-17. .doi: 10.1080/00220272.2020.1720299

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research, 74*(1), 59-109.doi: 10.3102/00346543074001059

Fredricks, J. A., Hofkens, T., Wang, M. T., Mortenson, E., & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching, 55*(2), 271-298.doi: 10.1002/tea.21419

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence, 20*(2), 507-537. doi: 10.1111/j.1532-7795.2010.00645.x

Fryer Jr, R. G. (2016). Information, non-financial incentives, and student achievement: Evidence from a text messaging experiment. *Journal of Public Economics, 144*, 109-121.doi: 10.1016/j.jpubeco.2016.10.009

Garson, G. D. (2013a). *Hierarchical linear modeling: Guide and applications*: Sage.

Garson, G. D. (2013b). Introductory guide to HLM with HLM 7 software. *Hierarchical linear modeling: Guide and applications*, 55-96. doi.org/10.4135/9781483384450.n3

Gasco, J., & Villarroel, J.-D. (2014). The Motivation of Secondary School Students in Mathematical Word Problem Solving. *Electronic Journal of Research in Educational Psychology, 12*(1), 83-106. doi: 10.14204/ejrep.32.13076

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of educational psychology, 107*(3), 663.doi: 10.1037/edu0000003

Gasparyan, M. V., & Smirnova, N. B. (2015). *Role of mathematics in modern economy of Russia.* Paper presented at the Culture and society: history and present (pp. 183-187). Retrived from https://www.elibrary.ru/item.asp?id=23799362

Geng, W. F. (2014). The Application of Structural Equation Modeling in the Applied Linguistics Research. *Applied Mechanics and Materials,* 687–691, 1532–1535. doi:10.4028/www.scientific.net/amm.687-691.1532

Gherasim, L. R., Butnaru, S., & Mairean, C. (2013). Classroom environment, achievement goals and maths performance: Gender differences. *Educational Studies, 39*(1), 1-12. doi:10.1080/03055698.2012.663480

Glynn, S. M. (2012). International assessment: A Rasch model and teachers' evaluation of TIMSS science achievement items. *Journal of Research in Science Teaching, 49*(10), 1321-1344. doi:10.1002/tea.21059

Golay, P., Thonon, B., Nguyen, A., Caroline, F., & Favrod, J. (2018). Confirmatory Factor Analysis of the French Version of the Savoring Beliefs Inventory. *Frontiers in Psychology, 9* (181),1-7. Retrieved from https://doi.org/10.3389/fpsyg.2018.00181

Goldstein, H. (2011). *Multilevel statistical models*: John Wiley & Sons.

Gómez, L. E., Arias, B., Verdugo, M. Á., & Navas, P. (2012). Application of the Rasch rating scale model to the assessment of quality of life of persons with intellectual disability. *Journal of Intellectual and Developmental Disability, 37*(2), 141-150. doi: 10.3109/13668250.2012.682647

Gomez, R. (2008). Parent ratings of the ADHD items of the disruptive behavior rating scale: Analyses of their IRT properties based on the generalized partial credit model. *Personality and Individual Differences, 45*(2), 181-186.doi. 10.1016/j.paid.2008.04.001

Gore, J., & Rosser, B. (2020). Beyond content-focused professional development: powerful professional learning through genuine learning communities across grades and subjects. *Professional Development in Education*, 1-15.doi:10.1080/19415257.2020.1725904

Gravemeijer, K., Stephan, M., Julie, C., Lin, F.-L., & Ohtani, M. (2017). What Mathematics Education May Prepare Students for the Society of the Future? *International journal of science and mathematics education, 15*(1), 105-123. doi: I 10.1007/s10763-017-9814-6

Green, C., Eady, M., & Andersen, P. (2018). Preparing quality teachers: Bridging the gap between tertiary experiences and classroom realities. *Teaching & Learning Inquiry, 6*(1), 104-125. Retrieved from https://ro.uow.edu.au/sspapers/3509

Grimbeek, P., & Nisbet, S. (2006). Surveying primary teachers about compulsory numeracy testing: Combining factor analysis with Rasch analysis. *Mathematics Education Research Journal, 18*(2), 27-39.

Grimes, P. W. (1994). Public versus private secondary schools and the production of economic education. *The Journal of Economic Education, 25*(1), 17-30.

Guilera, G., Gómez-Benito, J., & Hidalgo, M. D. (2010). Citation analysis in research on differential item functioning. *Quality & Quantity, 44*(6), 1249-1255. doi:10.1007/s11135-009-9274-3

Guo, X., Zhang, Z., Zhao, R., Wang, G., & Xi, J. (2018). Association between coal consumption and urbanization in a coal-based region: a multivariate path analysis. *Environmental Science and Pollution Research, 25*(1), 533-540.

Guthrie, G. (2014). The failure of progressive paradigm shift in Papua New Guinea. *Papua New Guinea Journal of Education, 41*(1), 3-16. Retrieved from https://pngnri.org/images/Publications/OP---201407---Kukari---PNG-Journal-of-Education

Guthrie, G. (2015). The formalistic education paradigm in Papua New Guinea. *Contemporary PNG Studies, 22*, 33. Retrieved from https://search.informit.com.au/documentSummary;dn=219595271094473;res=IELIND

Hackling, M. W. (2016). Think Piece: Preparing today's children for the workplaces of tomorrow: The critical role of STEM education. *International Journal of Innovation in Science and Mathematics Education 23*(3),61-63.

Hagquist, C., & Andrich, D. (2015). Determinants of artificial DIF–a study based on simulated polytomous data. *Psychological Test and Assessment Modeling, 57*(3), 342-376.

Hagquist, C., & Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and quality of life outcomes, 15*(181),1-8.doi: 10.1186/s12955-017-0755-0

Hailaya, W. M. (2014). *Teacher assessment literacy and student outcomes in the province of Tawi-Tawi, Philippines.* (Doctoral thesis, University of Adelaide, Adelaide SA, Australia). Retrieved from https://librarysearch.adelaide.edu.au/discovery/fulldisplay?docid

Hallinan, M., & Kubitschek, W. N. (2012). A comparison of academic achievement and adherence to the common school ideal in public and Catholic schools. *Sociology of Education, 85*(1), 1-22. doi:10.1177/0038040711431586

Hallman-Thrasher, A., Connor, J., & Sturgill, D. (2019). Strong discipline knowledge cuts both ways for novice mathematics and science teachers. *International journal of science and mathematics education, 17*(2), 253-272. doi: 10.1177/00224 87108324554.

Hambleton, R. K., & Swaminathan, H. (2013). Item Response Theory: Principles and Applications. In *Item Response Theory*. Springer Netherlands.

Hamrouni, A. D., & Akkari, I. (2012). Qualitative analysis using NVivo 7software. *International Journal of Business and Social Science, 3*(4), 189-205.

Han, S., & Carpenter, D. (2014). Construct validation of student attitude toward science, technology, engineering, and mathematics project-based learning: The case of Korean middle grade students. *Middle Grades Research Journal, 9*(3), 27-41.

Harris, D., Black, L., Hernandez-Martinez, P., Pepin, B., Williams, J., & Team, w. t. T. (2015). Mathematics and its value for engineering students: what are the implications for teaching? *International Journal of Mathematical Education in Science and Technology, 46*(3), 321-336. doi:10.1080/0020739X.2014.979893

Harrison, R. L., Reilly, T. M., & Creswell, J. W. (2020). Methodological Rigor in Mixed Methods: An Application in Management Studies. *Journal of Mixed Methods Research*, doi. 1558689819900585.

Hart, T., & Achterman, P. (2017). Qualitative Analysis Software. *The International Encyclopedia of Communication Research Methods*, 1-12. doi: 10.1002/9781118901731.iecrm0194

Hashim Ali, S. A. (2013). *The motivation of lifelong mathematics learning.* Paper presented at the AIP Conference Proceedings. Kuala Lumpur, Malaysia. Retrieved from https://doi.org/10.1063/1.4801203

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: Process versus structural equation modeling. *Australasian Marketing Journal (AMJ), 25*(1), 76-81. doi:10.1016/j.ausmj.2017.02.001

Heikkinen, H. L., Wilkinson, J., Aspfors, J., & Bristol, L. (2018). Understanding mentoring of new teachers: Communicative and strategic practices in Australia and Finland. *Teaching and Teacher Education, 71*, 1-11. doi: 10.1016/j.tate.2017.11.025

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the academy of marketing science, 43*(1), 115-135. doi: 10.1007/s11747-014-0403-8

Herges, R. M., Duffied, S., Martin, W., & Wageman, J. (2017). Motivation and Achievement of Middle School Mathematics Students. *The Mathematics Educator, 26*(1), 83–106

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational psychologist, 41*(2), 111-127. doi: 10.1207/s15326985ep4102_4

Hiebert, J. (2003). Teaching mathematics in seven countries: Results from the TIMSS 1999 video study: *Education Statistics Quarterly,* DIANE Publishing.Retrieved from https://librarysearch.adelaide.edu.au/discovery/openurl?institution.

Hiebert, J., & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS video study. *The Elementary School Journal, 101*(1), 3-20. Retrieved from https://www.jstor.org/stable/1002332

Hine, G. (2016). Investigating declining enrolments in secondary mathematics. Conference paper presented at the Education conference at The University of Notre Dame Australia. Retrived from http://www.merga.net.au/node/37

Hirshberg, M. J., Flook, L., Enright, R. D., & Davidson, R. J. (2020). Integrating mindfulness and connection practices into preservice teacher education improves classroom practices. *Learning and Instruction, 66* (101298). 1-11.doi: 10.1016/j.learninstruc.2019.101298

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management, 24*(5), 623-641. doi:10.1177/014920639802400504

Hogan, A., Thompson, G., Sellar, S., & Lingard, B. (2018). Teachers' and school leaders' perceptions of commercialisation in Australian public schools. *The Australian Educational Researcher, 45*(2), 141-160.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Articles*, *6* (1),53-60.

Hopfenbeck, T. N., & Kjaernsli, M. (2016). Students' test motivation in PISA: The case of Norway. *The Curriculum Journal, 27*(3), 406-422.

Horn, A. S., & Jang, S. T. (2017). The Impact of Graduate Education on Teacher Effectiveness: Does a Master's Degree Matter? MHEC Research Brief. *Midwestern Higher Education Compact*. Retrieved from https://files.eric.ed.gov/fulltext/ED587432.pdf.

Horrocks, C & King,N. (2010*). Interviews in Qualitative Research.*SAGE Publications,New Delhi,India.

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher–Student Dialogue During Classroom Teaching: Does It Really Impact on Student Outcomes? *Journal of the Learning Sciences, 28*(4-5), 462-512.doi: 10.1080/10508406.2019.1573730

Howell, D. C. (2019). Missing Values: How to Treat Them Appropriately. *A Guide to the Scientific Career: Virtues, Communication, Research and Academic Writing*, (pp.639-647).doi: 10.1002/9781118907283.ch69

Howes, S., Mako, A. A., Swan, A., Walton, G., Webster, T., & Wiltshire, C. (2014). A lost decade? Service delivery and reforms in Papua New Guinea 2002 – 2012, The National Research Institute and the Development Policy Centre, Canberra. Retrieved from https://openresearch-repository.anu.edu.au

Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. *The Oxford handbook of quantitative methods, 2*(1), 281-294.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications (*2nd ed.). New york and Hove: Routledge Taylor and Francis Group.

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*: New york and Hove: Routledge Taylor and Francis Group.

Hu, B. Y., Fan, X., Wu, Y., LoCasale-Crouch, J., & Song, Z. (2020). Teacher–child interaction quality and Chinese children's academic and cognitive development: New perspectives from piecewise growth modeling. *Early Childhood Research Quarterly, 51*, 242-255. Doi: 10.1016/j.ecresq.2019.10.003

Huggins-Manley, A. C., & Algina, J. (2015). The Partial Credit Model and Generalized Partial Credit Model as Constrained Nominal Response Models, With Applications in M plus. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(2), 308-318. doi:https://doi.org/10.1080/10705511.2014.937374

Hungi, N. (2005). Employing the Rasch model to detect biased items *Applied Rasch measurement: A book of exemplars* (pp. 139-157): Springer.

Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, 13* (10). Retrieved from http://epaa.asu.edu/epaa/v13n10/.

Innabi, H., & Dodeen, H. (2018). Gender differences in mathematics achievement in Jordan: A differential item functioning analysis of the 2015 TIMSS. *School Science and Mathematics, 118*(3-4), 127-137.doi: 10.1111/ssm.12269

Iqbal, M. Z., Mirza, M. S., & Shams, J. A. (2017). Changes in Students' Beliefs: A Case of Mathematics. *Bulletin of Education & Research, 39*(3), 93-103

Istance, D., & Kools, M. (2013). OECD work on technology and education: Innovative learning environments as an integrating framework. *European Journal of Education, 48*(1), 43-57.

Istance, D., & Paniagua, A. (2018). *Teachers as Designers of Learning Environments: The Importance of Innovative Pedagogies. Educational Research and Innovation*. OECD Publishing. 2, rue Andre Pascal, F-75775 Paris Cedex 16, France. Retrieved from https://eric.ed.gov/?id=ED582804

Jacob, R., Hill, H., & Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness, 10*(2), 379-407. doi:10.1080/19345747.2016.1273411

Johnsen, S. K. (2017). Test of Nonverbal Intelligence: A Language-Free Measure of Cognitive Ability *Handbook of Nonverbal Assessment* (pp. 185-206): Springer.

Joncas, M., & Foy, P. (2011). Sample design in TIMSS and PIRLS. *Methods and Procedures in TIMSS and PIRLS*, 1-21. Retrieved from http://timss.bc.edu/publications/timss

Joskin, A. M. (2013). Investigating the implementation process of a curriculum: A case study from Papua New Guinea. (Doctoral thesis, Victoria University of Wellington, New Zealand). Retrieved from https://www.wgtn.ac.nz/library.

Kadijević, Đ. (2006). Developing trustworthy TIMSS background measures: A case study on mathematics attitude. *The Teaching of Mathematics* (17), 41-51.

Kaleebu, N., Gee, A., Watson, A. H., Jones, R., & Jauk, M. (2017). SMS Story: A Case Study of a Controlled Trial in Papua New Guinea *Mobile Learning in Higher Education in the Asia-Pacific Region* (pp. 623-645): Springer.Singapore.

Kanyongo, G. Y., & Ayieko, R. (2017). Hierarchical Linear Modelling of Sixth-Grade Students' Socio-economic Status and School Factors On Mathematics Achievement: Case studies of Kenya and Zimbabwe. *African Journal of Research in Mathematics, Science and Technology Education, 21*(2), 187-199. doi:1080/18117295.2017.1334368

Kanzow, P., & Wiegand, A. (2020). Teaching of composite restoration repair: Trends and quality of teaching over the past 20 years. *Journal of Dentistry*, 103303. doi: 10.1016/j.jdent.2020.103303

Keeves, J. P., & Masters, G.N. (1999). Issues in educational measurement. In G.N. Masters & J.P.Keeves (Eds,), *Advances in measurement in educational research and assessment*, (pp.268-281). The Netherlands: Pergamon.

Kelcey, B., Hill, H. C., & Chin, M. J. (2019). Teacher mathematical knowledge, instructional quality, and student outcomes: a multilevel quantile mediation analysis. *School Effectiveness and School Improvement*,*30* (4), 398-431, doi: 10.1080/09243453.2019.1570944

Kennedy, J. P., Lyons, T., & Quinn, F. (2014). The continuing decline of science and mathematics enrolments in Australian high schools. *Teaching Science, 60*(2), 34-46. Retrieved from https://eprints.qut.edu.au/73153/

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural equation modeling, 10*(3), 333-351.doi: 10.1207/S15328007SEM1003_1

Khan, R. M. A., Iqbal, N., & Tasneem, S. (2015). The Influence of Parents Educational Level on Secondary School Students Academic Achievements in District Rajanpur. *Journal of Education and Practice, 6*(16), 76-79.

Khine, M. S. (2013). *Application of structural equation modeling in educational research and practice* (Vol. 7): Springer.

Khine, M. S., Fraser, B. J., Afari, E., Oo, Z., & Kyaw, T. T. (2018). Students' perceptions of the learning environment in tertiary science classrooms in Myanmar. *Learning Environments Research, 21*(1), 135-152. Doi:10.1007/s10984-017-9250-0

Khun-Inkeeree, H., Omar-Fauzee, M. S., & Othman, M. K. H. (2016). Students' attitude towards achievment in mathematics: A cross sectional study of year 6 students in Songklha Province,Thailand. *European Journal of Education Studies, 2*(4), 89-99. doi:10.5281/zenodo.154048

Kiemer, K., Gröschner, A., Pehmer, A.-K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and instruction, 35*, 94-103. doi:10.1016/j.learninstruc.2014.10.003

Kim, C. (2007). *Effects of motivation, volition, and belief change strategies on attitudes, study habits, and achievement in mathematics education*. (Doctoral thesis, The Florida State University) . Retrieved from https://www.learntechlib.org/p/124729/

Kim, C., Park, S. W., Cozart, J., & Lee, H. (2015). From Motivation to Engagement: The Role of Effort Regulation of Virtual High School Students in Mathematics Courses. *Educational Technology & Society, 18*(4), 261-272.

Kim, H.-J., & Abraham, I. (2017). Measurement of fatigue: Comparison of the reliability and validity of single-item and short measures to a comprehensive measure. *International journal of nursing studies, 65*, 35-43.doiu: 10.1016/j.ijnurstu.2016.10.012

Kiss, A. (2018). Investigating Young Children's Attitudes toward Mathematics: Improved Measurement and the Relation to Achievement (Doctoral thesis, The University of Minnesota, USA) Retrieved from http://hdl.handle.net/11299/199057.

Kiwanuka, H. N., Van Damme, J., Anumendem, D. N., Van Den Noortgate, W., & Namusisi, S. (2015). Factors affecting Mathematics achievement of first-year secondary school students in Central Uganda. *South African Journal of Education, 35*(3), 1-16.doi: 10.15700/saje.v35n3a1106

Kiwanuka, H. N., Van Damme, J., Van Den Noortgate, W., Anumendem, D. N., Vanlaar, G., Reynolds, C., & Namusisi, S. (2017). How do student and classroom characteristics affect attitude toward mathematics? A multivariate multilevel analysis. *School Effectiveness and School Improvement, 28*(1), 1-21. doi: 10.1080/09243453.2016.1201123

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd edition.). New York and London: Guilford Press.

Kline, P. (2014). *An easy guide to factor analysis*: Routledge.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*: Guilford publications.

Korbel, P. W. (2015). *Students' experiences of mathematics and how they affect further engagement with mathematics.* (Doctoral thesis, University of Adelaide, Australia). Retrieved from https://librarysearch.adelaide.edu.au

Kosgei, A., Mise, J. K., Odera, O., & Ayugi, M. E. (2013). Influence of teacher characteristics on students' academic achievement among secondary schools. *Journal of Education and Practice, 4*(3), 76-82.

Kozlowski, K. P. (2016). *Doing school: Learning behavior, classrom interactions, and the racial achievement gap.* (Doctoral thesis, The University of North Carolina, USA) Retrieved from https://doi.org/10.17615/2eda-6983

Kravia, G., & Owens, K. (2014). Design-Based Research for Professional Learning for" Cultural Mathematics". *Mathematics Education Research Group of Australasia*. Paper presented at the Annual Meeting of the Mathematics Education Research Group of Australasia (MERGA) (37th, Sydney, New South Wales, Australia). Retrieved from https://eric.ed.gov/?id=ED572631.

Krosnick, J. A. (2018). Questionnaire design *The Palgrave Handbook of Survey Research* (pp. 439-455): Springer.

