

PARCC

Final Technical Report for 2015 Administration

Educational Testing Service
Pearson
Measured Progress

April 18, 2016



Table of Contents

Executive Summary.....	1
Section 1: Introduction	7
1.1 Background	7
1.2 Purpose of the Operational Tests	8
1.3 Composition of Operational Tests	8
1.4 Intended Population	8
1.5 Groups and Organizations Involved with PARCC	9
1.6 Overview of the Technical Report.....	10
1.7 Glossary of Abbreviations	13
Section 2: Test Development.....	15
2.1 Overview of the PARCC Assessment, Claims, and Design.....	15
2.1.1 English Language Arts/Literacy (ELA/L) Assessments – Claims and Subclaims.....	15
2.1.2 Mathematics Assessments – Claims and Subclaims	16
2.2 Test Development Activities	17
2.2.1 Item Development Process.....	17
2.2.2 Item and Text Review Committees.....	18
2.2.3 Operational Test Construction.....	20
2.2.4 Linking Design of the Operational Test.....	24
2.2.5 Graphical Representation of PARCC Operational Test Linking Design	25
2.2.6 Field Test Data Collection Overview	27
Section 3: Test Administration.....	28
3.1 Testing Windows.....	28
3.2 Test Security and Administration Policies.....	29
3.3 Accessibility Features and Accommodations.....	31
3.3.1 Participation Guidelines for PARCC Assessments	31
3.3.2 PARCC Accessibility System.....	31
3.3.3 What are Accessibility Features?	32
3.3.4 Accommodations for Students with Disabilities and English Learners.....	32
3.3.5 Unique Accommodations.....	34
3.3.6 Emergency Accommodations	34

3.3.7 Student Refusal Form..... 34

3.4 Testing Irregularities and Security Breaches..... 34

3.5 Data Forensics Analyses..... 36

 3.5.1 Response Change Analysis..... 37

 3.5.2 Plagiarism Analysis..... 37

 3.5.3 Internet and Social Media Monitoring..... 37

 3.5.4 Off-Hours Testing Monitoring..... 38

3.6 Quality of Test Administration Studies..... 38

Section 4: Item Scoring..... 40

 4.1 Machine Scored Items..... 40

 4.1.1 Key Based Items..... 40

 4.1.2 Rule Based Items..... 40

 4.2 Human or Handscored Items..... 41

 4.2.1 Scorer Training..... 42

 4.2.2 Scorer Qualification..... 45

 4.2.3 Managing Scoring..... 46

 4.2.4 Monitoring Scoring..... 46

Section 5: Test Taker Characteristics..... 50

 5.1 Overview of Test Taking Population..... 50

 5.2 Composition of Operational Forms..... 50

 5.3 Rules for Inclusion of Students in Analyses..... 51

 5.4 Test Takers by Grade, Mode, and Gender..... 51

 5.5 Demographics..... 53

Section 6: Classical Item Analysis..... 54

 6.1 Overview..... 54

 6.2 Data Screening Criteria..... 54

 6.3 Description of Classical Item Analysis Statistics..... 55

 6.4 Summary of Classical Item Analysis Flagging Criteria..... 57

 6.5 Classical Item Analysis Results..... 57

Section 7: Differential Item Functioning..... 69

 7.1 Overview..... 69

 7.2 DIF Procedures..... 69

7.3 Operational Analysis DIF Comparison Groups	71
7.4 Operational Differential Item Functioning Results	73
Section 8: Reliability.....	74
8.1 Overview	74
8.2 Reliability and SEM Estimation	75
8.3 Reliability Results for Total Group	76
8.4 Reliability Results for Subgroups of Interest.....	78
8.5 Reliability Results for English Language Arts/Literacy Subscores	102
8.6 Reliability Results for Mathematics Subscores	105
8.7 Reliability of Classification	107
8.8 Inter-rater Agreement	112
Section 9: Validity	113
9.1 Overview	113
9.2 Evidence Based on Test Content.....	113
9.3 Evidence Based on Internal Structure.....	115
9.3.1 Intercorrelations	115
9.3.2 Reliability.....	128
9.3.3 Local Item Dependence	128
9.4 Evidence Based on Relationships to Other Variables	133
9.5 Evidence from the Special Studies	139
9.5.1 Benchmarking Study	139
9.5.2 Content Alignment Studies	140
9.5.3 Mode Comparability Study	141
9.5.4 Device Comparability Study	142
9.6 Evidence Based on Response Processes	142
9.7 Interpretations of Test Scores.....	143
9.8 Evidence Based on the Consequences to Testing.....	143
9.9 Summary	144
Section 10: IRT Calibration and Scaling in Operational Year One	146
10.1 Overview	146
10.2 IRT Data Preparation.....	146
10.2.1 Overview	146

10.2.2 Student Inclusion/Exclusion Rules	147
10.2.3 Items Excluded from IRT Sparse Matrices	148
10.2.4 Omitted, Not Reached and Not Presented Items	148
10.2.5 Quality Control of the IRT Sparse Matrix Data Files	148
10.3 Description of the Calibration Process.....	149
10.3.1 Special Studies to Inform the Operational Calibration and Scaling	149
10.3.2 IRT Item Exclusion Rules (Before Calibration).....	153
10.3.3 PARSCALE Calibration Procedures and Convergence Criteria	153
10.3.4 Calibration Quality Control	154
10.4 Model Fit Evaluation Criteria	155
10.5 Items Excluded from Score Reporting	157
10.5.1 Item Review Process	157
10.5.2 Count and Percentage of Items Excluded from Score Reporting	159
10.6 Scaling Parameter Estimates from Paper to Online.....	162
10.6.1 Scaling Procedures (PBT to CBT)	162
10.6.2 Comparability across Spanish and English Versions	163
10.6.3 Scaling Quality Control.....	164
10.7 Items Excluded from Spring 2015 Paper to Online Linking Sets	164
10.8 Correlations and Plots of Scaling Item Parameter Estimates	165
10.9 Scaling Constants	173
10.10 Summary Statistics and Distributions from IRT Analyses	174
10.10.1 IRT Summary Statistics for English Language Arts/Literacy	174
10.10.2 IRT Summary Statistics for Mathematics	177
10.11 Mean Linking Item AIS Effect Size Compared to Overall Scale Score Effect Size.....	181
Section 11: Performance Level Setting	185
11.1 Performance Standards	185
11.2 Performance Levels and Policy Definitions.....	185
11.3 Performance Level Setting Process for the PARCC Assessment System.....	187
11.3.1 PARCC Research Studies	188
11.3.2 PARCC Pre-Policy Meeting	188
11.3.3 Performance Level Setting Meetings	188
11.3.4 PARCC Post-Policy Reasonableness Review	190

Section 12: Scale Scores	191
12.1 Operational Test Content (Claims and Subclaims)	191
12.1.1 English Language Arts/Literacy	191
12.1.2 Mathematics	193
12.2 Establishing the Reporting Scales	193
12.2.1 Full Summative Score Scale and Performance Levels	194
12.2.2 ELA/L Reading and Writing Claim Scale	195
12.2.3 Subclaims Scale	196
12.3 Creating Conversion Tables	196
12.4 Score Distributions	202
12.4.1 Score Distributions for ELA/L	202
12.4.2 Score Distributions for Mathematics	211
12.5 Interpreting Claim Scores and Subclaim Scores	216
12.5.1 Interpreting Claim Scores	216
12.5.2 Interpreting Subclaim Scores	216
Section 13: Quality Control Procedures	218
13.1 Quality Control of the Item Bank	218
13.2 Quality Control of Test Form Development	219
13.3 Quality Control of Test Materials	220
13.4 Quality Control of Scanning	221
13.5 Quality Control of Image Editing	221
13.6 Quality Control of Answer Document Processing and Scoring	222
13.7 Quality Control of Psychometric Processes	223
13.7.1 ETS Psychometric Quality Control Process	223
13.7.2 Pearson Psychometric Quality Control Process	225
13.7.3 Measured Progress Psychometric Quality Control Process	226
References	229
Appendices	232
Appendix 5: Test Takers by Grade and Mode, for Each State	232
Appendix 7: Summary of Differential Item Function (DIF) Results	270
Appendix 8: Reliability of Classification by Content and Grade Level	292
Appendix 10.1: IRT Results for Spring 2015 English Language Arts/Literacy (ELA/L)	313

Appendix 10.2: IRT Results for Spring 2015 Mathematics.....	325
Appendix 12.1: Form Composition	349
Appendix 12.2: Scaling Constants and Associated Information	355
Appendix 12.3: Raw-to-Scale Conversion Tables for Performance Level Setting (PLS) Forms	360
Appendix 12.4: IRT Test Characteristic Curves, Information Curves and CSEM Curves	423
Appendix 12.5: Subgroup Scale Score Performance.....	486
Addendum: Statistical Summary of the Fall 2014 Administration.....	516
Addendum 5: Test Taker Characteristics	516
Addendum 6: Classical Item Analysis.....	523
Addendum 7: Differential Item Functioning.....	525
Addendum 8: Reliability.....	529
Addendum 9: Validity	538
Addendum 10: IRT Calibration and Scaling for the Fall 2014 Tests.....	542
Addendum 12: Scale Scores.....	553

List of Tables

Table 1. Defined Summative Scale Scores and Cut Scores	5
Table 2. Defined Scaled Scores and Cut Scores for Reading and Writing Claim Scores	6
Table 1.1 Glossary of PARCC Abbreviations and Acronyms.....	13
Table 2.1 Number of Core Operational Forms per Grade/Subject by Component and Mode for ELA/L and Mathematics	21
Table 3.1 PARCC Fall 2014 and Spring 2015 Testing Windows.....	28
Table 4.1 Training Materials Used During Scoring.....	44
Table 4.2 Mathematics Qualification Requirements	46
Table 4.3 Scoring Validity Agreement Requirements	48
Table 4.4 Inter-rater Agreement Expectations and Results.....	49
Table 5.1 ELA/L Test Takers by Grade and Mode: All States Combined	52
Table 5.2 Mathematics Test Takers by Grade and Mode: All States Combined.....	52
Table 5.3 Spanish-Language Mathematics Test Takers, by Grade and Mode: All States Combined.....	53
Table 6.1 Summary of p Values for ELA/L Operational Items by Grade, Component, and Mode.....	59
Table 6.2 Summary of p Values for Mathematics Operational Items by Grade, Component, and Mode	61
Table 6.3 Summary of Item-total Polyserial Correlations for ELA/L Operational Items by Grade, Component, and Mode	64
Table 6.4 Summary of Item-total Correlations for Mathematics Operational Items by Grade, Component, and Mode	66
Table 7.1 DIF Categories for Dichotomous Selected Response and Constructed Response Items	71
Table 7.2 DIF Categories for Polytomous Constructed Response Items.....	71
Table 7.3 Traditional DIF Comparison Groups	72
Table 8.1 Summary of ELA/L Test Reliability Estimates for Total Group	77
Table 8.2 Summary of Mathematics Test Reliability Estimates for Total Group	78
Table 8.3 Summary of Test Reliability Estimates for Subgroups: Grade 3 ELA/L.....	81
Table 8.4 Summary of Test Reliability Estimates for Subgroups: Grade 4 ELA/L.....	82
Table 8.5 Summary of Test Reliability Estimates for Subgroups: Grade 5 ELA/L.....	83
Table 8.6 Summary of Test Reliability Estimates for Subgroups: Grade 6 ELA/L.....	84
Table 8.7 Summary of Test Reliability Estimates for Subgroups: Grade 7 ELA/L.....	85
Table 8.8 Summary of Test Reliability Estimates for Subgroups: Grade 8 ELA/L.....	86
Table 8.9 Summary of Test Reliability Estimates for Subgroups: Grade 9 ELA/L.....	87
Table 8.10 Summary of Test Reliability Estimates for Subgroups: Grade 10 ELA/L.....	88
Table 8.11 Summary of Test Reliability Estimates for Subgroups: Grade 11 ELA/L.....	89
Table 8.12 Summary of Test Reliability Estimates for Subgroups: Grade 3 Mathematics	90
Table 8.13 Summary of Test Reliability Estimates for Subgroups: Grade 4 Mathematics	91
Table 8.14 Summary of Test Reliability Estimates for Subgroups: Grade 5 Mathematics	92
Table 8.15 Summary of Test Reliability Estimates for Subgroups: Grade 6 Mathematics	93
Table 8.16 Summary of Test Reliability Estimates for Subgroups: Grade 7 Mathematics	94
Table 8.17 Summary of Test Reliability Estimates for Subgroups: Grade 8 Mathematics	95
Table 8.18 Summary of Test Reliability Estimates for Subgroups: Algebra I	96

Table 8.19 Summary of Test Reliability Estimates for Subgroups: Geometry	97
Table 8.20 Summary of Test Reliability Estimates for Subgroups: Algebra II	98
Table 8.21 Summary of Test Reliability Estimates for Subgroups: Integrated Mathematics I	99
Table 8.22 Summary of Test Reliability Estimates for Subgroups: Integrated Mathematics II	100
Table 8.23 Summary of Test Reliability Estimates for Subgroups: Integrated Mathematics III	101
Table 8.24 Descriptions of ELA/L Claims and Subclaims.....	102
Table 8.25 Average ELA/L Reliability Estimates for Total Test and Subscores	104
Table 8.26 Average Mathematics Reliability Estimates for Total Test and Subscores	106
Table 8.27 Reliability of Classification: Summary for ELA/L.....	108
Table 8.28 Reliability of Classification: Grade 3 ELA/L.....	109
Table 8.29 Reliability of Classification: Summary for Mathematics	111
Table 8.30 Inter-rater Agreement Expectations and Results.....	112
Table 9.1 Average Intercorrelations and Reliability between Grade 3 ELA/L Subclaims.....	117
Table 9.2 Average Intercorrelations and Reliability between Grade 4 ELA/L Subclaims.....	117
Table 9.3 Average Intercorrelations and Reliability between Grade 5 ELA/L Subclaims.....	118
Table 9.4 Average Intercorrelations and Reliability between Grade 6 ELA/L Subclaims.....	118
Table 9.5 Average Intercorrelations and Reliability between Grade 7 ELA/L Subclaims.....	119
Table 9.6 Average Intercorrelations and Reliability between Grade 8 ELA/L Subclaims.....	119
Table 9.7 Average Intercorrelations and Reliability between Grade 9 ELA/L Subclaims.....	120
Table 9.8 Average Intercorrelations and Reliability between Grade 10 ELA/L Subclaims.....	120
Table 9.9 Average Intercorrelations and Reliability between Grade 11 ELA/L Subclaims.....	121
Table 9.10 Average Intercorrelations and Reliability between Grade 3 Mathematics Subclaims	121
Table 9.11 Average Intercorrelations and Reliability between Grade 4 Mathematics Subclaims	122
Table 9.12 Average Intercorrelations and Reliability between Grade 5 Mathematics Subclaims	122
Table 9.13 Average Intercorrelations and Reliability between Grade 6 Mathematics Subclaims	123
Table 9.14 Average Intercorrelations and Reliability between Grade 7 Mathematics Subclaims	123
Table 9.15 Average Intercorrelations and Reliability between Grade 8 Mathematics Subclaims	124
Table 9.16 Average Intercorrelations and Reliability between Algebra I Subclaims	124
Table 9.17 Average Intercorrelations and Reliability between Geometry Subclaims	125
Table 9.18 Average Intercorrelations and Reliability between Algebra II Subclaims	125
Table 9.19 Average Intercorrelations and Reliability between Integrated Mathematics I Subclaims	126
Table 9.20 Average Intercorrelations and Reliability between Integrated Mathematics II Subclaims	126
Table 9.21 Average Intercorrelations and Reliability between Integrated Mathematics III Subclaims	127
Table 9.22 Conditions used in LID Reliability Investigation and Results.....	131
Table 9.23 Summary of Q_3 Values for ELA/L Grade 4 and Integrated Mathematics II	131
Table 9.24 Average Correlations between ELA/L and Mathematics for Grade 3	134
Table 9.25 Average Correlations between ELA/L and Mathematics for Grade 4	134
Table 9.26 Average Correlations between ELA/L and Mathematics for Grade 5	135
Table 9.27 Average Correlations between ELA/L and Mathematics for Grade 6	135
Table 9.28 Average Correlations between ELA/L and Mathematics for Grade 7	136
Table 9.29 Average Correlations between ELA/L and Mathematics for Grade 8	136
Table 9.30 Average Correlations between ELA/L and Mathematics for High School.....	137

Table 9.31 Average Correlations between ELA/L Reading and Mathematics for High School.....	137
Table 9.32 Average Correlations between ELA/L Writing and Mathematics for High School.....	138
Table 10.1 N-Counts, Percent of Students, and Number of Items in the ELA/L IRT Calibration Files	148
Table 10.2 N-Counts, Percent of Students, and Number of Items in the Mathematics IRT Calibration Files	149
Table 10.3 Number and Percentage of ELA/L Items Excluded from IRT Calibration	160
Table 10.4 Number and Percentage of Mathematics Items Excluded from IRT Calibration	161
Table 10.5 WRMSD Flagging Criteria for Inspection and Possible Removal of Linking Items	163
Table 10.6 Number of ELA/L Items Excluded from the CBT/PBT Linking Sets.....	164
Table 10.7 Number of Mathematics Items Excluded from the CBT/PBT Linking Sets.....	165
Table 10.8 Number of Items, Number of Points and Correlations for ELA/L Linking Items	165
Table 10.9 Number of Items, Number of Points and Correlations for Mathematics Linking Items	166
Table 10.10 Scaling Constants Spring 2015 PBT to CBT for ELA/L	173
Table 10.11 Scaling Constants Spring 2015 PBT to CBT for Mathematics	174
Table 10.12 CBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade.....	175
Table 10.13 PBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade.....	175
Table 10.14 CBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade	176
Table 10.15 PBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade.....	176
Table 10.16 CBT IRT Model Fit for All Items for ELA/L by Grade	177
Table 10.17 PBT IRT Model Fit for All Items for ELA/L by Grade	177
Table 10.18 CBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject..	178
Table 10.19 PBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject..	179
Table 10.20 CBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject	179
Table 10.21 PBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject	180
Table 10.22 CBT IRT Model Fit for All Items for Mathematics by Grade/Subject.....	180
Table 10.23 PBT IRT Model Fit for All Items for Mathematics by Grade/Subject.....	181
Table 10.24 Average Item Score Effect Size Compared to Overall Scale Score Effect Size for ELA/L.....	183
Table 10.25 Average Item Score Effect Size Compared to Overall Scale Score Effect Size for Mathematics	184
Table 11.1 PARCC PLS Committee Meetings and Dates	190
Table 12.1 Form Composition for ELA/L Grade 3	192
Table 12.2 Contribution of Prose Constructed Response Items to ELA/L	193
Table 12.3 Mathematics Form Composition for Grade 3	193
Table 12.4 Defined Summative Scale Scores	194
Table 12.5 Defined Scaled Scores for Reading and Writing Claim Scores	195
Table 12.6 Calculating Scaling Constants for Reading and Writing Claim Scores.....	195
Table 12.7 Subgroup Performance for ELA/L: Grade 3.....	208
Table 12.8 Subgroup Performance for ELA/L: Grade 9.....	210
Table 12.9 Subgroup Performance for Mathematics Scale Scores: Grade 3	214
Table 12.10 Subgroup Performance for Mathematics Scale Scores: Algebra I	215

Table 12.11 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics I.....	216
Table A.5.1 ELA/L Test Takers, by State, Grade, and Mode.....	232
Table A.5.2 Mathematics Test Takers, by State, Grade, and Mode.....	236
Table A.5.3 Spanish-Language Mathematics Test Takers, by State, Grade, and Mode.....	240
Table A.5.4 All States Combined: ELA/L Test Takers by Grade, Mode, and Gender.....	244
Table A.5.5 All States Combined: All Mathematics Test Takers by Grade, Mode, and Gender	245
Table A.5.6 All States Combined: Spanish-Language Mathematics Test Takers, by Grade, Mode, and Gender	247
Table A.5.7 Demographic Information for Grade 3 ELA/L, Overall and by State.....	249
Table A.5.8 Demographic Information for Grade 4 ELA/L, Overall and by State.....	250
Table A.5.9 Demographic Information for Grade 5 ELA/L, Overall and by State.....	251
Table A.5.10 Demographic Information for Grade 6 ELA/L, Overall and by State.....	252
Table A.5.11 Demographic Information for Grade 7 ELA/L, Overall and by State.....	253
Table A.5.12 Demographic Information for Grade 8 ELA/L, Overall and by State.....	254
Table A.5.13 Demographic Information for Grade 9 ELA/L, Overall and by State.....	255
Table A.5.14 Demographic Information for Grade 10 ELA/L, Overall and by State.....	256
Table A.5.15 Demographic Information for Grade 11 ELA/L, Overall and by State.....	257
Table A.5.16 Demographic Information for Grade 3 Mathematics, Overall and by State	258
Table A.5.17 Demographic Information for Grade 4 Mathematics, Overall and by State	259
Table A.5.18 Demographic Information for Grade 5 Mathematics, Overall and by State	260
Table A.5.19 Demographic Information for Grade 6 Mathematics, Overall and by State	261
Table A.5.20 Demographic Information for Grade 7 Mathematics, Overall and by State	262
Table A.5.21 Demographic Information for Grade 8 Mathematics, Overall and by State	263
Table A.5.22 Demographic Information for Algebra I, Overall and by State	264
Table A.5.23 Demographic Information for Geometry, Overall and by State	265
Table A.5.24 Demographic Information for Algebra II, Overall and by State	266
Table A.5.25 Demographic Information for Integrated Mathematics I, Overall and by State	267
Table A.5.26 Demographic Information for Integrated Mathematics II, Overall and by State	268
Table A.5.27 Demographic Information for Integrated Mathematics III, Overall and by State	269
Table A.7.1 Differential Item Functioning for ELA/L Grade 3	270
Table A.7.2 Differential Item Functioning for ELA/L Grade 4	271
Table A.7.3 Differential Item Functioning for ELA/L Grade 5	272
Table A.7.4 Differential Item Functioning for ELA/L Grade 6	273
Table A.7.5 Differential Item Functioning for ELA/L Grade 7	274
Table A.7.6 Differential Item Functioning for ELA/L Grade 8	275
Table A.7.7 Differential Item Functioning for ELA/L Grade 9	276
Table A.7.8 Differential Item Functioning for ELA/L Grade 10	277
Table A.7.9 Differential Item Functioning for ELA/L Grade 11	278
Table A.7.10 Differential Item Functioning for Mathematics Grade 3	279
Table A.7.11 Differential Item Functioning for Mathematics Grade 4	280
Table A.7.12 Differential Item Functioning for Mathematics Grade 5	281
Table A.7.13 Differential Item Functioning for Mathematics Grade 6	282

Table A.7.14 Differential Item Functioning for Mathematics Grade 7	283
Table A.7.15 Differential Item Functioning for Mathematics Grade 8	284
Table A.7.16 Differential Item Functioning for Algebra I	285
Table A.7.17 Differential Item Functioning for Geometry	286
Table A.7.18 Differential Item Functioning for Algebra II	287
Table A.7.19 Differential Item Functioning for Integrated Mathematics I	288
Table A.7.20 Differential Item Functioning for Integrated Mathematics II	289
Table A.7.21 Differential Item Functioning for Integrated Mathematics III	290
Table A.7.22 Differential Item Functioning for Mathematics: Spanish-Language vs. English-Language Forms	291
Table A.8.1 Reliability of Classification: Grade 3 ELA/L	292
Table A.8.2 Reliability of Classification: Grade 4 ELA/L	293
Table A.8.3 Reliability of Classification: Grade 5 ELA/L	294
Table A.8.4 Reliability of Classification: Grade 6 ELA/L	295
Table A.8.5 Reliability of Classification: Grade 7 ELA/L	296
Table A.8.6 Reliability of Classification: Grade 8 ELA/L	297
Table A.8.7 Reliability of Classification: Grade 9 ELA/L	298
Table A.8.8 Reliability of Classification: Grade 10 ELA/L	299
Table A.8.9 Reliability of Classification: Grade 11 ELA/L	300
Table A.8.10 Reliability of Classification: Grade 3 Mathematics	301
Table A.8.11 Reliability of Classification: Grade 4 Mathematics	302
Table A.8.12 Reliability of Classification: Grade 5 Mathematics	303
Table A.8.13 Reliability of Classification: Grade 6 Mathematics	304
Table A.8.14 Reliability of Classification: Grade 7 Mathematics	305
Table A.8.15 Reliability of Classification: Grade 8 Mathematics	306
Table A.8.16 Reliability of Classification: Algebra I	307
Table A.8.17 Reliability of Classification: Geometry	308
Table A.8.18 Reliability of Classification: Algebra II	309
Table A.8.19 Reliability of Classification: Integrated Mathematics I	310
Table A.8.20 Reliability of Classification: Integrated Mathematics II	311
Table A.8.21 Reliability of Classification: Integrated Mathematics III	312
Table A.10.1 CBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade	313
Table A.10.2 PBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade	315
Table A.10.3 CBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade	317
Table A.10.4 PBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade	319
Table A.10.5 CBT IRT Model Fit for All Items for ELA/L by Grade	321
Table A.10.6 PBT IRT Model Fit for All Items for ELA/L by Grade	323
Table A.10.7 CBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject.	325
Table A.10.8 PBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject.	329
Table A.10.9 CBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject	333

Table A.10.10 PBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject	337
Table A.10.11 CBT IRT Model Fit for All Items for Mathematics by Grade/Subject	341
Table A.10.12 PBT IRT Model Fit for All Items for Mathematics by Grade/Subject	345
Table A.12.1 Form Composition for ELA/L Grade 3	349
Table A.12.2 Form Composition for ELA/L Grade 4	349
Table A.12.3 Form Composition for ELA/L Grade 5	349
Table A.12.4 Form Composition for ELA/L Grade 6	350
Table A.12.5 Form Composition for ELA/L Grade 7	350
Table A.12.6 Form Composition for ELA/L Grade 8	350
Table A.12.7 Form Composition for ELA/L Grade 9	351
Table A.12.8 Form Composition for ELA/L Grade 10	351
Table A.12.9 Form Composition for ELA/L Grade 11	351
Table A.12.10 Form Composition for Mathematics Grade 3	352
Table A.12.11 Form Composition for Mathematics Grade 4	352
Table A.12.12 Form Composition for Mathematics Grade 5	352
Table A.12.13 Form Composition for Mathematics Grade 6	352
Table A.12.14 Form Composition for Mathematics Grade 7	353
Table A.12.15 Form Composition for Mathematics Grade 8	353
Table A.12.16 Form Composition for Algebra I	353
Table A.12.17 Form Composition for Geometry	353
Table A.12.18 Form Composition for Algebra II	354
Table A.12.19 Form Composition for Integrated Mathematics I	354
Table A.12.20 Form Composition for Integrated Mathematics II	354
Table A.12.21 Form Composition for Integrated Mathematics III	354
Table A.12.22 Threshold Scores and Scaling Constants for ELA/L Grades 3 to 8	355
Table A.12.23 Threshold Scores and Scaling Constants for Mathematics Grades 3 to 8	356
Table A.12.24 Threshold Scores and Scaling Constants for High School ELA	357
Table A.12.25 Threshold Scores and Scaling Constants for High School Mathematics	358
Table A.12.26 Scaling Constants for Reading and Writing Grades 3 to 8	359
Table A.12.27 Conversion Table for Performance Level Setting Form: ELA/L Grade 3	360
Table A.12.28 Conversion Table for Performance Level Setting Form: ELA/L Grade 4	363
Table A.12.29 Conversion Table for Performance Level Setting Form: ELA/L Grade 5	366
Table A.12.30 Conversion Table for Performance Level Setting Form: ELA/L Grade 6	369
Table A.12.31 Conversion Table for Performance Level Setting Form: ELA/L Grade 7	373
Table A.12.32 Conversion Table for Performance Level Setting Form: ELA/L Grade 8	377
Table A.12.33 Conversion Table for Performance Level Setting Form: ELA/L Grade 9	381
Table A.12.34 Conversion Table for Performance Level Setting Form: ELA/L Grade 10	385
Table A.12.35 Conversion Table for Performance Level Setting Form: ELA/L Grade 11	389
Table A.12.36 Conversion Table for Performance Level Setting Form: Mathematics Grade 3	393
Table A.12.37 Conversion Table for Performance Level Setting Form: Mathematics Grade 4	395
Table A.12.38 Conversion Table for Performance Level Setting Form: Mathematics Grade 5	397

Table A.12.39 Conversion Table for Performance Level Setting Form: Mathematics Grade 6	399
Table A.12.40 Conversion Table for Performance Level Setting Form: Mathematics Grade 7	401
Table A.12.41 Conversion Table for Performance Level Setting Form: Mathematics Grade 8	403
Table A.12.42 Conversion Table for Performance Level Setting Form: Algebra I	405
Table A.12.43 Conversion Table for Performance Level Setting Form: Algebra II	408
Table A.12.44 Conversion Table for Performance Level Setting Form: Geometry	411
Table A.12.45 Conversion Table for Performance Level Setting Form: Integrated Mathematics I	414
Table A.12.46 Conversion Table for Performance Level Setting Form: Integrated Mathematics II	417
Table A.12.47 Conversion Table for Performance Level Setting Form: Integrated Mathematics III	420
Table A.12.48 Subgroup Performance for ELA/L Scale Scores: Grade 3	486
Table A.12.49 Subgroup Performance for ELA/L Scale Scores: Grade 4	488
Table A.12.50 Subgroup Performance for ELA/L Scale Scores: Grade 5	490
Table A.12.51 Subgroup Performance for ELA/L Scale Scores: Grade 6	492
Table A.12.52 Subgroup Performance for ELA/L Scale Scores: Grade 7	494
Table A.12.53 Subgroup Performance for ELA/L Scale Scores: Grade 8	496
Table A.12.54 Subgroup Performance for ELA/L Scale Scores: Grade 9	498
Table A.12.55 Subgroup Performance for ELA/L Scale Scores: Grade 10	500
Table A.12.56 Subgroup Performance for ELA/L Scale Scores: Grade 11	502
Table A.12.57 Subgroup Performance for Mathematics Scale Scores: Grade 3	504
Table A.12.58 Subgroup Performance for Mathematics Scale Scores: Grade 4	505
Table A.12.59 Subgroup Performance for Mathematics Scale Scores: Grade 5	506
Table A.12.60 Subgroup Performance for Mathematics Scale Scores: Grade 6	507
Table A.12.61 Subgroup Performance for Mathematics Scale Scores: Grade 7	508
Table A.12.62 Subgroup Performance for Mathematics Scale Scores: Grade 8	509
Table A.12.63 Subgroup Performance for Mathematics Scale Scores: Algebra I	510
Table A.12.64 Subgroup Performance for Mathematics Scale Scores: Geometry	511
Table A.12.65 Subgroup Performance for Mathematics Scale Scores: Algebra II	512
Table A.12.66 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics I	513
Table A.12.67 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics II	514
Table A.12.68 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics III	515
Table ADD.5.1 State Participation in ELA/L Fall 2014 Operational Tests, by Grade	516
Table ADD.5.2 State Participation in Mathematics Fall 2014 Operational Tests, by Grade	517
Table ADD.5.3 All States Combined: Fall 2014 ELA/L Test Takers by Grade and Gender	517
Table ADD.5.4 All States Combined: Fall 2014 Mathematics Test Takers by Grade and Gender	517
Table ADD.5.5 Fall 2014 ELA/L Test Takers by State, Grade, and Gender	518
Table ADD.5.6 Fall 2014 Mathematics Test Takers by State, Content Area, and Gender	519
Table ADD.5.7 Demographic Information for Fall 2014 Grade 9 ELA/L, Overall and by State	519
Table ADD.5.8 Demographic Information for Fall 2014 Grade 10 ELA/L, Overall and by State	520
Table ADD.5.9 Demographic Information for Fall 2014 Grade 11 ELA/L, Overall and by State	520
Table ADD.5.10 Demographic Information for Fall 2014 Algebra I, Overall and by State	521
Table ADD.5.11 Demographic Information for Fall 2014 Geometry, Overall and by State	521
Table ADD.5.12 Demographic Information for Fall 2014 Algebra II, Overall and by State	522

Table ADD.6.1 Summary of p Values for Fall 2014 ELA/L and Mathematics Operational Items, by Grade and Component	523
Table ADD.6.2 Summary of Polyserial Correlations for Fall 2014 ELA/L and Mathematics Operational Items, by Grade and Test	524
Table ADD.7.1 Differential Item Functioning for Fall 2014 High School ELA/L.....	525
Table ADD.7.2 Differential Item Functioning for Fall 2014 High School Mathematics.....	527
Table ADD.8.1 Summary of Test Reliability Estimates for Fall 2014 Total Group.....	529
Table ADD.8.2 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Grade 9 ELA/L.....	530
Table ADD.8.3 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Grade 10 ELA/L.....	531
Table ADD.8.4 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Grade 11 ELA/L.....	532
Table ADD.8.5 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Algebra I	533
Table ADD.8.6 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Geometry	534
Table ADD.8.7 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Algebra II	535
Table ADD.8.8 Fall 2014 Average ELA/L Reliability Estimates for Reading Total and Subscores.....	536
Table ADD.8.9 Fall 2014 Average ELA/L Reliability Estimates for Writing Total and Subscores.....	536
Table ADD.8.10 Fall 2014 Average Mathematics Reliability Estimates for Total Test and Subscores.....	537
Table ADD.9.1 Average Intercorrelations and Reliability between Grade 9 ELA/L Reading Subclaims ...	538
Table ADD.9.2 Average Intercorrelations and Reliability between Grade 10 ELA/L Reading Subclaims .	538
Table ADD.9.3 Average Intercorrelations and Reliability between Grade 11 ELA/L Reading Subclaims .	539
Table ADD.9.4 Average Intercorrelations and Reliability between Algebra I Subclaims.....	539
Table ADD.9.5 Average Intercorrelations and Reliability between Geometry Subclaims	539
Table ADD.9.6 Average Intercorrelations and Reliability between Algebra II Subclaims	540
Table ADD.9.7 Average Correlations between ELA/L and Mathematics for High School.....	540
Table ADD.9.8 Average Correlations between Reading and Mathematics for High School.....	540
Table ADD.9.9 Average Correlations between Writing and Mathematics for High School.....	541
Table ADD.10.1 Number of Students and Number and Percent of Fall block Items Excluded from Calibration.....	543
Table ADD.10.2 Number of Fall Block Items Excluded from the Fall Block/Spring Linking Sets.....	544
Table ADD.10.3 Number of Items, Points, and Correlations for Fall Block/Spring Linking Items.....	544
Table ADD.10.4 Scaling Constants 2014 Fall Block to Spring 2015.....	545
Table ADD.10.5 2014 Fall Block IRT Summary Parameter Estimates by Grade/Subject	545
Table ADD.10.6 2014 Fall Block IRT Standard Errors of Parameter Estimates by Grade/Subject	545
Table ADD.10.7 2014 Fall Block IRT Model Fit by Grade/Subject	546
Table ADD.10.8 PBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade	547
Table ADD.10.9 PBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade	548
Table ADD.10.10 PBT IRT Model Fit for All Items for ELA/L by Grade	549
Table ADD.10.11 PBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject	550
Table ADD.10.12 PBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject	551
Table ADD.10.13 PBT IRT Model Fit for All Items for Mathematics by Grade/Subject	552
Table ADD.12.1 Fall 2014 Subgroup Performance for ELA/L Scale Scores: Grade 9	553

Table ADD.12.2 Fall 2014 Subgroup Performance for ELA/L Scale Scores: Grade 10	555
Table ADD.12.3 Fall 2014 Subgroup Performance for ELA/L Scale Scores: Grade 11	557
Table ADD.12.4 Fall 2014 Subgroup Performance for Mathematics Scale Scores: Algebra I.....	559
Table ADD.12.5 Fall 2014 Subgroup Performance for Mathematics Scale Scores: Geometry.....	560
Table ADD.12.6 Fall 2014 Subgroup Performance for Mathematics Scale Scores: Algebra II.....	561

List of Figures

Figure 2.1 PARCC ELA/L CBT Linking Design (Grades 6-8, and 11).....	26
Figure 2.2 PARCC Mathematics CBT Linking Design (Grades 3-8, and Algebra II)	26
Figure 3.1 Configuration of Units within the 2014-2015 PARCC Assessments (Sample for Grade 7 and High School)	29
Figure 3.2 Claims from the Administration Phase of the PARCC Theory of Action.	39
Figure 9.1 Comparison of Score Reliability by Item Reliability and Cluster (Testlet) Reliability Approaches	130
Figure 9.2 Distribution of Q_3 Values for Grade 4 ELA/L	132
Figure 9.3 Distribution of Q_3 Values for Integrated Mathematics II	132
Figure 10.1 Grade 3 ELA/L 3-Category Item, 2 PL/GPC Model, N-count 124,090, chi-square 305.0, adjusted fit = 0.050.	156
Figure 10.2 Algebra I Dichotomous (2-Category) Item, 2 PL Model, N-count 42,630, chi-square 5167.6, adjusted fit = 0.329.	157
Figure 10.3 Algebra I Dichotomous Item, after the 3PL model was applied using the Weighted Least Squares approach, N-count 42,630.	159
Figure 10.4 ELA/L Grade 3 Transformed New a - vs. Reference a -Parameter Estimates.....	167
Figure 10.5 ELA/L Grade 3 Transformed New b - vs. Reference b -Parameter Estimates.....	167
Figure 10.6 ELA/L Grade 3 Transformed New and Reference TCC for Linking Items	168
Figure 10.7 ELA/L Grade 11 Transformed New a - vs. Reference a -Parameter Estimates.....	168
Figure 10.8 ELA/L Grade 11 Transformed New b - vs. Reference b -Parameter Estimates.....	169
Figure 10.9 ELA/L Grade 11 Transformed New and Reference TCC for Linking Items	169
Figure 10.10 Mathematics Grade 8 Transformed New a - vs. Reference a -Parameter Estimates.....	170
Figure 10.11 Mathematics Grade 8 Transformed New b - vs. Reference b -Parameter Estimates.....	170
Figure 10.12 Mathematics Grade 8 Transformed New and Reference TCC for Linking Items	171
Figure 10.13 Integrated Mathematics III Transformed New a - vs. Reference a -Parameter Estimates ...	171
Figure 10.14 Integrated Mathematics III Transformed New b - vs. Reference b -Parameter Estimates ...	172
Figure 10.15 Integrated Mathematics III Transformed New and Reference TCC for Linking Items	172
Figure 12.1 Test Characteristic Curves for ELA/L Grade 3	199
Figure 12.2 Information Curves for ELA/L Grade 3	200
Figure 12.3 Conditional Standard Error of Measurement Curve for ELA/L Grade 3.....	201
Figure 12.4 Distributions of ELA/L Scale Scores: Grades 3 to 11	202
Figure 12.4 (continued) Distributions of ELA/L Scale Scores: Grades 3 to 11	203
Figure 12.5 Distributions of Reading Scale Scores: Grades 3 to 11	204
Figure 12.5 (continued) Distributions of Reading Scale Scores: Grades 3 to 11.....	205
Figure 12.6 Distributions of Writing Scale Scores: Grades 3 to 11	206
Figure 12.6 (continued) Distributions of Writing Scale Scores: Grades 3 to 11.....	207
Figure 12.7 Distributions of Mathematics Scale Scores: Grades 3 to 8.....	212
Figure 12.8 Distributions of Mathematics Scale Scores: High School.....	213
Figure A.12.1 Test Characteristic Curves ELA/L Grade 3.....	423
Figure A.12.2 Test Characteristic Curves ELA/L Grade 4.....	424

Figure A.12.3 Test Characteristic Curves ELA/L Grade 5.....	425
Figure A.12.4 Test Characteristic Curves ELA/L Grade 6.....	426
Figure A.12.5 Test Characteristic Curves ELA/L Grade 7.....	427
Figure A.12.6 Test Characteristic Curves ELA/L Grade 8.....	428
Figure A.12.7 Test Characteristic Curves ELA/L Grade 9.....	429
Figure A.12.8 Test Characteristic Curves ELA/L Grade 10.....	430
Figure A.12.9 Test Characteristic Curves ELA/L Grade 11.....	431
Figure A.12.10 Test Characteristic Curves Mathematics Grade 3	432
Figure A.12.11 Test Characteristic Curves Mathematics Grade 4	433
Figure A.12.12 Test Characteristic Curves Mathematics Grade 5	434
Figure A.12.13 Test Characteristic Curves Mathematics Grade 6	435
Figure A.12.14 Test Characteristic Curves Mathematics Grade 7	436
Figure A.12.15 Test Characteristic Curves Mathematics Grade 8	437
Figure A.12.16 Test Characteristic Curves Algebra I	438
Figure A.12.17 Test Characteristic Curves Geometry	439
Figure A.12.18 Test Characteristic Curves Algebra II	440
Figure A.12.19 Test Characteristic Curves Integrated Mathematics I	441
Figure A.12.20 Test Characteristic Curves Integrated Mathematics II	442
Figure A.12.21 Test Characteristic Curves Integrated Mathematics III	443
Figure A.12.22 Test Information Curves ELA/L Grade 3	444
Figure A.12.23 Test Information Curves ELA/L Grade 4	445
Figure A.12.24 Test Information Curves ELA/L Grade 5	446
Figure A.12.25 Test Information Curves ELA/L Grade 6	447
Figure A.12.26 Test Information Curves ELA/L Grade 7	448
Figure A.12.27 Test Information Curves ELA/L Grade 8	449
Figure A.12.28 Test Information Curves ELA/L Grade 9	450
Figure A.12.29 Test Information Curves ELA/L Grade 10	451
Figure A.12.30 Test Information Curves ELA/L Grade 11	452
Figure A.12.31 Test Information Curves Mathematics Grade 3	453
Figure A.12.32 Test Information Curves Mathematics Grade 4	454
Figure A.12.33 Test Information Curves Mathematics Grade 5	455
Figure A.12.34 Test Information Curves Mathematics Grade 6	456
Figure A.12.35 Test Information Curves Mathematics Grade 7	457
Figure A.12.36 Test Information Curves Mathematics Grade 8	458
Figure A.12.37 Test Information Curves Algebra I	459
Figure A.12.38 Test Information Curves Geometry	460
Figure A.12.39 Test Information Curves Algebra II	461
Figure A.12.40 Test Information Curves Integrated Mathematics I	462
Figure A.12.41 Test Information Curves Integrated Mathematics II	463
Figure A.12.42 Test Information Curves Integrated Mathematics III	464
Figure A.12.43 CSEM Curves ELA/L Grade 3	465
Figure A.12.44 CSEM Curves ELA/L Grade 4	466

Figure A.12.45 CSEM Curves ELA/L Grade 5	467
Figure A.12.46 CSEM Curves ELA/L Grade 6	468
Figure A.12.47 CSEM Curves ELA/L Grade 7	469
Figure A.12.48 CSEM Curves ELA/L Grade 8	470
Figure A.12.49 CSEM Curves ELA/L Grade 9	471
Figure A.12.50 CSEM Curves ELA/L Grade 10	472
Figure A.12.51 CSEM Curves ELA/L Grade 11	473
Figure A.12.52 CSEM Curves Mathematics Grade 3	474
Figure A.12.53 CSEM Curves Mathematics Grade 4	475
Figure A.12.54 CSEM Curves Mathematics Grade 5	476
Figure A.12.55 CSEM Curves Mathematics Grade 6	477
Figure A.12.56 CSEM Curves Mathematics Grade 7	478
Figure A.12.57 CSEM Curves Mathematics Grade 8	479
Figure A.12.58 CSEM Curves Algebra I	480
Figure A.12.59 CSEM Curves Geometry	481
Figure A.12.60 CSEM Curves Algebra II	482
Figure A.12.61 CSEM Curves Integrated Mathematics I	483
Figure A.12.62 CSEM Curves Integrated Mathematics II	484
Figure A.12.63 CSEM Curves Integrated Mathematics III	485

Executive Summary

The Partnership for Assessment of Readiness for College and Careers (PARCC) is a state-led consortium designed to create next-generation assessments that, compared to traditional K-12 assessments, more accurately measure student progress toward college and career readiness. The PARCC assessments are aligned to the Common Core State Standards (CCSS) and were administered operationally for the first time in the 2014-2015 academic year. PARCC comprises assessments in both English Language Arts/Literacy (ELA/L) and Mathematics in grades 3 to 8 and high school.

The information provided in this technical report is intended for use by those who evaluate tests, interpret scores, or use test results in making educational decisions. It is assumed that the reader has technical knowledge of test construction and measurement procedures, as stated in *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). The purpose of this technical report is to describe the first operational administration of the PARCC assessments and includes the following topics:

- Background and purpose of the assessments,
- Test development of items and forms,
- Test administration, security, and scoring,
- Test taker characteristics,
- Classical item analyses and differential item functioning,
- Reliability and validity of scores,
- Item response theory (IRT) calibration and scaling,
- Performance level setting,
- Development of the score reporting scales and student performance, and
- Quality control procedures.

Background and Purpose

Assessments for the first operational administration were constructed in 2014. Eleven states and the District of Columbia participated in the first administration of the PARCC assessments during the 2014-2015 school year. A small subset of students were tested in fall 2014. ELA/L grades 9, 10, and 11, and Algebra I, Geometry, and Algebra II were administered in the fall; these assessments were administered on paper only. The majority of students tested during the spring 2015 window when all grades and content areas were administered online and on paper.

The PARCC assessments are designed to achieve several purposes. First, the tests are intended to provide evidence to determine whether students are on track for college- and career-readiness. Second, the tests are structured to access the full range of CCSS and measure the total breadth of student performance. Finally, the tests are designed to provide data to help inform classroom instruction, student interventions and professional development.

The two operational administrations, Fall 2014 and Spring 2015, of the PARCC assessment each included two separate components: the Performance-Based Assessment (PBA) and the End-of-Year (EOY) assessment. Both components were administered computer-based tests (CBT) and paper-based tests (PBT) during the spring testing window; tests administered in the fall were paper-based only. A valid score in both the PBA and EOY assessments was required for a student to receive a summative score.

Item Types

The tests contain selected response, brief and extended constructed response, technology-enabled, and technology-enhanced items (TEI), as well as performance tasks. Technology-enabled items are single-response or constructed-response items that involve some type of digital stimulus or open-ended response box with which the students engage in answering questions. Technology-enhanced items involve specialized student interactions for collecting performance data. Therefore, the act of performing the task is the way in which data are collected. Students may be asked, among other tasks, to categorize information, organize or classify data, order a series of events, plot data, generate equations, highlight text, or fill in a blank. One example of a TEI is an interaction in which students are asked to drag response options onto a Venn diagram to show the relationship among ideas.

Classical and IRT Item Analysis

Classical item analyses and differential item functioning analyses were performed on the data to evaluate the psychometric characteristics of the operational test items after items were administered and before scores were reported. The two-parameter logistic/generalized partial credit (2PL/GPC) IRT models were used for calibrations and scaling.¹ Multiple operational core forms were administered for each grade in ELA/L and mathematics. The forms included sets of embedded common items to provide data to support horizontal linking across test forms within a grade and content area.

The purpose of the IRT calibration and scaling was to place all operational items for a single grade/subject onto a common scale. The results of the 2014 field test dimensionality study indicated that multidimensional models, based on predetermined test structures (e.g., PBA versus EOY, and ELA/reading versus ELA/L writing), did not provide significantly better model fit compared to a unidimensional model, for both ELA/L and mathematics. A mode comparability study based on the 2014 field test data did not provide evidence to assume that scores resulting from PBT and CBT forms were strictly comparable between modes, particularly for PBA. Based on the findings from these two studies, the operational data were calibrated concurrently across the PBA and EOY components, and calibrations were conducted separately for PBT and CBT response data using IRT models consistent with mixed format data.

¹ IRT model comparison analyses conducted using the 2014 field test data clearly indicated that the item fit was better for the 2PL/GPC than the 1PL/PC model combination, for both ELA/L and for mathematics. Consequently, the recommendation by ETS and Pearson to use the 2PL/GPC models for operational calibrations and scaling was approved by PARCC. No ELA/L items and relatively few mathematics items were single select multiple choice items (approximately 15% to 23%). Consequently, the 3PL/GPC model did not markedly improve fit when compared to the fit of the 2PL/GPC model.

After calibration and model fit evaluation was completed, a master list of all items flagged as problematic was compiled and brought to the PARCC Priority Alert Task Force.² The Task Force reviewed each item, its content and the statistical properties, and made decisions about whether to include the item in the operational scores. Sometimes, an item was rejected because it appeared to have content issues, and sometimes an item was excluded because it could not be calibrated or showed extremely poor IRT model fit. Ultimately the decision about whether to keep or exclude each flagged item was made by the Task Force. The goals of the Task Force were to: a) minimize the number of items excluded from the operational test forms, and b) avoid advantaging or disadvantaging any test takers.

Once the item response data from the computer-based tests (CBT) and the paper-based tests (PBT) were calibrated for all grades and content areas, all available item parameter estimates of common items across modes, were used to transform the PBT item parameter estimates onto the CBT scales. The software program **STUIRT** (Kim & Kolen, 2004) was used to obtain Stocking and Lord (1983) transformation values to link the PBT scales to the CBT scales.

The PBT forms for all grades and content areas were generated using items from the CBT forms. In response to several practical constraints based on the number of forms constructed for each mode and to meet the blueprints (e.g., inclusion of TEI on CBT forms), there was no single CBT form that was administered intact in the paper delivery mode at any grade level. For example, TEI from online forms were replaced in the paper forms with items having similar content, but appropriate for paper-based testing. However, for both ELA/L and mathematics, the content on PBT forms significantly overlapped content on the CBT forms. A mode comparability study was conducted in 2015 and the results are presented as a separate special report. The study evaluated the extent to which scores from CBT and PBT forms could be considered as comparable with regard to psychometric characteristics. A major finding was that score comparability was inconsistent across the content domains and grade levels investigated.

Overall Scale Scores, Claim Scores, and Subclaim Scores

The PARCC ELA/L and mathematics scores are expressed as various types of scale scores (both total scores and claim scores, related to the claims structures described below), as well as by performance levels used to describe how well students meet the academic standards for their grade level. On the basis of a student's total score, an inference is drawn about how much knowledge and skill in the content area the student has acquired. The total score is also used to classify students in terms of the level of knowledge and skill in the content area as students progress in their K-12 education. These levels are called performance levels and are reported as:

- Level 5: Exceeded expectations
- Level 4: Met expectations
- Level 3: Approached expectations
- Level 2: Partially met expectations
- Level 1: Did not yet meet expectations

² The Priority Alert Task Force comprised Parcc Inc. staff, state leads, and state staff.

Students classified as either Level 4 or Level 5 are meeting or exceeding the grade level expectations. Additionally, information on more specific skills is provided and is reported as *Below Expectations*, *Nearly Meets Expectations*, and *Meets or Exceeds Expectations*.

PARCC has developed performance level descriptors (PLDs) to assist with the understanding and interpretations of the ELA/L and mathematics scores (<http://www.parcconline.org/assessments/test-design/ela-literacy/ela-performance-level-descriptors> and <http://www.parcconline.org/assessments/test-design/mathematics/math-performance-level-descriptors>). Additionally, resource information is available online to educators, parents, and students (<http://avocet.pearson.com/PARCC/Home#10829>), which includes information on understanding and interpreting the ELA/L and mathematics score reports.

The claim structures for ELA/L and for mathematics, grounded in the Common Core State Standards, informs the design and development of the summative assessments.

Claim Structure for ELA/L

Master Claim. The master claim is the overall performance goal for the PARCC ELA/L Assessment System—students must demonstrate that they are college- and career-ready or on track to readiness as demonstrated through reading and comprehending of grade-level texts of appropriate complexity and writing effectively when using and/or analyzing sources.

Major Claims: 1) reading and comprehending a range of sufficiently complex texts independently, and 2) writing effectively when using and/or analyzing sources.

Subclaims: The subclaims further explicate what is measured on the PARCC assessments and include claims about student performance on the standards and evidences outlined in the PARCC evidence tables for reading and writing (<http://www.parcconline.org/assessments/test-design/ela-literacy/test-specifications-documents>). The claims and evidences are grouped into the following categories.

1. Vocabulary, Interpretation, and Use
2. Reading Literature
3. Reading Informational Text
4. Written Expression
5. Knowledge of Language and Conventions

Claim Structure for Mathematics

Master Claim. The degree to which a student is college- or career-ready or on track to being ready in mathematics. The student solves grade-level/course-level problems aligned to the Standards for Mathematical Content with connections to the Standards for Mathematical Practice.

Subclaims: The subclaims further explicate what is measured on the PARCC assessments and include claims about student performance on the standards and evidences outlined in the PARCC evidence statement tables for mathematics (<http://www.parcconline.org/assessments/test-design/mathematics/math-test-specifications-documents>). The claims and evidence are grouped into the following categories.

Subclaim A: Major Content with Connections to Practices

Subclaim B: Additional and Supporting Content with Connections to Practices

Subclaim C: Highlighted Practices with Connections to Content: Expressing Mathematical Reasoning by constructing viable arguments, critiquing the reasoning of others, and/or attending to precision when making mathematical statements

Subclaim D: Highlighted Practice with Connections to Content: Modeling/Application by solving real-world problems by applying knowledge and skills articulated in the standards.

Score Scales

Scale scores were defined for each test as a linear transformation of the IRT theta (Θ) scale. The test characteristic curves associated with the performance level setting forms were used to identify the theta values associated with the Level 2 and Level 4 point scores. By defining Level 2 and 4 scale scores to be 700 and 750, respectively, the linear relationship between theta and scale scores was established.