Kyriakides, L., Creemers, B. P., & Panayiotou, A. (2018). Using educational effectiveness research to promote quality of teaching: The contribution of the dynamic model. *ZDM*, 50:381–393. doi:10.1007/s11858-018-0919-3

Lagani, W., Mokela, D., Saweri, W., Kiromat, M., Ripa, P., Vince, J.,Subhi, R. (2010). Papua New Guinea: real progress towards MDG 4 and real challenges. *International health, 2*(3), 186-196. doi: 10.1016/j.inhe.2010.05.001

Lahdenperä, J., Postareff, L., & Rämö, J. (2019). Supporting quality of learning in university mathematics: A comparison of two instructional designs. *International Journal of Research in Undergraduate Mathematics Education, 5*(1), 75-96. doi:10.1007/s40753-018-0080-y

Lamb, R. L., Annetta, L., Meldrum, J., & Vallett, D. (2012). Measuring science interest: Rasch validation of the science interest survey. *International journal of science and mathematics education, 10*(3), 643-668. doi. 10.1007/s10763-011-9314-z

Law, H. (2018). Gender and mathematics: Pathways to mathematically intensive fields of study in Australia. *Advances in Life Course Research, 37*, 42-56.doi: 0.1016/j.alcr.2018.07.002

Le Fanu, G., & Kelep-Malpo, K. (2015). Papua New Guinea: Inclusive education *Education in Australia, New Zealand and the Pacific* (pp. 219-242): Bloomsbury Academic London.

Le, L. (2006). *Analysis of differential item functioning.* Paper presented at the annual meeting of American Educational Research Association, San Francisco CA. Retrieved from https://www.acer.org/files/analysis

Leder, G. C., Forgasz, H. J., & Solar, C. (1996). Research and intervention programs in mathematics education: A gendered issue *International handbook of mathematics education* (pp. 945-985): Springer. Retrieved from https://doi.org/10.1007/978-94-009-1465-0_26

Lee, S. W. (2018). Pulling back the curtain: Revealing the cumulative importance of high-performing, highly qualified teachers on students' educational outcome. *Educational Evaluation and Policy Analysis, 40*(3), 359-381. doi: 10.3102/0162373718769379

Lefebvre, P., Merrigan, P., & Verstraete, M. (2011). Public subsidies to private schools do make a difference for achievement in mathematics: Longitudinal evidence from Canada. *Economics of Education Review, 30*(1), 79-98. doi:10.1016/j.econedurev.2010.07.005

Leke, D. K. (2010). *Vocational education and training in secondary schools: investigation into a VET model for Papua New Guinea (Doctoral thesis,University of Adelaide,Australia).* Retrieved from https://librarysearch.adelaide.edu.au

Leong, Y. H., Kaur, B., Lee, N. H., & Toh, T. L. (2019). Exemplary Practices of Mathematics Teachers *Mathematics Education in Singapore* (pp. 385-404): Springer.doi: 10.1007/978-981-13-3573-0_16

Leppink, J., & Pérez-Fuster, P. (2017). We need more replication research–A case for test-retest reliability. *Perspectives on medical education, 6*(3), 158-164. doi: 10.1007/s40037-017-0347-z

Li, M., Zhang, Y., Liu, H., & Hao, Y. (2018). Gender differences in mathematics achievement in Beijing: A meta-analysis. *British Journal of Educational Psychology, 88*(4), 566-583. doi:10.1111/bjep.12203

Liao, F., Murphy, D., & Barrett-Lennard, G. (2018). A confirmatory factor analysis of the Mandarin-Chinese version of the Barrett-Lennard Relationship Inventory. *Asia Pacific Journal of Counselling and Psychotherapy, 9*(1), 20-45. doi:10.1080/21507686.2017.1419269

Lim, D. (2018). *Quality Assurance in Higher Education: A Study of Developing Countries: A Study of Developing Countries*: Routledge.

Linacre, J. (2002). What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions, 16 (2), 878.

Linacre, J. (2009). The efficacy of Warm's weighted mean likelihood estimate (WLE) correction to maximum likelihood estimate (MLE) bias. *Rasch Measurement Transactions, 23*(1), 1188-1189.

Liou, P.Y., & Jessie Ho, H.-N. (2018). Relationships among instructional practices, students' motivational beliefs and science achievement in Taiwan using hierarchical linear modelling. *Research Papers in Education, 33*(1), 73-88. doi: 10.1080/02671522.2016.1236832

Los, R. E. (2016). The interaction between student motivation and the instructional environment on academic outcome: A hierarchical linear model. *Social Psychology of Education* 22:471-500.doi:10.1007/s11218-019-09487-5

Luke, D. A. (2004). *Multilevel modeling* (Vol. 143): Sage Publication Inc, New York.

Lye, C. Y. (2016). *Towards a comprehensive model of English as a second language learning and teaching in the Malaysian independent Chinese secondary schools.* (Doctoral thesis, University of Adelaide, Australia). Retrieved from https://librarysearch.adelaide.edu.au/discovery/fulldisplay?docid

MacCallum, R. C. (1995). *Model specification: Procedures, strategies, and related issues.* In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (p. 16–36). Sage Publications, Inc.

Macdonald, C. (2016). The political economy of mineral resource governance and children's rights in Papua New Guinea, UNRISD Working Paper, No. 2016-16, United Nations Research Institute for Social Development (UNRISD), Geneva.Retrieved from http://hdl.handle.net/10419/186106

Mack, M. R., Stanich, C. A., & Goldman, L. M. (2019). Math Self-Beliefs Relate to Achievement in Introductory Chemistry Courses *It's Just Math: Research on Students' Understanding of Chemistry and Mathematics* (pp. 81-104): ACS Publications.doi: 10.1021/bk-2019-1316.ch006

Magis, D., & Verhelst, N. (2017). On the finiteness of the weighted likelihood estimator of ability. *Psychometrika, 82*(3), 637-647. doi: 10.1007/s11336-016-9518-9

Malik, R. H., & Rizvi, A. A. (2018). Effect of Classroom Learning Environment on Students' Academic Achievement in Mathematics at Secondary Level. *Bulletin of Education and Research, 40*(2), 207-218.

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Craven, R. G. (2020). Confirmatory factor analysis (cfa), exploratory structural equation modeling (esem), and set-esem: Optimal balance between goodness of fit and parsimony. *Multivariate behavioral research, 55*(1), 102-119. doi:10.1080/00273171.2019.1602503

Martin, M. O., & Kelly, D. L. (1996). *Third International Mathematics and Science Study. Technical Report. Volume I: Design and Development*: ERIC. Boston College, Center for the Study of Testing, Evaluation. Retrieved from http://wwwcsteep.bc.edu/timss

Martin, O., & Mullis, I. V. S. (2013). Relationships among reading, mathematics and science achievement at the fourth grade-implications for early learning. *Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College*.

Martin, M. O., Mullis, I. V., Foy, P., & Arora, A. (2011). Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales. *Methods and Procedures in TIMSS and PIRLS*, 1-11. Retrieved from timssandpirls.bc.edu/methods/pdf/TP11_Context_Q_Scales.pdf.

Martin, M. O., Mullis, I. V., & Hooper, M. (2017). Methods and Procedures in PIRLS 2016. *International Association for the Evaluation of Educational Achievement*. Retrieved from https://eric.ed.gov/?id=ED580352.

Martin, O., Mullis, I. V., & Foy, P. (2015). TIMSS 2015 assessment design. *TIMSS*, 85-99. Retrieved from https://timssandpirls.bc.edu/timss2015.

Martynova, E., West, S. G., & Liu, Y. (2017). Review of Principles and Practice of Structural Equation Modeling, Structural Equation Modeling: A Multidisciplinary Journal, 25(2), 325-329, doi: 10.1080/10705511.2017.1401932

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrica, 47*(2), 149-174.

Matane, P. (1986). Ministerial committee report: *A philosophy of education for Papua New Guinea*: Ministerial Report. Port Moresby: University of Papua New Guinea. PNG. Retrieved from https://searchworks.stanford.edu/view/4050063.

Matang, R. (2002). The role of ethnomathematics in mathematics education in Papua New Guinea: Implications for mathematics curriculum. *Directions: Journal of Educational Studies, 24*(1), 27-37.

McCormick, A. (2011). Some partners are more equal than others: EFA and Civil Society in Papua New Guinea and Vanuatu education policy processes. *International Education Journal: Comparative Perspectives, 10*(2). Retrieved from http://iejcomparative.org

McMillan, J. H. (2012). *SAGE handbook of research on classroom assessment*: Sage Publications. Los Angeles, London, New Delhi, Singapore, Washington DC

Meade, A. W., & Fetzer, M. (2009). Test bias, differential prediction, and a revised approach for determining the suitability of a predictor in a selection context. *Organizational Research Methods, 12*(4), 738-761. doi:10.1177/1094428109331487

Mel, M. A. (2007). Quality assurance and assessment in education in Papua New Guinea. *Educational Research for Policy and Practice, 6*(3), 219-228. doi: 10.1007/s10671-007-9032-z

Mertens, W., Pugliese, A., & Recker, J. (2016). *Quantitative Data Analysis*: A Companion for Accounting and Information Systems Research. In *Quantitative Data Analysis*. Springer International Publishing AG. doi.10.1007/978-3-319-42700-3

Meyer, H.-D., & Benavot, A. (2014). PISA, power, and policy: The emergence of global educational governance*. International Rev Education. 60*:591–596.doi: 10.1007/s11159-014-9429-x

Miserandino, A. (2019). The funding and future of Catholic education in the United States. *British Journal of Religious Education, 41*(1), 105-114. doi.10.1080/01416200.2017.1352484

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series, 17*(4), 59-82.

Mohamad, M. M., Sulaiman, N. L., Sern, L. C., & Salleh, K. M. (2015). Measuring the validity and reliability of research instruments. *Procedia-Social and Behavioral Sciences, 204*, 164-171.doi. 10.1016/j.sbspro.2015.08.129

Mohammadpour, E., & Shekarchizadeh, A. (2015). Mathematics achievement in high-and low-achieving secondary schools. *Educational Psychology, 35*(6), 689-713. doi.org/10.1080/01443410.2013.864753

Moloney, R., & Xu, H. L. (2018). *Teaching and Learning Chinese in Schools Case Studies in Quality Language Education* (1st ed. 2018.). Springer International Publishing. doi:10.1007/978-3-319-89372-3

Mueller, G., Mylonas, D., & Schumacher, P. (2018). Quality assurance of the clinical learning environment in Austria: Construct validity of the Clinical Learning Environment, Supervision and Nurse Teacher Scale (CLES+T scale). *Nurse Education Today*, *66*, 158–165. doi:10.1016/j.nedt.2018.04.02

Mujtaba, T., Reiss, M. J., & Hodgson, A. (2014). Motivating and supporting young people to study mathematics: A London perspective. *London Review of Education, 12*(1), 121-142. doi.org/10.18546/LRE.12.1.11

Mullis, I. V., & Chrostowski, S. J. (2004). *TIMSS 2003 Technical Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*: Lynch School of Education, Boston College. Retrived from https://eric.ed.gov/?id=ED494653.

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics* (pp. 139-171). Chestnut Hill, MA: TIMSS & PIRLS International Study Center. Retrived from https://eric.ed.gov/?id=ED494653.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counseling psychology, 38*(1), 30. doi: 10.1037/0022-0167.38.1.30

Murray, S. (2011). Declining participation in post-compulsory secondary school mathematics: students' views of and solutions to the problem. *Research in Mathematics Education, 13*(3), 269-285.doi: 10.1080/14794802.2011.624731

Mursidi, A., & Soeharto, S. (2017). An Introduction: Evaluation Of Quality Assurance For Higher Educational Institutions Using Rasch Model. *JETL (Journal Of Education, Teaching and Learning), 1*(1), 1-6.

Musau, L. M., & Migosi, J. A. (2014). Instructional resources and girls' academic achievement in Science, Mathematics and Technology subjects: experiences from Kenya. *International Journal of Education Economics and Development, 5*(2), 140-151.doi: 10.1504/IJEED.2014.061452

Muthén, L., & Muthén, B. (2003). *Mplus. he comprehensive modeling program for applied researchers user guide.* Los Angeles, CA: Muthén & Muthén.

Nassiri, V., Lovik, A., Molenberghs, G., & Verbeke, G. (2018). On using multiple imputation for exploratory factor analysis of incomplete data. *Behavior research methods, 50*(2), 501-517.doi: 10.3758/s13428-017-1013-4

Neale, D. C. (1969). The role of attitudes in learning mathematics. *The Arithmetic Teacher, 16*(8), 631-640. Retrieved from https://librarysearch.adelaide.edu.au

Neofa, Z. (2010). *A case study of how primary school teachers in Papua New Guinea understand outcomes-based education (* Doctoral thesis, Queensland University of Technology, Australia). Retrieved from https://eprints.qut.edu.au/45482/

Ng, B., Wang, C. J., & Liu, W. C. (2017). Self-regulated learning in Singaporean context: a congeneric approach of confirmatory factor analysis. *International Journal of Research & Method in Education, 40*(1), 91-107. doi:10.1080/1743727X.2015.1075125

Ng, B. L., Liu, W., & Wang, J. C. (2016). Student motivation and learning in mathematics and science: A cluster analysis. *International Journal of Science and Mathematics Education, 14*(7), 1359-1376. doi:10.1007/s10763-015-9654-1

Nguyen, B. M. D., Alcantar, C. M., Curammeng, E. R., Hernandez, E., Kim, V., Paredes, A. D.Teranishi, R. T. (2017). The Racial Heterogeneity Project: Implications for Educational Research, Practice, and Policy. *ACT, Inc.* Retrieved from https://files.eric.ed.gov/fulltext/ED583600

Nicholas, J., Poladian, L., Mack, J., & Wilson, R. (2015). Mathematics preparation for university: entry, pathways and impact on performance in first year science and mathematics subjects. *International Journal of Innovation in Science and Mathematics Education (formerly CAL-laborate International), 23*(1), 37-51.

Niehaus, E., Campbell,C. M., & Inkelas, K.K. (2014). HLM behind the curtain:Unveiling decisions behind the use and interpretation of HLM in higher education research. *Research in Higher Education, 55*(1),101-122.

Norton, S., & Zhang, Q. (2013). Chinese Students' Engagement with Mathematics Learning. *International Journal for Mathematics Teaching & Learning*. 23(1), Retrieved from https://research-repository.griffith.edu.au/handle/10072/58928.

NSW Department of Education and Training (2003). *Quality teaching in NSW schools.* Retrieved from: http://www.darcymoore.net/wpcontent/uploads/2012/02/qt_EPSColor.pdf

Nyman, R. (2017). *Interest and Engagement: Perspectives on Mathematics in the Classroom (*Doctoral thesis, University of Gothenburg Sweden). Retrieved from https://gupea.ub.gu.se/handle/2077/51917

O'Donoghue, T. A. (1994). Transnational knowledge transfer and the need to take cognisance of contextual realities: a Papua New Guinea case study. *Educational review, 46*(1), 73-88. doi:10.1080/0013191940460108

Ode, E. O. (2014). Impact of audio-visual (avs) resources on teaching and learning in some selected private secondary schools in makurdi. *International journal of Research in humanities, arts and literature, 2*(5), 195-202.

OECD (2014). *TALIS 2013 Results: An International Perspective on Teaching and Learning*: OECD Publishing. Retrieved from https://doi.org/10.1787/9789264196261-en.

OECD. (2011). Lessons from PISA for Mexico,Strong Performers and Successful Reformers in Education OECD Publishing, Paris, Retrieved from https://doi.org/10.1787/9789264107243-en.

OECD (2005). PISA 2009 Technical Report: OECD Publishing Paris. Retrieved from https://www.oecd.org/pisa/sitedocument/PISA-2005-technical-report-final

OECD.(2017).PISA 2015 Technical Report: OECD Publishing Paris. Retrieved from https://www.oecd.org/pisa/sitedocument/PISA-2015-technical-report-final

OECD (2004). Learning for Tomorrow's World: First Results from PISA 2003*, OECD Publications, Paris.*

OECD (2009). *Top of the Class-High performers in science in PISA 2006.*Paris, France: OECD Publications.

OECD. (2010). Education Today 2010: The OECD Perspective. Retrieved from *http://www.oecd.org/education/educationtoday2010theoecdperspective.htm*

OECD (2011), Education at a Glance 2011: Highlights, OECD Publishing. doi.org/10.1787/eag_highlights-2011-en

Oermann, M. H., & Gaberson, K. B. (2016). Evaluation and testing in nursing education. In *Evaluation and testing in nursing education* (4th ed.). Springer Publishing Company.

Oktiningrum, W., Zulkardi, Z., & Hartono, Y. (2016). Developing PISA-like Mathematics Task with Indonesia Natural and Cultural Heritage as Context to Assess Students Mathematical Literacy. *Journal on Mathematics Education, 7*(1), 1-8. doi:10.22342/jme.7.1.2812.1-8

Okumu, I. M., Nakajjo, A., & Isoke, D. (2008). Socioeconomic determinants of primary school dropout: the logistic model analysis. Retrieved from: http: //mpra.ub.uni-muenchen.de/7851

Olayinka, A.-R. B. (2016). Effects of Instructional Materials on Secondary Schools Students' Academic Achievement in Social Studies in Ekiti State, Nigeria. *World Journal of Education, 6*(1), 32-39. doi:10.5430/wje.v6n1p32

Oliveira, S. M. d. S., Fernandes, D. C., & Sisto, F. F. (2014). Analysis of the school anxiety inventory in Brazil using the Rasch rating scale model. *Psychological reports, 115*(1), 165-178. doi:10.2466/08.03.10.PR0.115c15z3

Olson, A. M., & Stoehr, K. J. (2019). From numbers to narratives: Preservice teachers experiences' with mathematics anxiety and mathematics teaching anxiety. *School Science and Mathematics, 119*(2), 72-82. doi:10.1111/ssm.12320

Onwuegbuzie, A. J., & Combs, J. P. (2011). Data analysis in mixed research: A primer. *International Journal of Education, 3*(1), 1-25. doi:10.5296/ije.v3i1.618

Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International journal of social research methodology, 8*(5), 375-387. doi: 10.1080/13645570500402447

Onwuegbuzie, A. J., & Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures. *The Qualitative Report : an Online Journal Dedicated to Qualitative Research Since 1990.*, *11*(3). Doi:10.1017/S0018246X06005280

Osteen, P. (2010). An introduction to using multidimensional item response theory to assess latent factor structures. *Journal of the Society for Social Work and Research, 1*(2), 66-82. doi:10.5243/jsswr.2010.6

Owens, K. (2014). The impact of a teacher education culture-based project on identity as a mathematically thinking teacher. *Asia-Pacific Journal of Teacher Education, 42*(2), 186-207. doi:10.1080/1359866X.2014.892568

Owens, K. (2015). Culture at the forefront of mathematics research at the University of Goroka: The Glen Lean Ethnomathematics Centre. *South Pacific Journal of Pure and Applied Mathematics, 2*(3), 117-130.

Owens, K., & Kaleva, W. (2008). *Case studies of mathematical thinking about area in Papua New Guinea.* Paper presented at the PME 32 and PME-NA XXX: Joint Meeting of the Internation Group for the Psychology of Mathematics Education and North American chapter. Retrieved from https://researchoutput.csu.edu.au

Oyerinde, B. A. (2014). *Correlating parental involvement and mathematics achievement of african american eighth-grade students*. (Doctoral thesis, Walden University, USA). Retrieved from https: https://librarysearch.adelaide.edu.au

Özsoy, G., & Ataman, A. (2017). The effect of metacognitive strategy training on mathematical problem solving achievement. *International Electronic Journal of Elementary Education, 1*(2), 67-82.