The result is 201 defined full summative scale score points for each ELA/L and mathematics assessment, ranging from 650 to 850. A scale score of 700 is always the minimum for Level 2 performance, a scale score of 750 is always the minimum for Level 4 performance.

The thresholds for summative performance levels on the scale score metric recommended by the scale score task force are described in Table 1.

Table 1. Defined Summative Scale Scores and Cut Scores

	Lowest Obtainable Scale Score	Cut Score Level 2	Cut Score Level 4	Highest Obtainable Scale Score
Full Summative	650	700	750	850

As with the full summative scores, scale scores for Reading and Writing were defined for each test as a linear transformation of the IRT theta (Θ) scale. The same IRT theta scale was used for Reading and Writing as was used for the ELA/L full summative scores. The theta values associated with the Level 2 and Level 4 performance levels were identified using the test characteristic curves associated with the performance level setting forms. Parallel to the full summative scores, the relationship between theta and scale scores was established with Level 2 and 4 theta scores and the corresponding predefined scale scores.

The result was 81 defined scale score points for Reading, ranging from 10 to 90. A scale score of 30 is the cut score for minimum Level 2 performance, a scale score of 50 is the cut score for minimum Level 4 performance. There are 51 defined scale score points for Writing, ranging from 10 to 60. A scale score of 25 is the cut score for minimum Level 2 performance, a scale score of 35 is the cut score for minimum Level 4 performance. The threshold Reading and Writing performance levels on the scale score metric recommended by the scale score task force are described in Table 2.

Table 2. Defined Scaled Scores and Cut Scores for Reading and Writing Claim Scores

	Lowest Obtainable Scale Score	Cut Score Level 2	Cut Score Level 4	Highest Obtainable Scale Score
Reading	10	30	50	90
Writing	10	25	35	60

Regarding the subclaim scores, the Level 4 cut is defined as *Meets or Exceeds Expectations* because high school students at Level 4 or above are likely to have the skills and knowledge to meet the definition of career and college readiness. Subclaim outcomes center on that performance level and are reported at three levels:

- *Below Expectations;*
- *Nearly Meets Expectations; or*
- *Meets or Exceeds Expectations.*

Quality Control

To ensure IRT calibrations, scaling and conversion tables were produced accurately, Pearson replicated the IRT calibrations and scale score transformations carried out by ETS, and the generation of the score conversion tables. While ETS used **PARSCALE**, Pearson independently conducted the same calibrations using **IRTPRO** (Cai, Thissen & du Toit, 2011) calibration software. ETS and Pearson both used **STUIRT** software to transform their PBT item parameter estimates onto the CBT scales for each grade/subject. Pearson's scaling constants were compared to those generated by ETS and found to be consistent. Measured Progress (MP) performed independent quality control comparisons between the ETS and Pearson item parameter estimates to identify any differences. In addition, MP independently made certain that the same items were excluded from the CBT/PBT linking sets, and compared transformed PBT parameter estimates computed by ETS and Pearson. If items had large differences across modes, the items were discussed and any remaining issues resolved. Measured Progress prepared reports documenting their findings. Exact matches were found between all ETS and Pearson conversion tables before scores were reported.

Section 1: Introduction

1.1 Background

States associated with the Partnership for Assessment of Readiness for College and Careers (PARCC) came together in early 2010 with a shared vision of ensuring that all students - regardless of income, family background or geography - have equal access to a world-class education that will prepare them for success after high school in college and/or careers. The PARCC goal was to develop new assessments that tie into more rigorous academic expectations and help prepare students for success in college and the workforce, as well as to provide information back to teachers and parents about where students are on their path to success. Calling on the expertise of thousands of teachers, higher education faculty and other educators in multiple states, the PARCC assessment system is a high quality set of summative assessments, diagnostic assessments, formative tasks, and other support materials for teachers including professional development and communications tools.

The PARCC consortium is designed to develop and administer next-generation assessments that, compared to traditional K-12 assessments, more accurately measure student progress toward college and career readiness. The assessments are aligned to the Common Core State Standards (CCSS) and include both English Language Arts/Literacy (ELA/L) assessments (grades three through eleven) and mathematics assessments (grades three through eight, and high school). Compared to traditional standardized tests, these assessments are intended to measure more complex skills like critical-thinking, persuasive writing, and problem-solving.

In 2013, the PARCC Governing Board launched Parcc Inc., a non-profit organization designed to support the successful delivery of the tests in 2014-15, and the long-term success of the multi-state partnership. States continue to govern decisions about the assessment system; the non-profit organization is their “agent” for overseeing the many vendors involved in the PARCC assessment system, coordinating the multiple work groups and committees (including Governing Board meetings), managing the PARCC intellectual property, overseeing the research agenda and the Technical Advisory Committee, and developing and launching the multiple non-summative tools.

Summative assessments for the first operational administration were constructed in 2014. Eleven states and the District of Columbia participated in the first administration of the PARCC assessments during the 2014-2015 school year. A small subset of students tested in Fall 2014. ELA/L grades 9, 10, and 11, and Algebra I, Geometry, and Algebra II were administered in the fall; these assessments were administered on paper only. The majority of students tested during the Spring 2015 window when all grades and content areas were administered online and on paper.

The purpose of this technical report is to describe the first operational administration of the PARCC summative assessments, including test form construction, test administration, item scoring, test taker characteristics, classical item analysis results, reliability results, evidence of validity, item response theory (IRT) calibrations and scaling, performance level setting procedure, and quality control procedures.

1.2 Purpose of the Operational Tests

The PARCC assessments are designed to achieve several purposes. First, the tests are intended to provide evidence to determine whether students are on track for college- and career-readiness. Second, the tests are structured to access the full range of CCSS and measure the total breadth of student performance. Finally, the tests are designed to provide data to help inform classroom instruction, student interventions, and professional development.

1.3 Composition of Operational Tests

Each operational test form was constructed to reflect the full test blueprint in terms of content, standards measured, and item types. Sets of common items, included to provide data to support horizontal linking across test forms within a grade and content area, were proportionally representative of the operational test blueprint.

The two operational administrations, Fall 2014 and Spring 2015, of the PARCC assessment each included two separate components: the Performance-Based Assessment (PBA) and the End-of-Year (EOY) assessment. Both components were administered as computer-based tests (CBT) and as paper-based tests (PBT) during the spring testing window; tests administered in the fall were paper-based only. A valid score in both the PBA and EOY assessments was required for a student to receive a summative score.

The PBA and EOY components utilized somewhat different item types. The PBA was administered after approximately 75 percent of instructional time was complete. The purpose of the PBA component was to measure critical thinking, reasoning, and the ability to apply skills and knowledge in reading, writing, and mathematics. The ELA/L PBA component comprised three types of tasks: literary analysis, narrative writing, and research simulation. For each task, students were instructed to read one or more texts, answer several brief questions, and then write an essay based on the material they read. The mathematics PBA consisted of tasks designed to assess a student's ability to use mathematics to solve real-life problems. Some of the tasks required that students describe how they solved a problem, while other tasks measured conceptual understanding and ability to apply concepts by means of selected-response or technology-enhanced items.

The EOY administration occurred after approximately 90 percent of instruction was complete. Students were required to demonstrate their skills and knowledge by answering innovative selected-response and short-answer questions that measured concepts and skills. The ELA/L EOY assessment had between two and four literary and informational texts; each text had five or six brief comprehension and vocabulary questions. The mathematics EOY assessment contained tasks that measured a combination of conceptual understanding, applications, skills, and procedures.

1.4 Intended Population

The PARCC tests are intended for students taking ELA/L and/or mathematics in grades 3 through 11, as well as students taking high school mathematics (i.e., Algebra I, Geometry, Algebra II, Integrated

Mathematics I – III). For these students, the PARCC tests measured whether students were meeting state academic standards and mastering the knowledge and skills needed to progress in their K-12 education and beyond.

1.5 Groups and Organizations Involved with PARCC

- Parcc Inc. is a nonprofit organization that assumes the responsibility for management of the PARCC consortium, as well as the development and implementation of PARCC assessments.

Parcc Inc. created a number of committees of educators and state leads to help manage and lead the day-to-day work of the consortium. These committees include:

- the PARCC consortium Governing Board that makes major policy and operational decisions,
 - the Technical Advisory Committee that helps ensure all assessments will provide reliable results to inform valid instructional and accountability decisions,
 - the K-12 State Leads that coordinates all aspects of development of the PARCC assessment system and serves as the conduit to the Technical Advisory Committee and the Governing Board,
 - the Advisory Committee on College Readiness which includes higher education executive officers from PARCC states and other state- and nationally-recognized leaders in the postsecondary community, and
 - the Higher Education Leadership Team which is responsible for coordinating higher education engagement in the PARCC assessment system and works closely with the Advisory Committee on College Readiness.
- Test and item development activities were conducted by Pearson, ETS, and WestEd under the guidance and oversight of PARCC leadership.³
 - Pearson served as the primary contractor for the PARCC operational administration and was responsible for developing test forms, production of all testing materials, packaging and distribution, receiving and scanning of materials, and scoring, as well as program management and customer service.
 - Educational Testing Service (ETS) served as a subcontractor and was responsible for all psychometric analyses of the PARCC operational test data. This included classical item analyses,

³ PARCC leadership includes the following groups: PARCC Governing Board, K-12 State Leads, Higher Education Leadership Team, Technical Advisory Committee, Operational Working Group members from each of the member states, and staff members from Parcc, Inc., the project management partner for the PARCC Consortium.

differential item functioning (DIF) analyses, item calibrations based on item response theory (IRT), scaling, and development of all conversion tables.

- Measured Progress (MP) served as a subcontractor to conduct external evaluations; they were responsible for reviewing and comparing the psychometric IRT calibrations performed by ETS, which were replicated by Pearson. MP also provided comparisons of results obtained independently from ETS and from Pearson for raw-to-theta (RST) conversion tables, summative and claim scale scores, performance level classifications, and subclaim performance level classifications.

1.6 Overview of the Technical Report

This report begins by providing explanations of the test form construction process, test administration, and scoring of the test items. Subsequent sections of the report present descriptions of test taker characteristics, results of classical item analyses, results of reliability analyses, evidence of validity, item response theory (IRT) calibrations and scaling, performance level setting procedure, and quality control procedures.

The technical report contains the following sections:

- Section 2 – Test Development

This section describes the PARCC test design and the procedures followed during the development of operational test forms.

- Section 3 – Test Administration

This section presents the operational administration schedule, information regarding test security and confidentiality, accessibility features and accommodations, and testing irregularities and security breaches.

- Section 4 – Scoring of the Items

The key-based and rule-based processes for machine scored items, as well as the training and monitoring processes for human scored items are provided in this section.

- Section 5 – Test Taker Characteristics

This section describes the composition of test forms, rules for inclusion of students in analyses, distributions of test takers by grade, mode, and gender, and distributions of demographic variables of interest.

- Section 6 – Classical Item Analyses

The classical item-level statistics calculated for the operational test data, the flagging criteria used to identify items that performed differently than expected, and the results of these analyses are presented in this section.

- Section 7 – Differential Item Functioning

In this section, the methods for conducting differential item functioning analyses as well as corresponding flagging criteria are described. This is followed by definitions of the comparison groups and subsequent results for the comparison groups.

- Section 8 - Reliability

The results of internal consistency reliability analyses and corresponding standard errors of measurement, for each grade, content area, and mode (CBT or PBT) for all test takers, and for subgroups of interest, is provided in this section. This is followed by reliability results for subscores and reliability of classification (i.e., decision accuracy and decision consistency). Finally, expectations and results for interrater agreement for hand scored items are summarized.

- Section 9 – Validity

Validity evidence based on analyses of the internal structure of the tests is provided in this section. Correlations between subscores are reported by grade, content area, and mode (CBT or PBT) for all test takers.

- Section 10 - IRT Calibration and Scaling

This section presents the information related to the calibration and scaling of item response data including: data preparation, the calibration process, model fit evaluation, and items excluded from score reporting. In addition, the scaling process (paper to online) is described and evaluated.

- Section 11 – Performance Level Setting (PLS) Procedure and Results

Performance levels and policy definitions, as well as the processes followed to establish performance level thresholds are described in this section.

- Section 12 - Scale Scores

This section provides an overview of the claims and subclaims, describes the development of the reporting scales and conversion tables, and presents scale score distributions. Finally, information regarding the interpretation of claim scores and subclaim scores is presented.

- Section 13 – Quality Control Procedures

All aspects of quality control are presented in this section. These activities range from quality assurance of item banking, test form construction, and all testing materials to quality control of scanning, image editing, and scoring. This is followed by a detailed description of the steps taken to ensure that all psychometric analyses were of the highest quality.

- References
- Appendices

To facilitate utility, tables in the appendices are numbered sequentially according to the section represented by the tables. For example, the first appendix table for Section 5 is numbered A.5.1, the second appendix table for Section 5 is numbered A.5.2, and so on.

- Addendum

The addendum presents the results of analyses for the Fall 2014 operational administration. These results are reported separately from the Spring 2015 results because fall testing involved a nonrepresentative subset of students testing only on paper and only for ELA/L grades 9, 10, and 11, as well as Algebra I, Geometry, and Algebra II.

To organize the addendum, tables are numbered sequentially according to the section represented by the tables. For example, the first addendum table for Section 5 is numbered ADD.5.1, the second addendum table for section 5 is numbered ADD.5.2, and so on.

1.7 Glossary of Abbreviations

Table 1.1 Glossary of PARCC Abbreviations and Acronyms

Abbreviation/Acronym	Definition
1PL/PC	One-parameter/Partial Credit Model
2PL/GPC	Two-parameter Logistic/Generalized Partial Credit Model
3PL/GPC	Three-parameter Logistic/Generalized Partial Credit Model
AAF	Accessibility, Accommodations, and Fairness
ABBI	Assessment Banking for Building and Interoperability
AERA	American Educational Research Association
AIS	Average Item Score
AIQ	Assessment and Information Quality
APA	American Psychological Association
ASC	Additional and Supporting Content (Mathematics)
ASL	American Sign Language
ATA	Automatic Test Assembler
CBT	Computer-Based Test
CCSS	Common Core State Standards
CSEM	Conditional Standard Error of Measurement
DIF	Differential Item Functioning
DPL	Digital Production Line
DPP	Digital Pre-press
EBSS	Evidence-based Standard Setting
ELA/L	English Language Arts/Literacy
EL	English Learners
EOC	End-of-Course
EOY	End-of-Year
ePEN2	Electronic Performance Evaluation Network
ESEA	Elementary and Secondary Education Act
ETS	Educational Testing Service
FRL	Free or Reduced-price Lunch
FS	Full Summative
FT	Field Test
IA	Item Analysis
ICC	Item Characteristic Curve
IDEA	Individuals with Disabilities Education Act
IEP	Individualized Education Program
INF	Information Curve
IRA	Inter-rater Agreement
IRF	Item Response File
IRT	Item Response Theory
IRS	Individual Student Report
K-12	Kindergarten to Grade 12
LEA	Local Education Agency
LID	Local Item Dependence
MAD	Mean Absolute Difference
MC	Major Content (Mathematics)
MH	Mantel-Haenszel
MP	Measured Progress

MP	Modeling Practice (Mathematics)
MR	Mathematical Reasoning
NAEP	National Assessment of Educational Progress
NCLB	No Child Left Behind
NCME	National Council on Measurement in Education
NSLP	National School Lunch Program
OE responses	Open-ended responses
OMR	Optical Mark Reading
OWG	Operational Working Group
PARCC	Partnership for Assessment of Readiness for College and Careers
PBA	Performance-Based Assessment
PBT	Paper-Based Test
PCR	Prose Constructed Response (ELA/L)
PEJ	Postsecondary Educators' Judgment
PLD	Performance Level Descriptor
PLS	Performance Level Setting
PV	Product Validation
QA	Quality Assurance
RD	Reading (ELA/L)
RI	Reading Information (ELA/L)
RL	Reading Literature (ELA/L)
RMSD	Root Mean Square Difference
RV	Reading Vocabulary (ELA/L)
RST	Raw-score-to-theta
SD	Standard Deviation
SDF	Student Data File
SE	Standard Error
SEJ	Standard Error of Judgment
SEM	Standard Error of Measurement
SIRB	Scored Item Response Block
SMD	Standardized Mean Difference
SSMC	Single Select Multiple Choice
SWD	Students with Disabilities
TCC	Test Characteristic Curve
TTS	Text to Speech
UIN	Unique Item Number
WE	Writing Written Expression (ELA/L)
WKL	Writing Knowledge Language and Conventions (ELA/L)
WLS	Weighted Least Squares
WR	Writing (ELA/L)
WRMSD	Weighted Root Mean Square Difference

Section 2: Test Development

2.1 Overview of the PARCC Assessment, Claims, and Design

Aligned to the Common Core State Standards (CCSS) as articulated in the PARCC Model Content Frameworks, the PARCC assessments are designed to determine whether students are college- and career-ready or on track, assess the full range of the CCSS, measure the full range of student performance, and provide data to help inform instruction, interventions, and professional development. Test development is an ongoing process involving educators, researchers, psychometricians, subject matter professionals, and assessment experts who participate in the development of the PARCC test design and its underlying foundational documents; develop and review passages and items used to build the PARCC assessments; monitor the program for quality, accessibility, and fairness for all students; and construct, review, and score the assessments

The PARCC summative assessments include both English Language Arts/Literacy (ELA/L) and mathematics assessments in grades 3 to 8 and high school. The high school mathematics tests include traditional mathematics and integrated mathematics course pathways. Tests contain selected response, brief and extended constructed response, technology-enabled and technology-enhanced items (TEI), as well as performance tasks. Technology-enabled items are single-response or constructed-response items that involve some type of digital stimulus or open-ended response box with which the students engage in answering questions. Technology-enhanced items involve specialized student interactions for collecting performance data. In other words, the act of performing the task is the way in which data is collected. Students may be asked, among other interactions, to categorize information, organize or classify data, order a series of events, plot data, generate equations, highlight text, or fill in a blank. One example of a TEI is an interaction in which students are asked to drag response options onto a Venn diagram to show the relationship among ideas.

The PARCC assessments offer a wide range of accessibility features for all students and accommodations for students with disabilities (e.g., screen reader, assistive technology, braille, large print, text-to-speech, and ASL video versions of the test, as well as response accommodations that allow students to respond to test items using different formats). For English learners who are native Spanish speakers, PARCC offers a paper-based edition of the mathematics assessment in Spanish, and both large print and Text-to-Speech versions of the test in Spanish (refer to the PARCC Accessibility Features and Accommodations Manual for in-depth information).

2.1.1 English Language Arts/Literacy (ELA/L) Assessments – Claims and Subclaims

The ELA/L summative assessment at each grade level includes two parts: The Performance Based Assessment (PBA) and the End-of-Year assessment (EOY). The PBA consists of three task types: Literary Analysis, Research Simulation, and Narrative Writing. For each performance-based task, students are asked to read or view one or more texts, answer comprehension and vocabulary questions, and write an extended response that requires them to draw evidence from text(s). The EOY contains literary and informational reading passages with comprehension and vocabulary questions.

The claim structure, grounded in the CCSS, undergirds the design and development of the ELA/L summative assessments.

Master Claim. The master claim is the overall performance goal for the PARCC ELA/Literacy Assessment System—students must demonstrate that they are college- and career-ready or on track to readiness as demonstrated through reading and comprehending of grade-level texts of appropriate complexity and writing effectively when using and/or analyzing sources.

Major Claims: 1) reading and comprehending a range of sufficiently complex texts independently, and 2) writing effectively when using and/or analyzing sources.

Sub Claims: The sub claims further explicate what is measured on the PARCC assessments and include claims about student performance on the standards and evidences outlined in the PARCC evidence tables for reading and writing (refer to *PARCC Test Specifications Documents*). The claims and evidences are grouped into the following categories.

1. Vocabulary Interpretation and Use
2. Reading Literature
3. Reading Informational Text
4. Written Expression
5. Knowledge of Language and Conventions

2.1.2 Mathematics Assessments – Claims and Subclaims

The summative mathematics assessment at each grade level includes both short- and extended-response questions focused on applying skills and concepts to solve problems that require demonstration of the mathematical practices from the Common Core State Standards with a focus on modeling and reasoning with precision. The assessments also include performance-based short-answer questions focused on conceptual understanding, procedural skills, and application.

The claim structure, grounded in the CCSS, undergirds the design and development of the summative assessments.

Master Claim. The degree to which a student is college- or career-ready or on track to being ready in mathematics. The student solves grade-level/course-level problems aligned to the Standards for Mathematical Content with connections to the Standards for Mathematical Practice.

Sub Claims: The sub claims further explicate what is measured on the PARCC assessments and include claims about student performance on the standards and evidences outlined in the PARCC evidence statement tables for mathematics (refer to *PARCC Test Specifications Documents*). The claims and evidence are grouped into the following categories.

Subclaim A: Major Content with Connections to Practices.

Subclaim B: Additional and Supporting Content with Connections to Practices.

Subclaim C: Highlighted Practices with Connections to Content: Expressing Mathematical Reasoning by constructing viable arguments, critiquing the reasoning of others, and/or attending to precision when making mathematical statements.

Subclaim D: Highlighted Practice with Connections to Content: Modeling/Application by solving real-world problems by applying knowledge and skills articulated in the standards.

2.2 Test Development Activities

Test development activities began with the standards and model content frameworks. From these, PARCC, in collaboration with more than 2,000 educators, researchers, and psychometricians, has developed the PARCC test specifications documents that guide the development of test items and the composition of the tests. These documents include the College- and Career-Ready determinations and Performance-Level Descriptions, Claim Structure, Evidence Statement Tables, Blueprints, Informational Guides, Passage Selection Guidelines, Mathematics Sequencing Guidelines, Task Generation Models, Fairness and Sensitivity Guidelines, Text Selection Guidelines, and the Style Guide. Refer to the PARCC website for further information about these documents.

2.2.1 Item Development Process

PARCC test and item development activities were conducted by Pearson, ETS, and WestEd under the guidance and oversight of PARCC leadership, including the PARCC Governing Board, the K-12 State Leads, the Higher Education Leadership Team, the Technical Advisory Committee, the Operational Working Group members from each of the member states, the PARCC State Text and Content Review Committees, and staff members from Parcc, Inc., the project management partner for the PARCC Consortium.

Developing high quality assessment content with authentic stimuli for computer-based tests (CBT) and paper-based tests (PBT) measuring rigorous standards is a complex process involving the services of many experts including assessment designers, psychometricians, managers, trainers, content providers, content experts, editors, artists, programmers, technicians, human scorers, advisors, and members of the PARCC Operational Working Groups.

Bank Analysis and Item Development Plan

The PARCC summative item bank houses passages and items at each assessed grade level and subject. The bank supports the administration of the assessments, along with item release and practice tests. Items are developed and field tested annually. Prior to the annual item development cycle, the item development teams, in conjunction with members of the Operational Working Groups (OWGs) for ELA/L and mathematics, evaluated the strengths of the bank and considered the needs for future tests to establish an item development plan.

Text Selection for ELA/L

Using the PARCC Passage Selection Guidelines, English language arts subject matter experts were trained to search for appropriate passages to support an annual pool of passages for consideration. Guided by the PARCC test specifications documents, Pearson, ETS, and WestEd recruited, trained, and managed the contracted subject matter experts to deliver the number of texts specified in the annual item development plan. The Passage Selection Guidelines provided a text complexity framework, and guidance on selecting of a variety of text types and passages that allow for a range of standards/evidences to be demonstrated to meet the PARCC claims. PARCC ELA/L tests are based on authentic texts, including multi-media stimulus. Authentic texts are grade-appropriate texts that are not developed for the purposes of the assessment or to achieve a particular readability metric, but reflect the original language of the authors. Pearson, ETS, and WestEd content experts reviewed the passages for adherence to the PARCC passage selection guidelines (passage selection guidelines can be found on PARCC website) and meet to the annual item development plan described above in the number and distribution of genres and topics prior to review and consideration by the State Text Review Committee. ELA/L item development was not conducted until after texts were approved by the State Text Review committee.

Item Development

Guided by the PARCC foundational documents, Pearson, ETS, and WestEd recruited and trained the item writers and managed the item writing to develop the number of items specified in the annual item development plan. Prior to further committee reviews, the assessment teams at Pearson, ETS, and WestEd reviewed the items - for content accuracy, alignment to the standards, range of difficulty, adherence to universal design principles (which maximize the participation of the widest possible range of students), bias and sensitivity, and copy edit to enable the accurate measurement of the PARCC standards.

2.2.2 Item and Text Review Committees

Members of the PARCC OWGs for ELA/L and mathematics, state-level experts, local educators, postsecondary faculty, and community members from the PARCC states conducted rigorous reviews of every item and passage being developed for the PARCC assessment system to ensure all test items are of the highest quality, aligned to the standards, and fair for all student populations. All PARCC reviewers were nominated by their state education agency. The purpose of the educator reviews was to provide feedback to Pearson, ETS, WestEd, and PARCC on the quality, accuracy, alignment, and appropriateness of the test passages and items developed annually for the summative PARCC assessments. The meetings were conducted either in person or virtually and included large group training on the expectations and processes of each meeting, followed by break outs into grade/subject working committees where additional training was provided.

State Text Review

The State Text Review is a review and approval by the State Text Review Committee of the texts eligible for item development. Participants reviewed and provided feedback to Pearson, ETS, WestEd, and PARCC about the grade-level appropriateness, content, and potential bias concerns, and reached consensus about which texts would move forward for development. The State Text Review Committee was made up of both State Content and Bias and Sensitivity committee members.

State Content Item Review

During State Content Item Review, committees reviewed and edited test items for adherence to the PARCC foundational documents, basic universal design principles, PARCC Accessibility Guidelines, associated item metadata, and PARCC Style Guide. Committees accessed the item content within the Pearson Assessment Banking for Building and Interoperability (ABBI) system that previews how the passages and items will be displayed in an operational online environment. Committees also verified that the appropriate scoring rule had been applied to each item. The Content Review committees were made up of Operational Working Group members and educators nominated by PARCC member states.

State Bias and Sensitivity Review

Educators and community members make up the committee that reviews items and tasks to confirm that there are no bias or sensitivity issues that would interfere with a student's ability to achieve his or her best performance. The committee reviewed items and tasks to evaluate adherence to the Fairness and Sensitivity Guidelines, and to ensure that items and tasks do not unfairly advantage or disadvantage one student or group of students over another. Bias and Sensitivity Committee members made edits and modifications to items and passages to eliminate sources of bias and improve accessibility for all students.

Editorial Review

The PARCC editorial review committee is comprised of state-level editors who reviewed up to 10 percent of the items and tasks. The committee reviewed the items for copy edit, clarity, and adherence to the PARCC Style Guide.

Data Review Committee

Following the field test, educator and bias committee members met to evaluate test items and associated performance data with regard to appropriateness, level of difficulty, and potential gender, ethnic, or other bias, then recommended acceptance or rejection of each field-test item for inclusion on an operational assessment. The committee also made recommendations that items be revised and re-field tested. Items that were approved by the committee are eligible for use on operational summative assessments.

2.2.3 Operational Test Construction

Under the guidance in the Operational Test Form Creation Specifications, Pearson constructed the operational forms for the two test components that contributed to the summative score--the Performance-Based Assessment (PBA) and the End-of-Year (EOY) assessment to adhere to the test blueprints and the assessment goals outlined in the form creation specifications. These goals were:

- Test forms designed to measure well across the full range of student ability;
- Scores that are comparable among forms and across test administrations;
- Scales that support classification of students into performance levels;
- The number of parallel forms for PBA and EOY are maximized;
- Overexposure of items is minimized; and
- Adherence to standards for validity, reliability, and fairness (*Standards for Educational and Psychological Testing, 2014*).

Each content-area and grade-level assessment was based on a specific test blueprint that guided how each test is built. Test blueprints determined the range and distribution of content, and the distribution of points across the PARCC subclaims and task types.

Multiple operational forms were constructed for each component/grade/subject. These forms were designed to facilitate psychometric equating through a common item linking strategy (described in Section 2.2.4) and to be constructed as “parallel” as possible from a content and test-taking experience. Evaluation criteria for parallelism included adherence to blueprint; sequencing of content across the forms; statistical averages and distributions for difficulty (e.g., p value) and discrimination (e.g., polyserial correlation); item type and cognitive complexity; and passage characteristics for ELA/L including genre, topics, word count, and text complexity.

Core forms are the operational test forms consisting of only those items that will count towards a student’s score. Core forms are constructed to meet the blueprint and psychometric properties outlined in the test construction specifications. PARCC creates multiple core forms for a given assessment to enhance test security and to support opportunity for item release. The number of core operational forms per component/grade/subject is provided in Table 2.1. Additionally, appropriate forms were identified as accessibility and accommodated forms; and the core forms for all assessments except the ELA/L-PBA included embedded field test items. Accessibility and Accommodated forms and embedded field testing are described later in this section.

Table 2.1 Number of Core Operational Forms per Grade/Subject by Component and Mode for ELA/L and Mathematics

Grade/ Subject	ELA/L				Mathematics			
	CBT		PBT		CBT		PBT	
	PBA	EOY	PBA	EOY	PBA	EOY	PBA	EOY
Grade 3	6	6	4	4	6	6	4	4
Grade 4	6	6	4	4	6	6	4	4
Grade 5	6	6	4	4	6	6	4	4
Grade 6	6	6	4	4	6	6	4	4
Grade 7	6	6	4	4	6	6	4	4
Grade 8	6	6	4	4	6	6	4	4
Grade 9	8	8	4	4				
Grade 10	8	8	4	4				
Grade 11	6	6	4	4				
Algebra I					8	8	4	4
Geometry					8	9*	4	4
Algebra II					6	6	4	4
Integrated Mathematics I					2	2	2	2
Integrated Mathematics II					2	2	2	2
Integrated Mathematics III					2	2	1	2

Note: * For Geometry CBT EOY there are 9 core forms, instead of 8, because one item is different on two versions.

Test Construction Activities

After the Data Review Meetings and prior to the Test Construction Meetings, Pearson assessment specialists constructed initial versions of all of the core forms, as depicted in Table 2.1. The construction model varied slightly between the two subject areas.

For ELA/L, content specialists constructed the initial core forms shown in Table 2.1 based on the support documents and specific processes to achieve fair parallel forms. The following steps were used to construct the operational core ELA/L form inputs taken to the Test Construction Committee for review.

1. Constructed the online EOY forms to match blueprint and test construction specifications
2. Constructed the paper EOY forms to match the blueprint and test construction specifications
3. Identified EOY Accommodated and Accessibility Forms by evaluating the constructed forms for eligibility
4. Repeated for PBA

The ELA/L construction process included iterative steps between content specialists and psychometricians. Custom PARCC test construction reports (i.e., SAS Reports) generated by the Pearson psychometric team provided information on adherence to blueprint and statistical averages/distributions of item difficulty and discrimination describing the forms and allowing comparison of the forms. These reports facilitated content changes to better achieve the test construction goals.

For mathematics, Pearson employed the use of an automatic test assembler (ATA) to select the items for the initial forms. Based on the blueprints and other test construction goals and specifications, the ATA was able to create sets of items best satisfying the statistical parameters outlined in the test construction specifications; however the ATA was unable to sequence the items as required by the PARCC Mathematics Sequencing Guidelines. Sequencing was conducted by assessment specialists who ordered the items according to the sequencing guidelines. To achieve the appropriate linking design, assessment specialists created linking item sets from an ATA-generated linking blueprint; these sets are shared across forms using the strategy described later in Section 2.2.4. The following steps were used to construct the linking sets and operational core form inputs taken to the Test Construction Committee for review.

1. ATA pulls EOY linking online blueprint
2. Blueprint sequenced and linking item sets created
3. ATA uses linking sets and pulls online EOY forms
4. Construct the online EOY forms
5. Construct the paper EOY forms
6. Identify EOY Accommodated and Accessibility Forms
7. Repeat for PBA

Similar to the ELA/L construction process, mathematics included iterative steps between assessment specialists and psychometricians. Custom PARCC test construction reports (i.e., SAS Reports) generated by the Pearson psychometric team provided information on adherence to blueprint and statistical averages/distributions of item difficulty and discrimination allowing a comparison of the forms and facilitating content changes to better achieve the test construction goals. Since the mathematics forms were generated by the ATA, psychometricians could also generate the SAS reports prior to content experts reviewing the forms.

Pearson assessment specialists identified forms for each grade/subject suitable for use as the accommodated forms. The content of these forms was also reviewed by Pearson accessibility specialists allowing for content changes prior to the Test Construction Meetings.

These test construction activities provided significant inputs to commence the Test Construction Meetings including:

- The proposed items for the initial operational core forms and the accommodated forms described above
- SAS reports describing each form and comparing parallel forms
- Recommended accommodated forms

Test Construction Meeting to Review Test Construction Inputs

Members of the State Item Content Committees and the Accessibility, Accommodations, and Fairness (AAF) Operational Working Group (OWG) participated in the building of operational core forms that met

PARCC assessment blueprints for the performance-based and end-of-year components of the summative assessments. In that process, they met in an in-person meeting to review and made recommendations for changes so that test forms conformed to both the content and psychometric requirements of the assessment.

Accommodated Form Review Process

In addition to participating in many of the development activities including the State Text Review and the State Bias and Sensitivity Review meetings, the Accessibility, Accommodations, and Fairness (AAF) Operational Working Group (OWG) reviewed the proposed accommodated forms at the Test Construction Meeting for accessibility to make sure that the content can be accommodated for students with disabilities and English learners without changing the underlying measured construct.

Forms were identified to support the following accommodations:

Group 1

- Braille (Tactile Graphics available)
- Large Print
- Refreshable braille (ELA/L only) (Tactile Graphics available)
 - Also supports Screen Reader Assistive Technology
- Screen Reader Assistive Technology (Mathematics)
- Spanish Paper (Mathematics only)
 - One additional item group will be needed for a second form

Group 2

- Closed Captioning (ELA/L only)
- Descriptive Video (ELA/L only)
 - Not for Spring 2015
- Signing: ASL (ELA/L only)
- Online TTS (Text and Graphics only)
 - ELA/L
 - Mathematics
 - Need to support text only, and text and graphics
- Online Spanish/TTS (Mathematics only)

At the conclusion of the meetings, all test forms were constructed to meet test blueprints and PARCC requirements, and to the extent possible, reflect the operational linking design. Each test form reflected the test blueprint in terms of content, item types, and test length, as well as *expected* difficulty and performance along the ability continuum. Linking sets were proportionally representative of the operational test blueprint. The operational core forms, linking set forms, and field test forms were reviewed by PARCC Forms Review committees and approved prior to the test administration.

Spanish-Language Assessments for Mathematics

For English learners, PARCC offers a paper-based edition of the mathematics assessment in Spanish, as well as large print and Text-to-Speech versions of the test in Spanish. Once the operational form was

approved, the form was sent to Pearson's subcontractor, Teneo, for transadaptation of the items. Transadaptation differs from translation in that it takes into consideration the grade-level appropriateness of the words, as well as the linguistic and cultural differences that exist between speakers of two different languages. Accounting for these differences allows the item to measure the achievement of Spanish language speakers in the same way that the original version of the item does for native speakers of English. The PARCC Spanish Glossary provided guidance to the translator conducting the transadaptation in grade-level and culturally-appropriate ways of transadapting the items. For the Spanish language text-to-speech form, the alternate text (used for description and/or text in art and graphics) was transadapted from the alternate text for the English language version of the Text-to-Speech form. Phonetic mark-up, which guides how the text-to-speech reader pronounces content-specific words and phrases, was also applied in this process.

In addition to the expert review of potential content for all accommodated forms conducted by the AAF OWG with assistance from content experts at the Test Construction Meetings, the transadapted forms underwent three additional quality checks: a Pearson Spanish copy edit services review and approval, an AAF OWG review and approval, and a Spanish DIF analysis after the administration.

2.2.4 Linking Design of the Operational Test

This section begins with a discussion of special considerations for selection of linking items, followed by two examples of the graphical representations of the linking designs for ELA/L and for mathematics. To support the goal of score comparability within and across administrations and years, PARCC implemented a hybrid approach that incorporated the strengths of common item linking and randomly equivalent groups. The use of repeated operational core items was leveraged for common item linking. In addition, all forms were available throughout the operational administrations, with spiraling at the student level, leveraged to support linking through randomly equivalent groups.

The PARCC operational test forms involved various types of linking: horizontal linking, testing mode linking, and across administration linking. Horizontal linking consisted of linking items, or common items, included in multiple forms in a single administration. The horizontal linking was achieved through a daisy-chain strategy. This strategy links multiple operational forms together in a ring; where each operational form shares some items with a preceding form and some items with a following form, and the last form also shares some items with the first form. Together, all the shared items make up the horizontal linking set. All forms for the grade and subject are connected, but not identical (e.g., A is connected to B, B is connected to C, and C is connected to A). Testing mode linking consisted of common items placed in computer-based forms and paper-based forms within an administration to support the development of scores on the same reporting scale. Across administration linking, or year-to-year linking, consisted of common items included in two different administrations. For Spring 2015, no linking to the Spring 2014 field test was required because 2015 is the first year of operational administration; however linking to Fall 2014 was achieved, as described below. The placement of linking items across forms or administrations supports the development of comparable scores.

Linking item sets can be internal or external linking sets. Internal linking sets consist of common items in operational positions such that the items contribute to the students' scores. External linking sets consist

of common items in positions resulting in the items not contributing to students' scores. The 2014-2015 linking designs included both external and internal linking sets. The horizontal linking across forms within an administration and the testing mode linking included internal linking sets. The across administration linking between the Fall 2014 administration and the Spring 2015 administration included both internal and external linking sets.

For ELA/L, the horizontal linking designs for the Spring 2015 online test forms were based on the number of unique test forms constructed for a grade. After constructing the unique test forms, the test forms were divided into sections and sections were dispersed across additional forms such that each section appeared on two forms. As a result, the operational linking sets represented full test blueprints. This means that linking items were selected to reflect the content balance, task models, types of items, and cognitive complexity of the full PARCC assessment.

For mathematics, the ATA pulled an initial blueprint linking set that was divided into item sets and distributed across the Spring 2015 online forms following a daisy-chain strategy, as depicted below in Figure 2.2.

The paper forms for both subjects were generated from the online forms. In response to several practical constraints based on the number of forms constructed for each mode and to meet the blueprints (e.g., inclusion of technology enhanced items in CBT forms), there was no one online form that was administered intact in the paper delivery mode at any grade level. For example, technology enhanced items from online forms were replaced in the paper forms with items from similar content, but appropriate for paper-based testing. However, for both subjects, the content on paper forms significantly overlaps with that on the online forms.

2.2.5 Graphical Representation of PARCC Operational Test Linking Design

This section includes two examples of graphical representations to illustrate the horizontal linking designs described above. Designs for mode comparability linking and across administrations linking are not included in these graphs.

Figure 2.1 illustrates the linking design for ELA/L CBT for grades 6-8, and 11. The numbers in the cells represent the unique texts. For each grade level, three unique CBT forms (identified as the shaded Forms 1, 3, and 5 in Figure 2.1) were developed, each including four texts/sections. The three unique test blueprints were distributed across three overlap forms (identified as Forms 2, 4, and 6 in Figure 2.1). For example, Form 3 consists of Text/Sections 5, 6, 7, 8. These sections are also on Form 4 (Text/Sections 5 and 7) and Form 2 (Text/Sections 6 and 8). For ELA/L the linking design resulted in all sections of the tests serving as linking sets. The linking design for grades 3-5 and 9-10 are similar with some exceptions. Grades 3-5 do not have as many sections and grades 9-10 have additional forms.

Sections	Form 1	Form 2	Form 3	Form 4	Form 5	Form 6
Text 1:	1	1	5	5	9	9
Text 2:	2	6	6	10	10	2
Text 3:	3	3	7	7	11	11
Text 4:	4	8	8	12	12	4

Figure 2.1 PARCC ELA/L CBT Linking Design (Grades 6-8, and 11)

Figure 2.2 illustrates the linking design for mathematics for grades 3-8, and Algebra II. The linking test blueprints were distributed across six forms by sections. The colors in the chart below represent the sets of shared common items among forms. Each section was administered on two adjacent forms, as shown by the colored blocks representing the common item sets. For example, a common item set is shared in section 1 of Forms 1 and 2, section 2 of forms 1 and 6, section 3 of forms 5 and 6, and so on. Other items, possibly unique to one form, were then selected to complete each of the six forms. For mathematics the linking design resulted in approximately one third of the items within a test form serving as linking sets divided across two adjacent forms creating the daisy-chain design. The linking design for Algebra I, Geometry, and Integrated Mathematics I, II, and III are similar with some exceptions. Mathematics I, II, and III and have fewer forms and Algebra I and Geometry have additional forms.

Sections	Form 1	Form 2	Form 3	Form 4	Form 5	Form 6
Section 1:						
Section 2:						
Section 3:						
Section 4:						
Section 5:						
Section 6:						

Figure 2.2 PARCC Mathematics CBT Linking Design (Grades 3-8, and Algebra II)

2.2.6 Field Test Data Collection Overview

Field test items were embedded in the 2015 spring operational forms to collect data for psychometric analysis necessary to support the assessment system for future administrations. Both PBA and EOY field test administrations entailed paper and computer administration modes, with computer administration as the dominant mode. The ELA/L PBA did not include embedded field test items.

The initial data collection design entailed two conditions. Condition one, which comprised the mathematics PBA and EOY assessments and the ELA/LEOY assessment, was an embedded census field test model in which all students taking the summative assessment participated in the field test. Field test sets were constructed to balance the expected cognitive load and difficulty across forms, reflected in the number of points, distribution of task types, and balance of passages for ELA/L. Forms for each content area were spiraled at the student level.

Under Condition 2, which comprised the ELA/literature PBA assessment, a voluntary task tryout was conducted. District classrooms were recruited by PARCC to participate in a task tryout of the ELA/L PBA assessments, taking a mini set of items composed of a narrative, research, or literature assessment task paired with one literary or informational text and its associated set of questions. Participation under Condition 2 was unsatisfactory to establish satisfactorily reliable field test statistics, but did provide some performance information that was used by the PARCC ELA/L OWG to evaluate the ways in which students might respond to the performance tasks.

Section 3: Test Administration

3.1 Testing Windows

The 2014-2015 operational administration of PARCC assessments included a Fall/Winter Block administration in the fall of 2014 (Fall 2014) as well as a spring administration in the spring of 2015 (Spring 2015).⁴ Both of these administrations included two separate testing windows: the Performance-Based Assessment (PBA) and the End-of-Year (EOY) assessment. A student must have participated in both the PBA and EOY testing windows to receive a complete score and individual score report.

For both Fall 2014 and Spring 2015, the PBA testing window occurred after approximately 75 percent of instructional time was completed. The ELA/L PBA focused on writing effectively when analyzing text while the mathematics PBA focused on applying skills and concepts, and on understanding multistep problems that require abstract reasoning, precision, perseverance, and strategic use of tools. The EOY testing window occurred after approximately 90 percent of instructional time was completed. For both the ELA/L and mathematics EOY assessments, students demonstrated their acquired skills and knowledge by answering selected-response questions.

Table 3.1 PARCC Fall 2014 and Spring 2015 Testing Windows

Administration	Performance-Based Assessment (PBA)	End-of-Year (EOY) Assessment
Fall/Winter Block 2014	December 1, 2014 – December 19, 2014	December 15, 2014 – January 16, 2015
Spring 2015	February 16, 2015 – May 8, 2015	April 13, 2015 – June 5, 2015

Each PARCC assessment comprised multiple units, and with certain mathematics units there were separate sections to moderate calculator usage during testing; that is, within the same unit, students were allowed to use a calculator in some sections, but not in others. Figure 3.1 shows configuration of the units within each of the components (PBA and EOY) of the PARCC mathematics and ELA/L assessments for the 2014-2015 operational administration.

⁴ The 2014 Fall administration was in paper-based format only, while the 2015 Spring administration was administered in both paper- and computer-based formats.

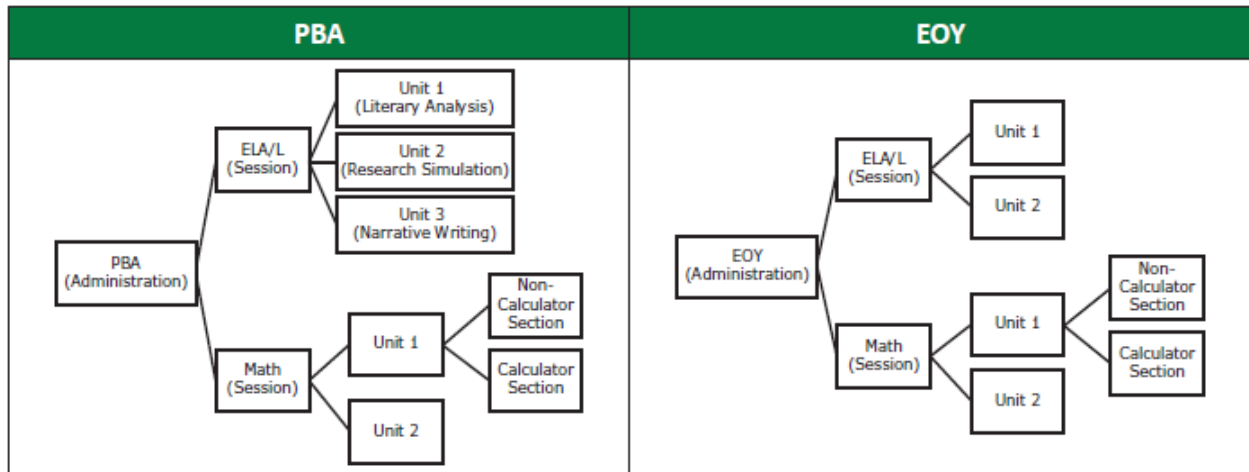


Figure 3.1 Configuration of Units within the 2014-2015 PARCC Assessments (Sample for Grade 7 and High School)

3.2 Test Security and Administration Policies

The administration of any PARCC assessment is a secure testing event. Maintaining the security of test materials before, during, and after the test administration is crucial to obtaining valid and reliable results. School Test Coordinators are responsible for ensuring that all personnel with authorized access to secure materials are trained in—and subsequently act in accordance with—all security requirements.

School Test Coordinators must implement chain-of-custody requirements for specified materials. School Test Coordinators are responsible for the distribution and collection of materials to and from Test Administrators, returning test materials to a secure location after testing each day, and for destroying specified materials once all testing has been completed.

The administration of PARCC assessment includes both secure and non-secure materials, and these materials are further delineated by whether they are “scorable” or “nonscorable,” depending on whether the assessments were administered via paper/pencil (e.g., paper-based assessments) or online (e.g., computer-based assessments). For the 2014-2015 paper-based administration, students responded directly in the paper-based test booklet (i.e., consumable test booklets) at all grade levels.

Secure vs. Non-Secure Materials

PARCC defines secure materials as those that must be closely monitored and tracked to prevent unauthorized access to or prohibited use or distribution of secure content such as test items, reading passages, student work, etc. For paper-based tests, secure materials include both used and unused test booklets and used scratch paper while for computer-based tests, secure materials include student testing tickets, seal codes for accessing sections of the online tests, and used scratch paper. PARCC defines non-secure materials as any authorized testing materials that do not include secure content

(e.g., test items or student work). These include test administration manuals, unused scratch paper, mathematics reference sheets that have not been written upon, etc.

Scorable vs. Nonscorable Materials

Paper-based assessments have both scorable and nonscorable materials while computer-based assessments only have nonscorable materials. Scorable materials for paper-based assessments comprise used (e.g., includes student work) consumable test booklets only. Scorable materials must be returned to the vendor to be scored. All other materials for paper-based testing, such as blank (e.g., unused) test booklets, test administration manuals, scratch paper, mathematics reference sheets, etc., are deemed non-scorable. For computer-based tests, there are no scorable materials as student work is submitted electronically for scoring; thus there are no physical materials to return.

Students taking the paper-based test may not have access to scorable or nonscorable secure test content before or after testing. Scorable secure materials that are to be provided by Test Administrators to students include Test Booklets. Nonscorable secure materials that are distributed by Test Administrators to paper-based testing students include Large Print Test Booklets, braille Test Booklets, scratch paper (paper used by students to take notes and work through items), and printed Mathematics Reference Sheets (grades 5–8 and high school).

Students taking the computer-based test may not have access to secure test materials before testing, including printed Student Testing Tickets and printed Seal Codes. Printed Mathematics Reference Sheets (if applicable) and scratch paper must be new and unmarked.

School Test Coordinators are required to maintain a tracking log to account for collection and destruction of test materials, including mathematics reference sheets written on by students and scratch paper written on by students. As part of the test administration policy, schools are required to maintain the Chain-of-Custody Form or tracking log of secure materials for at least three years unless otherwise directed by state policy. Copies of the Chain-of-Custody Form for paper-based testing are included in each Local Education Agency (LEA) or school's test materials shipment.

Test Administrators are not to have extended access to test materials before or after administration. Test Administrators must document the receipt and return of all secure test materials (used and unused) to the School Test Coordinator immediately after testing.

All PARCC test security and administration policies are found in the *PARCC Test Coordinator Manual* and the *PARCC Test Administrator Manuals*. Archived versions of test administration manuals from past administration years can be found at: <http://parconline.org/assessments/administration/archived-testing-manuals>. State security and administration policies may exceed that of the PARCC policies. State-specific policies are included in **Appendix C** of the *Test Coordinator Manual*.

3.3 Accessibility Features and Accommodations

3.3.1 Participation Guidelines for PARCC Assessments

All students, including students with disabilities and English learners, are required to participate in statewide assessments and have their assessment results be part of the state's accountability systems, with narrow exceptions for English learners in their first year in a U.S. school, and certain students with disabilities who have been identified by the Individualized Education Program (IEP) team to take their state's alternate assessment. All eligible students will participate in the PARCC ELA/L and mathematics assessments. Federal laws governing student participation in statewide assessments include the No Child Left Behind Act of 2001 (NCLB), the Individuals with Disabilities Education Improvement Act of 2004 (IDEA), Section 504 of the Rehabilitation Act of 1973 (reauthorized in 2008), and the Elementary and Secondary Education Act (ESEA) of 1965, as amended. All students can receive accessibility features on PARCC assessments.

Four distinct groups of students may receive accommodations on PARCC assessments:

1. **Students with disabilities who have an Individualized Education Program (IEP);**
2. **Students with a Section 504 plan** who have a physical or mental impairment that substantially limits one or more major life activities, have a record of such an impairment, or are regarded as having such an impairment, but who do not qualify for special education services;
3. **Students who are English learners;** and
4. **Students who are English learners with disabilities who have an IEP or 504 plan.** These students are eligible for accommodations intended for both students with disabilities and English learners.

Testing accommodations for students with disabilities or students who are English learners (EL) must be documented according to the guidelines and requirements outlined in the PARCC Accessibility Features and Accommodations Manual, Archived versions of past editions of the Accessibility Features and Accommodations Manual can be found at:

<http://parconline.org/assessments/administration/archived-testing-manuals>.

3.3.2 PARCC Accessibility System

Through a combination of universal design principles and accessibility features, PARCC has designed an inclusive assessment system by considering accessibility from initial design through item development, field testing, and implementation of the assessments for all students, including students with disabilities, English learners, and English learners with disabilities. Accommodations may still be needed for some students with disabilities and English learners to assist in demonstrating what they know and can do. However, the accessibility features available to students should minimize the need for

accommodations during testing and ensure the inclusive, accessible, and fair testing of the diverse students being assessed.

3.3.3 What are Accessibility Features?

On the PARCC computer-based assessments, accessibility features are tools or preferences that are either built into the assessment system or provided externally by Test Administrators, and may be used by any student taking the PARCC assessments (i.e., students with and without disabilities, gifted students, English learners, and English learners with disabilities). Since accessibility features are intended for all students, they are not classified as accommodations. Students should have the opportunity to select and practice using them prior to testing to determine which are appropriate for use on the PARCC assessment. Consideration should be given to the supports a student finds helpful and consistently uses during instruction. Practice tests that include accessibility features are available for teacher and student use throughout the year. Practice tests are available at parcc.pearson.com.

3.3.4 Accommodations for Students with Disabilities and English Learners

It is important to ensure that performance in the classroom and on assessments is influenced minimally, if at all, by a student's disability or linguistic/cultural characteristics that may be unrelated to the content being assessed. For PARCC assessments, accommodations are considered to be adjustments to the testing conditions, test format, or test administration that provide equitable access during assessments for students with disabilities and students who are English learners. In general, the administration of the assessment should not be the first occasion on which an accommodation is introduced to the student. To the extent possible, accommodations should:

- provide equitable access during instruction and assessments;
- mitigate the effects of a student's disability;
- not reduce learning or performance expectations;
- not change the construct being assessed; and
- not compromise the integrity or validity of the assessment.

Accommodations are intended to reduce and/or eliminate the effects of a student's disability and/or English language proficiency level; however, **accommodations should never reduce learning expectations by reducing the scope, complexity, or rigor of an assessment.** Moreover, accommodations provided to a student on the PARCC assessments must be generally consistent with those provided for classroom instruction and classroom assessments. There are some accommodations that may be used for instruction and for formative assessments that are not allowed for the summative assessment because they impact the validity of the assessment results; for example, allowing a student to use a thesaurus or access the Internet during a PARCC assessment. There may be consequences (e.g., excluding a student's test score) for the use of non-allowable accommodations during PARCC

assessments. It is important for educators to become familiar with PARCC policies regarding accommodations used for assessments.