Paek, I., & Cai, L. (2014). A comparison of item parameter standard error estimation procedures for unidimensional and multidimensional item response theory modeling. *Educational and Psychological Measurement, 74*(1), 58-76. doi.10.1177/0013164413500277

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. *Perception, 11*, 239-266. Retrieved from https://librarysearch.adelaide.edu.au/

Pal Pandi, A., Paranitharan, K., & Jeyathilagar, D. (2018). Implementation of IEQMS model in engineering educational institutions–a structural equation modelling approach. *Total Quality Management & Business Excellence, 29*(1-2), 29-57.doi:10.1080/14783363.2016.1154431

Pantziara, M., & Philippou, G. N. (2015). Students' motivation in the mathematics classroom. Revealing causes and consequences. *International Journal of Science and Mathematics Education, 13*(2), 385-411.

PNG NDOE. (2004). *Achieveing a better future – A National Plan for Education 2005-2014*, Waigani, Papua New Guinea Department of Education. Retriieved from https://planipolis.iiep.unesco.org/en/2004/achieving-better-future-national-plan-education-2005-2014-papua-new-guinea-4678.

PNG NDOE. (2016). National Education Plan 2015-2019: Quality leraning for all. Retrieved from https://www.education.gov.pg/documents/NEP_2015-2019_DoE.pdf

PNG NDOE. (2009). *Upper Secondary Curriculum Implementation handbook.* Retrieved from https://www.education.gov.pg/TISER/documents/curriculum/implementation-support-booklet-upper-secondary.pdf

PNG NDOE. (2006a). *Mathemtaics Lower Secondary Syllabus*. Retrieved from https://www.education.gov.pg/quicklinks/secondary-syllabus/lower/syllabus-lower-secondary-mathematics.pdf

PNG NDOE. (2006b). *Mathematics Upper Secondary Syllabus.* Retrieved from http://www.educationpng.gov.pg/Teachers/secondary/upper/teachers-guide-upper-secondary-advanced-mathematics.pdf

PNG DHERST. (2006). *National Selection Report*, Waigani, Papua New Guinea Office of Higher Education. Retrieved from https://web.dherst.gov.pg

Park, D., Gunderson, E. A., Tsukayama, E., Levine, S. C., & Beilock, S. L. (2016). Young children's motivational frameworks and math achievement: Relation to teacher-reported instructional practices, but not teacher theory of intelligence. *Journal of educational psychology, 108*(3), 300. doi:10.1037/edu0000064

Penfield, R. D., Myers, N. D., & Wolfe, E. W. (2008). Methods for assessing item, step, and threshold invariance in polytomous items following the partial credit model. *Educational and Psychological Measurement, 68*(5), 717-733. doi: 10.1177/0013164407312602

Pennell, C. (2020). *Evaluating the K–12 Literacy Curriculum: A Step by Step Guide for Auditing Programs, Materials, and Instructional Approaches*: Routledge.

Pepin, B., & Roesken-Winter, B. (2015). *From beliefs to dynamic affect systems in mathematics education*: Exploring a mosaic of relationships and interactions. Springer.

Pianta, R. C., Whittaker, J. E., Vitiello, V., Ruzek, E., Ansari, A., Hofkens, T., & DeCoster, J. (2020). Children's school readiness skills across the pre-K year: Associations with teacher-student interactions, teacher practices, and exposure to academic content. *Journal of Applied Developmental Psychology, 66*, 101084. Retrieved from https://doi.org/10.1016/j.appdev.2019.101084

PISA, O. (2013*). Results*: Ready to learn-students' engagement, drive and self-beliefs (Volume III). PISA: OECD Publishing.

PISA, O. (2016). *Results (Volume I)*: *Excellence and Equity in Educatio*n, PISA: OECD Publishing, Paris.

Pitsia, V., Biggart, A., & Karakolidis, A. (2017). The role of students' self-beliefs, motivation and attitudes in predicting mathematics achievement: A multilevel analysis of the Programme for International Student Assessment data. *Learning and Individual Differences, 55*, 163-173. doi:10.1016/j.lindif.2017.03.014

Plouffe, R. A., Paunonen, S. V., & Saklofske, D. H. (2017). Item properties and the convergent validity of personality assessment: A peer rating study. *Personality and Individual Differences, 111*, 96-105. doi:10.1016/j.paid.2017.01.051

Pokropek, A. (2015). Phantom effects in multilevel compositional analysis: Problems and solutions. *Sociological Methods & Research, 44*(4), 677-705. doi: 10.1177/0049124114553801

Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and Policy Analysis, 36*(4), 399-416.doi: 10.3102/0162373714531851

Polit, D. F & Beck, C. T. (2006).The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29 (5),*489-497.doi.10.1002/nur.20147.

Pollock, K., & Mindzak, M. (2018). Specialist teachers: A review of the literature: Retrieved from https://www.etfo.ca/DefendingWorkingConditions/IssuesInEducation/SpecialistTeachersDocument

Polly, D., Wang, C., Martin, C., Lambert, R., Pugalee, D., & Middleton, C. (2018). The Influence of Mathematics Professional Development, School-Level, and Teacher-Level Variables on Primary Students' Mathematics Achievement. *Early Childhood Education Journal, 46*(1), 31-45. doi:10.1007/s10643-017-0837-y

Pourghebleh, B., & Navimipour, N. J. (2017). Data aggregation mechanisms in the Internet of things: A systematic review of the literature and recommendations for future research. *Journal of Network and Computer Applications, 97*, 23-34. doi:10.1016/j.jnca.2017.08.006

Prendergast, M., Hongning, Z., & Block, A. (2016). A comparative study of students attitudes towards mathematics in two different school systems. *International Journal for Mathematics Teaching and Learning, 17*(2), 1-24.

Prendergast, M., & O'Donoghue, J. (2014). 'Students enjoyed and talked about the classes in the corridors': pedagogical framework promoting interest in algebra. *International Journal of Mathematical Education in Science and Technology, 45*(6), 795-812. doi:10.1080/0020739X.2013.877603

Preti, A., Vellante, M., Gabbrielli, M., Lai, V., Muratore, T., Pintus, E., Tronci, D. (2013). Confirmatory factor analysis and measurement invariance by gender, age and levels of psychological distress of the short TEMPS-A. *Journal of affective disorders, 151*(3), 995-1002.doi: 10.1016/j.jad.2013.08.025

Prieto-Rodriguez, E., Gore, J., & Holmes, K. (2016). Exploring quality teaching in the online environment using an evidence-based approach. *Australian Journal of Teacher Education, 41*(8), 22-39. doi:10.14221/ajte.2016v41n8.2

Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education, 4*(1),1-11. doi.10.1080/2331186X.2017.1301013

Ramlall, I. (2016). Applied Structural Equation Modelling for Researchers and Practitioners: Using R and Stata for Behavioural Research. In *Applied Structural Equation Modelling for Researchers and Practitioners*. Emerald Publishing Limited.

Raudenbush, S., Bryk, A., Cheong, Y., & Congdon Jr, R. (2004). HLM 6: *Hierarchical Linear and Nonlinear Modeling.* Scientific Software International,Inc.USA.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.).London: Sage Publications, Inc.

Ray, A., & Margaret, W. (2003). *PISA Programme for international student assessment (PISA) PISA 2000 technical report: PISA 2000 technical report*: OCED Publishing.

Reckase, M. D. (2009). Multidimensional item response theory models. In *Multidimensional item response theory* (pp. 79-112). Springer, New York, NY.

Reed, H., Kirschner, P. A., & Jolles, J. (2015). Self-beliefs mediate mathematical performance between primary and lower secondary school: A large scale longitudinal cohort study. *Frontline Learning Research*, *3*(1), 36–54.

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating gender differences in mathematics and science: Results from the 2011 Trends in Mathematics and Science Survey. *Research in Science Education, 49*(1), 25-50. doi: 10.1007/s11165-017-9630-6

Reise, S. P. (2000). Using multilevel logistic regression to evaluate person-fit in IRT models. *Multivariate Behavioral Research, 35*(4), 543-568.

Rena, R. (2011). Challenges for quality primary education in Papua New Guinea-A case study. *Education Research International, 2011*. 1–11.Retrieved from https://doi.org/10.1155/2011/485634

Rosenthal, S. (2017). Data imputation. *The International Encyclopedia of Communication Research Methods*, 1-12. Retrieved from https://doi.org/10.1002/9781118901731.iecrm0058

Ross, S. P. (2008). *Motivation correlates of academic achievement: Exploring how motivation influences academic achievement in the PISA 2003 dataset.* (Doctoral thesis, University of Victoria, Australia). Retrieved from **:** http://hdl.handle.net/1828/3209

Russo, J., Bobis, J., Sullivan, P., Downton, A., Livy, S., McCormick, M., & Hughes, S. (2020). Exploring the relationship between teacher enjoyment of mathematics, their attitudes towards student struggle and instructional time amongst early years primary teachers. *Teaching and Teacher Education, 88*, 102983. doi: 10.1016/j.tate.2019.102983

Rutherford, T., Liu, A. S., Lam, A. S., & Schenke, K. (2020). Impact on mathematics self-beliefs from a mastery-based mathematics software. *Journal of Research on Technology in Education, 52*(1), 79-94. doi:10.1080/15391523.2019.1689210

Saadati, F., Cerda, G., Giaconi, V., Reyes, C., & Felmer, P. (2019). Modeling Chilean mathematics teachers' instructional beliefs on problem solving practices. *International Journal of Science and Mathematics Education, 17*(5), 1009-1029. doi:10.1007/s10763-018-9897-8

Sabah, S., & Hammouri, H. (2010). Does subject matter matter? Estimating the impact of instructional practices and resources on student achievement in science and mathematics: Findings from TIMSS 2007. *Evaluation & Research in Education, 23*(4), 287-299,. doi: 10.1080/09500790.2010.509782

Sakarneh, M. (2014). Quality Teaching Practices in the Jordanian Inclusive Primary Classrooms. *Asian Social Science, 10*(19), 26-40.

Sakarneh, M., Paterson, D., & Minichiello, V. (2016). The Applicability of the New South Wales Quality Teaching Model to the Jordanian Primary School Context. *Dirasat: Educational Sciences, 62*(4019), 1-17.

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in nursing & health, 23*(3), 246-255. doi:10.1002/1098-240X(200006)23:3<246::AID-NUR9>3.0.CO;2-H

Sarı, M., & Cenkseven, F. (2008). Quality of school life and self-concept among elementary school students. *Journal of Human Sciences, 5*(2),1-16.

Sarouphim, K. M., & Chartouny, M. (2017). Mathematics education in Lebanon: gender differences in attitudes and achievement. *Educational Studies in Mathematics, 94*(1), 55-68. doi:0.1007/s10649-016-9712-9

Sauceda, A. K. (2017). *The Effect of Teacher Degree Level, Teacher Certification, and Years of Teacher Experience on Student Achievement in Middle School Mathematics*: (Doctoral thesis,Baker University). Retrived from http://www.bakeru.edu.

Saxe, G. B., & Esmonde, I. (2012). *Cultural development of mathematical ideas: Papua New Guinea studies*: Cambridge University Press.

Schaaps, P. (2019). Explicating the South African Psychological Ownership Questionnaire's confirmatory factor analysis model fit: A Bayesian structural equation modelling approach. *SA Journal of Industrial Psychology, 45*(1), 1-16.doi:10.4102/sajip.v45i0.1643

Scheerens, J., Luyten, H., & Van Ravens, J. (2011). *Perspectives on educational quality: Illustrative outcomes on primary and secondary schooling in the Netherlands*: Springer Science & Business Media.

Schenke, K., Lam, A. C., Conley, A. M., & Karabenick, S. A. (2015). Adolescents' help seeking in mathematics classrooms: Relations between achievement and perceived classroom environmental influences over one school year. *Contemporary Educational Psychology, 41*, 133-146. doi: 10.1016/j.cedpsych.2015.01.003

Scheuneman, J., & Bleistein, C. (1999). *Item bias*. *Advances in measurement in educational research and assessment,* Pergamon.Turkey.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*: psychology press.

Schulz, W. (2004). Scaling Procedures for Likert-type Items on Student's Concepts,Attitudes, and Actions. In W. Schulz & H. Sibbems (Eds). *IEA Civic Education study report* (pp. 93-126). The Netherlands: the International Association for the Evaluation of Educational Assessment.

Seals, C. (2018). *Teacher Beliefs: Effects of a Teacher Based Mindset Intervention on Math Student Motivation and Achievement*: (Doctoral thesis Michigan State University,USA) Retrived from https://lib.msu.edu.

Shen, J. (2016). Principles and applications of multilevel modeling in human resource management research. *Human Resource Management, 55*(6), 951-965. doi:10.1002/hrm.21666

Shernoff, D. J., Ruzek, E. A., & Sinha, S. (2017). The influence of the high school classroom environment on learning as mediated by student engagement. *School Psychology International, 38*(2), 201-218. doi:10.1177/0143034316666413

Shi, D., Lee, T., Fairchild, A. J., & Maydeu-Olivares, A. (2020). Fitting ordinal factor analysis models with missing data: A comparison between pairwise deletion and multiple imputation. *Educational and Psychological Measurement, 80*(1), 41-66.

Shin, S.-H., Slater, C. L., & Ortiz, S. (2017). Leader perceptions and student achievement: An examination of reading and mathematics international test results in Korea and the USA. *International Journal of Educational Management, 31*(7), 1103-1118.

Shmis, T., Ustinova, M., & Chugunov, D. (2019). Learning Environments and Learning Achievement in the Russian Federation*:* How School Infrastructure and Climate Affect Student Success: The World Bank. doi: 10.1596/978-1-4648-1499-0

Shoraka, M., Arnold, R., Kim, E. S., Salinitri, G., & Kromrey, J. (2016). Parental characteristics and the achievement gap in mathematics: Hierarchical Linear Modeling Analysis of Longitudinal Study of American Youth (LSAY). *Alberta Journal of Educational Research, 61*(3), 280-293.

Shouse, R. (2018). Academic press, sense of community, and student achievement *Redesigning American education* (pp. 60-86): Routledge.

Siahaan, U. M. J., & de Mello, L. A. (2020). Values of Matematics: How Implementation of Values of Mathematics in School. *Education and Information Technology, 4*(1), 1-7. doi: 10.11648/j.ajeit.20200401.11

Siddiqui, N., & Gorard, S. (2017). Comparing government and private schools in Pakistan: The way forward for universal education. *International Journal of Educational Research, 82*, 159-169. doi:0.1016/j.ijer.2017.01.007

Simzar, R., Domina, T., & Tran, C. (2016). Eighth-grade algebra course placement and student motivation for mathematics. *AERA Open, 2*(1), doi.0.02332858415625227.

Singh, J. (2016). Effect of school and home factors on learning outcomes at elementary school level: a hierarchical linear model. *Education 3-13, 44*(2), 116-139. doi:10.1080/03004279.2014.899383

Singh, R., & Sarkar, S. (2015). Does teaching quality matter? Students learning outcome related to teaching quality in public and private primary schools in India. *International Journal of Educational Development, 41*, 153-163. doi: 0.1016/j.ijedudev.2015.02.009.

Siniscalco, M., & Auriat, N. (2012). Quantitative research methods in educational planning, *Sample design for educational survey research*.Series editor: Kenneth N. Ross: UNESCO, International Institute for Educational Planning/UNESCO, France. Retrived from http://www.unesco.org/iiep

Sivakumar, S. (2014). *Examining the relationship between moral reasoning, cognitive reasoning and learning preferences of high-school students in selected South Australian schools.* (Doctoral thesis, Univeristy of Adelaide, Australia). Retrieved from https://librarysearch.adelaide.edu.au

Skrodal, S. (2010). *Virtual classroom simulation: design and trial in a preservice teacher education program.* (Doctoral thesis, Univeristy of Adelaide, Australia). Retrieved from https://librarysearch.adelaide.edu.au

Snijders,T. A. B.,& Bosker, R.J. (2012).Multilevel analysis: *An introduction to basic and advanced multilevel modeling* (2nd ed.).Los Angelas,London,New Delhi,Singapore,Washington DC:SAGE publications Ltd.

Soni, A., & Kumari, S. (2017). The role of parental math anxiety and math attitude in their children's math achievement. *International journal of science and mathematics education, 15*(2), 331-347. doi: 10.1007/s10763-015-9687-5

Spark, C. (2010). Changing lives: Understanding the barriers that confront educated women in Papua New Guinea. *Australian Feminist Studies, 25*(63), 17-30. doi: 10.1080/08164640903499901

Starkey, J., & Rymarz, R. (2018). Teaching catholic studies: a study of religious education in senior high school. *Journal of Religious Education, 66*(1), 65-75.

Stelmack, J., Szlyk, J. P., Stelmack, T., & Babcock-Parziale, J. (2004). Use of Rasch person-item map in exploratory data analysis: a clinical perspective. *Journal of rehabilitation research and development, 41*(2), 233–242.

Stephen, R., & Anthony, B. (2002). *Hierarchical linear models*: Sage Publications, Thousand Oaks, CA.

Stephens, M., Landeros, K., Perkins, R., & Tang, J. H. (2016). Highlights from TIMSS and TIMSS Advanced 2015: Mathematics and Science Achievement of US Students in Grades 4 and 8 and in Advanced Courses at the End of High School in an International Context. NCES 2017-002. *National Center for Education Statistics*. Retrieved from https://eric.ed.gov/?id=ED570893.

Stevans, L. K., & Sessions, D. N. (2000). Private/public school choice and student performance revisited. *Education Economics, 8*(2), 169-184. doi.10.1080/096452900410730

Strobl, C., Kopf, J., & Zeileis, A. (2015). Rasch trees: A new method for detecting differential item functioning in the Rasch model. *Psychometrika, 80*(2), 289-316.

Stronge, J. H. (2018). *Qualities of effective teachers*. (3rd Ed.,) ASCD. Williamsburg, Virginia.

Sukthankar, N. (1995). Gender and mathematics education in Papua New Guinea. *Equity in mathematics education: Influences of feminism and culture*, *51*(2)135-140.

Sukthankar, N. (2006). Mathematics education in Papua New Guinea: A challenge. *Challenging mathematics in and beyond the classroom, 16*, 1-8. Retrieved from http://www.wfnmc.org/icmis16ppngsukthankar.pdf.

Susanti, D., & Rosyida, I. (2020). Student's Mathematical Reasoning Ability Viewed from Self-Confidence in Mathematical Modeling with Open-ended Approach Learning. *Unnes Journal of Mathematics Education Research, 9*(1), 114-122. Retrieved from http://journal.unnes.ac.id/sju/index.php/ujmer

Taasoobshirazi, G., & Wang, S. (2016). The performance of the SRMR, RMSEA, CFI, and TLI: An examination of sample size, path size, and degrees of freedom. *Journal of Applied Quantitative Methods, 11*(3). Retrieved from http://jaqm.ro/issues/volume-11,issue-3/pdfs/2_GI_SH_.pdf.

Tan, T. X., Zhou, Y., & Li, G. (2020). Maternal education and Chinese first Graders' performance in language and literacy and math: Role of home environment. *Early Childhood Education Journal, 48*(2), 243-252. doi: 10.1007/s10643-019-00986-w

Tapo, M. F. (2004). *National standards/local implementation: case studies of differing perceptions of national education standards in Papua New Guinea.* (Doctoral thesis, Queensland University of Technology, Australia). Retrieved from https://www.library.qut.edu.au.

Tejada, A. J. R., Luque, M. N., Rojas, O. M. L., & Moreno, P. J. P. (2011). Prejudiced attitude measurement using the Rasch rating scale model. *Psychological reports, 109*(2), 553-572. doi: 10.2466/07.17.PR0.109.5.553-572

Tella, A. (2017). Teacher variables as predictors of academic achievement of primary school pupils mathematics. *International Electronic Journal of Elementary Education, 1*(1), 16-33.