To the extent possible, accommodations should adhere to the following principles:

- Accommodations enable students to participate more fully and fairly in instruction and assessments and to demonstrate their knowledge and skills.
- Accommodations should be based upon an individual student's needs rather than on the category of a student's disability, level of English language proficiency alone, level of or access to grade-level instruction, amount of time spent in a general classroom, current program setting, or availability of staff.
- Accommodations should be based on a documented need in the instruction/assessment setting and should not be provided for the purpose of giving the student an enhancement that could be viewed as an unfair advantage.
- Accommodations for students with disabilities must be described and documented in the student's appropriate plan (i.e., either a 504 plan or an approved IEP); and must be provided if they are listed.
- Accommodations for English learners should be described and documented.
- Students who are English learners with disabilities are eligible to receive accommodations for both students with disabilities and English learners.
- Accommodations should become part of the student's program of daily instruction as soon as possible after completion and approval of the appropriate plan.
- Accommodations should not be introduced for the first time during the testing of a student.
- Accommodations should be monitored for effectiveness.
- Accommodations used for instruction should also be used, if allowable, on local district assessments and state assessments.

In the following scenarios, the school must follow each state's policies and procedures for notifying the state assessment office:

- A student **was provided a test accommodation that was not listed** in his or her IEP/504 plan/documentation for an English learner, or
- A student **was not provided a test accommodation that was listed** in his or her IEP/504 plan/documentation for an English learner.

3.3.5 Unique Accommodations

PARCC provides a comprehensive list of accessibility features and accommodations in the PARCC Accessibility Features and Accommodations Manual that are designed to increase access to PARCC assessments and that will result in valid, comparable assessment scores. However, students with disabilities or English learners may require additional accommodations that are not already listed. PARCC states individually review requests for unique accommodations in their respective states and provide a determination as to whether the accommodation would result in a valid score for the student, and if so, would approve the request.

3.3.6 Emergency Accommodations

An emergency accommodation may be appropriate for a student who incurs a temporary disabling condition that interferes with test performance shortly before or during the PARCC assessment window. A student, whether or not they already have an IEP or 504 plan, may require an accommodation as a result of a recently-occurring accident or illness. Cases include students who have a recently-fractured limb (e.g., arm, wrist, or shoulder); whose only pair of eyeglasses has broken; or a student returning to school after a serious or prolonged illness or injury. An emergency accommodation should be given only if the accommodation will result in a valid score for the student (i.e., does not change the construct being measured by the test[s]). If the principal (or designee) determines that a student requires an emergency accommodation on the PARCC assessment, an Emergency Accommodation Form must be completed and maintained in the student's assessment file. If required by a PARCC state, the school may need to consult with the state or district assessment office for approval. **The parent must be notified that an emergency accommodation was provided.** If appropriate, the Emergency Accommodation Form may also be submitted to the district assessment coordinator to be retained in the student's central office file. Requests for emergency accommodations will be approved after it is determined that use of the accommodation would result in a valid score for the student.

3.3.7 Student Refusal Form

If a student refuses an accommodation listed in his or her IEP, 504 plan, or if required by the PARCC member state, an English learner plan, the school should document in writing that the student refused the accommodation, and the accommodation must be offered and remain available to the student during testing. This form must be completed and placed in the student's file and a copy must be sent to the parent on the day of refusal. Principals (or designee) should work with Test Administrators to determine who, if any others, should be informed when a student refuses an accommodation documented in an IEP, 504 plan, or (if required by the PARCC member state) English learner plan.

3.4 Testing Irregularities and Security Breaches

Any action that compromises test security or score validity is prohibited. These may be classified as testing irregularities or security breaches. Below are examples of activities that compromise test security or score validity (note that these lists are not exhaustive). It is highly recommended that School

Test Coordinators discuss other possible testing irregularities and security breaches with Test Administrators during training.

Examples of test security breaches and irregularities include but are not limited to:

- **Electronic Devices**
 - Using a cell phone or other prohibited handheld electronic device (e.g., smartphone, iPod, smart watch, personal scanner) while secure test materials are still distributed, while students are testing, after a student turns in his or her test materials, or during a break.
 - Exception: Test Coordinators, Technology Coordinators, Test Administrators, and Proctors are permitted to use cell phones in the testing environment only in cases of emergencies or when timely administration assistance is needed. LEAs may set additional restrictions on allowable devices as needed.

- **Test Supervision**
 - Coaching students during testing, including giving students verbal or nonverbal cues, hints, suggestions, or paraphrasing or defining any part of the test.
 - Engaging in activities (e.g., grading papers, reading a book, newspaper, or magazine) that prevent proper student supervision at all times while secure test materials are still distributed or while students are testing.
 - Leaving students unattended for any period of time while secure test materials are still distributed or while students are testing.
 - Deviating from testing time procedures.
 - Allowing cheating of any kind.
 - Providing unauthorized persons with access to secure materials.
 - Unlocking a test in PearsonAccess^{next} during non-testing times.
 - Failing to provide a student with a documented accommodation or providing a student with an accommodation that is not documented and therefore is not appropriate.
 - Allowing students to test before or after the state's test administration window.

- **Test Materials**
 - Losing a student test booklet or answer document.
 - Losing a student testing ticket.
 - Leaving test materials unattended or failing to keep test materials secure at all times.
 - Reading or viewing the passages or test items before, during, or after testing.
 - Exception: Administration of a Human Reader/Signer accessibility feature for mathematics or accommodation for English language arts/literacy which requires a Test Administrator to access passages or test items.

- Copying or reproducing (e.g., taking a picture of) any part of the passages or test items or any secure test materials or online test forms.
 - Revealing or discussing passages or test items with anyone, including students and school staff, through verbal exchange, email, social media, or any other form of communication.
 - Removing secure test materials from the school's campus or removing them from locked storage for any purpose other than administering the test.
- **Testing Environment**
 - Allowing unauthorized visitors in the testing environment.
 - Failing to follow administration directions exactly as specified in the *Test Administrator Manual*.
 - Displaying testing aids in the testing environment (e.g., a bulletin board containing relevant instructional materials) during testing.

All instances of security breaches and testing irregularities must be reported to the School Test Coordinator immediately. The Form to Report a Testing Irregularity or Security Breach must be completed within two school days of the incident.

If any situation occurs that could cause any part of the test administration to be compromised, schools refer to the *PARCC Test Coordinator Manual* for each state's policy and immediately follow those steps.

Occasionally, individuals will contact state department of education office with allegations of testing irregularities or security breaches. State procedures in these instances may vary. In these cases, the state's designee may contact the School Test Coordinator or LEA Test Coordinator and ask that individual to investigate the allegation and report back to the appropriate state level organization.

Instructions for the School Test Coordinator or LEA Test Coordinator to report a testing irregularity or security breach (the Form to Report a Testing Irregularity or Security Breach is available in **Appendix D** of the *PARCC Test Coordinator Manual*).

3.5 Data Forensics Analyses

Maintaining the validity of test scores is essential in any high-stakes assessment program, and misconduct represents a serious threat to test score validity. When used appropriately, data forensic analyses can serve as an integral component of a wider test security protocol. The results of these data forensic analyses may be instrumental in identifying potential cases of misconduct for further follow-up and investigation.

In 2014-2015, PARCC conducted the following four data forensics analyses on its operational assessments:

- Response Change Analysis

- Plagiarism Analysis
- Internet and Social Media Monitoring
- Off Hours Testing Monitoring

An overview of each data forensics analysis method is provided next.

3.5.1 Response Change Analysis

Response change analysis looks at how often student answers are changed, focusing specifically on an excessive number of wrong answers changed to right answers. In traditional paper-based, multiple-choice testing programs, this is sometimes referred to as “erasure analysis⁵.” The rationale for erasure analysis is that a teacher or administrator who is intent on improving classroom performance might be motivated to change student responses after the answer sheets are collected. A clustered number of student answer documents from the same school or classroom with unusually high numbers of answers changed from wrong to right might provide evidence to support follow-up investigation. PARCC’s response change analysis extended the traditional erasure method to account for issues specific to computer-based testing as well as the variety of items types on the PARCC assessments, such as partial-credit, multi-part, and multiple-select items.

3.5.2 Plagiarism Analysis

Plagiarism analysis compares the responses given for a group of written composition items, looking for high degrees of similarity. For the PARCC assessments, the primary item type of interest was the prose constructed response (PCR) tasks in the English Language Arts and Literacy (ELA/L) content area. This analysis was conducted for PCR tasks administered online using some of the same artificial intelligence (AI) techniques that are applied in automated essay scoring. Specifically, this method was based on Latent Semantic Analysis (LSA) technology to detect possible plagiarism. Using LSA, the content of each constructed response was compared against the content of every other constructed response and a measure that indicated the degrees of similarity was generated for each pair of response comparison. Because LSA provided a semantic representation of language, rather than a syntactic or word-based representation, it allowed the detection of potential copying behaviors, even when test takers or administrators substituted synonymous words or phrases.

3.5.3 Internet and Social Media Monitoring

Internet and social media monitoring was being carried out by Caveon, LLC. Caveon’s team monitored English language websites and searchable forums that were publicly available for suspected proxy testing solicitations and website postings that contain, or appear to contain, infringements of PARCC’s

⁵ The term “erasure analysis” is sometimes objected to because it is inferential rather than descriptive. A more descriptive term is “mark discrimination analysis” which recognizes that the scanning approach makes discriminations among the darkness of selected answer choices when multiple responses to a multiple-choice item are detected during answer sheet processing.

protected operational test content. The Internet and social media outlets monitored included popular websites (such as Facebook and Twitter), blogs, discussion forums, video archives, document archives, brain dumps, auction sites, media outlets, peer-to-peer servers, etc. Caveon's process generated regular updates that categorize identified threats by level of actual or potential risk based upon the representations made on the web sites, or actual analysis of the proffered content. For example, categorizations typically ranged from "cleared" (lowest risk but bookmarked for continued monitoring) to "severe" (highest risk). Note that this process only considered potential breaches of secure item content, not violations of testing administration policies. Summary reports describing the threats were provided to PARCC through notification emails.

3.5.4 Off-Hours Testing Monitoring

Off-hours testing monitoring checks for suspicious testing activities at test administration locations occurring outside of the set windows for computer-based testing sessions. PARCC states established set start and end time for administering computer-based assessments. Based on these hours, authorized users (that is, users with the State Role) were allowed to override the start and end times for a test session. The off-hours testing monitoring process tracked such occurrences and logged them in an operational report, which listed the sessions within an organization that selected to test outside the set window. PARCC states could use this report to follow-up with the organizations identified in the report.

3.6 Quality of Test Administration Studies

As part of the planned research agenda for the 2014-2015 PARCC administration, HumRRO conducted a set of studies to investigate the quality of the administration of the PARCC assessments during the first operational administration. The purpose of the studies was to identify potential threats to the validity of PARCC test administration so that they could be addressed and reconciled for future test administrations.

The research studies were an extension of the Quality of Test Administration study from the spring 2014 PARCC field test. The guiding framework for both Quality of Test Administration studies comes from the Theory of Action (TOA) in the PARCC validity studies memorandum (Thacker, Sinclair, Wise, & Becker, 2014). The claims from the Test Administration phase of the TOA are listed in Figure 3.2. These were the claims investigated in the 2014-2015 studies.

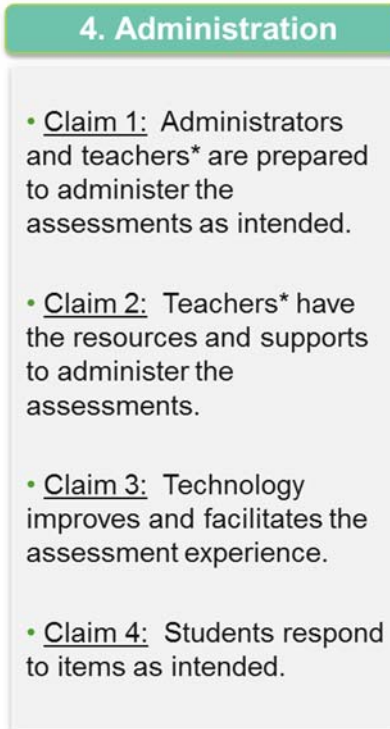


Figure 3.2 Claims from the Administration Phase of the PARCC Theory of Action.

Overall, the 2014-2015 studies found notable improvements in the validity evidence collected from the 2014 field test. In particular, the findings indicate stronger support for all four of the test administration claims in the Theory of Action.

The primary threats to the veracity of these claims, as identified through studies were:

- Continued problems of a local nature with logging students on and with students getting kicked-off the system
- The amount of time and staff required for administering PARCC assessments was overly burdensome on schools' resources
- Many students were having difficulty understanding the directions on the test for answering test questions

These concerns should be addressed for future operational assessments, as threats to the veracity of the test administration claims may undermine the validity of subsequent goals in the PARCC theory of action (i.e., valid scoring, reporting, and use of test scores). Several recommendations for addressing these threats were provided in the study report, which may be requested from Parcc Inc.

Section 4: Item Scoring

4.1 Machine Scored Items

4.1.1 Key Based Items

Pearson performed a key review prior to the test administration to verify that the scoring (answer) keys were correct for each form. Once the forms were constructed and approved by PARCC for publication, an independent key review was performed by experienced and trained content staff. The content staff reviewed each item and confirmed that the key was correct. If discrepancies were identified, a senior content specialist or content manager reviewed the flagged item(s) and worked with the item developers to resolve the issue.

4.1.2 Rule Based Items

Rule based scoring refers to item types that use various scoring models. PARCC uses QTI (Question and Test Interoperability) item type implementation based on scoring model rules. Examples of these item types include “choice interaction” which presents a set of choices where one or more choices can be selected; text entry, where the response is entered in a text box; hot spot or text interaction, where an area in a graph or text in a paragraph (for example) can be highlighted or match interaction, where an association can be made between pairs of choices in a set. These items include the scoring rules and correct responses as part of their item XML (markup language) coding.

During the initial stages of item development for PARCC, Pearson staff worked closely with PARCC to first delineate the rules for the scoring rubrics and then to adjust those rules based on student responses. During the Item Tryout⁶ planning phase, Pearson content staff received input from PARCC staff to develop a thorough rule based scoring process that met PARCC needs.

Pearson worked with the item developers to review initial scoring rules created during the item development. Once the rule based scoring process was approved by PARCC, and prior to test construction, Pearson content staff worked closely with the item developers to finalize scoring rubrics for items to be scored via the rule based scoring method. The proposed scoring rubrics were sent to PARCC for review, and if any additional changes were needed or new rules added, Pearson documented and applied the requested edits.

During test construction, Pearson monitored and evaluated the scoring and updated the scoring keys/ scoring rules in the item bank. After the tryout items were scored, Pearson prepared a frequency distribution of student responses for each item or task scored using a rule based approach and compared this to the expected response based on correct answers to ensure that scoring keys and rules were appropriately applied. The content team does this by analyzing the student response data to determine if scoring is acceptable using the item metadata and the student response file in conjunction with any potential item issues as flagged by psychometrics. These frequency distributions included an indication of right/wrong and other identifying information defined by PARCC and those items that

⁶ The item tryout was a set of item studies conducted in spring 2015.

showed a statistical anomaly, whereby the frequency distribution was outside of the expected range, were sent to content experts to verify that the items were coded with the correct key.

Following the Rule Based Scoring Educator Committee's review, which occurred prior to year 1 test construction, Pearson analyzed the feedback from the committees and made recommendations about adjustments to the scoring rubrics, based on the results of the reviews. Upon submission of the results, Pearson worked with PARCC staff to discuss these findings and determine next steps prior to the completion of scoring. In subsequent years as scoring inquiries arise throughout the process of test construction, forms creation, testing, scoring, and psychometric analysis, items with scoring discrepancies are brought before the PARCC Task Force for resolution. This committee consists of representatives from each state as well as PARCC and Pearson content specialists.

Following the initial development of the PARCC rule based scoring rubrics, Pearson has continued to monitor and evaluate new item development to ensure the scoring rules established are maintained within all item types as approved.

Pearson continues to use several avenues to monitor scoring each year. Prior to testing a third party key review by which reviewers check operational and field test items for correct keys. Any disputed items go to a 2nd review with Pearson content experts and anything still in question is taken before the PARCC task force for review and possible key change. During testing, Pearson creates early testing files for frequency distribution analysis whereby items where an incorrect key receives a high distribution of responses are further evaluated for accuracy. After testing during psychometric analysis all responses are again evaluated for distribution of responses and potential scoring abnormalities. Any change in scoring that may be requested as a result of the psychometric analysis is also taken before the PARCC task force for decisions. These processes are the same for both paper and online modes of testing.

4.2 Human or Handscored Items

PARCC 2015 constructed-response items were scored by human scorers in a process referred to as handscoring. Online training units are used to train all scorers. The online training units include prompts, passages, rubrics, anchors, practice responses, and qualification sets. Scorers who successfully complete the training and qualify, demonstrating they can correctly score student responses based on the guidelines in the online training units, are permitted to score student responses using the ePEN2 (electronic Performance Evaluation Network, second generation) scoring platform. The ePEN2 system is used to score online and paper responses. The first score assigned to a response is the score reported for that response. A response may have a second, third, or fourth score as part of quality monitoring. Scorer quality is monitored throughout scoring.

Pearson staff roles and responsibilities were as follows:

- Scorers are individuals who applied scores to student responses.
- Scoring Supervisors monitored the work of a team of 12 to 24 scorers through review of scorer statistics and backreading, which is a review of responses scored by each scorer. When

backreading, a supervisor sees the scores applied by scorers, which helps the supervisor provide additional coaching or instruction to the scorer being backread.

- Scoring Directors managed the scoring quality of a subset of items and monitored the work of scoring supervisors and scorers for their assigned items. Scoring directors backread responses scored by supervisors and scorers as part of their quality monitoring duties.
- Content Specialists managed the scoring quality of three mathematics grades/courses or a task type across all ELA/Literacy grades. Content specialists monitored the work of the scoring directors.
- Project Managers documented the procedures, identified risks, and managed day-to-day administrative matters.
- A Program Manager provides oversight for the entire scoring process.

All Pearson employees involved in the scoring or the supervision of scoring possessed a four-year college degree.

4.2.1 Scorer Training

Key steps in the development of scorer training materials were Rangefinding and Rangefinder Review meetings where educators and administrators from PARCC states met to interpret the rubrics and consensus score student responses. Rangefinding meetings are held prior to scoring field test items, and Rangefinder Review meetings are held prior to scoring operational items.

At Rangefinding meetings, educators and administrators from PARCC states review student responses and use scoring rubrics to determine consensus scores. Those responses scored in rangefinding are used to create field test scorer training sets. After items are selected for operational testing, PARCC educators and administrators attend rangefinder review meetings to review and approve proposed operational training sets.

When developing scorer training materials, Pearson scoring directors carefully reviewed detailed notes and records from PARCC rangefinding and rangefinder review committee meetings. Training sets were completed using the responses scored by the committees and additional suitable student response samples (as needed). PARCC reviewers reviewed and approved all scorer training sets prior to use for scorer training.

During training, scorers review training sets of scored student responses with annotations that explain the rationale for the score assigned. The anchor set is the primary reference for scorers as they internalize the rubric during training. Each anchor set consists of responses that are clear examples of student performance at each score point. The responses selected were representative of typical approaches to the task and arranged to reflect a continuum of performance. All scorers had access to the anchor set whenever they were training and scoring and were directed to refer to it regularly during

scoring. Scorers were trained on the anchor sets during initial training and used them as a reference throughout scoring.

Practice sets were used in training to help trainees practice applying the scoring guidelines. Scorers reviewed the anchor sets, scored the practice sets, and then were able to compare their assigned scores for the practice sets to the actual PARCC assigned scores to help them learn.

Qualification sets are used to confirm that scorers understand how to accurately score student responses. Qualification sets are composed of responses that are clear examples of score points. Scorers must meet specified agreement percentages on qualification sets in order to score student responses.

Pearson developed two types of training sets to train scorers; prototype and abbreviated sets. "Prototype" training sets are complete training sets consisting of anchor, practice and qualification sets (refer to 4.2.2 for information on qualification process). In English language arts/literacy (ELA/L) there was one prototype training set per task type (Research Simulation Task, Literary Analysis Task, and Narrative Writing Task) at each of the nine grade levels (grades 3-11). In mathematics, a prototype training set was built for a grouping of items under one or more evidence statements for a total of approximately six to ten prototype sets per grade level or course.

The prototype training approach promoted consistency in scoring, as each subsequent abbreviated training set for the ELA/L task type or mathematics item grouping was based on the prototype. Once a prototype was chosen, full training materials were developed for that item, and at each grade level, scorers were trained to score a particular task type using the prototype training materials for that task type.

Abbreviated training sets were prepared for all items not selected for prototype training sets. The abbreviated training sets included an anchor set and two practice scoring sets so scorers could internalize the scoring standards for these new items, which were similar to prototype items they had previously scored.

Anchor and practice sets for both prototype and abbreviated items included annotations for each response. Annotations are formal written explanations of the score for each student response.

The core elements of the training materials used to train and qualify scorers are the anchors, practice sets, and qualification sets. Items are designated as prototype or abbreviated. Training materials for prototype items are used to train scorers to score student responses to the prototype items and to qualify scorers to score designated item types. Training materials for abbreviated items are used to train scorers on the additional items they will score following qualification.

The table below details the composition of the anchors, practice sets, and qualification sets.

Table 4.1 Training Materials Used During Scoring

Training Set Development	
Description	Specification
Anchor Set	
<p>The anchor set is the primary reference for scorers as they internalize the rubric during training. All scorers have access to the anchor set whenever they are training and scoring, and are directed to refer to it regularly.</p> <p>The anchor set comprises clear examples of student performance at each score point. The responses selected may be representative of typical approaches to the task or arranged to reflect a continuum of performance.</p>	<p>The anchor set for mathematics prototype items comprises 3 annotated responses per score point.</p> <p>The anchor set for subsequent abbreviated items for mathematics comprise 1-3 annotated responses per score point.</p> <p>The anchor sets for ELA/L prototype items comprise 3 annotated responses per score point. Anchor sets for prototype items include separate complete anchor sets for each applicable scoring trait (Reading Comprehension, Written Expression, and Conventions).</p>
Practice Sets	
<p>Practice sets are used to help trainees develop experience in independently applying the scoring guide (the rubric) to student responses. Some of these responses clearly reinforce the scoring guidelines presented in the anchor set. Other responses are selected because they are more difficult to evaluate, fall near the boundary between two score categories, or represent unusual approaches to the task.</p> <p>The practice sets provide guidance and practice for trainees in defining the line between score categories, as well as applying the scoring criteria to a wider range of types of responses</p>	<p>The practice sets for mathematics prototype and abbreviated items include two sets of ten annotated responses.</p> <p>ELA/L practice sets for prototype items include three sets of ten annotated responses.</p> <p>The subsequent ELA/L practice sets for abbreviated items include two sets of ten annotated responses.</p>
Qualification Sets	
	<p>The qualification sets for mathematics prototype items include 3 sets of 10 responses each (not annotated).</p> <p>The subsequent mathematics abbreviated items for mathematics do not include qualification sets.</p>

Table 4.1 Training Materials Used During Scoring

Training Set Development	
Description	Specification
<p>Qualification sets are used to confirm that scorer trainees understand the scoring criteria and are able to assign scores to student responses accurately. The responses in these sets are selected to reinforce the application of the scoring criteria illustrated in the anchor set.</p> <p>Scorer trainees must demonstrate acceptable performance on these sets by meeting a pre-determined standard for accuracy in order to qualify to score. Pearson scoring staff define and document qualifying standards in conjunction with PARCC prior to scoring. The qualification sets for mathematics prototype items include 3 sets of 10 responses each (not annotated).</p> <p>The subsequent mathematics abbreviated items for mathematics do not include qualification sets.</p> <p>The qualification sets for ELA/L prototype items include 3 sets of 10 responses each (not annotated).</p> <p>The subsequent ELA/L abbreviated items do not include qualification sets.</p>	<p>The qualification sets for ELA/L prototype items include 3 sets of 10 responses each (not annotated).</p> <p>The subsequent ELA/L abbreviated items do not include qualification sets.</p>

4.2.2 Scorer Qualification

In order to score items, scorers are required to show that they are able to accurately apply PARCC scoring methodology through a qualification process. Scorers are asked to apply scores to three qualification sets consisting of 10 responses each. ELA/L scorers apply a score for each trait to each response in the qualification sets. Literary Analysis and Research Simulation Tasks each had three traits, Reading Comprehension, Written Expression, and Conventions. The Narrative Writing Task had two traits, Written Expression and Conventions. Mathematics scorers apply a score for each part that is a constructed response. The number of constructed-response parts for each mathematics prompt ranged from one to four. Scorers must match the PARCC approved score at a percentage agreed to by PARCC.

For ELA/L qualification, scorers were required to meet the following three conditions:

1. On at least one of the three qualifying sets, at least 70% of the ratings on each of the three scoring traits (considered separately), must agree exactly with the PARCC-approved scores.
2. On at least two of the three qualifying sets, at least 70% of the ratings (combined across the three scoring traits) must agree exactly with the PARCC-approved scores.
3. Combining over the three qualifying sets and across the three scoring traits, at least 96% of the ratings must be within one point of the PARCC-approved scores.

For mathematics qualification, the requirements were based on the item types and score point ranges. Because mathematics items can have one or more scoring traits, a scorer needed to achieve the following requirements separately for each scoring trait (when applicable to the item):

Table 4.2 Mathematics Qualification Requirements

Category	Score Point Range	Perfect Agreement	Within One Point
2-category	0-1	90%	100%
3-category	0-2	80%	96%
4-category	0-3	70%	96%
5-category	0-4	70%	96%
7-category	0-6	70%	96%

On at least two of the three qualifying sets, the perfect agreement ratings indicated above for each category must agree exactly with the PARCC-approved scores. Over the three qualifying sets, each scoring trait must have at least 96% of the ratings within one point of the PARCC-approved scores. The average is exclusive to each trait so an item with multiple scoring traits would have multiple trait rating averages within one point of the PARCC-approved score.

4.2.3 Managing Scoring

Pearson created a Handscoring Specifications document that detailed the handscoring schedule, customer requirements, range finding plans, quality management plans, item information, and staffing plans for each scoring administration. This document is available as the appendix to Section 4.

4.2.4 Monitoring Scoring

Second Scoring. During scoring, Pearson’s ePEN2 scoring system automatically and randomly distributed a minimum of 10 percent of student responses for second scoring; scorers had no indication whether a response had been scored previously. Humans applied the second score for all mathematics

items. Second scoring for ELA/L was performed either by human scorers or by the Intelligent Essay Assessor. If the first and second scores applied were non-adjacent, a third and occasionally a fourth score was assigned to resolve scorer disagreements. When a resolution score (i.e. 3rd score) was nonadjacent to one and/or both of the first and second scores, the content specialist or scoring director would apply an adjudication score (4th score). The first score is the score to be reported, while the second, resolution, and adjudication scores are used to monitor scorer performance only. The first score is the reported score since approximately 90% of the responses receive only one score applied by a single scorer.