Thien, L. M., Darmawan, I. G. N., & Ong, M. Y. (2015). Affective characteristics and mathematics performance in Indonesia, Malaysia, and Thailand: what can PISA 2012 data tell us? *Large-scale Assessments in Education, 3*(1), 1-16. doi: 10.1186/s40536-015-0013-z

Thien, L. M., Razak, N. A., Keeves, J., & Darmawan, I. G. N. (2016). *What Can PISA 2012 Data Tell Us?: Performance and Challenges in Five Participating Southeast Asian Countries*: Springer.

Thomas, S. L., Schmidt, K. M., Erbacher, M. K., & Bergeman, C. S. (2016). What You Don't Know Can Hurt You: Missing Data and Partial Credit Model Estimates. *Journal of applied measurement, 17*(1), 14-34.

Thomson, S., Hillman, K., Wernert, N., Schmid, M., Buckley, S., & Munene, A. (2012). Monitoring Australian year 4 student achievement internationally : TIMSS and PIRLS 2011. Melbourne: Australian Council for Educational Research (ACER). Retrieved from https://research.acer.edu.au

Thomson, S., Wernert, N., Underwood, C., & Nicholas, M. (2008). TIMSS 2007: Taking a closer look at mathematics and science in Australia. https://research.acer.edu.au/timss_2007/2.

Thune, C. (2017). Quality assurance of higher education in Denmark *Global Perspectives on Quality in Higher Education* (pp. 82-102): Routledge.

TIMSS & PIRLS. (2016). *Methods and Procedues in TIMSS 2015*. Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA), Boston College,United States.

Tok, Ş. (2015). The Effects of Teaching Mathematics Creatively on Academic Achievement, Attitudes towards Mathematics, and Mathematics Anxiety. *International Journal of Innovation in Science and Mathematics Education (formerly CAL-laborate International), 23*(4),1-24.

Tortorella, G. L., Marodin, G. A., Miorando, R., & Seidel, A. (2015). The impact of contextual variables on learning organization in firms that are implementing lean: a study in Southern Brazil. *The International Journal of Advanced Manufacturing Technology, 78*(9), 1879-1892. doi: 10.1007/s00170-015-6791-1

Tshewang, R., Chandra, V., & Yeh, A. (2017). Students' and teachers' perceptions of classroom learning environment in Bhutanese eighth-grade mathematics classes. *Learning Environments Research, 20*(2), 269-288. doi: l 10.1007/s10984-016-9225-6

Tuomi, I. (2006). Open Educational Resources: What they are and why do they matter. Report prepared for the OECD. Retrieved from *http://www. meaningprocessing. com/personalPages/tuomi/articles/OpenEducationalResources_OECDreport. pdf*

Turner, M. M. (1990). *Papua New Guinea: The challenge of independence*: Penguin Books.

Turner, V. E. (2017). *Learning Environments in Mathematics.* (Doctoral thesis, Brenau University, Georgia, USA). Retrieved from https://eric.ed.gov/?id=ED576012

Ufer, S., Rach, S., & Kosiol, T. (2017). Interest in mathematics= interest in mathematics? What general measures of interest reflect when the object of interest changes. *ZDM, 49*(3), 397-409. doi:10.1007/s11858-016-0828-2

Ulandari, L., Amry, Z., & Saragih, S. (2019). Development of Learning Materials Based on Realistic Mathematics Education Approach to Improve Students' Mathematical Problem Solving Ability and Self-Efficacy. *International Electronic Journal of Mathematics Education, 14*(2), 375-383. doi:10.29333/iejme/5721

Urquiola, M. (2016). Competition among schools: Traditional public and private schools *Handbook of the Economics of Education* (Vol. 5, pp. 209-237): doi.10.1016/B978-0-444-63459-7.00004-X

Uwineza, I., Rubagiza, J., Hakizimana, T., & Uwamahoro, J. (2018). Gender attitudes and perceptions towards mathematics performance and enrolment in Rwandan secondary schools. *Rwandan Journal of Education, 4*(2), 44-56

van de Grift, W. J. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement, 25*(3), 295-311. doi.10.1080/09243453.2013.794845

van Vendeloo, S. N., Prins, D. J., Verheyen, C. C., Prins, J. T., van den Heijkant, F., van der Heijden, F. M., & Brand, P. L. (2018). The learning environment and resident burnout: a national study. *Perspectives on medical education*, 7,120–125. doi.10.1007/s40037-018-0405-1

Vanlaar, G., Kyriakides, L., Panayiotou, A., Vandecandelaere, M., McMahon, L., De Fraine, B., & Van Damme, J. (2016). Do the teacher and school factors of the dynamic model affect high-and low-achieving student groups to the same extent? A cross-country study. *Research Papers in Education, 31*(2), 183-211. doi.10.1080/02671522.2015.1027724

Varhol, A., Drageset, O. G., & Hansen, M. N. (2020). Discovering key interactions. How student interactions relate to progress in mathematical generalization. *Mathematics Education Research Journal*, 1-18. doi.10.1007/s13394-020-00308-z

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*: John Wiley & Sons.

Wang, J., & Wang, X. (2019). *Structural equation modeling: Applications using Mplus*: John Wiley & Sons.

Wang, W., Yin, H., Lu, G., & Zhang, Q. (2017). Environment matters: exploring the relationships between the classroom environment and college students' affect in mathematics learning in China. *Asia Pacific Education Review, 18*(3), 321-333. doi: 10.1007/s12564-017-9478-5

Wang, W. C., & Wu, S. L. (2011). The random-effect generalized rating scale model. *Journal of Educational Measurement, 48*(4), 441-456. doi.10.1111/j.1745-3984.2011.00154.x

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427-450. doi.1 0.1007/BF02294627

Watt, H. M., Carmichael, C., & Callingham, R. (2017). Students' engagement profiles in mathematics according to learning environment dimensions: Developing an evidence base for best practice in mathematics education. *School PsychologyInternational, 38*(2), 166-183. doi:10.1177/0143034316688373

Weeks, S. G. (1993). Education in Papua New Guinea 1973-1993: the late-development effect? *Comparative education, 29*(3), 261-273. doi:10.1080/0305006930290304

Weems, G. H., & Onwuegbuzie, A. J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement and Evaluation in Counseling and Development, 34*(3), 166-178.doi: : 10.1080/07481756.2002.12069033

Wei, S., Liu, X., & Jia, Y. (2014).Using Rasch measurement to valiadte the instrument of students' understanding of models in Science. *International journal of science and mathematics education, 12*(5), 1067-1082.

Wei, Y., Clifton, R. A., & Roberts, L. W. (2011). School resources and the academic achievement of Canadian students. *Alberta Journal of Educational Research, 57*(4), 460-478.

Wetzel, E., & Carstensen, C. H. (2014). Reversed thresholds in partial credit models: A reason for collapsing categories? *Assessment, 21*(6), 765-774. doi.10.1177/1073191114530775

Whannell, R., & Tobias, S. (2015). Improving mathematics and science education in rural Australia: A practice report. *Australian and International Journal of Rural Education, 25*(2), 91-97.

Wilkins, J. L. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education, 11*(2), 139-164. doi:10.1007/s10857-007-9068-2

Williamson, B. (2017). *Big data in education: The digital future of learning, policy and practice*: SAGE Publications.

Wilson, P. (2011). Disposition towards engagement in mathematics. *Proceedings of the British Society for Research into Learning Mathematics, 31*(2), 67-72.

Wilson, P. S., Cooney, T. J., & Stinson, D. W. (2005). What constitutes good mathematics teaching and how it develops: Nine high school teachers' perspectives. *Journal of Mathematics Teacher Education, 8*(2), 83-111. doi:10.1007/s10857-005-4796-7

Wilson, R., & Mack, J. (2014). Declines in high school mathematics and science participation: evidence of students' and future teachers' disengagement with maths. *International Journal of Innovation in Science and Mathematics Education , 22*(7),35-48.

Winheller, S., Hattie, J. A., & Brown, G. T. (2013). Factors influencing early adolescents' mathematics achievement: High-quality teaching rather than relationships. *Learning Environments Research, 16*(1), 49-69. doi:10.1007/s10984-012-9106-6

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913-934. doi:0.1177/0013164413495237

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology, 8*(1), 52-69.

Woodcock, R. W. (1999). What can Rasch-based scores convey about a person's test performance? . In S. E. E. S. L. Herschberger (Ed.), *The New Rules of Measurement: What Every Psychologist and Educator Should Know* (pp. 105-127). New Jersey: Lawrence Erlbaum.

Wright, B. D. (1996a). Comparing Rasch measurement and factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 3*(1), 3-24.doi:10.1080/10705519609540026

Wright, B. D., & Stone, M. (1999). *Measurement essentials* (2nd ed.). Wilmington, Delaware: Wide Range Inc.

Wu, M., & Adams, R. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach*: Educational Measurement Solutions Melbourne. Retrieved from http:www.edmeasuement.com.au

Wu, M., Tam, H. P., & Jen, T.-H. (2016). Educational Measurement for Applied Researchers. *Theory into Practice*. Springer Singapore. doi:10.1007/978-981-10-3302-5

Xiao, C., Bruner, D. W., Dai, T., Guo, Y., & Hanlon, A. (2019). A Comparison of Missing-Data Imputation Techniques in Exploratory Factor Analysis. *Journal of nursing measurement, 27*(2), 313-334. 10.1891/1061-3749.27.2.313

Yang, X. (2015). Rural junior secondary school students' perceptions of classroom learning environments and their attitude and achievement in mathematics in West China. *Learning Environments Research, 18*(2), 249-266. doi:10.1007/s10984-015-9184-3

Yao, S.-Y., Bull, R., Khng, K. H., & Rahim, A. (2018). Psychometric properties of the NEPSY-II affect recognition subtest in a preschool sample: a Rasch modeling approach. *The Clinical Neuropsychologist, 32*(1), 63-80. doi.10.1080/13854046.2017.1343865

Yasar, S., & Cogenli, A. G. (2014). Determining Validity and Reliability of Data Gathering Instruments Used by Program Evaluation Studies in Turkey. *Procedia-Social and Behavioral Sciences, 131*, 504-509. doi.10.1016/j.sbspro.2014.04.156

Yasin, R. M., Yunus, F. A. N., Rus, R. C., Ahmad, A., & Rahim, M. B. (2015). Validity and reliability learning transfer item using rasch measurement model. *Procedia-Social and Behavioral Sciences, 204*, 212-217.doi: 10.1016/j.sbspro.2015.08.143

Yasmin, M., Naseem, F., & Masso, I. C. (2019). Teacher-directed learning to self-directed learning transition barriers in Pakistan. *Studies in educational evaluation, 61*, 34-40. doi: 10.1016/j.stueduc.2019.02.003

Yavuz Mumcu, H., & Cansız Aktaş, M. (2020). Development of an attitude-towards-using-mathematics scale for high-school students and an analysis of student attitudes. *International Journal of Mathematical Education in Science and Technology, 51*(1), 3-25. doi: 10.1080/0020739X.2019.1679398

Yoo, Y. S. (2018). Modelling of factors influencing gender difference in mathematics achievement using TIMSS 2011 data for Singaporean eighth grade students. *Asia Pacific Journal of Education, 38*(1), 1-14. doi: 10.1080/02188791.2017.1334626

Yu, R., & Singh, K. (2018). Teacher support, instructional practices, student motivation, and mathematics achievement in high school. *The Journal of Educational Research, 111*(1), 81-94. doi: 10.1080/00220671.2016.1204260

Yumiati, Y., & Kusumah, Y. (2019). Interaction between Students' Learning and Early Mathematical Skills to Increase Mathematical Critical Thinking Skills. *Al-Jabar: Jurnal Pendidikan Matematika, 10*(1), 125-134. doi:10.24042/ajpm.v10i1.3712

Zainuri, N. A., Asshaari, I., Ariff, F. H. M., Razali, N., Othman, H., Hamzah, F. M., & Nopiah, Z. M. (2016). Item analysis for final exam questions of engineering mathematics course (Vector calculus) in UKM. *Journal of Engineering Science and Technology*, 11,53-60.

Zakharov, A., Tsheko, G., & Carnoy, M. (2016). Do “better” teachers and classroom resources improve student achievement? A causal comparative approach in Kenya, South Africa, and Swaziland. *International Journal of Educational Development, 50*, 108-124. doi:10.1016/j.ijedudev.2016.07.001

Zhang, D. (2008). The effects of teacher education level, teaching experience, and teaching behaviors on student science achievement. *Graduate Theses and Dissertations*, 155. Retrieved from https://digitalcommons.usu.edu/etd

Zhao, D., & Zhao, D. (2016). *Chinese Students' Higher Achievement in Mathematics*: Springer Singaropre.

Zhu, Y., Kaiser, G., & Cai, J. (2018). Gender equity in mathematical achievement: the case of China. *Educational Studies in Mathematics, 99*(3), 245-260. doi.10.1007/s10649-018-9846-z

Zlatkin-Troitschanskaia, O., Schmidt, S., Brückner, S., Förster, M., Yamaoka, M., & Asano, T. (2016). Macroeconomic knowledge of higher education students in Germany and Japan–a multilevel analysis of contextual and personal effects. *Assessment & Evaluation in Higher Education, 41*(5), 787-801. doi.10.1080/02602938.2016.1162279

Appendices

Appendix A

Permission/approval documents

THE UNIVERSITY
of ADELAIDE

RESEARCH SERVICES
OFFICE OF RESEARCH ETHICS, COMPLIANCE
AND INTEGRITY
THE UNIVERSITY OF ADELAIDE
LEVEL 4, RUNDLE MALL PLAZA
50 RUNDLE MALL
ADELAIDE SA 5000 AUSTRALIA

TELEPHONE +61 883135137
FACSIMILE +61 88313 3700
EMAIL hrec@adelaide.edu.au

CRICOS Provider Number 00123M

14 July 2017

Associate Professor S Alagumalai
School of Education

Dear Associate Professor Alagumalai

ETHICS APPROVAL No: H-2017-113

PROJECT TITLE: **Examining factors which affect mathematics achievement for grade 10 and 12 students in Papua New Guinea (PNG)**

The ethics application for the above project has been reviewed by the Low Risk Human Research Ethics Review Group (Faculty of Arts and Faculty of the Professions) and is deemed to meet the requirements of the *National Statement on Ethical Conduct in Human Research (2007)* involving no more than low risk for research participants. You are authorised to commence your research on **14 Jul 2017.**

Ethics approval is granted for three years and is subject to satisfactory annual reporting. The form titled Annual Report on Project Status is to be used when reporting annual progress and project completion and can be downloaded at http://www.adelaide.edu.au/research-services/oreci/human/reporting/. Prior to expiry, ethics approval may be extended for a further period.

Participants in the study are to be given a copy of the Information Sheet and the signed Consent Form to retain. It is also a condition of approval that you **immediately report** anything which might warrant review of ethical approval including:

- serious or unexpected adverse effects on participants,
- previously unforeseen events which might affect continued ethical acceptability of the project,
- proposed changes to the protocol; and
- the project is discontinued before the expected date of completion.

Please refer to the following ethics approval document for any additional conditions that may apply to this project.

Yours sincerely

DR JOHN TIBBY		**DR JOANNA HOWE**
Co-Convenor	*for*	**Co-Convenor**
Low Risk Human Research Ethics Review Group (Faculty of Arts and Faculty of the Profession		**Low Risk Human Research Ethics Review Group (Faculty of Arts and Faculty of the Professions)**

THE UNIVERSITY
of ADELAIDE

RESEARCH SERVICES
OFFICE OF RESEARCH ETHICS, COMPLIANCE AND INTEGR\IY
THE UNIVERSITY OF ADELAIDE

LEVEL 4, RUNDLE MALL PLAZA
50 RUNDLE MALL
ADELAIDE SA 5000 AUSTRALIA

TELEPHONE +61 883135137
FACSIMILE +61 88313 3700
EMAIL hrec@edelaide.edu.au

CRCOS Provider Number 23M

Applicant: **Associate Professor S Alagumalai**

School: **School of Education**

Project Title: **Examining factors which affect mathematics achievement for grade 10 and 12 students in Papua New Guinea (PNG)**

The University of Adelaide Human Research Ethics Committee
Low Risk Human Research Ethics Review Group (Faculty of Arts and Faculty of the Professions)

ETHICS APPROVAL No: H-2017-113 App. No.: 0000022410

APPROVED for the period: 14 Jul 2017 to 31 Jul 2020

Thank you for your responses, dated 13.07.17 and 14.07.17, to the matters raised. It is also noted that this project involves PhD student Jerome Oko.

DR JOHN TIBBY
Co-Convenor
Low Risk Human Research Ethics Review Group
(Faculty of Arts and Faculty of the Professions)

DR JOANNA HOWE
for **Co-Convenor**
Low Risk Human Research Ethics Review Group
(Faculty of Arts and Faculty of the Professions)

DEPARTMENT OF EDUCATION

Office of the Deputy Secretary - Policy & Corporate Services Directorate

TELEPHONE: (675) 301 3342/301 3343
FASCIMILE: (675) 323 1031/325 4648
EMAIL: Titus Hatagen@education.gov.pg

FINCORP HAUS, LEVEL 6
PO BOX 446
WAIGANI - N.C.D
PAPUA NEW GUINEA

Date: 30th August 2017
File: PPR 1-1-2

Jerome Oko
School of Education
University of Adelaide

Dear Mr Oko,

SUBJECT: APPROVAL TO CONDUCT RESEARCH AND ACCESS TO 2017 GRADE 10 AND 12 MATHEMATICS NATIONAL EXAMINATION RESULTS

Your research proposal titled ***"Examining factors which affect mathematics achievement of Grade 10 and 12 students in Papua New Guinea"*** have been approved by the Research, Evaluation and Monitoring Steering Committee (RESC).

While your research is approved to collect data in educational institution/s, it is also subject to approval by Provincial Research Committee (where applicable) and/or the Provincial Education Advisor or the principals of your nominated Institutions. It is your responsibility to ensure such is obtained prior to the fieldwork.

The RESC have also approved your request to have access to the Grade 10 and 12 standardized mathematics examination results with their raw scores. The Assistant Secretary for Measurement Services Division (MSD) will advised of this decision. The RESC committee have given the approval based on the importance of the Research study, as the results will provide DOE significant information and understanding of the factors, which affect mathematics results and may be used to improve Grade 10 and 12 students' mathematics achievement through focusing on the development of stronger mathematics education programs in PNG.

Use this letter as an approval for this purpose. In serious case of breach of ethical issues and DOE Research guidelines, the Department of Education reserves the right to inform your institution or sponsors directly and take necessary actions as deem necessary. Failure to observe the above conditions may lead to the withdrawal of the research approval.

I thank you and wish you good luck in your study.

Yours sincerely,

TITUS ROMANO HATAGEN
Deputy Secretary-Policy& Corporate Servi Directorate
Chairman – Research, Evaluation & Moni Steering Committee

cc: Assistant Secretary – Research & Eval Division

THE UNIVERSITY *of* ADELAIDE

Assistant Secretary

NCD Education Services

Department of Education

P O Box 446

Waigani

NCD

7th August, 2017

Dear Mr Lora,

REF: PERMISSION FOR DATA COLLECTION - PhD STUDY

I am Jerome Oko, a PhD student in the School of Education at the University of Adelaide. I am conducting a research on a project titled **'Examining factors which affect mathematics achievement of Grade 10 and 12 students in Papua New Guinea,** under the supervision of Associate Professor Sivakumar Alagumalai, Dr I. Gusti Damawan and Associate Professor Nicholas Buchdahl at the University of Adelaide, South Australia.

I would be most grateful if you would grant permission for selected Grade 10 and 12 classes and their mathematics teachers at NCD secondary schools to assist in the project by completing a questionnaire on **'Examining factors which affect mathematics achievement of Grade 10 and 12 students.** I will be undertaking this research on the site at the school during class patron time so as to minimize interference with your school's core class time. The research will be carried out on an agreeable time between August and September.

Grade 10 and 12 students will be asked to complete a questionnaire for 25 minutes in their class advisory /homeroom period and the next day they will do a mathematics test for 1 hour on the topics they have covered in class with their teachers.

(If the above time allocation is not possible, please suggest a most appropriate time for me to undertake this survey at your school).

Simultaneously, teachers' version of the questionnaires will be completed by Grade 10 and 12 mathematics teachers. Teachers should be able to complete the questionnaire in 30 minutes in their non-contact time in the school. A short interview with two-selected mathematics Grade 10 and 12 teachers will be carried out for 30-35 minutes. The interview will be centered upon certain aspects of the topic.

In conducting this study, **ethics are strictly observed (University of Adelaide HREC Project Approval Number: H-2017-113),** and requires participant consent. Any information provided will be treated in the strictest confidence and no participants or school will be individually identifiable in the resulting thesis, report or other publications. This research requires participant consent. Participation in this study is voluntary, and your students and teachers may withdraw at any stage or choose not to answer any questions. No negative consequence will follow from such withdrawal. Since participation is purely voluntary, non-participation will not affect students' academic progress or the teachers' professional progress in any way.

In this project, audio recording of mathematics teachers' interviews will be made. Therefore, will seek your selected mathematics teachers' consent, to audio record for the interview, on conditions that name or identities are not revealed or disclosed for any reason.

Should you require additional information regarding this research, please contact me by phone telephone on (+618) 8303-7196 or email jerome.oko@adelaide.edu.au. Should I be unavailable, my supervisor, Associate Professor Sivakumar Alagumalai on telephone (+618) 8313 5630 and email sivakumar.alagumalai@adelaide.edu.au.

This research project has been approved by the University of Adelaide Human Research Ethics Committee. The Secretary of this Committee can be contacted on (08) 8303-6028,fax (08) 8303 7325,email. hrec@adelaide.edu.au

Thank you for your assistance.

Sincerely yours,

Jerome OKO

PERMISSION REQUEST FORM FROM IEA

1. Requested IEA material

Number IEA-17-/145

To be completed by anyone seeking permission to reuse, reproduce, or translate IEA materia

Type of requested material (select all that apply): [ ] Abstract [ ] Text excerpt [ ] Report

[ ] Full article [ ] Chapter [ ] Figure/table [ ] restricted use items [**X**] Study instruments [ ] Other, please specify:

Please indicate the source of the IEA material:
Author/editor: IEA's Trends in International Mathematics and Science Study
Title: Survey Questionnaires. TIMSS 2011 Copyright © 2012 International Association for the Evaluation of Educational Achievement
ISBN: Date of publication: **2014**

Description of requested IEA material (provide exact page numbers, chapter name/author, figure/table numbers, item numbers, and URL address): **Student, Teacher, School & Curriculum Questionnaires: Student Item: 7,8,18,20,21,22 pp15-21, Teacher Items; 4,5,6,9,10,11,14,18,19,22,pp.2-12.School Items:7.9.10. 5.https:timssandpirls.bc.edu/timss2015/questionnaires/index.html**

NOTE: Please attach text or graphics from the IEA material you would like to use or reproduce.

2. How IEA material will be used

Information about your intended use (select all that apply): [ ] Non-commercial [ ] Commercial

[**X]** Thesis [ ] Secondary analysis [ ] Publication [**X**] Survey research [ ] Other, please specify:

Please provide a brief description of your intended use: To use for a Doctoral Thesis studies with the School of Education with the University of Adelaide

Information about where the requested content will appear/where you will use the requested content:

Author: **Jerome OKO**

Title: **Examining the Factors Affecting Mathematics Achievement for Grade 10 and 12 students in Papua New Guinea (PNG)** Language: **English**
Publisher or sponsor: **Unpublished Thesis**
Intended audience: **Policy makers, Students & Teachers, interested public**

International Association for the Evaluation of Educational Achievement
IEA Amsterdam I Keizersgracht 311, 1016 EE Amsterdam, The Netherlands I Tel. +31 (0)20 625 36 25 | Fax +31 (0)20 420 71 36 |

Format of the work where the requested content will appear/where you will use the requested content (select all that apply):

[**X**] Printed [ ] Electronic [ ] Other, please specify:
Number of copies for distribution: 600-1000 Retail price: ____________
Date to be released: **3 rd July, 2017**
Additional comments: **Will make questionnaires for survey**

3. Requestor information

First name: **Jerome**
Last name: **Oko**
Name of institution or organization: **The University of Adelaide, School of Education**
Address: **46 Ward Street, Pennington**
City and zip code: **Adelaide, 5013** Country: **Australia**
Phone: **0426932229** Email: **jerome.oko@adelaide.edu.au**

Signature: Date of request: **8th May, 2017**

To submit your request, please sign and return this form to IEA by mail (Keizersgracht, 1016 EE Amsterdam, The Netherlands), email (secretariat@iea.nl), or fax (+31 204207136).

In signing, you agree that the permissioned material will be distributed only as part of or for use along with the original work, where the primary value does not lie with the permissioned material itself.

Please note that by signing this form you also state that you have filled out this form truthfully and to the best of your knowledge and that you have read and will comply with all conditions. Providing IEA with incorrect or incomplete information will not only invalidate permission if granted, but can also hold you liable for any damage arising from your failure to comply with these requirements. In case you have any hesitations and/or reservations regarding the information you have to fill out on this form, please contact IEA. Please allow 4-5 weeks for the processing of your permission request.

> When quoting and/or citing from one or more publications and/or restricted use items from TIMSS, PIRLS and IEA for the sake of educational or research purposes, please print an acknowledgment of the source, including the year and name of that publication and/or restricted use item. Please use the following acknowledgment as an
>
> SOURCE: TIMSS 2007 Assessment. Copyright (0 2009 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
>
> Please note that citing without naming the source can or will constitute plagiarism for which you can and will be held accountable.

Number IEA 17-145 (to be filled by IEA)

4. Permission[**X**] Granted [ ] Denied

Terms of agreement: Permission is granted for non-exclusive rights to reproduce the material requested above upon the terms and solely for the purpose indicated.

Free [**X**]

Signature:

e [ ] Fee (payable to NL63ABNA0481961968)

ıre: ___ ___ Date: 22-5-2017

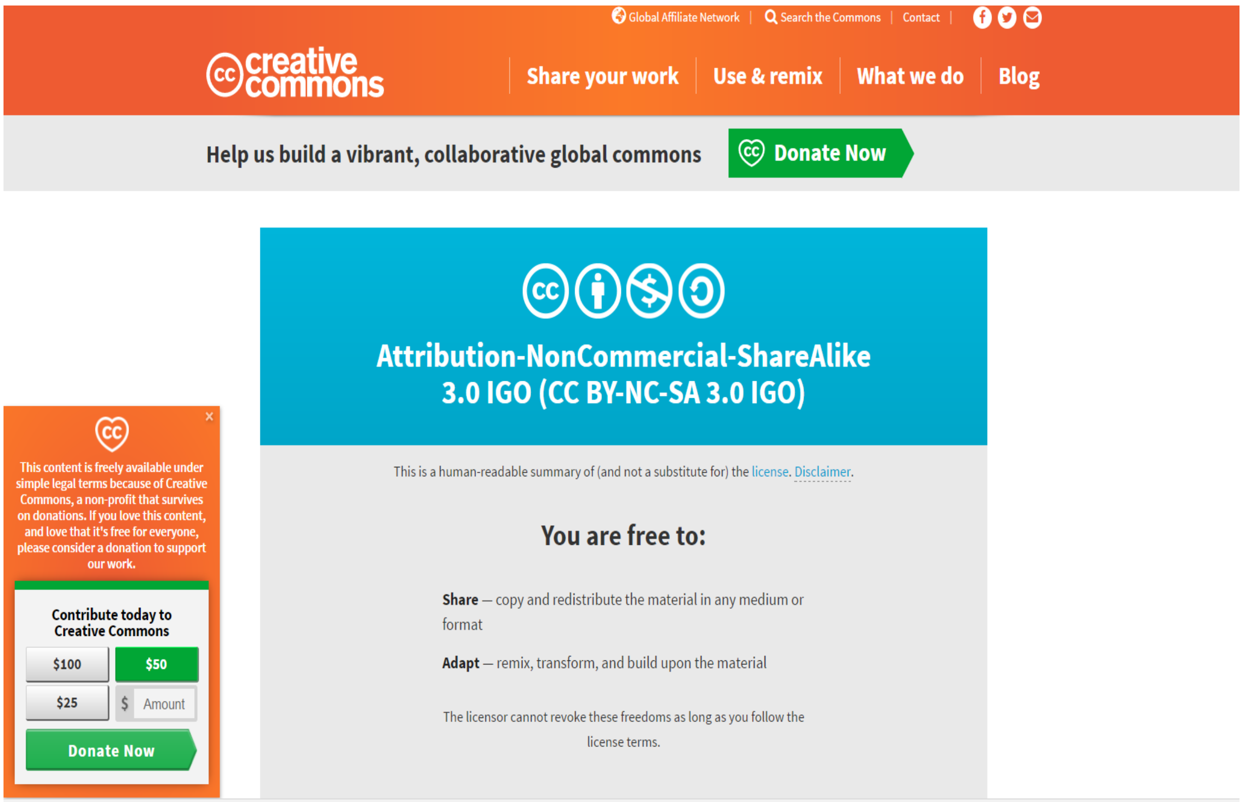
Name:

Title: Executive Director

Permission from OECD

Dear Jerome,
Thank you for your message. Please refer to the Creative Commons page available here https://creativecommons.org/licenses/by-nc-sa/3.0/igo/ and to the licence available here https://creativecommons.org/licenses/by-nc-sa/3.0/igo/legalcode.

Best regards,
Gersa



Appendix B

Student and teacher survey questionnaires

Examining factors which affect mathematics achievement
of Grade 10 and 12 students in Port Moresby, Papua New Guinea
(Student Questionnaire)

I. Information about this Questionnaire

This questionnaire is intended for Grades 10 and 12 students in secondary schools. It contains items that ask for general information about the student and his/her attitude and motivation towards learning mathematics. It has been organised into sections (A, B and C) to the factors under study.

Your response to the questionnaires contribute to improving the standard of mathematics education in PNG. It is vital that you respond to each of the items very carefully so that the information provided reflects your situation. Your answers will be collated with other responses from other students in which no individual students/schools can be identified. All your answers and your identity will be kept strictly confidential.

General Instructions to Student Participants

1. Please read each item carefully and answer them accurately as you can. For every section of the questionnaire, specific instructions of how to answer the items are given. If you make a mistake in responding to items that have the options, simply mark **X** on the previous choice and check another box corresponding to your new answer. **Please do not leave any item unanswered.**

2. The questionnaire must to be returned to your class patron/teacher as soon as it has been completed.

3. Complete the questionnaire during your class. Your class teacher/patron will help distribute and explain the instructions.

4. Time allocated to answer the questions is 25 minutes

NB. The researcher to communicate the information and the instructions to student participants.

School Name: ________________________________
School ID: __________________________________
School Type: ________________________________
Teacher ID: _________________________________

A. GENERAL INFORMATION ABOUT YOU

Instructions. Fill in the box and lines with words or sentences that correspond to your answer. For items that have the given choices, check only one box.

Student Name: [ ]

1. Gender.

☐ Male ☐ Female

2. Grade/Year Level: ☐ Grade 12

☐ Grade 10

3. Father's Occupation: ______________________________________

4. Father's highest education level: ☐ Certificate

☐ Diploma

☐ Bachelor Degree

☐ Postgraduate (Masters/PhD)

☐ Others: please specify ________________

5. Mother's Occupation: ________________________

6. Mother's highest education level: ☐ Certificate

☐ Diploma

☐ Bachelor's degree

☐ Postgraduate (Masters/PhD)

☐ Others: Please specify __________________

B. MOTIVATION TOWARDS MATHEMATICS (How motivated are you to learn mathematics?)

Instructions: This section is about your motivation towards learning mathematics. Read each item carefully and answer by putting a tick (√) only on ONE box. **Please answer all questions.**

SD=Strongly Disagree, D=Disagree, A=Agree, SA=Strongly Agree

		SD (1)	**D** (2)	**A** (3)	**SA** (4)
7.	I am prepared for my mathematics examinations.	☐	☐	☐	☐
8.	Jobs that require mathematics skills seems interesting to me.	☐	☐	☐	☐
9.	Learning mathematics will help me get ahead in the world.	☐	☐	☐	☐
10.	It is important to do well in my mathematics class.	☐	☐	☐	☐
11.	Doing well in mathematics will help me get into university/colleges.	☐	☐	☐	☐
12.	Learning mathematics will give me more job opportunities.	☐	☐	☐	☐
13.	Learning mathematics is worthwhile for me because it will improve my career prospects.	☐	☐	☐	☐
14.	I keep studying until I understand mathematics material.	☐	☐	☐	☐
15.	I take part in mathematics competitions.	☐	☐	☐	☐
16.	I do mathematics more than 2 hours a day outside of school.	☐	☐	☐	☐

SD=Strongly Disagree, D=Disagree, A=Agree, SA=Strongly Agree

		SD (1)	**D** (2)	**A** (3)	**SA** (4)
17.	I have my homework finished in time for mathematics class.	☐	☐	☐	☐
18.	I work hard on my mathematics homework.	☐	☐	☐	☐
19.	I keep my mathematics work well organized.	☐	☐	☐	☐
20.	I talk about mathematics problems with my friends.	☐	☐	☐	☐
21.	I listen and pay attention in mathematics class.	☐	☐	☐	☐
22.	I avoid distractions when I am studying mathematics.	☐	☐	☐	☐
23.	Mathematics is one of my favorite subjects.	☐	☐	☐	☐
24.	I am interested in the things I learn in mathematics.	☐	☐	☐	☐
25.	The teacher did not get students interested in the material.	☐	☐	☐	☐
26.	It is interesting to learn mathematics theory.	☐	☐	☐	☐

C. ATTITUDE TOWARDS MATHEMATICS

Instruction. This section is about your attitude in learning mathematics. Read each item carefully and answer by ticking (√) only ONE box. **Please answer all questions.**

SD=Strongly Disagree, D=Disagree, A=Agree, SA=Strongly Agree

		SD (1)	D (2)	A (3)	SA (4)
27.	I like studying for my mathematics class outside of school.	☐	☐	☐	☐
28.	I am studying mathematics because I like to learn new things.	☐	☐	☐	☐
29.	I look forward to my mathematics lessons.	☐	☐	☐	☐
30.	I enjoy thinking about the world in terms of mathematics relationships.	☐	☐	☐	☐
31.	I do mathematics because I enjoy it.	☐	☐	☐	☐
32.	I enjoy figuring out challenging mathematics.	☐	☐	☐	☐
33.	Learning mathematics is important because it stimulates my thinking.	☐	☐	☐	☐
34.	Learning mathematics is a worthwhile exercise.	☐	☐	☐	☐
35.	I am curious when learning mathematics.	☐	☐	☐	☐
36.	When I do mathematics problems, it completely gets my attention.	☐	☐	☐	☐

SD=Strongly Disagree, D=Disagree, A=Agree, SA=Strongly Agree

		SD (1)	D (2)	A (3)	SA (4)
37.	I get a sense of satisfaction when I solve mathematics problems.	☐	☐	☐	☐
38.	Mathematics is an important subject for me because I can use in my daily life.	☐	☐	☐	☐
39.	Sometimes the mathematics course material is too hard.	☐	☐	☐	☐
40.	I feel confident in mathematics class.	☐	☐	☐	☐
41.	I study mathematics regularly.	☐	☐	☐	☐
42.	Making an effort in mathematics is worth it because it will help me in the work that I will do later on.	☐	☐	☐	☐
43.	I want to do well in mathematics.	☐	☐	☐	☐
44.	I want to try harder in mathematics.	☐	☐	☐	☐
45.	I do badly in mathematics whether or not I study for my examinations.	☐	☐	☐	☐
46.	I am very good at solving mathematics problems.	☐	☐	☐	☐

Thank you very much for your time and effort in completing this questionnaire.

Examining factors which affect mathematics achievement
of Grade 10 and 12 students in Port Moresby, Papua New Guinea

(Teacher Questionnaire)

I. Information about this Questionnaire

This questionnaire is intended for Grades 10 and 12 mathematics teachers in secondary schools. It contains items that ask for general information, school curriculum, teaching quality and classroom environment. It has been organised into sections (A, B, C and D) to the factors under study.

Your response to the questionnaires contribute to improving the standards of mathematics education in Papua New Guinea. It is vital that you respond to each of the items very carefully so that the information provided reflects your situation. Your answers will be collated with other responses from other teachers in which no individual teachers/schools can be identified. All your answers and your identity will be kept strictly confidential.

II. General Instructions to Teacher Participants

1. Please read each item carefully and answer them accurately as you can. For every section of the questionnaire, specific instructions of how to answer the items are given. If you make a mistake in responding to items that have the options, simply mark **X** on the previous choice and check another box corresponding to your new answer. If you make and error in answering questions that require writing of numbers, words and sentences, simply cross out your previous response and write the new answer next to it. **Please do not leave any item unanswered.**

2. The questionnaire must to be returned to the researcher as soon as it has been completed.

3. Complete the questionnaire during your non-contact period in the school. The researcher will help distribute the questionnaire and explain the instructions.

For Researcher Use Only

School Name: ________________________________
School ID: ___________________________________
School Type: _________________________________
Teacher ID: __________________________________

GENERAL INFORMATION

Instruction: Fill in the box by ticking and lines with words or sentences that correspond to your answer. For items that have the given options, check only one box**.**

Teacher Name: ______________________________

Grade(s) Teaching: ______________________________

1.What is your gender?