Backreading. Backreading is one of the major responsibilities of Pearson scoring supervisors and a primary tool for proactively guarding against scorer drift where scorers score responses in comparison to one another instead of in comparison to the training responses. Scoring supervisory staff used the ePEN2 backreading tool to review scores assigned to individual student responses by any given scorer to confirm that the scores were correctly assigned and to give feedback and remediation to individual scorers. Pearson backread approximately five percent of the handscored responses. Backreading scores did not override the original score but were used to monitor scorer performance.

Validity. Validity responses are pre-scored responses strategically interspersed in the pool of live responses. These responses were not distinguishable from any other responses so that scorers were not aware they were scoring validity responses rather than live responses. The use of validity responses provided an objective measure that helped ensure that scorers were applying the same standards throughout the project. In addition, validity was at times shared with scorers in a process known as “validity as review.” Validity as review provided scorers automated, immediate feedback: a chance to review responses they mis-scored, with reference to the correct score and a brief explanation of that score. One validity response was sent to scorers for every 25 “live” responses scored.

PARCC validity agreement requirements for scorers are listed in Table 4.2. Scorers had to meet the required validity agreement percentages to continue working on the PARCC project. Scorers who did not maintain expected agreement statistics were given a series of interventions culminating in a targeted calibration set: a test of scorer knowledge. Scorers who did not pass targeted calibration were removed from scoring the item and all the scores they assigned were deleted.

Table 4.3 Scoring Validity Agreement Requirements

Subject	Score Point Range	Perfect Agreement	Within One Point
Mathematics	0-1	90%	96%*
Mathematics	0-2	80%	96%
Mathematics	0-3	70%	96%
Mathematics	0-4	65%	95%
Mathematics	0-6	65%	95%
ELA/L	Multi-trait	65%	96%

Note: * A 0 or 1 score compared to a blank score will have a disagreement greater than 1 point.

Calibration Sets. Calibration sets are special sets created during scoring to help train scorers on particular areas of concern or focus. Scoring directors used calibration sets to reinforce rangefinding standards, introduce scoring decisions, or address scoring issues and trends. Calibration was targeted either to correct a scoring issue or trend, or to continue scorer training by introducing a scoring decision. Calibration was administered regularly throughout scoring.

Inter-rater Agreement. Inter-rater agreement is the agreement between the first and second scores assigned to student responses and is the measure of how often scorers agree with each other. Pearson scoring staff used inter-rater agreement statistics as one factor in determining the needs for continuing training and intervention on both individual and group levels. PARCC inter-rater agreement expectations were as follows:

Table 4.4 Inter-rater Agreement Expectations and Results

Subject	Score Point Range	Perfect Agreement Expectation	Perfect Agreement Result	Within One Point Expectation	Within One Point Result
Mathematics	0-1	90%	95%	96%*	100%
Mathematics	0-2	80%	93%	96%	99%
Mathematics	0-3	70%	91%	96%	99%
Mathematics	0-4	65%	88%	95%	98%
Mathematics	0-6	65%	83%	95%	95%
ELA/L	Multi-trait	65%	65%	96%	98%

Note: * A 0 or 1 score compared to a blank score will have a disagreement greater than 1 point.

Pearson's ePEN2 scoring system included comprehensive inter-rater agreement reports that allowed supervisory personnel to monitor both individual and group performance. Based on reviews of these reports, scoring experts targeted individuals for increased backreading and feedback, and if necessary, retraining.

The perfect agreement rate for all mathematics responses scored by two scorers was 92% and the within one point rate was 99%. For all ELA/L responses scored by two scorers, the perfect agreement rate was 65% and the within one point rate was 98%.

Section 5: Test Taker Characteristics

5.1 Overview of Test Taking Population

Approximately five million students participated in the first operational administration of the PARCC assessments during the 2014–2015 school year in [Arkansas](#), [Colorado](#), [District of Columbia](#), [Illinois](#), [Louisiana](#), [Maryland](#), [Massachusetts](#), [Mississippi](#), [New Jersey](#), [New Mexico](#), [Ohio](#), and [Rhode Island](#). Not all participating states had students testing in all grades. Assessments were administered for English Language Arts/Literacy (ELA/L) in grades 3 through 11; mathematics assessments were administered in grades 3 through 8, as well as for traditional high school mathematics (Algebra I, Geometry, and Algebra II) and integrated high school mathematics (Integrated Mathematics I – III). A small subset of students tested on paper in ELA/L grades 9, 10, and 11, and Algebra I, Geometry, and Algebra II during fall of 2014. Test taker characteristics for this group are presented in an addendum.⁷ The majority of students tested during the Spring 2015 window when all grades and content areas were administered online and on paper.

5.2 Composition of Operational Forms

The two operational administrations, Fall 2014 and Spring 2015, of the PARCC assessment each included two separate components: the Performance-Based Assessment (PBA) and the End-of-Year (EOY) assessment. Both components were administered as computer-based tests (CBT) and paper-based tests (PBT). A student had to have valid scores in both the PBA and EOY components to receive a summative score.

The PBA and EOY components measured different types of knowledge and skills. The PBA was administered after approximately 75 percent of instructional time was complete. The PBA component consisted of relatively long questions, many of which required multiple steps. The purpose of this component was to measure critical thinking, reasoning, and the ability to apply skills and knowledge in reading, writing, and mathematics. The ELA/L PBA focused on writing effectively when analyzing text. The mathematics PBA focused on applying skills and concepts, and on understanding multistep problems that require abstract reasoning, precision, and perseverance.

The EOY administration occurred after approximately 90 percent of instruction was complete. Students were required to demonstrate their skills and knowledge by answering innovative selected-response and short-answer questions that measured concepts and skills. In the ELA/L EOY component students demonstrated their understanding of literary and informational passages. The mathematics EOY component required students to show their understanding of concepts, procedures, and short applications.

⁷ Addendum 5 presents a summary of the test taker characteristics for the Fall 2014 administration.

5.3 Rules for Inclusion of Students in Analyses

Criteria for inclusion of students were implemented prior to all operational analyses. These rules were established by ETS psychometricians in consultation with PARCC and Pearson to determine which, if any, student records should be removed from analyses. This data screening process resulted in higher quality, albeit slightly smaller, data sets.

Student response data were included in analyses if:

- 1) Valid form numbers were observed for both the PBA and the EOY assessment components,
- 2) Neither the PBA nor the EOY records were flagged as “void” (i.e., do not score) based on information received from Pearson as the scoring vendor, and
- 3) The student attempted at least 25% of the PBA items *and* at least 25% of the EOY items,

Additionally, in cases where students had more than one valid record for either PBA or EOY, the record with the higher raw score was chosen. Records for students with administration issues or anomalies, per information received from Pearson, were excluded from analyses.

5.4 Test Takers by Grade, Mode, and Gender

Table 5.1 presents, for each grade of ELA/L, the number and percentage of students who took the test in each mode (CBT or PBT), as well as students who took one component of the test (PBA or EOY) online and the other component on paper. The latter condition is referred to as ‘Mixed Mode’. This information is provided for all participating states combined. Table 5.2 presents the same type of information for all students who took the mathematics assessments, and Table 5.3 provides this information for students who took the mathematics assessments in Spanish.

Markedly more students tested online than on paper, across all grades for both content areas. For ELA/L the percentages of online test takers, for all states combined, ranged from 73.2% to 94% while the percentages of paper test takers ranged from 5.8% to 26.7%. For all mathematics test takers, the percentages of students testing online ranged from 69.6% to 92.9%, whereas the percentages of students testing on paper ranged from 7% to 30.2%. The percentages of mathematics online students taking Spanish-language forms ranged from 70.2% to 100% and the percentages of mathematics students taking paper Spanish-language forms ranged from 1.6% to 29.6%. Generally, the percentage of students who tested online increased steadily from the lower grades to the higher grades. For example, about 70% of the ELA/L grade 3 students tested online, while 94% of the grade 11 students tested online. Overall, fewer students tested at the higher grades for both content areas.

Table 5.1 ELA/L Test Takers by Grade and Mode: All States Combined

Grade	No. of Valid Cases	CBT		PBT		Mixed Modes*	
		N	%	N	%	N	%
3	518,857	379,768	73.2	138,309	26.7	780	0.2
4	632,312	479,742	75.9	151,501	24.0	1,069	0.2
5	638,239	496,795	77.8	140,505	22.0	939	0.1
6	632,506	512,391	81.0	118,840	18.8	1,275	0.2
7	624,008	513,949	82.4	108,963	17.5	1,096	0.2
8	617,564	505,003	81.8	111,477	18.1	1,084	0.2
9	413,098	344,594	83.4	67,600	16.4	904	0.2
10	267,159	244,345	91.5	22,363	8.4	451	0.2
11	171,171	160,948	94.0	9,938	5.8	285	0.2
Grand Total	4,514,914	3,637,535		869,496		7,883	

Note: Includes students taking accommodated forms of ELA/L.

* Students who took one test component (i.e., PBA or EOY) as CBT and the other component as PBT.

Table 5.2 Mathematics Test Takers by Grade and Mode: All States Combined

Grade	No. of Valid Cases	CBT		PBT		Mixed Modes*	
		N	%	N	%	N	%
3	650,262	452,907	69.6	196,140	30.2	1,215	0.2
4	635,075	475,918	74.9	158,116	24.9	1,041	0.2
5	640,102	492,387	76.9	146,743	22.9	972	0.2
6	632,127	513,755	81.3	117,147	18.5	1,225	0.2
7	608,990	504,839	82.9	103,054	16.9	1,097	0.2
8	504,561	413,073	81.9	90,517	17.9	971	0.2
A1	480,604	405,661	84.4	73,955	15.4	988	0.2
GO	205,219	184,494	89.9	20,310	9.9	415	0.2
A2	186,890	173,572	92.9	13,099	7.0	219	0.1
M1	30,217	26,953	89.2	3,251	10.8	13	0.0
M2	12,282	11,374	92.6	908	7.4	n/a	n/a
M3	8,396	7,240	86.2	1,156	13.8	n/a	n/a
Grand Total	4,594,725	3,662,173		924,396		8,156	

Note: Includes students taking mathematics in English, students taking Spanish-language forms for mathematics, and students taking accommodated forms. A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. n/a = not applicable.

* Students who took one test component (i.e., PBA or EOY) as CBT and the other component as PBT.

Table 5.3 Spanish-Language Mathematics Test Takers, by Grade and Mode: All States Combined

Grade	No. of Valid Cases	CBT		PBT		Mixed Modes*	
		N	%	N	%	N	%
3	5,338	3,749	70.2	1,581	29.6	8	0.1
4	2,862	2,251	78.7	603	21.1	8	0.3
5	2,245	1,897	84.5	339	15.1	9	0.4
6	1,716	1,581	92.1	129	7.5	6	0.3
7	2,048	1,847	90.2	188	9.2	13	0.6
8	1,996	1,844	92.4	145	7.3	7	0.4
A1	2,256	2,120	94.0	128	5.7	8	0.4
GO	887	848	95.6	39	4.4	n/a	n/a
A2	506	498	98.4	8	1.6	n/a	n/a
M1	171	167	97.7	4	2.3	n/a	n/a
M2	41	41	100.0	n/a	n/a	n/a	n/a
M3	11	11	100.0	n/a	n/a	n/a	n/a
Grand Total	20,077	16,854		3,164		59	

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. n/a = not applicable.

* Students who took one test component (i.e., PBA or EOY) as CBT and the other component as PBT.

The number and percentage of students with valid test scores in each content area, grade, and mode of assessment are presented, for all states combined and for each state separately, in **Appendix 5** as Tables A.5.1, A.5.2, and A.5.3, for ELA/L test takers, all mathematics test takers, and students taking the Spanish-language mathematics tests, respectively. Table A.5.4 presents the ELA/L distribution by grade, mode, and gender, for all states combined; Tables A.5.5, and A.5.6 present similar information for all mathematics test takers and for students taking the Spanish-language mathematics tests, respectively.

5.5 Demographics

Also presented in **Appendix 5** is student demographic information for the following characteristics: economically disadvantaged, students with disabilities, English learners (EL), gender, and race/ethnicity (American Indian/Alaska Native; Black/African American; Hispanic/Latino; White/Caucasian; Native Hawaiian or Other Pacific Islander; two or more races reported; race not reported). Student demographic information was provided by the states and district and captured in PearsonAccess by means of a student data upload. The demographic data was verified by the states and district prior to score reporting.

Tables A.5.7 through A.5.15 provide demographic information for students with valid ELA/L scores, and Tables A.5.16 through A.5.27 present demographics for students with valid mathematics scores. All tables of demographic information are organized by grade; the results are first aggregated across all PARCC states and then presented for each state. Percentages are not reported for any states in which fewer than 20 students tested in a grade/content area.

Section 6: Classical Item Analysis

6.1 Overview

This section describes the results of the classical item analysis conducted for data obtained from the operational test items. Item analysis for operational items serves two purposes: to inform item exclusion decisions for IRT analysis, and to provide item statistics for the item bank.

PARCC item analysis included data from the following types of items: key-based selected response items, rule-based machine-scored items, and hand-scored constructed response items. For each operational item, the analysis produced item difficulty and item discrimination, as well as item response frequencies.

6.2 Data Screening Criteria

In preparation for item analysis, student response files were processed to verify that the data were free of errors. ETS Data Quality Services staff created a SAS-based automated verification program specifically for PARCC. The program ran predefined checks on all data files and verified that all fields and data needed to perform the statistical analyses were present and within expected ranges.

Item analyses were conducted by test form for each of the four combinations based on administration modes and assessment components:

- Computer-based testing Performance-Based Assessment (CBT-PBA),
- Computer-based testing End-of-Year assessment (CBT-EOY),
- Paper-based testing Performance-Based Assessment (PBT-PBA), and
- Paper-based testing End-of-Year assessment (PBT-EOY).

Before beginning item analysis, ETS performed the following data screening operations:

1. All component records that had an invalid form number were excluded.
2. All component records that were flagged as “void” based on the student file layout provided by Pearson were excluded.
3. All records where the student attempted fewer than 25% of the PBA items and/or attempted fewer than 25% of the EOY items were excluded.
4. For students with more than one valid record for either PBA or EOY, the record with the higher raw score was chosen.
5. Records for students with administration issues or anomalies were excluded per information provided by Pearson.

6.3 Description of Classical Item Analysis Statistics

Item analysis involved computing, for every item in each form for each grade/subject, a set of statistics based on observed item scores. Each statistic was designed to provide primary information about the performance of each item from an empirical perspective.

The following statistics and associated flagging rules were used to identify items that were not performing as expected:

- 1. Classical item difficulty indices (*p* value and average item score).** When constructing PARCC tests, a wide range of item difficulties is desired (i.e., from easy to hard items) so that students of all ability levels can be assessed with precision. At the operational stage, item difficulty statistics are used by test developers to build forms that meet desired test difficulty targets. Some of the items proved to be unexpectedly difficult because of students' lack of familiarity with the item type or limited opportunity to learn the content represented in the item.

For dichotomously scored items, item difficulty is indicated by its *p* value, which is the proportion of test takers who answered that item correctly. The range for *p* values is from .00 to 1.00. Items with high *p* values are easy items and those with low *p* values are difficult items. Dichotomously scored items were flagged for review if the *p* value was above .95 (i.e., too easy) or below .25 (i.e., too difficult).

For polytomously scored items, difficulty is indicated by the average item score (AIS). The AIS can range from .00 to the maximum total possible points for an item. To facilitate interpretation, the AIS values for polytomously scored items are often expressed as percentages of the maximum possible score, which are equivalent to the *p* values of dichotomously scored items. The desired *p* value range for polytomously scored items is .30 to .80; items with values outside this range were flagged for review.

- 2. The percentage of students choosing each response option.** Selected response items on PARCC assessments refer primarily to single-select multiple choice items. These items require that the test taker select a response from a number of answer options. These statistics for single-select multiple choice items indicate the percentage of students who select each of the answer options and the percentage that omit the item. The percentages are also computed for the high-performing subgroup of students who scored at the top 20% on the total assessment component score (i.e., the PBA or EOY total score). Items were flagged for review if more high-performing test takers chose the incorrect option than the correct response. Such a result could indicate that the item has multiple correct answers or is miskeyed.
- 3. Item-total correlation.** This statistic describes the relationship between test takers' performance on a specific item and their performance on the total test. The item-total correlation is usually referred to as the item discrimination index. For PARCC operational item analysis, the total score on the assessment component (i.e., PBA or EOY score) was used as the total test score. The polyserial correlation was calculated for both selected response items and constructed response items as an

estimate of the correlation between an observed continuous variable and an unobserved continuous variable hypothesized to underlie the variable with ordered categories (Olsson, Drasgow, and Dorans, 1982). Item-total correlations can range from -1.00 to 1.00. Desired values are positive and larger than .20. Negative item-total correlations indicate that low ability test takers perform better on an item than high ability test takers, an indication that the item may be potentially flawed. Item-total correlations below .20 were flagged for review. Items with extremely low or negative values were considered for exclusion from IRT calibrations or linking (refer to Section 10 for details on item inclusion and exclusion criteria for IRT analyses).

4. **Distractor-total correlation.** For selected response Items, this estimate describes the relationship between selecting an incorrect response (i.e., a distractor) for a specific item and performance on the total test. The polyserial correlation is calculated (refer to #3 above) for the distractors. Items with distractor-total correlations above .00 were flagged for review as these items may have multiple correct answers, be miskeyed, or have other content issues.
5. **Percentage of students omitting or not reaching each item.** For both selected response and constructed response items, this statistic is useful for identifying problems with test features such as testing time and item/test layout. Typically, if students have an adequate amount of testing time, approximately 95 percent of students should attempt to answer each question on the test. A distinction is made between “omit” and “not reached” for items without responses:
 - a. An item is considered “omit” if the student responded to subsequent items.
 - b. An item is considered “not reached” if the student did not respond to any subsequent items.

Patterns of high omit or not reached rates for items located near the end of a test section may indicate that test takers did not have adequate time. Items with high omit rates were flagged. Omit rates for constructed response items tend to be higher than for selected response items. Therefore, the omit rate for flagging individual items was 5% for selected response items and 15% for constructed response items. If a test taker omitted an item, then the test taker received a score of ‘0’ for that item and was included in the N-count for that item. However, if an item was near the end of the test and classified as not reached, the test taker did not receive a score and was not included in the N-count for that item.

6. **Distribution of item scores.** For constructed response items, examination of the distribution of scores is helpful to identify how well the item is functioning. If no students’ responses are assigned the highest possible score point, this may indicate that the item is not functioning as expected (e.g., the item could be confusing, poorly worded, or just unexpectedly difficult), the scoring rubric is flawed, and/or test takers did not have an opportunity to learn the content. In addition, if all or most test takers score at the extreme ends of the distribution (e.g., 0 and 2 for a 3-category item), this may indicate that there are problems with the item or the rubric so that test takers can receive either full credit or no credit at all, but not partial credit.

The raw score frequency distributions for constructed response items were computed to identify items with few or no observations at any score points. Items with a low percentage (i.e., <3%) of test takers obtaining any score point were flagged. In addition, constructed response items were flagged if they had U-shaped distributions, with high frequencies for extreme scores and very low frequencies for middle score categories. Items with such response patterns may pose problems during the IRT calibrations and therefore need to be excluded (refer to Section 10 for more information).

6.4 Summary of Classical Item Analysis Flagging Criteria

In summary, items are flagged for review if the item analysis yielded any of the following results:

1. p value above .95 for dichotomous items and above .80 for polytomous items
2. p value below .25 for dichotomous items, and below .30 for polytomous items
3. Item-total correlation below .20
4. Any distractor-total correlation above .00
5. Greater number of high-performing students (top 20%) choosing a distractor than the keyed response
6. High percentage of omits: above 5% for selected response items, and above 15% for constructed response items
7. High percentage that did not reach the item: above 5% for selected response items, and above 15% for constructed response items
8. Constructed response items with a score value obtained by less than 3% percent of responses

ETS's psychometric staff carefully reviewed each of the flagged items and summarized the results for Pearson and PARCC with recommendations for subsequent analyses.

6.5 Classical Item Analysis Results

This section presents tables summarizing the item analysis results for the Spring 2015 operational items. The Fall 2014 results are included in an Addendum⁸ to this report.

- Tables 6.1 and 6.2 present p value information by grade, component, and mode for the ELA/L and mathematics operational items from the Spring 2015 operational administration.
- Tables 6.3 and 6.4 present item-total correlations by grade, component, and mode for the ELA/L and mathematics operational items from the Spring 2015 operational administration.

⁸ Addendum 6 provides a summary of the IA results for the Fall 2014 administration.

An operational item could appear on multiple test forms. The tables list only unique items in each test mode and component combination (e.g., CBT-PBA), and the reported item statistics are based on the weighted averages across multiple occurrences of an item.

Spoiled or 'do not score' items, based on information provided by Pearson as the scoring vendor, were excluded from the total test score for each form in item analysis. These items were removed from scoring because of item performance, technical scoring issues, content concerns, or multiple/no correct answers. Additionally, some items were dropped during item calibrations due to:

- A low weighted polyserial,
- A low p value (e.g., extremely difficult item), or
- Extremely poor IRT model fit or item not able to calibrate.

Tables 10.3 and 10.4 in section 10 present the count and percentage of CBT and PBT items excluded from IRT calibration along with the reasons the items were excluded for ELA/L and mathematics, respectively. The tables in this section and in Addendum 6 include only those items that were used for operational scoring.

Table 6.1 Summary of *p* Values for ELA/L Operational Items by Grade, Component, and Mode

Grade	Component	Mode	N of Unique Items	Mean <i>p</i> Value	SD <i>p</i> Value	Min <i>p</i> Value	Max <i>p</i> Value	Median <i>p</i> Value
3	PBA	CBT	63	0.43	0.16	0.15	0.82	0.44
3	PBA	PBT	40	0.45	0.14	0.16	0.68	0.45
3	EOY	CBT	39	0.42	0.15	0.10	0.70	0.40
3	EOY	PBT	25	0.38	0.15	0.10	0.70	0.36
4	PBA	CBT	71	0.45	0.14	0.16	0.80	0.43
4	PBA	PBT	51	0.44	0.15	0.15	0.72	0.43
4	EOY	CBT	44	0.43	0.15	0.20	0.73	0.40
4	EOY	PBT	24	0.41	0.14	0.22	0.73	0.37
5	PBA	CBT	79	0.45	0.15	0.18	0.80	0.42
5	PBA	PBT	49	0.45	0.15	0.23	0.80	0.39
5	EOY	CBT	40	0.40	0.10	0.22	0.64	0.39
5	EOY	PBT	24	0.38	0.12	0.15	0.64	0.38
6	PBA	CBT	80	0.47	0.16	0.20	0.77	0.47
6	PBA	PBT	46	0.47	0.14	0.20	0.81	0.50
6	EOY	CBT	78	0.41	0.14	0.15	0.77	0.40
6	EOY	PBT	44	0.44	0.14	0.19	0.77	0.41
7	PBA	CBT	68	0.44	0.15	0.21	0.81	0.42
7	PBA	PBT	44	0.45	0.14	0.21	0.72	0.44
7	EOY	CBT	67	0.47	0.14	0.18	0.78	0.46
7	EOY	PBT	44	0.49	0.15	0.18	0.78	0.48
8	PBA	CBT	73	0.47	0.15	0.13	0.84	0.44
8	PBA	PBT	46	0.49	0.16	0.13	0.84	0.48
8	EOY	CBT	70	0.43	0.15	0.05	0.85	0.41

Grade	Component	Mode	N of Unique Items	Mean <i>p Value</i>	SD <i>p Value</i>	Min <i>p Value</i>	Max <i>p Value</i>	Median <i>p Value</i>
8	EOY	PBT	43	0.45	0.16	0.17	0.86	0.42
9	PBA	CBT	95	0.43	0.15	0.17	0.80	0.43
9	PBA	PBT	46	0.47	0.14	0.23	0.80	0.44
9	EOY	CBT	93	0.39	0.14	0.12	0.81	0.38
9	EOY	PBT	44	0.42	0.15	0.12	0.81	0.44
10	PBA	CBT	94	0.40	0.13	0.15	0.78	0.36
10	PBA	PBT	46	0.38	0.11	0.15	0.66	0.36
10	EOY	CBT	91	0.41	0.13	0.11	0.70	0.41
10	EOY	PBT	49	0.44	0.15	0.20	0.81	0.46
11	PBA	CBT	83	0.38	0.13	0.13	0.72	0.36
11	PBA	PBT	53	0.37	0.11	0.13	0.67	0.36
11	EOY	CBT	75	0.31	0.13	0.08	0.68	0.32
11	EOY	PBT	54	0.33	0.11	0.12	0.68	0.32

Note: PBA = Performance-Based Assessment; EOY = End-of-Year assessment; CBT = computer-based testing (online); PBT = paper-based testing (paper).