☐ Male
☐ Female

2. What highest academic qualifications do you have? (Please check only one box)

☐ Diploma
☐ Bachelor degree
☐ Master's degree/units
☐ Others please specify: ______________________________

3. Major/Area of specialisation:______________________________

4. Minor/Area of specialisation:______________________________

5. How many years of experience do you have as a classroom teacher? (Check only one box)

☐ 1-5 years ☐ 2-25 years
☐ 6-10 years ☐ 26-30 years
☐ 16-20 years ☐ More than 30 years

B. QUALITY OF TEACHING.

Instruction. This section is about your teaching instructions. Read each item carefully and answer by checking (√) only on one box. **Please answer all questions.**

AN=Almost Never, S=Seldom, ST=Sometimes, O=Often, VO=Very Often

		AN (1)	**S** (2)	**ST** (3)	**O** (4)	**VO** (5)
6.	I inspire students to learn mathematics.	☐	☐	☐	☐	☐
7.	I demonstrate to students a variety of problem-solving strategies.	☐	☐	☐	☐	☐
8.	I provide challenging tasks for the highest achieving students.	☐	☐	☐	☐	☐
9.	I adapt my teaching to engage students' interest.	☐	☐	☐	☐	☐
10.	I assist students' comprehension of mathematics.	☐	☐	☐	☐	☐
11.	I improve understanding of struggling students.	☐	☐	☐	☐	☐
12.	I explain mathematics problem solving strategies step by step.	☐	☐	☐	☐	☐
13.	I assist in the development of students' higher-order thinking skills.	☐	☐	☐	☐	☐
14.	I explain new mathematics content appropriately.	☐	☐	☐	☐	☐
15.	I explain and demonstrate how to solve mathematics problems.	☐	☐	☐	☐	☐
16.	I ensure students work on problems individually or with peers with my guidance.	☐	☐	☐	☐	☐

17. Students work on problems while I am occupied by other tasks. ☐ ☐ ☐ ☐ ☐

18. I solve mathematics like the examples in students' textbook. ☐ ☐ ☐ ☐ ☐

19. I make mathematics relevant to students. ☐ ☐ ☐ ☐ ☐

20. I ask students to work on problems that there is no immediate obvious method of solution. ☐ ☐ ☐ ☐ ☐

C. CLASSROOM ENVIRONMENT

Instruction. This section is about the classroom environment. Read each item carefully and answer by checking (√) only on ONE box. **Please answer all questions.**

SD=Strongly Disagree, D=Disagree, A=Agree, SA=Strongly Agree

		SD (1)	D (2)	A (3)	SA (4)
21.	I have a good rapport with students.	☐	☐	☐	☐
22.	I take personal interest in students' engagement in learning.	☐	☐	☐	☐
23.	I help students with doing mathematics problems.	☐	☐	☐	☐
24.	I consider students feelings in learning.	☐	☐	☐	☐
25.	I decide where the students will sit.	☐	☐	☐	☐
26.	Student to teacher ratio is manageable.	☐	☐	☐	☐
27.	There is classroom discussion.	☐	☐	☐	☐
28.	All students in the class do the same work.	☐	☐	☐	☐
29.	I teach without questions from the students.	☐	☐	☐	☐
30.	I talk with each student in the classroom.	☐	☐	☐	☐
31.	There is adequate space for teaching in the classroom.	☐	☐	☐	☐
32.	Students ask the teacher question.	☐	☐	☐	☐
33	Students explain mathematics problems and graphs	☐	☐	☐	☐

D. INSTRUCTIONAL RESOURCES

Instructions: This section is about curriculum materials. Read each item carefully and answer by checking (√) only on ONE box. **Please answer all questions.**

SD=Strongly Disagree, D=Disagree, A=Agree, SA=Strongly Agree

		SD (1)	**D** (2)	**A** (3)	**SA** (4)
34.	Instructional materials such as mathematics textbooks are adequate for students.	☐	☐	☐	☐
35.	There are adequate library resources relevant for mathematics lessons.	☐	☐	☐	☐
36.	Students are provided calculators for mathematics lessons.	☐	☐	☐	☐
37.	Teachers have adequate instructional materials and supplies.	☐	☐	☐	☐
38.	Internet is available for research and plan lessons.	☐	☐	☐	☐
39.	There are sufficient number of Grade 10 and 12 mathematics teachers.	☐	☐	☐	☐
40.	Mathematics termly programs are evaluated at the end of each term.	☐	☐	☐	☐
41.	The mathematics curriculum is relevant and practical for the students learning.	☐	☐	☐	☐
42.	Objectives of the mathematics curriculum are measurable.	☐	☐	☐	☐
43.	All Grade $10 and 12 mathematics topics are covered before the national examinations.	☐	☐	☐	☐
44.	The mathematics curriculum is regularly revised.	☐	☐	☐	☐
45.	Mathematics teachers' guides are easy to follow in implementing the curriculum.	☐	☐	☐	☐

MATHEMATICS TEACHERS' INTERVIEW PROTOCOL

Fieldwork Stage : Interview session – Quality of teaching
Day : 1

Data Identity
Pseudonym name :
Date :
Site/Venue : School
Duration : 30-35 minutes

Interview Goal
To explore teacher's teaching pedagogy and content knowledge in mathematics

Type of Interview
Unstructured interview

Language Used
English

Nature of Interview Questions
The following seven questions will be elaborated for the teachers' responses to explore their perceptions on teaching mathematics.

No.	Questions	Note
1	**Teaching Strategies** • What are some specific methods you are using to teach mathematics in the classroom? • How are the students learning from those specific methods you are using in teaching mathematics in the classroom?	
2	**Content Knowledge** • What are some topics in mathematics that you are confident in teaching in Grade 10/Grade 12? • What are some topics in mathematics you are not confident and less comfortable in teaching in Grade 10 /Grade 12? What are your reasons? • Do you have professional development/training on the mathematics topics that you are not confident teaching in the school? • How do you encourage students in mathematics in your daily teaching?	

Appendix C

Grade 10 & 12 students mathematics tests

Examining factors which affect mathematics achievement of Grade 10 and 12 students in Port Moresby, Papua New Guinea

Grade 10 Mathematics Test Items

I. Information about this Test

This test is intended for Grade 10 students in secondary schools.

Your time in doing the test will contribute to improving the standards of mathematics education in Papua New Guinea. It is vital that you answer each of the questions very carefully. Your answers will be collated with other responses from other teachers in which no individual teachers/schools can be identified. All your answers and your identity will be kept strictly confidential.

II. General Instructions to Student Participants

1. There are 40 Multiple Choice Questions and will be done in **1 hour**. Please read each question carefully and answer accordingly. Since your answers will be marked electronically, make sure you shade in the boxes of your choice completely on the answer sheet. If you make a mistake in responding to items that have the options, simply mark **X** on the previous choice and check another box corresponding to your new answer. **Please do not leave any item unanswered.**

2. The test papers and the answer sheets **MUST** to be returned to the teacher as soon as it has been completed.

For Researcher Use Only

School Name: ________________________________
School ID: ___________________________________
School Type: _________________________________
Teacher ID: __________________________________

For each question, choose the best answer by writing A, B, C or D in the spaces provided on the ANSWER SHEET.

QUSETION 1

Solve this equation $\frac{3x+12}{7} = 12$

A. 16 B. 24 C. 26 D. 72

QUESTION 2

A discount of 20% is given to a TV set. This reduces the price by K190.00. What was the original price?

A. K152 B. K228 C. K760 D. K950

QUESTION 3

What is the next number in the following series?

$27,\ 9, 3,\ 1, \frac{1}{3}, \frac{1}{9}, \ldots$

A. $\frac{1}{27}$ B. $-\frac{1}{27}$ C. $\frac{2}{9}$ D. $-\frac{2}{9}$

QUESTION 4

In a right-angle triangle, the opposite side is 118m and the angle is $76°$.

Which of the following can be used to find the length of the adjacent side?

A. $118\tan 76°$ B. $\frac{118}{\tan 76°}$

C. $\frac{\tan 76°}{118}$ D. $\frac{\tan 14°}{118}$

QUESTION 5

What is the gradient of the line passing through $(3,4)$ and $(5,8)$

A. 2 B. -2 C. -6 D. $\frac{1}{6}$

QUESTION 6

Abigail, Bridget and Clive are to share 420 oranges equally. But Bridget gets half the share of Abigail and Clive. What percentage of the orange did Abigail and Clive get?

A. 66 B. 67 C. 17 D. 90

QUESTION 7

Which of the following is equal to p^{-4} ?

A. p^4 B. $-\frac{1}{p^4}$ C. $\frac{1}{p^4}$ D. $\frac{4}{p}$

Question 8 and 9 refer to the following information

The marks of ten students in a test are: 8, 4, 5, 10, 9, 8, 6, 5, 8 and 6.

QUESTION 8

What is the modal mark?

A. 10 B. 8.5 C. 8 D. 6.9

QUESTION 9

What is the median mark?

A. 8 B. 7 C. 9 D. 6

QUESTION 10

What percentage of K25 is 25t?

A. 10 B. 1 C. 0.1 D. 0.01

QUESTION 11

When simplified; $\frac{7ab}{9} \div \frac{5b}{6}$

A. $\frac{42a}{45}$ B. $\frac{14a}{15}$

C. $\frac{35ab^2}{54}$ D. $\frac{54}{35ab^2}$

QUESTION 12

What is 3789 written in scientific notation?

A. 3.789×10^2 B. 3.789×10^3

C. 3.789×100 D. 3.789×1000

QUESTION 13

$12ab-3a+9a^2$ fully factorised is,

A. B. $3a(4ab-a+3a^2)$

C. $a(12b-3+9a)$ D. $3a(4-1+3a)$

QUESTION 14

A machine fills 1000 bottles in 5 minutes. How many bottles will it fill in 2 minutes?

A. 200 B. 300 C. 400 D. 500

QUESTION 15

Calculate the radius of a circle in cm whose area is $12.56cm^2$. Use $\pi=3.14$

A.1.0 B.1.5 C. 2.0 D.2.5

QUESTION 16

Solve the simultaneous equation: $y=5x+1$ and $y+x=13$

A. -1, 11 B. 2,-11 C. -2,-1 D. 2, 11

QUESTION 17

The sum of the set of numbers 1, 8, 2, _____, 13, 7 is 42. What is the missing number?

A. 11 B. 12 C. 10 D. 5

QUESTION 18

Which of these numbers can be evenly divided by 2, 3 and 5?

A. 2390 B. 2380 C. 2370 D. 2360

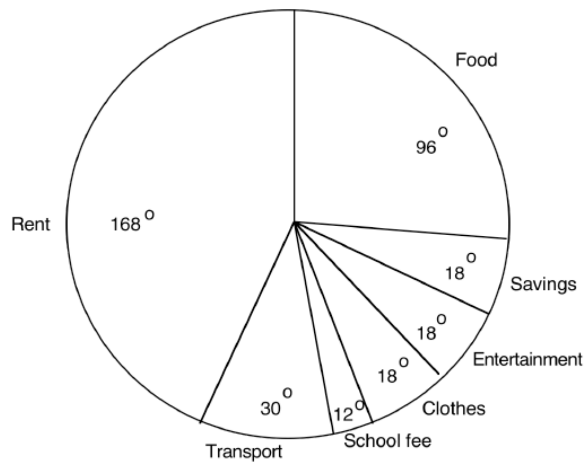
QUESTION 19

5.05kg, expressed in kilograms and grams is

	kilograms	grams	
A.	50	5	
B.	50	50	
C.	5	5	
D.	5	500	

QUESTION 20

The pie chart below shows how the Oliver family spends its fortnightly budget of K630.00.



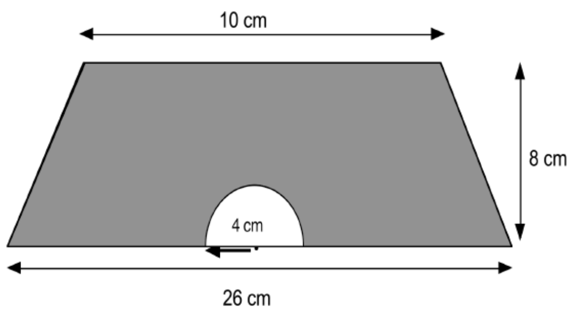
How much of the family's income is spent on school fees?

A. K210.00 B. K21.00

C. K20.10 D. K12.00

QUESTION 21

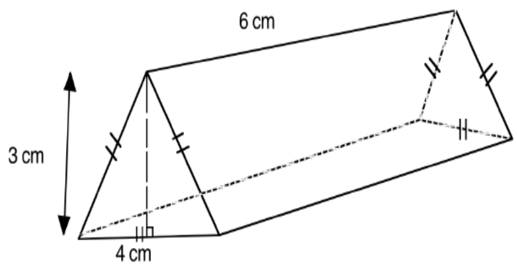
Find the shaded area of the compound shape in cm^2.
Use π = 3.14



A. 144 B. 118.88
C. 93.76 D. 25.12

QUESTION 22

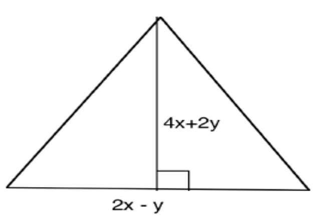
Find the total surface area in cm^2 of the triangular prism below.



A. 84 B. 72 C. 64 D. 36

QUESTION 23

The algebraic expression giving the area of the triangle is:



A. $\frac{(2x\quad y)(4x+2y)}{2}$ B. $\frac{(8x^2-8xy-2y^2)}{2}$

C. $8x^2-2y^2$ D. $\frac{4x^2-y^2}{2}$

Questions 24 and 25 refer to the information below.

The table shows the number of glasses of water drunk by a group of students.

Number of glasses of water	1	2	3	4	5	6
Frequency	10	3	5	2	6	4

QUESTION 24

Find the number of students who drank less than 3 glasses of water.

A. 13 B. 3 C. 5 D.18

QUESTION 25

Find the number of students who drank less than four glasses of water.

A. 18 B. 2 C. 7 D.20

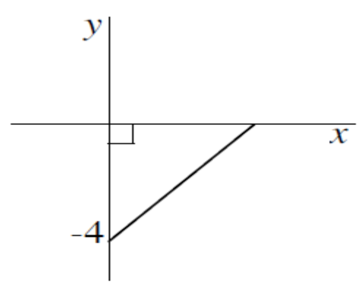
QUESTION 26

Susan ran 3000 m in exactly 8 minutes. What was her average speed in meters per second?

A. 3.75 B. 6.25 C. 16. D.37.5

QUESTION 27

The diagram shows an isosceles right-angled triangle on a set of axes.



What is the equation of the hypotenuse?

A. $y=-x-4$ B. $y=x-4$
C. $y=4-x$ D. $y=x-(-4)$

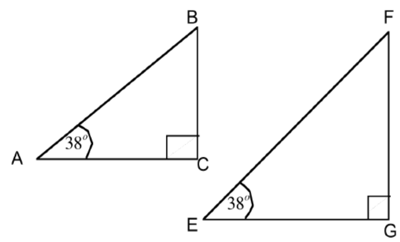
QUESTION 28

The selling price of a radio in January was K140.00. In June it was sold on a 20% discount price.
What is the new price?

A. K20 B. K28 C. K112 D. K120

QUESTION 29

Triangle ABC is similar to triangle EFG.

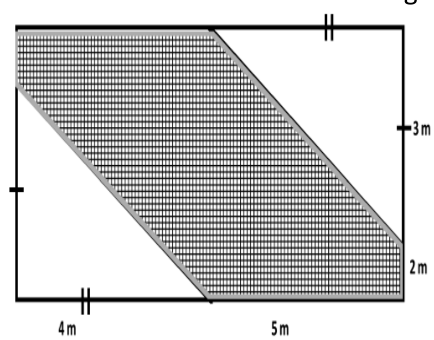


What is the length in centimetres of the corresponding side to AC?

A. 9 B. 12 C. 15 D. 25

QUESTION 30

Find the shaded area in this rectangle in m^2



A. 33 B.39 C. 45 D. 51

QUESTION 31

A length of fabric 46.8 metres long was divided in the ratio 2:3:1. Find the length of the smallest proportion of the ratio.

A. 7.8 m B. 11.2 m C. 15.6 m D. 23.4 m

QUESTION 32

A pencil and two biros cost K1.80. Two pencils and a biro cost K1.20. What is the cost of two pencils?

A. K1.80 B. 80t C. 40t D. 20t

Question 33

7, 8 8, 11, 16 are five numbers.
Find the mean of these five numbers.

A.50 B. 5 C. 8 D. 10

QUESTION 34

If $5x+10=2x+4$.The value of x is?
A. 2 B. -2 C. 2x D. 3x

QUESTION 35

If $p=4, q=-2$ and $r=3$
Find the value of $\frac{p-2q+2r}{p+r}$
A. 14 B. 7 C. 2 D. 3

QUESTION 36

What is the answer in positive indices for
$3m^{-3}n^5 \div 18m^5n^3$

A. $\frac{3m}{6n}$ B. $\frac{n^2}{6m^8}$ C. $\frac{m^3}{m^5}$ D. $\frac{n^2m^3}{6m^5}$

QUESTION 37

A salesman earns K200 per week plus K40 commission for each item he sells. How many items does he need to sell to earn a total of K2640 in two weeks?

A. 33 B. 56 C. 61 D. 66

QUESTION 38

Find the y-intercept of a straight line passing through the point (2, 4) whose gradient is 3.

A. 2   B. -2   C. 6   D. 3

QUESTION 39

A quadrilateral has diagonals of the same length and they bisect at right angles.
What is the name of the quadrilateral?

A. Rectangle   B. Parallelogram

C. Square   D. Rhombus

QUESTION 40

What is $12\frac{1}{2}$% of 10kg in grams?

A. 1250   B. 125   C. 12.5   D. 1.25

END OF TEST

Examining into factors that affect Grade 10 and 12 students' mathematics results in Port Moresby, Papua New Guinea (PNG)

Grade 12 Mathematics Test Items

I. Information about this Test

This test is intended for Grade 12 students in secondary schools.
Your time in doing the test will contribute to improving the standards of mathematics education in Papua New Guinea. It is vital that you answer each of the questions very carefully. Your answers will be collated with other responses from other teachers in which no individual teachers/schools can be identified. All your answers and your identity will be kept strictly confidential.

II. General Instructions to Student Participants

1. There are 40 Multiple Choice Questions and will be done in **1 hour**. Please read each question carefully and answer
2. accordingly. Since your answers will be marked electronically, make sure you shade in the boxes of your choice completely on the answer sheet. If you make a mistake in responding to items that have the options, simply mark **X** on the previous choice and check another box corresponding to your new answer. **Please do not leave any item unanswered.**

3. The test papers and the answer sheets **MUST** to be returned to the teacher as soon as it has been completed.

For Researcher Use Only

School Name: ________________________________
School ID: __________________________________
School Type: ________________________________
Teacher ID: _________________________________

For each question, choose the best answer by writing A, B, C, and D in the spaces provided on the ANSWER SHEET.

QUESTION 1

4 bananas and 3 coconuts cost K3.20. 12 bananas and 2 coconuts cost K4.00. Find the cost of each item.

A. 20t, 80t B. 20t, 60t

C. 80t, 50t D. K1, K2

QUESTION 2

Simplify $1-\frac{3x-1}{1-x}$

A. 2 B. 3x C. D.

QUESTION 3

The solution of the quadratic equation $x^2-2x-3=0$ is

A. $x=3, x=1$ B. $x=-1, x=3$

C. $x=-3, x=1$ D. $x=2, x=2$

QUESTION 4

A map has a scale of 1:25000.

What distance is represented by a length of 6.8cm on the map?

A. 170km B. 17.0km

C. 1.70km D. 0.17km

QUESTION 5

$\frac{(x-1)}{(x^2-1)}$ can be simplified to

A. $\frac{(x-1)}{(x+1)}$ B. $\frac{1}{(x+1)}$

C. 1 D. $x+1$

QUESTION 6

Which expression cannot be simplified any further?

A. $\frac{m-n}{n-m}$ B. $\frac{2x-2y}{y-x}$

C. D. $\frac{r+2s}{4s-2r}$

QUESTION 7

t^3-8 can be expressed as

A. $(t-8)^3$

B. $(t-2)(t^2-2t+4)$

C. $(t-2)(t^2+2t+4)$

D. $(t-2)(t^2-t+4)$

QUESTION 8

Which of these terms does not mean "data item'?

A. score B. value

C. average D. measurement

QUESTION 9

The spread of a frequency is measured by:

A. mean B. median

C. standard deviation D. average

QUESTION 10

Which of these is the best option?

A congruent shape has all corresponding

A. angles equal

B. sides equal

C. sides and angles equal

D. none of the above

QUESTION 11

Douglas, Renae and Lynette were given K235.00 and told to divide it in the ratio 2:1:3 amongst themselves in order of names listed.

To the nearest toea each will receive

A. K39.17, K120.00, K78.33

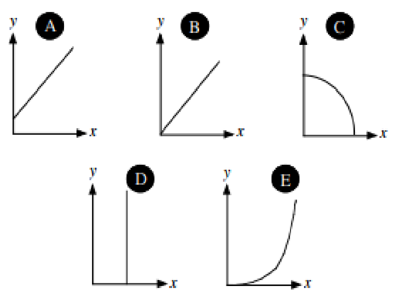
B. K78.33, K39.17, K117.50

C. K117.50, K39.17, K140.00

D. None of the above

Question 12, 13 and 14 refers to the graphs below.