Table 6.2 Summary of *p* Values for Mathematics Operational Items by Grade, Component, and Mode

Grade	Component	Mode	<i>N</i> of Unique Items	Mean <i>p</i> Value	SD <i>p</i> Value	Min <i>p</i> Value	Max <i>p</i> Value	Median <i>p</i> Value
3	PBA	CBT	84	0.41	0.28	0.03	0.98	0.33
3	PBA	PBT	48	0.42	0.27	0.03	0.98	0.38
3	EOY	CBT	191	0.50	0.27	0.03	0.98	0.49
3	EOY	PBT	110	0.50	0.25	0.07	0.98	0.49
4	PBA	CBT	66	0.46	0.30	0.04	0.94	0.37
4	PBA	PBT	46	0.47	0.28	0.08	0.89	0.38
4	EOY	CBT	144	0.49	0.24	0.05	0.92	0.53
4	EOY	PBT	89	0.50	0.23	0.05	0.92	0.52
5	PBA	CBT	66	0.38	0.23	0.03	0.88	0.34
5	PBA	PBT	41	0.38	0.19	0.03	0.82	0.38
5	EOY	CBT	160	0.43	0.24	0.06	0.92	0.37
5	EOY	PBT	94	0.48	0.21	0.08	0.92	0.49
6	PBA	CBT	73	0.38	0.24	0.05	0.90	0.34
6	PBA	PBT	43	0.40	0.24	0.07	0.90	0.36
6	EOY	CBT	136	0.39	0.23	0.03	0.84	0.35
6	EOY	PBT	89	0.39	0.20	0.03	0.93	0.37
7	PBA	CBT	79	0.32	0.22	0.02	0.80	0.25
7	PBA	PBT	47	0.35	0.23	0.02	0.80	0.28
7	EOY	CBT	168	0.31	0.24	0.01	0.81	0.26
7	EOY	PBT	83	0.35	0.21	0.01	0.81	0.32
8	PBA	CBT	63	0.30	0.25	0.02	0.75	0.19
8	PBA	PBT	46	0.31	0.22	0.02	0.75	0.26
8	EOY	CBT	146	0.28	0.21	0.02	0.76	0.20

Grade	Component	Mode	N of Unique Items	Mean <i>p Value</i>	SD <i>p Value</i>	Min <i>p Value</i>	Max <i>p Value</i>	Median <i>p Value</i>
8	EOY	PBT	70	0.32	0.22	0.02	0.76	0.27
A1	PBA	CBT	96	0.22	0.17	0.01	0.68	0.16
A1	PBA	PBT	51	0.24	0.19	0.01	0.68	0.22
A1	EOY	CBT	191	0.25	0.19	0.01	0.69	0.25
A1	EOY	PBT	82	0.28	0.18	0.02	0.69	0.31
GO	PBA	CBT	96	0.24	0.20	0.01	0.71	0.18
GO	PBA	PBT	46	0.27	0.21	0.01	0.71	0.22
GO	EOY	CBT	188	0.28	0.21	0.01	0.88	0.25
GO	EOY	PBT	89	0.35	0.21	0.03	0.88	0.37
A2	PBA	CBT	83	0.18	0.15	0.00	0.66	0.13
A2	PBA	PBT	51	0.15	0.14	0.01	0.66	0.12
A2	EOY	CBT	137	0.24	0.18	0.00	0.82	0.20
A2	EOY	PBT	84	0.30	0.16	0.00	0.82	0.29
M1	PBA	CBT	32	0.22	0.16	0.01	0.53	0.18
M1	PBA	PBT	23	0.25	0.16	0.03	0.53	0.29
M1	EOY	CBT	59	0.29	0.21	0.03	0.83	0.32
M1	EOY	PBT	42	0.32	0.19	0.03	0.83	0.34
M2	PBA	CBT	34	0.19	0.16	0.00	0.63	0.15
M2	PBA	PBT	26	0.25	0.19	0.02	0.63	0.24
M2	EOY	CBT	58	0.27	0.19	0.01	0.65	0.22
M2	EOY	PBT	48	0.34	0.17	0.01	0.65	0.39
M3	PBA	CBT	36	0.20	0.17	0.02	0.63	0.14
M3	PBA	PBT	20	0.18	0.16	0.02	0.61	0.12
M3	EOY	CBT	63	0.25	0.19	0.03	0.78	0.21

Grade	Component	Mode	N of Unique Items	Mean <i>p Value</i>	SD <i>p Value</i>	Min <i>p Value</i>	Max <i>p Value</i>	Median <i>p Value</i>
M3	EOY	PBT	46	0.28	0.17	0.03	0.78	0.28

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. PBA = Performance-Based Assessment; EOY = End-of-Year assessment; CBT = computer-based testing (online); PBT = paper-based testing (paper).

Table 6.3 Summary of Item-total Polyserial Correlations for ELA/L Operational Items by Grade, Component, and Mode

Grade	Component	Mode	N of Unique Items	Mean Polyserial	SD Polyserial	Min Polyserial	Max Polyserial	Median Polyserial
3	PBA	CBT	63	0.24	0.10	0.05	0.38	0.28
3	PBA	PBT	40	0.25	0.09	0.05	0.38	0.28
3	EOY	CBT	39	0.28	0.07	0.11	0.37	0.29
3	EOY	PBT	25	0.26	0.07	0.11	0.37	0.29
4	PBA	CBT	71	0.22	0.08	0.05	0.35	0.24
4	PBA	PBT	51	0.20	0.08	0.05	0.35	0.22
4	EOY	CBT	44	0.28	0.05	0.17	0.41	0.29
4	EOY	PBT	24	0.28	0.05	0.17	0.36	0.29
5	PBA	CBT	79	0.22	0.08	0.05	0.36	0.24
5	PBA	PBT	49	0.21	0.08	0.05	0.33	0.23
5	EOY	CBT	40	0.27	0.05	0.18	0.34	0.28
5	EOY	PBT	24	0.27	0.05	0.17	0.34	0.27
6	PBA	CBT	80	0.21	0.08	0.04	0.33	0.22
6	PBA	PBT	46	0.23	0.09	0.04	0.33	0.25
6	EOY	CBT	78	0.26	0.06	0.12	0.40	0.25
6	EOY	PBT	44	0.26	0.06	0.12	0.40	0.26
7	PBA	CBT	68	0.21	0.08	0.04	0.35	0.22
7	PBA	PBT	44	0.20	0.08	0.04	0.35	0.21
7	EOY	CBT	67	0.28	0.06	0.12	0.36	0.28
7	EOY	PBT	44	0.28	0.05	0.12	0.36	0.28
8	PBA	CBT	73	0.20	0.09	0.04	0.36	0.22
8	PBA	PBT	46	0.21	0.09	0.04	0.33	0.23
8	EOY	CBT	70	0.26	0.06	0.11	0.36	0.26

Grade	Component	Mode	N of Unique Items	Mean Polyserial	SD Polyserial	Min Polyserial	Max Polyserial	Median Polyserial
8	EOY	PBT	43	0.26	0.06	0.08	0.35	0.27
9	PBA	CBT	95	0.22	0.08	0.04	0.33	0.24
9	PBA	PBT	46	0.22	0.08	0.04	0.33	0.24
9	EOY	CBT	93	0.27	0.06	0.09	0.36	0.28
9	EOY	PBT	44	0.26	0.06	0.09	0.34	0.28
10	PBA	CBT	94	0.21	0.08	0.04	0.34	0.22
10	PBA	PBT	46	0.21	0.08	0.04	0.33	0.21
10	EOY	CBT	91	0.28	0.05	0.16	0.39	0.29
10	EOY	PBT	49	0.29	0.05	0.17	0.39	0.31
11	PBA	CBT	83	0.21	0.09	0.01	0.33	0.23
11	PBA	PBT	53	0.20	0.09	0.01	0.33	0.21
11	EOY	CBT	75	0.27	0.05	0.10	0.37	0.27
11	EOY	PBT	54	0.27	0.05	0.10	0.37	0.27

Note: PBA = Performance-Based Assessment; EOY = End-of-Year assessment; CBT = computer-based testing (online); PBT = paper-based testing (paper).

Table 6.4 Summary of Item-total Correlations for Mathematics Operational Items by Grade, Component, and Mode

Grade	Component	Mode	N of Unique Items	Mean Polyserial	SD Polyserial	Min Polyserial	Max Polyserial	Median Polyserial
3	PBA	CBT	84	0.42	0.22	0.13	0.83	0.37
3	PBA	PBT	48	0.42	0.21	0.13	0.80	0.38
3	EOY	CBT	191	0.58	0.14	0.10	0.79	0.60
3	EOY	PBT	110	0.57	0.14	0.14	0.79	0.61
4	PBA	CBT	66	0.43	0.23	0.11	0.78	0.37
4	PBA	PBT	46	0.42	0.22	0.12	0.78	0.33
4	EOY	CBT	144	0.55	0.15	0.12	0.79	0.58
4	EOY	PBT	89	0.52	0.16	0.12	0.79	0.56
5	PBA	CBT	66	0.34	0.17	0.12	0.77	0.31
5	PBA	PBT	41	0.34	0.17	0.12	0.77	0.31
5	EOY	CBT	160	0.52	0.14	0.10	0.79	0.55
5	EOY	PBT	94	0.52	0.15	0.10	0.79	0.56
6	PBA	CBT	73	0.42	0.22	0.12	0.78	0.38
6	PBA	PBT	43	0.41	0.21	0.12	0.78	0.37
6	EOY	CBT	136	0.52	0.15	0.15	0.81	0.56
6	EOY	PBT	89	0.51	0.16	0.15	0.81	0.54
7	PBA	CBT	79	0.39	0.21	0.13	0.80	0.35
7	PBA	PBT	47	0.36	0.18	0.13	0.70	0.29
7	EOY	CBT	168	0.49	0.18	0.11	0.83	0.51
7	EOY	PBT	83	0.46	0.17	0.12	0.83	0.41
8	PBA	CBT	63	0.39	0.20	0.10	0.82	0.35
8	PBA	PBT	46	0.37	0.20	0.11	0.77	0.33
8	EOY	CBT	146	0.47	0.15	0.10	0.73	0.50

Grade	Component	Mode	N of Unique Items	Mean Polyserial	SD Polyserial	Min Polyserial	Max Polyserial	Median Polyserial
8	EOY	PBT	70	0.47	0.15	0.13	0.71	0.49
A1	PBA	CBT	96	0.34	0.19	0.09	0.74	0.29
A1	PBA	PBT	51	0.32	0.18	0.10	0.69	0.24
A1	EOY	CBT	191	0.36	0.17	0.06	0.74	0.33
A1	EOY	PBT	82	0.35	0.16	0.06	0.72	0.33
GO	PBA	CBT	96	0.38	0.22	0.11	0.75	0.32
GO	PBA	PBT	46	0.36	0.22	0.12	0.75	0.24
GO	EOY	CBT	188	0.44	0.18	0.15	0.83	0.41
GO	EOY	PBT	89	0.42	0.18	0.16	0.82	0.39
A2	PBA	CBT	83	0.35	0.21	0.08	0.74	0.22
A2	PBA	PBT	51	0.32	0.19	0.08	0.65	0.25
A2	EOY	CBT	137	0.40	0.17	0.05	0.77	0.35
A2	EOY	PBT	84	0.36	0.19	0.04	0.77	0.34
M1	PBA	CBT	32	0.32	0.18	0.11	0.72	0.24
M1	PBA	PBT	23	0.31	0.17	0.11	0.68	0.24
M1	EOY	CBT	59	0.37	0.18	0.13	0.78	0.33
M1	EOY	PBT	42	0.38	0.19	0.14	0.78	0.31
M2	PBA	CBT	34	0.38	0.21	0.11	0.71	0.39
M2	PBA	PBT	26	0.38	0.22	0.11	0.71	0.39
M2	EOY	CBT	58	0.39	0.18	0.09	0.80	0.36
M2	EOY	PBT	48	0.40	0.18	0.13	0.80	0.38
M3	PBA	CBT	36	0.33	0.19	0.08	0.68	0.23
M3	PBA	PBT	20	0.28	0.17	0.08	0.59	0.22
M3	EOY	CBT	63	0.38	0.17	0.09	0.80	0.32

Grade	Component	Mode	N of Unique Items	Mean <i>Polyserial</i>	SD <i>Polyserial</i>	Min <i>Polyserial</i>	Max <i>Polyserial</i>	Median <i>Polyserial</i>
M3	EOY	PBT	46	0.34	0.17	0.04	0.80	0.32

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. PBA = Performance-Based Assessment; EOY = End-of-Year assessment; CBT = computer-based testing (online); PBT = paper-based testing (paper).

Section 7: Differential Item Functioning

7.1 Overview

Differential item functioning (DIF) analyses were conducted using the data obtained from the operational items. If an item performs differentially across identifiable subgroups (e.g., gender or ethnicity) when students are matched on ability, the item may be measuring something other than the intended construct (i.e., possible evidence of DIF). It is important, however, to recognize that item performance differences flagged for DIF might be related to actual differences in relevant knowledge or skills (item impact) or statistical Type I error. As a result, DIF statistics are used to identify *potential* item bias. Subsequent reviews by content experts and bias/sensitivity committees are required to determine the source and meaning of performance differences.

7.2 DIF Procedures

Dichotomous Items. The Mantel-Haenszel (MH) DIF statistic was calculated for selected-response items and for dichotomously-scored constructed-response items. In this method, test takers are classified to relevant subgroups of interest (e.g., gender or ethnicity). Using the PARCC Performance-Based Assessment (PBA) score, or the end-of-year (EOY) assessment score as the total criteria, test takers in a certain total score category in the focal group (e.g., females) are compared with examinees in the same total score category in the reference group (e.g., males). One each item, test takers in the focal group are also compared to test takers in the reference group who performed equally well on the test as a whole. The common odds ratio is estimated across all categories of matched test taker ability using the following formula (Dorans & Holland, 1993), and the resulting estimate is interpreted as the relative likelihood of success on a particular item for members of two groups when matched on ability.

$$\hat{\alpha}_{MH} = \frac{\sum_{s=1}^S R_{rs} W_{fs} / N_{ts}}{\sum_{s=1}^S R_{fs} W_{rs} / N_{ts}}, \quad (7-1)$$

in which:

S = the number of score categories,

R_{rs} = the number of test takers in the reference group who answer the item correctly,

W_{fs} = the number of test takers in the focal group who answer the item incorrectly,

R_{fs} = the number of test takers in the focal group who answer the item correctly,

W_{rs} = the number of test takers in the reference group who answer the item incorrectly, and

N_{ts} = the total number of test takers.

To facilitate the interpretation of MH results, the common odds ratio is frequently transformed to the delta scale using the following formula (Holland & Thayer, 1988):

$$MH\ D-DIF = -2.35 \ln(\hat{\alpha}_{MH}) \quad (7-2)$$

Positive values indicate DIF in favor of the focal group (i.e., positive DIF items are differentially easier for the focal group), whereas negative values indicate DIF in favor of the reference group (i.e., negative DIF items are differentially easier for the reference group).

Polytomous Items. For polytomously scored constructed-response items, the MH D-DIF statistic is not calculated; instead the standardization DIF (Dorans & Schmitt, 1991; Zwick, Thayer & Mazzeo, 1997; Dorans, 2013), in conjunction with the Mantel chi-square statistic (Mantel, 1963; Mantel & Haenszel, 1959), is used to identify items with DIF.

The standardization DIF compares the item means of the two groups after adjusting for differences in the distribution of test takers across the values of the matching variable (i.e., total test score) and is calculated using the following formula:

$$STD - EISDIF = \frac{\sum_{s=1}^S N_{fs} * E_f(Y|X = s)}{\sum_{s=1}^S N_{fs}} - \frac{\sum_{s=1}^S N_{rs} * E_r(Y|X = s)}{\sum_{s=1}^S N_{rs}} = \frac{\sum_{s=1}^S D_s}{\sum_{s=1}^S N_{fs}}, \quad (7-3)$$

in which:

X = the total score

Y = the item score

S = the number of score categories on X ,

N_{rs} = the number of test takers in the reference group in score category s ,

N_{fs} = the number of test takers in the focal group in score category s ,

E_r = the expected item score for reference group, and

E_f = the expected item score for reference group.

A positive *STD-EISDIF* value means that, conditional on the total test score, the focal group has a higher mean item score than the focal group. In contrast, a negative *STD-EISDIF* value means that, conditional on the total test score, the focal group has a lower mean item score than the reference group.

Classification. Based on the DIF statistics and significance tests, items are classified into three categories and assigned values of A, B, or C. Category A items contain negligible DIF, Category B items exhibit slight to moderate DIF, and Category C items possess moderate to large DIF values. Positive values indicate that, conditional on the total score, the focal group has a higher mean item score than the reference group. In contrast, negative DIF values indicate that, conditional on the total test score, the focal group has a lower mean item score than the reference group. The flagging criteria for dichotomously scored items are presented in Table 7.1; the flagging criteria for polytomously scored constructed response items are provided in Table 7.2.

Table 7.1 DIF Categories for Dichotomous Selected Response and Constructed Response Items

DIF Category	Criteria
A (negligible)	Absolute value of the MH D-DIF is not significantly different from zero, or is less than one.
B (slight to moderate)	1. Absolute value of the MH D-DIF is significantly different from zero but not from one, and is at least one; OR 2. Absolute value of the MH D-DIF is significantly different from one, but is less than 1.5. Positive values are classified as “B+” and negative values as “B-”.
C (moderate to large)	Absolute value of the MH D-DIF is significantly different from one, and is at least 1.5. Positive values are classified as “C+” and negative values as “C-”.

Table 7.2 DIF Categories for Polytomous Constructed Response Items

DIF Category	Criteria
A (negligible)	Mantel Chi-square <i>p value</i> > 0.05 or $ STD-EISDIF/SD \leq 0.17$
B (slight to moderate)	Mantel Chi-square <i>p value</i> < 0.05 and $ STD-EISDIF/SD > 0.17$
C (moderate to large)	Mantel Chi-square <i>p value</i> < 0.05 and $ STD-EISDIF/SD > 0.25$

Note: *STD-EISDIF* = standardized DIF; *SD* = total group standard deviation of item score.

7.3 Operational Analysis DIF Comparison Groups

Traditional Comparisons. DIF analyses were conducted on each test component form (i.e., each PBA or EOY form) for designated comparison groups defined on the basis of demographic variables including: gender, race/ethnicity, economic disadvantage, and special instructional needs such as students with disabilities (SWD) or English learners (EL). Student demographic information was provided by the states and district and captured in PearsonAccess by means of a student data upload. The demographic data was verified by the states and district prior to score reporting. These comparison groups are specified in Table 7.3.

Table 7.3 Traditional DIF Comparison Groups

Grouping Variable	Focal Group	Reference Group
Gender	Female	Male
Ethnicity	American Indian/Alaska Native (AmerIndian)	White
	Asian	White
	Black or African American	White
	Hispanic/Latino	White
	Native Hawaiian or Pacific Islander	White
	Multiple Race Selected	White
Economic Status*	Economically Disadvantaged (EcnDis)	Not Economically Disadvantaged (NoEcnDis)
Special Instructional Needs	English Learner (ELY)	Non English Learner (ELN)
	Students with Disabilities (SWDY)	Students without Disabilities (SWDN)

Note: * Economic status was based on participation in National School Lunch Program: receipt of free or reduced-price lunch.

Comparison across Languages. DIF analyses were also conducted for Spanish-language forms vs. English-language forms in mathematics. At each grade level, one computer-based test (CBT) form and one paper-based test (PBT) form of the mathematics test were transadpated into Spanish (refer to Section 2 for more information on the development of Spanish-language forms). The purpose of the Spanish vs. English DIF analysis was to evaluate how similarly the items functioned between the two languages because the data from the Spanish-language forms were not separately calibrated using IRT. The item parameter estimates based on the English speaking test takers were used to generate conversion tables for the Spanish-language forms. Spanish-language mathematics items flagged for C-DIF were reviewed by content specialists and the PARCC Priority Alert Task Force to decide if the items were problematic and should be excluded from scoring. An item could be dropped from a Spanish-language form but remain in the English-language form if no other issues were detected; in those cases separate conversion tables were generated for the two versions of the form which had different numbers of items.

Sample Size Requirement. DIF analyses were conducted when the following sample size requirements were met:

- The smaller group, reference or focal, had at least 100 students, and
- The combined group, reference and focal, had at least 400 students.

Spanish-language vs. English-language DIF analyses were not conducted for Integrated Mathematics II and III because of insufficient sample sizes.

7.4 Operational Differential Item Functioning Results

Appendix 7 presents tables summarizing the DIF results for the Spring 2015 operational items, with one table prepared for each content and grade level (e.g., ELA/L Grade 3). The Fall 2014 results are included in an Addendum⁹ to this report.

Spoiled or ‘do not score’ items were excluded from the total test score for each form in DIF analysis. These items were removed from scoring because of item performance, technical scoring issues, content concerns, multiple correct answers, or no correct answers. However, the tables in this section may include items for certain grade levels that were excluded from scoring based on later analyses (refer to Section 10.5 Items Excluded from Score Reporting for more information).

In the DIF results tables, the column “DIF Comparisons” identifies the focal and reference groups for the analysis performed; the column “Mode” identifies the test delivery mode. “Total N of Unique Items” reports the number of unique items included in the analysis, whereas “Total N of Item Occurrences” reports the number of times items were used on test forms. An item could be used in multiple test forms; therefore, items were counted according to the occurrences. For example, if the same item appeared in five test forms, it was counted as five occurrences; if this item was classified as B+ on one form and C+ on another form, both occurrences were reported in the corresponding columns. “Total N of Item Occurrences Included in DIF Analysis” reports the number of occurrences with sufficient sample sizes to be included in DIF analyses. In addition, “0” indicates that the DIF analysis did not classify any items in the particular DIF category, while “n/a” indicates that the DIF analysis was not performed due to insufficient sample sizes.

⁹ Addendum 7 provides a summary of DIF results for the Fall 2014 administration.

Section 8: Reliability

8.1 Overview

Reliability focuses on the extent to which differences in test scores reflect true differences in the knowledge, ability, or skill being tested rather than fluctuations due to chance. Thus, reliability measures the consistency of the scores across conditions that can be assumed to differ at random, especially which form of the test the test taker is administered and which persons are assigned to score responses to constructed-response questions. In statistical terms, the variance in the distributions of test scores, essentially the differences among individuals, is partly due to real differences in the knowledge, skill, or ability being tested (true variance) and partly due to random errors in the measurement process (error variance). Reliability is an estimate of the proportion of the total variance that is true variance.

There are several different ways of estimating reliability. The type of reliability estimate reported here is an internal-consistency measure, which is derived from analysis of the consistency of the performance of individuals across items within a test. It is used because it serves as a good estimate of alternate forms reliability, but it does not take into account form-to-form variation due to lack of test form parallelism, nor is it responsive to day-to-day variation due to, for example, the examinee's state of health or the testing environment.

Reliability coefficients range from 0 to 1. The higher the reliability coefficient for a set of scores, the more likely individuals would be to obtain very similar scores upon repeated testing occasions, if the students do not change in their level of the knowledge or skills measured by the test. The reliability estimates in the tables to follow attempt to answer the question, "How consistent would the scores of these test takers be over replications of the entire testing process?"

Reliability of classification estimates the proportion of students who are accurately classified into proficiency levels. There are two kinds of classification reliability statistics: decision accuracy and decision consistency. Decision accuracy is the agreement between the classifications actually made and the classifications that would be made if the test scores were perfectly reliable. Decision consistency is the agreement between the classifications that would be made on two independent different forms of the test.

Another index is inter-rater reliability for the human scored constructed-response items, which measures the agreement between individual raters (scorers). The inter-rater reliability coefficient answers the question, "How consistent would the scores of these test takers be over replication of scoring of the same responses by different scorers?"

Standard error of measurement (SEM) quantifies the amount of error in the test scores. SEM is the extent by which test takers' scores tend to differ from the scores they would receive if the test were perfectly reliable. The larger the SEM, the more the variability of a student's observer scores across repeated testing. Observed scores with large SEMs pose a challenge to the valid interpretation of a single test score.

Reliability and SEM estimates were calculated at the full assessment level (both PBT and CBT), and at the claim and subclaim levels.

8.2 Reliability and SEM Estimation

Coefficient alpha (Cronbach, 1951), which measures internal consistency reliability, is the most commonly used measure of reliability. Coefficient alpha is estimated by substituting sample estimates for the parameters in the formula below:

$$\alpha = \frac{n}{n-1} \left[1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_x^2} \right], \quad (8-1)$$

where n is the number of items, σ_i^2 is the variance of scores on the i -th item, and σ_x^2 is the variance of the total score (sum of scores on the individual items). Other things being equal, the more items a test includes, the higher the internal consistency reliability.

Since PARCC test forms have mixed item types (dichotomous and polytomous items), it is more appropriate to report stratified alpha (Feldt & Brennan, 1989). Stratified alpha is a weighted average of coefficient alphas for item sets with different maximum score points or “strata.” Stratified alpha is a reliability estimate computed by dividing the test into parts (“strata”), computing alpha separately for each part, and using the results to estimate a reliability coefficient for the total score. Stratified alpha is used here because different parts of the test consist of different item types and may measure different skills. The formula for the stratified alpha is:

$$\text{strata } \rho = 1 - \frac{\sum \sigma_{x_j}^2 (1 - \alpha_j)}{\sigma_x^2} \quad (8-2)$$

where $\sigma_{x_j}^2$ is the variance for part j of the test, σ_x^2 is the variance of the total scores, and α_j is coefficient alpha for part j of the test. Estimates of stratified alpha are computed by substituting sample estimates for the parameters in the formula.

The formula for the standard error of measurement is:

$$\sigma_E = \sigma_X \sqrt{1 - \rho_{XX'}} \quad (8-3)$$

where σ_X is the standard deviation of the test score, either total raw score or scale scores, and $\rho_{XX'}$ is the reliability estimated by substitution of appropriate statistics for the parameters in equation 8-1 or 8-2.

8.3 Reliability Results for Total Group

Tables 8.1 and Table 8.2 summarize test reliability estimates for the total testing group for English Language Arts/Literacy (ELA/L) and mathematics, respectively. The section includes only Spring 2015 results. The Fall 2014 results are located in the Addendum.¹⁰ The tables provide the average reliability, which is estimated by averaging the internal consistency estimates computed for all the individual forms of the test, and both the raw score and scale score SEMs, separately for the computer-based and paper-based tests within each grade level. In addition, the number of forms, the average sample size, and the average maximum possible score for each set of tests are provided.

English Language Arts/Literacy

The average reliability estimates for the CBT tests for grades 3-11 English Language Arts/Literacy (ELA/L) range from a low of .90 to a high of .93. The average reliability estimates for the PBT tests for ELA/L grades 3-11 range from a low of .89 to a high of .92. The tests for grades 3-5 have fewer maximum possible points than for the grades 6-11 tests. The average reliability estimates of the grades 3-5 tests are .90, and eleven of the twelve reliability estimates associated with grades 6-11 are at least .91.

The average raw score SEM is consistently between a very reasonable 5% and 6% of the maximum possible score. The scale score SEMs are lowest for grade 6 and highest for grade 3. Across the nine grade levels, the raw score SEMs for the PBT assessments are higher than for the CBT assessments. However, the scale score SEMs for the CBT tests are higher than for the corresponding PBT test for grade 6 and grades 8-10.

¹⁰ Addendum 8 provides a summary of reliability information for the Fall 2014 administration.

Table 8.1 Summary of ELA/L Test Reliability Estimates for Total Group

Grade Level	Testing Mode	Number of Forms	Average Sample Size	Average Maximum Possible Score	Average Reliability	Average Raw Score SEM	Average Scale Score SEM
3	CBT	36	10276	100	0.90	5.51	12.20
	PBT	16	8042	100	0.90	5.73	12.47
4	CBT	36	13108	105	0.91	5.83	9.84
	PBT	16	8942	105	0.89	6.14	10.61
5	CBT	36	13579	106	0.91	5.67	9.67
	PBT	16	8351	105	0.89	6.02	10.20
6	CBT	36	14105	137	0.92	6.75	8.54
	PBT	16	7129	137	0.92	7.23	8.31
7	CBT	36	14134	135	0.93	6.83	9.70
	PBT	16	6559	135	0.91	7.47	10.13
8	CBT	36	13846	136	0.92	6.98	10.40
	PBT	16	6735	136	0.92	7.16	10.11
9	CBT	64	5152	137	0.93	6.55	9.64
	PBT	16	3821	137	0.92	6.91	9.47
10	CBT	64	3628	137	0.93	7.12	11.64
	PBT	16	1228	137	0.93	7.29	11.34
11	CBT	36	4169	136	0.92	6.89	10.73
	PBT	16	428	136	0.91	7.52	11.11

Mathematics

The average reliability estimates for the grades 3-8 end-of-year (EOY) mathematics assessments are higher than for the six end-of-course (EOC) mathematics assessments. Of the 12 reliability estimates associated with the EOY mathematics assessments, 10 are either .93 or .94. The median average reliability for the EOC assessments is .91. Except for the Integrated Mathematics II and III tests, the average reliability estimates of the CBT mode and the PBT mode assessments are within .01.