Study the sketches of the 5 graphs below.



QUESTION 12

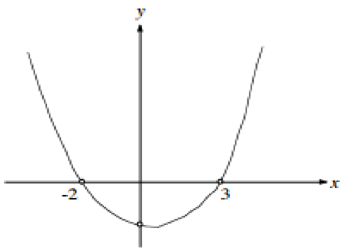
Which graph represents $y=2x^2$

QUESTION 13
Which graph represents $y-2=3$

QUESTION 14
Which graph represents $X=4$

Question 15 and 16 refers to the graph below

The graph below is a parabola.



QUESTION 15

Write the equation of the graph in the form

A. $y=x^2+x+6$ B. $y=x^2+x-6$

C. $y=x^2-5x+6$ D. $y=x^2-x-6$

QUESTION 16

What are the coordinates of the y intercept of the parabola?

A. 3 B. -2 C. -6 D. 5

Question 17, 18 and 19 refers to the table below.

Year	School A	School B	School C
2015	684	417	209
2016	782	483	241
2017	811	461	222

QUESTION 17
Which school is likely to be a new school?

A. School A B. School B

C. School C D. All the schools

QUESTION 18
What is the mean enrolment of school B for the 3 years?

A. 454 B. 483 C. 1,900 D. 1,904

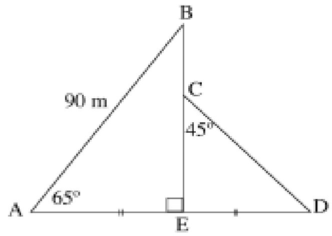
QUESTION 19

Of all the data provided, which is the median enrolment record?

A. 461 B. 417 C. 811 D.209

Questions 20, 21, 22 refers to the diagram below.

In the diagram below AB= 90cm; AE is equal to ED, $\angle EAB=65°$ and $\angle ECD=45°$



QUESTION 20

What is the length of AE to nearest meter?

A. 82 B. 38 C. 90 D. 45

QUESTION 21

Find the area of triangle ABE to nearest square meter.

A. 90 B. 45 C. 1551 D.1556

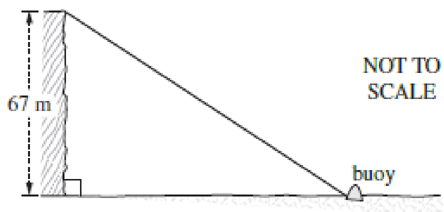
QUESTION 22

What is the perimeter of triangle ABE to the nearest meter?

A.210 B. 82 C.39 D.90

QUESTION 23

From the top of a cliff 67 metres above sea level, the angle of depression of a buoy is 42°.



How far is the buoy from the base of the cliff, to the nearest metre?

A. 60 m B. 74m C. 90m D. 100m

QUESTION 24
What amount must be invested now at 4% per annum, compounded quarterly, so that in five years it will have grown to K60000?

A. K8919 B. K11156

C. K49173 D. K49316

QUESTION 25

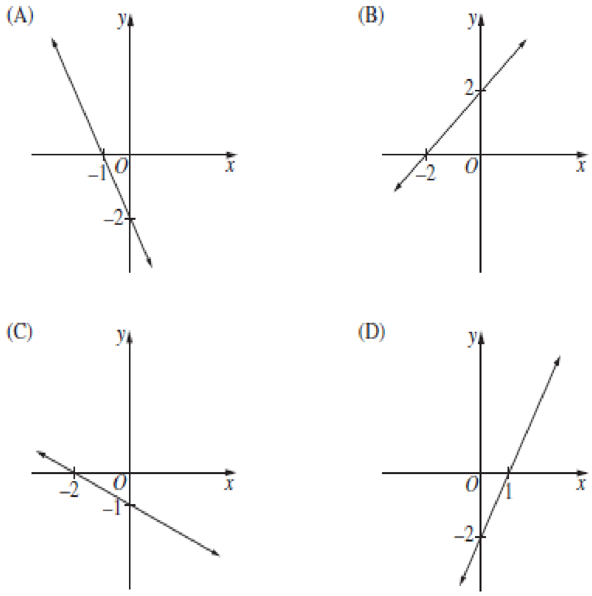
Jane sells jewellery. Her commission is based on a sliding scale of 6% on the first K2000 of her sales, 3.5% on the next K1000, and 2% thereafter.
What is Jane's commission when her total sales are K5670?

A. K188.40 B. K208.40

C. K 321.85 D. K652.05

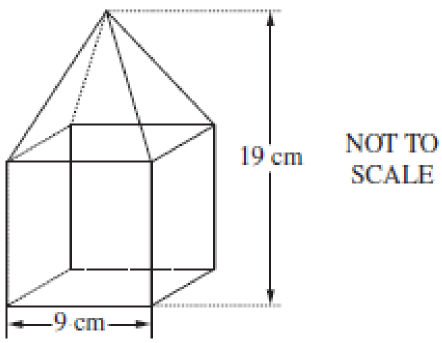
Question 26

Which of the following is the graph of y = 2x − 2?



QUESTION 27

A square pyramid fits exactly on top of a cube to form a solid.



What is the volume of the solid?

A.513 cm^3 B. 999 cm^3

C. 1242 cm^3 D. 1539 cm^3

QUESTION 28

In an experiment, a standard six-sided die was rolled 72 times. The results are shown in the table.

Number of die	Frequency
1	6
2	12
3	10
4	20
5	19
6	15

Which number on the die was obtained the expected number of times?

A. 1 B. 2 C. 3 D. 6

QUESTION 29

What is the slope of the line with equation $2x - 4y + 3 = 0$?

A. -2 B. 2 C. $-\frac{1}{2}$ D. $\frac{1}{2}$

QUESTION 30

The first three terms of an arithmetic series are 3, 7 and 11.
What is the 15th term of this series?

A. 59 B.63 C.465 D. 495

QUESTION 31

What is 0.00523359 written in scientific notation, correct to 4 significant figures?

A. 5.2336×10^{-2} B. 5.234×10^{-2}

C. 5.2336×10^{-3} D. 5.234×10^{-3}

QUESTION 32

On a school report, a student's record of completing homework is Graded using the following codes.

C = consistently
U = usually
S = sometimes
R = rarely
N = never

What type of data is this?

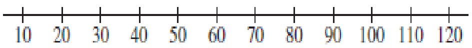
A. Categorical, ordinal

B. Categorical, nominal

C. Quantitative, continuous
D. Quantitative, discrete

QUESTION 33

The times, in minutes, that a large group of students spend on exercise per day are presented in the box-and-whisker plot.

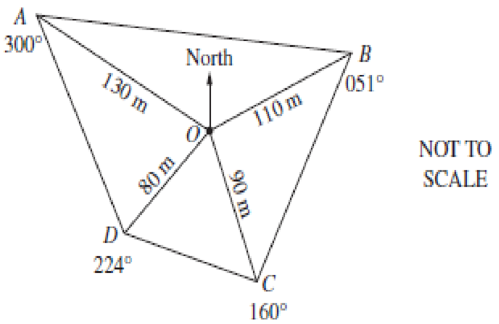


What percentage of these students spend between 40 minutes and 60 minutes per day on exercise?

A. 17% B. 20% C. 25% D. 50%

QUESTION 34

The diagram shows a radial survey of a field *ABCD*.



In triangle *AOB*, what is the size of *AOB?*

A. 51° B. 111° C. 125° D. 249°

QUESTION 35

The probability of winning a game is $\frac{7}{10}$

Which expression represents the probability of winning two consecutive games?

A. $\frac{7}{10}\times\frac{6}{9}$ B. $\frac{7}{10}\times\frac{6}{10}$

C. $\frac{7}{10}\times\frac{7}{9}$ D. $\frac{7}{10}\times\frac{7}{10}$

QUESTION 36

A Student Representative Council (SRC) consists of five members. Three of the members are being selected to attend a conference.
In how many ways can the three members be selected?

A. 10 B. 20 C. 30 D. 60

QUESTION 37

Which of the following is $3x^0 + 5x$ in its simplest form?
A. 6*x* B. 8*x* C. 1 + 5*x* D. 3 + 5*x*

QUESTION 38

Which of the following is $4x + 3y - x - 5y$ in its simplest form?
A. $3x - 2y$ B. $3x + 8y$

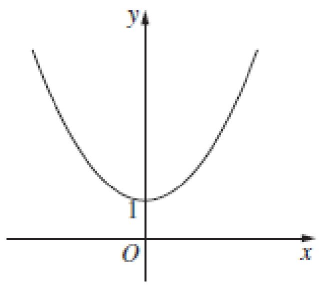
C. $5x - 2y$ D. $5x + 8y$

QUESTION 39

A measurement of 72 cm is increased by 20% and then the result is decreased by 20%. What is the new measurement, correct to the nearest centimetre?
A. 46 cm B. 69 cm C. 72 cm D. 104 cm

QUESTION 40

The diagram shows the graph of an equation.



A. $y = \frac{x}{3} + 1$ B. $y = 3^x + 1$

C. $y = 3x^2 + 1$ D. $y = 3x^3 + 1$

END OF TEST

Grade 10 and 12 test Answer Keys

GRADE 10 ANSWER KEY					GRADE 12 ANSWER KEY			
1	**B**	21	**B**		1	**A**	21	**C**
2	**D**	22	**D**		2	**C**	22	**A**
3	**A**	23	**A**		3	**B**	23	**B**
4	**B**	24	**A**		4	**A**	24	**C**
5	**A**	25	**D**		5	**B**	25	**B**
6	**C**	26	**B**		6	**A**	26	**D**
7	**C**	27	**B**		7	**C**	27	**B**
8	**C**	28	**C**		8	**C**	28	**B**
9	**B**	29	**B**		9	**C**	29	**D**
10	**B**	30	**A**		10	**C**	30	**A**
11	**A**	31	**A**		11	**B**	31	**D**
12	**B**	32	**C**		12	**B**	32	**A**
13	**A**	33	**D**		13	**A**	33	**C**
14	**C**	34	**B**		14	**D**	34	**B**
15	**C**	35	**C**		15	**D**	35	**D**
16	**D**	36	**B**		16	**C**	36	**A**
17	**A**	37	**B**		17	**C**	37	**D**
18	**C**	38	**A**		18	**B**	38	**A**
19	**D**	39	**C**		19	**A**	39	**B**
20	**D**	40	**A**		20	**B**	40	**C**

Lower and Upper Secondary School Units/Topics

Grade 9 units/topics	Grade 10 units/topics
9.1 Mathematics in our community • Number and operations • Ration and percentage • Rates • Money • Measurement • Data	**10.1 Managing your money** • Percentage and money • Spending money • Saving money • Borrowing money • Earning money
9.2 Patterns of change • Directed numbers • Indices • Equations	**10.2 Functions and graphs** • Directed numbers • Indices, scientific notation • Basic algebra • Graphs
9.3 Working with data • Recording data • Sorting and organising data • Measures of central tendency and measures • Misleading statistical graphs • Experimental probabilities • Probabilities based on geometry	**10.3 Trigonometric application** • Triangle property and terminology • Similar triangles • Pythagoras theorem • Introduction to trigonometric functions • Solving right triangles using trigonometry
9.4 Design in 2D and 3D Geometry • Properties of plane figure • Angles and lines • Surface area and volume	

Grade 11 general mathematics units/topics	Grade 12 general mathematics units/topics
11.1 Number and application (10 weeks) • Basic numeracy • Unit of measurement • Ratio and proportion • Basic Algebra **11.2 Managing Money 1 (8 weeks)** • Earnings and spending • Budgeting and loans • Equations **11.3 Statistics (6 weeks)** • Exploring data • Analysing data **11.4 Geometry (8 weeks)** • Lines, angles, triangles, and regular polygons • Geometric construction • Circles **11.5 Trigonometry (8 weeks)** • Trigonometry • Vectors	**12.1 Measurement (6 weeks)** • Scales and dimensions • Surveying **12.2 Managing Money 2 (8 weeks)** • Interest and Inflation • Consumer credit • Investments • Insurance • Using simple algebraic manipulation of financial formulae **12.3 Probability and Statistics (6 weeks)** • Probability • Correlation and regression **12.4 Algebra and Graphs (6 weeks)** • Equations • Graphs and functions **12.5 Applying Geometry in Papua New Guinean Arts (4 weeks)** • Tessellations and polyhedra

Grade 11 advance mathematics units/topics	Grade 12 advance mathematics units/topics
11.1 Number and application (10 weeks) • Basic numeracy • Unit of measurement • Ratio and proportion • Basic Algebra **11.2 Graphs and Functions (8 weeks)** **11.3 Geometry (6 weeks)**	**12.1 Patterns and Algebra (6 weeks)** **12.2 Trigonometry and Vectors 2 (8 weeks)** **12.3 Calculus (6 week)**

Appendix D

Information sheets, and consent and complaint form

RESEARCH PROJECT INFORMATION SHEET

(For Grade 10 and 12 students)

I am Jerome Oko, a PhD student in the School of Education at the University of Adelaide. I am conducting a research on a project titled **'Examining factors which affect mathematics achievement of Grade 10 and 12 students in Port Moresby,Papua New Guinea"** under the supervision of Associate Professor Sivakumar Alagumalai, Dr I. Gusti Damawan and Associate Professor Nicholas Buchdahl at the University of Adelaide, South Australia.

You will be invited to complete a questionnaire in the specified timeframes during your homeroom class/classroom advisory period and do a mathematics test on a different specified time and location in the school. The survey questionnaire will be collected after the homeroom class/classroom advisory period.

Student level factors questionnaire (with appropriate time allocation for each section shows):

- General information items - 5 minutes
- Items about motivation in learning mathematics - 10 minutes
- Items about attitude towards mathematics - 10 minutes

NOTE: Participant in this study is voluntary and is not related to assessment. If you cannot complete the questionnaire during homeroom/advisory class time, you can take the questionnaire home and can be completed after-hours and please make sure you complete it and return it to the submission box in the classroom next day. However, there is not an option for non-participation

The mathematics test will be done in 1 hour and will be collected by your teacher.

The research aims are to examine the factors that are influencing mathematics achievement of students. This is a major issue in many countries and it is no different to PNG's situation. The specific aim is to identify the school, teacher and student level factors and their relationships that might influence mathematics results. The study will identify the different factors that may have influence on students' mathematics achievement. These factors would provide useful information on how they affect Grade 10 and 12 students' mathematics achievement. The results of this study should provide educationists and policymakers a significant amount of information and a deeper understanding of the factors which affect~~ing~~ mathematics results.

In conducting this study**, ethics are strictly observed (University of Adelaide HREC Project Approval Number: H-2017-113)**, and requires participant consent. Participation in this study is completely voluntary, and you may withdraw at any stage or choose not to answer any questions. No negative consequence will follow from such withdrawal. No findings which will identify any individual participant or school will be published, and your privacy will be protected at all stages of the research. Hence, all student, teacher and school will be de-identified.

Should you require additional information regarding this research, please contact me by email jerome.oko@adelaide.edu.au. Should I be unavailable, my supervisor, Associate Professor Sivakumar Alagumalai on telephone (+618) 8313 5630 and email sivakumar.alagumalai@adelaide.edu.au.

Thank you for considering this request.

Sincerely yours,
Jerome Oko

THE UNIVERSITY *of* ADELAIDE

RESEARCH PROJECT INFORMATION SHEET

(For Grade 10 and 12 mathematics teachers)

I am Jerome Oko, a PhD student in the School of Education at the University of Adelaide. I am conducting a research leading to the production of thesis on the subject **'Examining factors which affect mathematics achievement of Grade 10 and 12 students in Port Moresby, Papua New Guinea**, under the supervision of Associate Professor Sivakumar Alagumalai, Dr I. Gusti Damawan and Associate Professor Nicholas Buchdahl at the University of Adelaide, South Australia.

You will be asked to complete a questionnaire in the specified timeframes during your non- contact period at the school (Term 3).

Teacher level questionnaires (with appropriate time allocation for each section shows):

- General Information items - 5 minutes
- Items about teaching quality -10 minutes
- Items about classroom environment -10 minutes
- Items about curriculum materials -10 minutes

Interviews about mathematics teaching methods and practices (this will happen to get more comprehensive data on teaching strategies in mathematics, and will take place after a week's administration of survey questionnaires) 30 to 35 minutes after official hours at school.

Questionnaires will be collected after the teachers have filled out the questionnaires on their non-contact period. If you cannot complete the questionnaire during that time, you can take the questionnaire home and can be completed after-hours and please make sure you complete it and return it to the researcher next day. However, there is not an option for non-participation

The research aim is to examine the factors that are influencing mathematics achievement of students. This is a major issue in many countries including PNG. The specific aim is to identify the school, teacher and student levels factors and their relationships that might influence mathematics results. The study will identify the different factors that may have influence on students' mathematics achievement. These factors would provide useful information on how they affect Grade 10 and 12 students' mathematics achievement. The results of this study should provide educationists and policymakers a significant amount of information and a deeper understanding of the factors which affecting mathematics results.

In conducting this study**, ethics are strictly observed (University of Adelaide HREC Project Approval Number: H-2017-113)**, and requires participant consent. Participation in this study is completely voluntary, and you may withdraw at any stage or choose not to answer any questions. No negative consequence will follow from such withdrawal. No findings which will identify any individual participant or school will be published, and your privacy will be protected at all stages of the research. Hence, all student, teacher and school will be de-identified.

Should you require additional information regarding this research, please contact me by email-jerome.oko@adelaide.edu.au. Should I be unavailable, my supervisor, Associate Professor Sivakumar Alagumalai on telephone (+618) 8313 5630 and email sivakumar.alagumalai@adelaide.edu.au.

Thank you for considering this request.

Sincerely yours,

Jerome Oko

Dear Parents/Guardian,

I am Jerome Oko, a PhD student in the School of Education at the University of Adelaide. I am conducting a research leading to the production of thesis on the subject **'Examining factors which affect mathematics achievement of Grade 10 and 12 students in Port Moresby, Papua New Guinea**, under the supervision of Associate Professor Sivakumar Alagumalai, Dr I. Gusti Damawan and Associate Professor Nicholas Buchdahl at the University of Adelaide, South Australia.

I would be most grateful if you would grant permission for your child to assist in this project by completing a questionnaire on 'Examining factors affecting Grade 10 and 12 students' mathematics achievements'. I will be undertaking this research on site at the school during class patron's period so as to minimize interference with your child's class time.

Your child will be asked to complete a questionnaire in 25 minutes during his/her homeroom period and do mathematics test for 1 hour on the topics covered with their teachers in the school.

NOTE: ***If completing the questionnaire in class patron time is not possible, your child will be allowed to take the questionnaire home but please encourage him/her to complete it and then return it to his/her mathematics teacher the next day.***

The research aims are to examine the factors that are influencing mathematics achievement of students. This is a major issue in many countries including PNG. The specific aim is to identify the school, teacher and student levels factors and their relationships that might influence mathematics results. The study will identify the different factors that may have influence on students' mathematics achievement. These factors would provide useful information on how they affect Grade 10 and 12 students' mathematics achievement. The results of this study should provide educationists and policymakers a significant amount of information and a deeper understanding of the factors which affecti~~ng~~ mathematics results.

In conducting this study**, ethics are strictly observed (University of Adelaide HREC Project Approval Number: H-2017-113)**, and requires participant consent. Participation in this study is completely voluntary, and your child may withdraw at any stage or choose not to answer any questions. No negative consequence will follow from such withdrawal. No findings which could identify any individual participant will be published, and your privacy will be protected at all stages of the research. Any information provided will be treated in the strictest confidence and neither participants nor school will be individually identifiable in the resulting thesis, report or other publications.

Should you require additional information regarding this research, please contact me by phone telephone on (+618) 8303-7196 email jerome.oko@adelaide.edu.au. Should I be unavailable, my supervisor, Associate Professor Sivakumar Alagumalai on telephone (+618) 8313 5630 and email sivakumar.alagumalai@adelaide.edu.au.

This research project has been approved by the University of Adelaide Human Research Ethics Committee. The Secretary of this Committee can be contacted on (08) 8303-6028, fax (08) 8303-7325, email hrec@adelaide.edu.au

Thank you for your attention and assistance.

Sincerely yours,

Jerome Oko

Consent form

Human Research Ethics Committee (HREC)

CONSENT FORM FOR STUDENTS AND TEACHERS

- I have read the attached Information Sheet and agree to take part in the following research project:

Title:	**Examining factors which affect Grade 10 and 12 students' mathematics achievement in Port Moresby,Papua New Guinea (PNG)**
Ethics Approval Number:	**(To be inserted)**

- I have had the project, so far as it affects me, fully explained to my satisfaction by the research worker. My consent is given freely.
- Although I understand the purpose of the research project it has also been explained that involvement may not be of any benefit to me.
- I have been informed that, while information gained during the study may be published, I will not be identified and my personal results will not be divulged.
- I understand that I am free to withdraw from the project at any time.

  A decision not to participate or withdraw will not affect the teachers' status of employment at the affiliated school/institution and students' academic progress and result.
- I agree to the interview being audio recorded. Yes ☐ No ☐
- I am aware that I should keep a copy of this Consent Form, when completed, and the attached Information Sheet.

Participant to complete:

Name: ______________________________ Signature: _________________________________ Date: _______________

Researcher/Witness to complete:

I have described the nature of the research to ____

(print name of participant)

and in my opinion, she/he understood the explanation.

Signature: ___________________________ Position: __________________________________ Date: _______________

The University of Adelaide

Human Research Ethics Committee (HREC)

CONTACTS FOR INFORMATION ON PROJECT AND INDEPENDENT COMPLAINTS PROCEDURE

The following study has been reviewed and approved by the University of Adelaide Human Research Ethics Committee:

Project Title:	**Examining factors which that are affect Grade 10 and 12 students' mathematics achievement in Port Moresby, Papua New Guinea (PNG)**
Approval Number:	**(To be inserted)**

The Human Research Ethics Committee monitors all the research projects which it has approved. The committee considers it important that people participating in approved projects have an independent and confidential reporting mechanism which they can use if they have any worries or complaints about that research.

This research project will be conducted according to the NHMRC National Statement on Ethical Conduct in Human Research (see http://www.nhmrc.gov.au/publications/synopses/e72syn.htm).

1. If you have questions or problems associated with the practical aspects of your participation in the project, or wish to raise a concern or complaint about the project, then you should consult the project co-ordinator:

Name:	**Associate Professor Sivakumar Alagumalai**
Phone:	+61 8 8313 5630
Email:	sivakumar.alagumalai@adelaide.edu.au

2. If you wish to discuss with an independent person matters related to:
   - making a complaint, or
   - raising concerns on the conduct of the project, or
   - the University policy on research involving human participants, or
   - your rights as a participant,

   contact the Human Research Ethics Committee's Secretariat on phone (08) 8313 6028 or by email to hrec@adelaide.edu.au

Appendix E

Letters of request for permission to conduct the study

To the School Principal

Dear Sir/Madam,

I am Jerome Oko, a PhD student in the School of Education at the University of Adelaide. I am conducting a research on a project titled **'Examining factors which affect mathematics achievement of Grade 10 and 12 students in Port Moresby,Papua New Guinea**, under the supervision of Associate Professor Sivakumar Alagumalai, Dr I. Gusti Damawan and Associate Professor Nicholas Buchdahl at the University of Adelaide, South Australia.
I would be most grateful if you would grant permission for selected Grade 10 and 12 classes and their mathematics teachers at your school to assist in the project by completing a questionnaire on **'Examining factors which affect mathematics achievement of Grade 10 and 12 students.** I will be undertaking this research on the site at the school during class patron time so as to minimize interference with your school's core class time.

Grade 10 and 12 students will be asked to complete a questionnaire for 25 minutes in their class advisory /homeroom period and the next day they will do a mathematics test for 1 hour on the topics they have covered in class with their teachers.

(If the above time allocation is not possible, please suggest a most appropriate time for me to undertake this survey at your school).

Simultaneously, teachers' version of the questionnaires will be completed by Grade 10 and 12 mathematics teachers. Teachers should be able to complete the questionnaire in 30 minutes in their non-contact time in the school. A short interview with two selected mathematics teachers will be carried out after a week for 30-35 minutes. The interview will be centered upon certain aspects of the topic.

In conducting this study**, ethics are strictly observed (University of Adelaide HREC Project Approval Number: H-2017-113)**, and requires participant consent. Any information provided will be treated in the strictest confidence and no participants or school will be individually identifiable in the resulting thesis, report or other publications. This research requires participant consent. Participation in this study is completely voluntary, and your students and teachers may withdraw at any stage or choose not to answer any questions. No negative consequence will follow from such withdrawal. Since participation is purely voluntary, non- participation will not affect students' academic progress or the teachers' professional progress in any way.

In this project, audio recording of mathematics teachers' interviews will be made. Therefore, will seek your selected mathematics teachers' consent, on the form attached to audio record for the interview, on conditions that name or identities are not revealed or disclosed for any reason.

Should you require additional information regarding this research, please contact me by phone telephone on (+618) 8303-7196, email jerome.oko@adelaide.edu.au. Should I be unavailable, my supervisor, Associate Professor Sivakumar Alagumalai on telephone (+618) 8313 5630 and email sivakumar.alagumalai@adelaide.edu.au.

Please also see the attached independent complaints procedure form should your teachers or students have any complaints about the project. This research project has been approved by the University of Adelaide Human Research Ethics Committee. The Secretary of this Committee can be contacted on (08) 8303-6028, fax (08) 8303-7325,email. hrec@adelaide.edu.au

Thank you for your attention and assistance.

Sincerely yours,

Jerome Oko

Dr Kombra
The Secretary for Education
National Department for Education (NDOE)
P.O Box
Waigani,NCD

Dear Dr Kombra,

I am Jerome Oko, a PhD student in the School of Education at the University of Adelaide. I am conducting a research leading to the production of thesis on the tittle **'Examining factors which affect mathematics achievement of Grade 10 and 12 students in Papua New Guinea**, under the supervision of Associate Professor Sivakumar Alagumalai, Dr I. Gusti Damawan and Associate Professor Nicholas Buchdahl at the University of Adelaide, South Australia.

The research aims is to examine the factors that are influencing mathematics achievement of students. This is a major issue in many countries including PNG. The specific aim is to identify the school, teacher and student levels factors and their relationships that might influence mathematics results. The study will identify the different factors that may have influence on students' mathematics achievement. These factors would provide useful information on how they affect Grade 10 and 12 students' mathematics achievement. The results of this study should provide educationists and policymakers a significant amount of information and a deeper understanding of the factors which affecting mathematics results.

In this regard, may I request your office for approval to conduct this research in selected secondary schools in the four regions (Highlands, Momase, Southern and New Guinea Islands) of Papua New Guinea. The researcher is also requesting for your approval to have access to Grade 10 and 12 standardized mathematics national examination results with their raw scores of the examinations towards the end of the academic year 2017 (December) from the Measurement Service Division (MSD). The data collected from the schools from the teachers and students with the standardized mathematics national examination and raw scores of the examinations will play a significant role in the project.

In conducting this study**, ethics are strictly observed (University of Adelaide HREC Project Approval Number: H-2017-113)**, and requires participant consent. Any information provided will be treated in the strictest confidence and neither participants nor school will be individually identifiable in the resulting thesis, report or other publications. Participation in this study is completely voluntary, and they may withdraw at any stage or choose not to answer any questions. No negative consequence will follow from such withdrawal.

This research project has been approved by the University of Adelaide Human Research Ethics Committee. The Secretary of this Committee can be contacted on (08) 8303-6028,fax (08) 8303-7325,email hrec@adelaide.edu.au

Should you require additional information regarding this research, please contact me by phone telephone on (+618) 8303-7196 or email jerome.oko@adelaide.edu.au. Should I be unavailable, my supervisor, Associate Professor Sivakumar Alagumalai on telephone (+618) 8313 5630 and email sivakumar.alagumalai@adelaide.edu.au.

Please see the attached independent complaints procedure form should you have any complaints about the project.

Thank you for considering this request

Sincerely yours, NOTED:

Jerome Oko

Professor Faye McCallum
Head School of Education

Appendix F

PNG MSU Grade 10 & 12 Mathematics Table of Specification 2017

Grade 10 Mathematics Table of Specification 2017

The knowledge and skills students have learnt in Mathematics can be assessed in a three hour examination. The exam will assess knowledge of Mathematics and students' understanding of the application of this knowledge. Students taking this exam would have achieved the broad learning outcomes over the two years of study.

The Mathematics examination paper contains the following parts;
Part A: Multiple Choice questions worth 25 marks
Part B Short answer questions worth 20 marks
Part C Extended response worth 5 marks

The examination will consist of:

1. Simple recall questions - 20%
2. Comprehension questions - 50%
3. Application, Analysis, Synthesis, Evaluation questions - 30%

Items will be written based on the following;

Unit	Topic	Knowledge 20 %			Comprehension 50 %			High Order 30 %		
		MC 22	SA 17	ER 1	MC 55	SA 41	ER 5	MC 33	SA 25	ER 2
9.1	Measurement	1			3	2		2	1	
	Rates	1	1		2	2		2	1	
	Ratio & Proportion	1	1		3	1	1	2	1	
	Number & Operations	1		1	2	2		2	1	
9.2	Equations	1	1		3	2		1	1	1
	Directed Numbers	1	1		2	2		2	1	
	Patterns of Change	1			2	2	1	1	1	
9.3	Presenting Data	1	1		3	2		1	1	
	Central Tendency and Spread	1	1		2	2		2	1	

	Experimental Probabilities	1			2	2		1	1	
9.4	Surface Area and Volume	1	1		3	2		2	1	
	Deductive Reasoning	1	1		2	2		1	1	
	Design in 2D & 3D Geometry	1	1		3	2	1	2	1	
10.1	Spending Money	1	1		2	2		1	1	1
	Managing your Money	1			2	2		2	1	
	Earning Money	1	1		3	2		1	2	
10.2	Quadratic Expressions	1	1		2	2		1	1	
	Basic Algebra	1	1		2	2	1	1	1	
	Indices and Scientific Notations	1			2	2		1	2	
	Simultaneous Equations	1	1		3	2		2	1	
10.3	Trigonometric Applications	1	1		2	2		1	1	
	Navigation		1		2			1	1	
	Pythagoras Theorem	1	1		3		1	1	1	
		22	**17**	**1**	**55**	**41**	**5**	**33**	**25**	**2**
	Total	**40**			**100**			**60**		

Items written will be based on the following Broad Learning Outcomes (BLO) and Unit Learning Outcomes (ULO).

BLO	Unit Learning Outcome (ULO)
2. Identify and apply mathematic skills in everyday life	
	9.1.1identify everyday situations where the basic operations can be applied
	9.1.2 make calculations using a range of methods and be aware of whether or not the result is reasonable
	9.2.1 identify and create representation of patterns to solve equations
	9.3.2 estimate and calculate probabilities
	9.4.2 classify shapes into families and their subgroups and justify reasoning
	10.1.1 apply percentages in a range of financial transactions and be aware whether or not the result is reasonable
	10.2.1 interpret and develop linear and quadratic equations from information provided in a given context
	10.2.2 plot and sketch graphs of linear and quadratic equations
	10.3.3 identify and apply calculations and be aware of whether or not the result is reasonable
3. Investigate and solve mathematical problems	9.2.2 create, investigate and interpret equations, explain the effect of order of operations, and justify solutions to equations
	9.3.1 represent, interpret, analyse and solve problems using discrete and continuous data
	9.4.3 interpret, analyse and solve measurement problems and justify selections and applications of formulae
	10.1.2 determine the costs and benefits of simple credit and investment or saving schemes
	10.3.1 demonstrate understanding of the basic concepts of similar figures and trigonometric ratios

	10.3.2 apply Pythagoras' theorem and trigonometric ratios to solve right- angle triangles and find lengths and angles in simple real problems
4. Communicate mathematical processes and results both orally and in writing	9.1.4 communicate mathematical processes and results
	9.3.2 communicate mathematical processes and results in writing
	9.3.3 communicate mathematical processes and results both orally and in writing
	9.4.3 communicate mathematical processes and results in writing
	10.1.3 communicate mathematical processes and results both orally and in writing
	10.2.2 communicate mathematical processes and results in writing
	10.3.4 communicate mathematical processes and results

UPPER SECONDARY SCHOOL CERTIFICATE EXAMINATIONS

EXAMINATION SPECIFICATION

Name of Subject: **General Mathematics**

The Upper Secondary School Certificate Examinations in General Mathematics will have two papers, Paper1 and Paper 2. Paper 1 will have the following parts:

Part A. Multiple Choice (30 Marks)

There shall be 30 Multiple Choice Questions worth 1 mark each.
Questions 1-15 Low Order Questions
Questions 16-25 Middle Order Questions
Questions 26-30 High Order Questions

Part B Short Answers (20 marks)

There shall be 20 Short Answer Questions worth 1 mark each.

Total Marks for Paper 1 = 50

The Examination Time for General Mathematics- Paper 1 is 2 hours and 30 minutes

Subject	Core Units	Content	Duration (WEEKS)	Number of Questions PART A (30)	LOW	MED	HIGH	PART B NO.OF QUES	Total Number of Questions (50)
Grade 11: General Mathematics	**11.1 Number and Application**	Basic numeracy Units of measurement Ratio and proportion Basic algebra	10	**4**	2	1	1	**3**	**7**
	11.2 Managing Money 1	Earnings and spending Budgeting and loans	8	**3**	2	1	-	**3**	**6**
	11.3 Statistics	Exploring data Analysis of data Probability	6	**3**	1	1	1	**1**	**4**
	11.4 Geometry	Lines, angles, triangles, and regular polygons Geometric construction Circles	8	**3**	2	1	-	**3**	**6**
	11.5 Trigonometry	Trigonometry Vectors	8	**3**	1	1	1	**3**	**6**
Grade 12: General Mathematics	**12.1 Measurements**	Scales and dimensions Survey	6	**3**	2	1	-	**1**	**4**
	12.2 Managing Money 2	Interest and inflation Consumer credit Investments Insurance	**8**	**3**	1	1	1	**3**	**6**

		Using simple algebraic manipulation of financial formulae							
	12.3 Probability and Statistics	Probability Correlation and regression	6	**3**	2	1	-	**1**	**4**
	12.4 Algebra and Graphics	Equations	6	**3**	1	1	1	**1**	**4**
	12.5 Applying Geometry in PNG Arts	Tessellations and polyhedra	4	**2**	1	1		**1**	**3**
		TOTAL	**70**	**30**	**15**	**10**	**5**	**20**	**50**

General Mathematics Examination- Paper 2
This examination will consist of 10 questions that will require working out. The working out must be shown in order to obtain marks. Each question is worth 5 marks.

Total Mark for Paper 2 = 50.

The Examination Time for **General Mathematics Paper 2** is also 2 hours and 30 minutes.

Subject	**Core Units**	**Content**	**Duration (WEEKS)**	**Number of Questions (10)**
Grade 11: General Mathematics	**11.1 Number and Application**	Basic numeracy Units of measurement Ratio and proportion Basic algebra	10	**2**
	11.2 Managing Money 1	Earnings and spending Budgeting and loans	8	**1**
	11.3 Statistics	Exploring data Analysis of data Probability	6	**1**
	11.4 Geometry	Lines, angles, triangles, and regular polygons	8	**1**

		Geometric construction		
		Circles		
	11.5 Trigonometry	Trigonometry	8	**1**
		Vectors		
Grade 12: General Mathematics	**12.1 Measurements**	Scales and dimensions	6	**1**
		Survey		
	12.2 Managing Money 2	Interest and inflation	**8**	**1**
		Consumer credit		
		Investments		
		Insurance		
		Using simple algebraic manipulation of financial formulae		
	12.3 Probability and Statistics	Probability	6	**1**
		Correlation and regression		
	12.4 Algebra and Graphics	Equations	6	**1**
	12.5 Applying Geometry in PNG Arts	Tessellations and polyhedra	4	**-**
		TOTAL	**70**	**10**

Examiners are required to conform to this specification and set questions based on the unit of work in the New Upper Secondary General Mathematics syllabus.

UPPER SECONDARY SCHOOL CERTIFICATE EXAMINATIONS

EXAMINATION SPECIFICATION

Name of Subject: **Advanced Mathematics**

The Upper Secondary School Certificate Examinations in Advanced Mathematics will have two papers, Paper1 and Paper 2. Paper 1 will have the following parts:

Part A. Multiple Choice (30 Marks)

There shall be 30 Multiple Choice Questions worth 1 mark each.
Questions 1-15 Low Order Questions

Subject	Core Units	Content	Duration (WEEKS)	Number of Questions PART A (30)	LOW	MED	HIGH	PART B NO.OF QUES	Total Number of Questions (50)
Grade 11 Advance Maths	**11.1 Number and Application**	Basic Numeracy Units of Measurement Ratio and proportion Basic Algebra	10	**4**	2	1	1	**3**	**7**
	11.2 Graphs and Functions	Algebraic Expressions Graphs and Functions	12	**6**	3	1	1	**3**	**9**
	11.3 Managing Data	Statistics Permutations and combination Probability	10	**5**	2	2	0	**2**	**7**
	11.4 Geometry	Congruency, similarity and construction Circles	8	**3**	1	1	1	**3**	**6**
Grade 12 Advanced	**12.1 Patterns and Algebra**	Sets Sequence and	10	**4**	2	2	0	**3**	**7**

Maths		Series							
		Binomial Theorem							
		Determinants							
	12.2 Trigonometry and Vectors	Trigonometry	10	**4**	2	2	1	**3**	**7**
		Vectors							
	12.3 Calculus	Differentiation	10	**4**	3	1	1	**3**	**7**
		Integration							
			70	**30**	15	10	5	**20**	**50**

Questions 16-25 Middle Order Questions
Questions 26-30 High Order Questions

Part B Short Answers (20 marks)

There shall be 20 Short Answer Questions worth 1 mark each.

Total Marks for Paper 1 = 50

The Examination Time for Advanced Mathematics- Paper 1 is 2 hours and 30 minutes

Advanced Mathematics Examination- Paper 2

This examination will consist of 10 questions that will require working out. The working out must be shown in order to obtain marks. Each question is worth 5 marks.

Total Mark for Paper 2 = 50.

The Examination Time for **Advanced Mathematics Paper 2** is also 2 hours and 30 minutes.

Examiners are required to conform to this specification and set questions based on the unit of work in the New Upper Secondary Advanced Mathematics syllabus

Subject	Core Units	Content	Duration (WEEKS)	Number of Questions
Grade 11 Advance Maths	**11.1 Number and Application**	Basic Numeracy	10	**2**
		Units of Measurement		
		Ratio and proportion		
		Basic Algebra		
	11.2 Graphs and Functions	Algebraic Expressions	12	**2**
		Graphs and Functions		
	11.3 Managing Data	Statistics	10	**1**
		Permutations and combination		
		Probability		
	11.4 Geometry	Congruency, similarity and construction	8	**1**
		Circles		
Grade 12 Advanced Maths	**12.1 Patterns and Algebra**	Sets	10	**2**
		Sequence and Series		
		Binomial Theorem		
		Determinants		
	12.2 Trigonometry and Vectors	Trigonometry	10	**1**
		Vectors		
	12.3 Calculus	Differentiation	10	**1**
		Integration		
		TOTAL	**70**	**10**