The SEM as percentage of total score consistently range from 4% to 5% of the maximum score. The SEMs for the scale scores are highest for Integrated Mathematics III, Algebra II, and grade 8 and lowest for Geometry. With the exception of Integrated Mathematics I and II, the scale score SEMs are higher for the PBT mode when compared to the CBT mode.

Table 8.2 Summary of Mathematics Test Reliability Estimates for Total Group

Grade Level	Testing Mode	Number of Forms	Average Sample Size	Average Maximum Possible Score	Average Reliability	Average Raw Score SEM	Average Scale Score SEM
3	CBT	36	12438	81	0.94	3.66	8.08
	PBT	16	10993	82	0.94	3.76	8.27
4	CBT	36	13080	82	0.94	3.60	7.36
	PBT	16	9137	82	0.94	3.79	7.64
5	CBT	37	13105	82	0.93	3.71	7.58
	PBT	16	8507	82	0.94	4.00	7.66
6	CBT	37	13585	82	0.94	3.73	7.07
	PBT	16	6768	82	0.94	3.92	7.42
7	CBT	36	13791	82	0.93	3.59	7.20
	PBT	16	5629	80	0.93	3.82	7.28
8	CBT	36	11240	82	0.91	3.48	10.22
	PBT	16	5129	81	0.92	3.79	10.36
A1	CBT	63	6058	96	0.91	3.84	9.67
	PBT	16	3810	97	0.91	4.06	9.80
GO	CBT	69	2493	96	0.93	3.72	6.79
	PBT	16	1083	96	0.93	4.32	6.82
A2	CBT	36	4606	104	0.92	3.90	10.30
	PBT	16	539	104	0.91	4.33	10.98
M1	CBT	4	6507	93	0.91	4.02	10.53
	PBT	4	678	90	0.92	4.16	9.81
M2	CBT	4	2775	95	0.90	3.55	9.42
	PBT	4	205	96	0.95	4.13	8.07
M3	CBT	4	1743	98	0.91	3.85	11.28
	PBT	2	356	103	0.85	4.32	12.49

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

8.4 Reliability Results for Subgroups of Interest

When sample size was sufficient, score reliability and scale score SEM were estimated for the groups identified for DIF analysis. Estimates were calculated only for groups of 100 or more students administered a specific test form.

Tables 8.3 through 8.11 summarize test reliability for groups of interest for English Language Arts/Literacy grades 3-11, and Tables 8.12 through 8.23 summarize test reliability for groups of interest for mathematics grades/subjects. Note that reliability estimates are dependent on score variance, and subgroups with smaller variance are likely to have lower reliability estimates than the total group.

Gender

English Language Arts/Literacy

Both the average reliability estimates and average SEMs for males and females are similar to the corresponding reliabilities and SEMs for the total group. Ten of the eighteen reliabilities are .01 higher for males than for females. The SEMs for females are all higher than for males.

Mathematics

As with the English Language Arts/Literacy test components, the reliability estimates and SEMs for males and females reflect the corresponding reliabilities and SEMs for the total group. Typically, the reliabilities are .01 higher for males than for females. The SEMs are generally very similar for females and males.

Ethnicity

English Language Arts/Literacy

The majority of the reliabilities for the ethnicity groups are .01-.02 lower than for the total group. There is not a consistent difference among the test reliabilities for White, African-American, Asian/Pacific Islander, Hispanic, and multiple ethnicity students, with the majority of the reliabilities between .89 and .91. However, the majority of the reliabilities for American Indian/Alaskan Native students range from .85 to .88. In general, the SEMs are similar to the total group SEMs. Nevertheless, for each most levels, the SEMs are highest for Asian/Pacific Islander students.

Mathematics

As with the English Language Arts/Literacy reliabilities, the reliabilities for ethnicity groups are marginally lower than for the total group of students. Once again the average SEMs reflect the total group SEMs. While there is variation across tests, the average reliabilities are generally highest for Asian/Pacific Islander students. For the EOC tests, the African-American group has the lowest reliabilities and for the EOY tests the American Indian/Alaskan Native groups has the lowest reliabilities.

Special Education Needs

English Arts/Literacy

The reliabilities for five groups of students (Economically Disadvantaged, Not Economically Disadvantaged, Non English Learner, Students with Disabilities, and Students without Disabilities) are generally .01 to .02 less than those for the total group of students. The majority of the reliabilities range from .89 to .91. The average reliabilities for English Learner students are lower, most often ranging from .86 to .88. The SEMs are generally similar to the total group SEMs, however, for 17 of the 18 sets of SEMs, the lowest SEM is for Students with Disabilities.

Mathematics

The average reliabilities for the larger student groups (Not Economically Disadvantaged, Non English Learner, and Students without Disabilities) are quite similar to the total group of students. For Economically Disadvantaged, English Learners, and Students with Disabilities, the average reliabilities average .05 lower than for the total group reliabilities. For the grade 3-8 tests, the Students with Disabilities have the highest SEMs. For the EOC courses, the majority of the highest SEMs are for English Learners.

Students Taking Accommodated Forms

English Arts/Literacy

Two of the four accommodation form types had sufficient sample sizes to allow for estimation of reliability and SEM. The other two groups did not have at least 100 students take any specific form. Within grades, the reliabilities of the Closed Caption forms, which range from .92 to .95, are higher than the average reliabilities for total group. For the Text-to-Speech forms, the reliabilities, which range from .83 to .87, are lower than for the total group reliabilities.

Mathematics

Only the Text-to-Speech forms had sufficient sample sizes for reliability and SEM estimation. With the exception of the Integrated Mathematics I and III courses, the Text-to-Speech reliabilities are very close to the total group reliabilities. The corresponding SEMs were somewhat higher than for the total group SEMs.

Students Taking Translated Forms

Mathematics

With the exception of Integrated Mathematics II and III, there were sufficient numbers of students taking the Spanish Language form for reliability and SEM estimation. For the six EOY forms, the reliabilities average .09 less than for the total group, with the largest differences being for grades 7 and 8. The corresponding SEMs are generally higher for the students administered the Spanish language forms. For the EOC forms, the reliability and SEM differences with the total group estimates are considerably greater than for the EOY forms.

Table 8.3 Summary of Test Reliability Estimates for Subgroups: Grade 3 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	10276	100	0.90	12.20	8042	100	0.90	12.47
Gender								
Male	5245	100	0.90	12.00	4076	100	0.90	12.19
Female	5031	100	0.90	12.39	3966	100	0.90	12.73
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	5149	100	0.89	12.24	3160	100	0.89	12.56
African American	1652	100	0.89	11.94	2084	100	0.89	12.35
Asian/Pacific Islander	599	100	0.89	13.03	332	100	0.89	13.12
American Indian/Alaska Native	123	100	0.88	11.37	308	100	0.85	12.49
Hispanic	2423	100	0.89	12.01	2079	100	0.89	12.30
Multiple	188	100	0.91	12.39	115	100	0.92	12.54
Special Instructional Needs								
Economically Disadvantaged	4516	100	0.88	11.97	4457	100	0.89	12.31
Not-economically Disadvantaged	5079	100	0.89	12.44	3047	100	0.89	12.73
English Learner (EL)	1215	100	0.86	11.94	1300	100	0.86	12.28
Non English Learner	8380	100	0.90	12.28	6208	100	0.91	12.54
Students with Disabilities (SWD)	943	100	0.91	11.40	584	100	0.89	11.73
Students without Disabilities	6873	100	0.90	12.29	3703	100	0.90	12.60
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	162	100	0.92	12.21	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	1435	100	0.85	10.88	-	-	-	-

Table 8.4 Summary of Test Reliability Estimates for Subgroups: Grade 4 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13108	105	0.91	9.84	8942	105	0.89	10.61
Gender								
Male	6688	105	0.91	9.69	4523	105	0.89	10.45
Female	6420	105	0.91	9.99	4419	105	0.89	10.76
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	6969	105	0.90	9.82	4255	105	0.88	10.56
African American	2090	105	0.89	9.78	1916	105	0.86	10.67
Asian/Pacific Islander	674	105	0.90	10.34	347	105	0.89	10.93
American Indian/Alaska Native	115	105	0.90	9.62	220	106	0.83	10.66
Hispanic	2756	105	0.90	9.84	1599	105	0.87	10.64
Multiple	270	105	0.91	9.84	148	105	0.88	10.52
Special Instructional Needs								
Economically Disadvantaged	5770	105	0.89	9.77	4518	105	0.87	10.60
Not-economically Disadvantaged	6600	105	0.90	9.96	3545	105	0.87	10.66
English Learner (EL)	923	105	0.87	9.89	680	105	0.83	10.92
Non English Learner	11473	105	0.91	9.85	7208	105	0.89	10.60
Students with Disabilities (SWD)	1315	105	0.91	9.56	746	105	0.87	10.46
Students without Disabilities	9158	105	0.90	9.86	4183	105	0.88	10.59
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	240	106	0.94	9.15	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	981	105	0.86	9.93	-	-	-	-

Table 8.5 Summary of Test Reliability Estimates for Subgroups: Grade 5 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13579	106	0.91	9.67	8351	105	0.89	10.20
Gender								
Male	6932	106	0.91	9.51	4229	105	0.89	9.98
Female	6647	106	0.90	9.82	4122	105	0.89	10.40
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7217	106	0.90	9.62	3721	105	0.88	10.15
African American	2141	106	0.89	9.66	1840	105	0.87	10.30
Asian/Pacific Islander	701	106	0.90	9.96	302	105	0.89	10.59
American Indian/Alaska Native	120	106	0.88	9.62	206	105	0.85	10.02
Hispanic	2765	106	0.89	9.76	1434	105	0.87	10.19
Multiple	247	106	0.91	9.70	121	105	0.88	10.25
Special Instructional Needs								
Economically Disadvantaged	5840	106	0.89	9.72	4208	105	0.87	10.23
Not-economically Disadvantaged	6936	106	0.90	9.69	3383	105	0.87	10.22
English Learner (EL)	741	106	0.85	9.84	478	105	0.83	10.36
Non English Learner	12060	106	0.91	9.67	6931	105	0.89	10.21
Students with Disabilities (SWD)	1354	106	0.89	9.24	757	105	0.87	9.88
Students without Disabilities	9243	106	0.90	9.66	4167	105	0.88	10.20
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	238	106	0.93	9.28	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	944	106	0.85	8.74	-	-	-	-

Table 8.6 Summary of Test Reliability Estimates for Subgroups: Grade 6 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	14105	137	0.92	8.54	7129	137	0.92	8.31
Gender								
Male	7198	137	0.92	8.40	3625	137	0.92	8.08
Female	6907	137	0.92	8.67	3504	137	0.91	8.52
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7320	137	0.91	8.48	3480	137	0.91	8.27
African American	2280	137	0.91	8.55	1609	137	0.91	8.33
Asian/Pacific Islander	711	137	0.92	8.88	264	137	0.92	8.67
American Indian/Alaska Native	113	137	0.91	8.52	192	137	0.88	8.69
Hispanic	3011	137	0.91	8.63	797	137	0.91	8.35
Multiple	233	137	0.92	8.54	123	137	0.92	8.13
Special Instructional Needs								
Economically Disadvantaged	6138	137	0.91	8.58	3209	137	0.91	8.29
Not-economically Disadvantaged	7127	137	0.91	8.56	3220	137	0.91	8.37
English Learner (EL)	673	137	0.88	8.74	266	137	0.87	8.66
Non English Learner	12611	137	0.92	8.53	5965	137	0.92	8.32
Students with Disabilities (SWD)	1286	137	0.91	8.32	772	137	0.90	8.15
Students without Disabilities	9197	137	0.91	8.52	4354	137	0.91	8.34
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	208	137	0.95	8.61	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	945	137	0.88	8.25	-	-	-	-

Table 8.7 Summary of Test Reliability Estimates for Subgroups: Grade 7 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	14134	135	0.93	9.70	6559	135	0.91	10.13
Gender								
Male	7216	135	0.93	9.54	3324	135	0.91	9.93
Female	6918	135	0.92	9.83	3235	135	0.90	10.27
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7393	135	0.92	9.67	3328	135	0.90	10.07
African American	2404	135	0.91	9.60	1276	135	0.89	10.27
Asian/Pacific Islander	683	135	0.93	10.04	254	135	0.91	10.32
American Indian/Alaska Native	121	135	0.91	9.67	151	135	0.86	10.32
Hispanic	2885	135	0.92	9.74	792	135	0.90	10.21
Multiple	217	135	0.93	9.64	118	135	0.91	9.89
Special Instructional Needs								
Economically Disadvantaged	6084	135	0.91	9.70	2764	135	0.89	10.18
Not-economically Disadvantaged	7189	135	0.92	9.74	3138	135	0.90	10.13
English Learner (EL)	659	135	0.88	9.81	264	135	0.86	10.87
Non English Learner	12633	135	0.93	9.70	5429	135	0.91	10.10
Students with Disabilities (SWD)	1292	135	0.91	9.40	648	135	0.88	10.25
Students without Disabilities	9309	135	0.92	9.68	3842	135	0.90	10.08
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	276	135	0.94	9.60	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	2008	135	0.87	9.52	-	-	-	-

Table 8.8 Summary of Test Reliability Estimates for Subgroups: Grade 8 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13846	136	0.92	10.40	6735	136	0.92	10.11
Gender								
Male	7092	136	0.92	10.24	3418	136	0.92	9.95
Female	6754	136	0.91	10.52	3317	136	0.91	10.25
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7230	136	0.91	10.39	3476	136	0.91	10.06
African American	2366	136	0.91	10.29	1333	136	0.90	10.09
Asian/Pacific Islander	683	136	0.91	10.81	250	136	0.92	10.34
American Indian/Alaska Native	112	136	0.90	10.11	157	136	0.85	11.19
Hispanic	2758	136	0.91	10.38	734	136	0.91	10.18
Multiple	192	136	0.92	10.39	133	136	0.91	9.91
Special Instructional Needs								
Economically Disadvantaged	5834	136	0.91	10.31	2687	136	0.90	10.07
Not-economically Disadvantaged	7087	136	0.91	10.49	3374	136	0.90	10.13
English Learner (EL)	637	136	0.87	10.44	263	136	0.86	10.71
Non English Learner	12304	136	0.92	10.40	5576	136	0.91	10.09
Students with Disabilities (SWD)	1237	136	0.90	10.07	668	136	0.90	9.91
Students without Disabilities	8942	136	0.91	10.41	4155	136	0.90	10.16
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	257	137	0.93	10.39	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	1214	136	0.86	10.07	-	-	-	-

Table 8.9 Summary of Test Reliability Estimates for Subgroups: Grade 9 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	5152	137	0.93	9.64	3821	137	0.92	9.47
Gender								
Male	2622	137	0.93	9.51	1910	137	0.92	9.37
Female	2531	137	0.92	9.72	1910	137	0.91	9.54
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	2693	137	0.92	9.61	2020	137	0.91	9.42
African American	648	137	0.91	9.59	498	137	0.90	9.48
Asian/Pacific Islander	262	137	0.93	10.06	130	137	0.93	9.74
American Indian/Alaska Native	130	135	0.85	9.92	-	-	-	-
Hispanic	1128	137	0.91	9.69	255	137	0.91	9.66
Multiple	106	136	0.94	9.56	106	137	0.93	9.39
Special Instructional Needs								
Economically Disadvantaged	2123	137	0.91	9.62	1355	137	0.91	9.47
Not-economically Disadvantaged	2870	137	0.92	9.69	1993	137	0.91	9.47
English Learner (EL)	266	137	0.87	9.82	137	137	0.86	9.72
Non English Learner	4716	137	0.92	9.62	3098	137	0.91	9.47
Students with Disabilities (SWD)	471	137	0.88	9.59	340	137	0.89	9.46
Students without Disabilities	3518	137	0.92	9.63	2278	137	0.91	9.39
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	169	135	0.93	8.99	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	1275	136	0.83	9.81	-	-	-	-

Table 8.10 Summary of Test Reliability Estimates for Subgroups: Grade 10 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	3628	137	0.93	11.64	1228	137	0.93	11.34
Gender								
Male	1847	137	0.93	11.44	583	137	0.93	11.13
Female	1781	137	0.92	11.79	645	137	0.92	11.50
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	1757	137	0.93	11.73	491	137	0.93	11.27
African American	658	137	0.91	11.37	557	137	0.91	11.35
Asian/Pacific Islander	221	137	0.93	12.22	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	792	137	0.92	11.50	125	137	0.94	11.17
Multiple	-	-	-	-	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	1305	137	0.92	11.41	425	137	0.92	11.35
Not-economically Disadvantaged	2011	137	0.93	11.85	496	137	0.93	11.39
English Learner (EL)	157	137	0.89	10.98	-	-	-	-
Non English Learner	3382	137	0.93	11.66	1179	137	0.93	11.35
Students with Disabilities (SWD)	333	137	0.90	10.83	173	137	0.89	10.83
Students without Disabilities	2746	137	0.92	11.76	738	137	0.93	11.41
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	110	137	0.95	10.53	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	507	137	0.85	10.76	-	-	-	-

Table 8.11 Summary of Test Reliability Estimates for Subgroups: Grade 11 ELA/L

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	4169	136	0.92	10.73	428	136	0.91	11.11
Gender								
Male	2127	136	0.92	10.59	192	136	0.91	10.86
Female	2042	136	0.92	10.79	236	136	0.90	11.24
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	2184	136	0.92	10.75	226	136	0.90	11.01
African American	432	136	0.90	10.74	123	135	0.90	10.62
Asian/Pacific Islander	250	136	0.93	11.19	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	1079	136	0.91	10.63	-	-	-	-
Multiple	-	-	-	-	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	1563	136	0.90	10.63	199	136	0.89	11.04
Not-economically Disadvantaged	2517	136	0.92	10.82	217	136	0.91	11.14
English Learner (EL)	161	136	0.86	10.55	-	-	-	-
Non English Learner	3919	136	0.92	10.74	400	136	0.91	11.06
Students with Disabilities (SWD)	363	136	0.88	10.38	125	135	0.92	10.11
Students without Disabilities	2814	136	0.92	10.81	149	136	0.89	11.31
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	349	136	0.85	9.63	-	-	-	-

Table 8.12 Summary of Test Reliability Estimates for Subgroups: Grade 3 Mathematics

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	12438	81	0.94	8.08	10993	82	0.94	8.27
Gender								
Male	6347	81	0.94	8.09	5552	82	0.94	8.29
Female	6090	81	0.93	8.05	5441	82	0.94	8.23
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	6596	81	0.93	7.98	4953	82	0.93	8.15
African American	1946	81	0.92	8.16	2521	82	0.92	8.32
Asian/Pacific Islander	648	81	0.94	8.12	397	82	0.94	8.34
American Indian/Alaska Native	395	82	0.91	8.14	123	82	0.92	8.30
Hispanic	2658	81	0.92	8.13	2167	82	0.92	8.31
Multiple	269	81	0.94	8.06	223	82	0.94	8.15
Special Instructional Needs								
Economically Disadvantaged	5557	81	0.92	8.15	5755	82	0.92	8.33
Not-economically Disadvantaged	6136	81	0.93	8.00	4246	82	0.93	8.16
English Learner (EL)	1405	81	0.92	8.15	1462	82	0.92	8.33
Non English Learner	10311	81	0.94	8.08	8345	82	0.94	8.27
Students with Disabilities (SWD)	1179	81	0.93	8.37	908	82	0.92	8.64
Students without Disabilities	8688	81	0.94	7.97	5829	82	0.94	8.15
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	6267	82	0.94	8.17	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	3888	82	0.90	8.59	1528	81	0.90	8.86

Table 8.13 Summary of Test Reliability Estimates for Subgroups: Grade 4 Mathematics

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13080	82	0.94	7.36	9137	82	0.94	7.64
Gender								
Male	6687	82	0.94	7.33	4614	82	0.94	7.62
Female	6393	82	0.94	7.37	4523	82	0.93	7.66
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	6916	82	0.93	7.34	4322	82	0.93	7.65
African American	2121	82	0.92	7.39	1882	82	0.91	7.57
Asian/Pacific Islander	677	82	0.94	7.48	357	82	0.94	7.84
American Indian/Alaska Native	404	82	0.91	7.40	171	82	0.90	7.78
Hispanic	2764	82	0.92	7.36	1723	82	0.92	7.59
Multiple	266	82	0.94	7.32	174	82	0.93	7.66
Special Instructional Needs								
Economically Disadvantaged	5813	82	0.92	7.39	4567	82	0.92	7.59
Not-economically Disadvantaged	6538	82	0.94	7.37	3653	82	0.93	7.73
English Learner (EL)	969	82	0.91	7.36	773	82	0.91	7.54
Non English Learner	11408	82	0.94	7.37	7278	82	0.94	7.66
Students with Disabilities (SWD)	1320	82	0.92	7.47	766	82	0.92	7.61
Students without Disabilities	9111	82	0.94	7.31	4406	82	0.93	7.64
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	6375	82	0.94	7.37	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	2301	82	0.88	7.59	608	82	0.90	7.54

Table 8.14 Summary of Test Reliability Estimates for Subgroups: Grade 5 Mathematics

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13105	82	0.93	7.58	8507	82	0.94	7.66
Gender								
Male	6705	82	0.94	7.57	4300	82	0.94	7.65
Female	6400	82	0.93	7.56	4207	82	0.93	7.67
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7117	82	0.93	7.48	3807	82	0.93	7.63
African American	2170	82	0.91	7.91	1775	82	0.91	7.68
Asian/Pacific Islander	700	82	0.94	7.35	311	82	0.95	7.80
American Indian/Alaska Native	397	82	0.90	8.09	145	82	0.88	7.67
Hispanic	2675	82	0.91	7.85	1541	82	0.91	7.69
Multiple	241	82	0.94	7.53	158	82	0.93	7.54
Special Instructional Needs								
Economically Disadvantaged	5685	82	0.91	7.88	4208	82	0.92	7.69
Not-economically Disadvantaged	6646	82	0.93	7.43	3510	82	0.93	7.66
English Learner (EL)	751	82	0.90	8.09	535	82	0.90	7.70
Non English Learner	11604	82	0.93	7.56	7010	82	0.94	7.67
Students with Disabilities (SWD)	1345	82	0.91	8.16	761	82	0.91	7.69
Students without Disabilities	8880	82	0.93	7.45	4393	82	0.94	7.63
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	5840	82	0.94	7.67	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	1895	81	0.87	8.59	329	82	0.85	7.73

Table 8.15 Summary of Test Reliability Estimates for Subgroups: Grade 6 Mathematics

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13585	82	0.94	7.07	6768	82	0.94	7.42
Gender								
Male	6950	82	0.94	7.05	3436	82	0.94	7.41
Female	6635	82	0.94	7.06	3332	82	0.94	7.40
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7213	82	0.93	6.95	3385	82	0.93	7.31
African American	2323	82	0.91	7.45	1322	82	0.91	7.64
Asian/Pacific Islander	705	82	0.95	6.94	255	82	0.95	7.44
American Indian/Alaska Native	333	82	0.93	7.37	103	82	0.84	8.25
Hispanic	2877	82	0.92	7.30	847	82	0.92	7.53
Multiple	232	82	0.94	7.04	176	82	0.93	7.41
Special Instructional Needs								
Economically Disadvantaged	5997	82	0.92	7.35	2903	82	0.92	7.55
Not-economically Disadvantaged	6815	82	0.94	6.93	3171	82	0.94	7.33
English Learner (EL)	678	82	0.91	7.57	313	82	0.91	7.76
Non English Learner	12436	82	0.94	7.03	5698	82	0.94	7.40
Students with Disabilities (SWD)	1295	82	0.92	7.54	707	82	0.90	7.74
Students without Disabilities	8805	82	0.94	6.94	4206	82	0.94	7.37
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	6202	82	0.94	7.18	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	1584	81	0.88	7.96	118	82	0.80	8.52

Table 8.16 Summary of Test Reliability Estimates for Subgroups: Grade 7 Mathematics

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	13791	82	0.93	7.20	5629	80	0.93	7.28
Gender								
Male	7052	82	0.93	7.18	2856	80	0.93	7.26
Female	6740	82	0.92	7.18	2773	80	0.92	7.29
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	7141	82	0.92	6.98	2962	80	0.93	7.13
African American	2441	82	0.89	7.88	936	80	0.88	7.85
Asian/Pacific Islander	605	82	0.94	6.73	219	80	0.95	7.08
American Indian/Alaska Native	337	82	0.89	8.34	107	82	0.84	8.37
Hispanic	2836	82	0.90	7.65	764	80	0.90	7.60
Multiple	204	82	0.92	7.20	161	82	0.91	7.49
Special Instructional Needs								
Economically Disadvantaged	6089	82	0.90	7.70	2327	80	0.89	7.69
Not-economically Disadvantaged	6822	82	0.93	6.91	2793	80	0.93	7.07
English Learner (EL)	690	82	0.87	8.23	310	80	0.87	8.02
Non English Learner	12189	82	0.93	7.14	4755	80	0.93	7.23
Students with Disabilities (SWD)	1306	82	0.88	8.31	579	80	0.87	8.23
Students without Disabilities	8903	82	0.92	6.97	3318	80	0.93	7.10
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	6152	82	0.93	7.60	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	1885	82	0.77	10.21	174	82	0.77	9.76

Table 8.17 Summary of Test Reliability Estimates for Subgroups: Grade 8 Mathematics

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	11240	82	0.91	10.22	5129	81	0.92	10.36
Gender								
Male	5829	82	0.92	10.19	2619	81	0.92	10.38
Female	5411	82	0.91	10.20	2511	81	0.91	10.31
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	5567	82	0.91	10.05	2663	81	0.92	10.16
African American	2201	82	0.87	10.91	952	81	0.88	10.92
Asian/Pacific Islander	417	82	0.94	9.46	201	81	0.94	9.79
American Indian/Alaska Native	294	82	0.86	10.80	-	-	-	-
Hispanic	2424	82	0.89	10.75	688	81	0.90	10.74
Multiple	154	82	0.91	9.96	195	80	0.90	10.55
Special Instructional Needs								
Economically Disadvantaged	5291	82	0.88	10.83	2152	81	0.88	10.87
Not-economically Disadvantaged	5101	82	0.92	9.91	2470	81	0.92	10.08
English Learner (EL)	650	82	0.84	11.23	290	81	0.87	11.03
Non English Learner	9705	82	0.91	10.18	4271	81	0.92	10.30
Students with Disabilities (SWD)	1205	82	0.84	11.35	579	81	0.85	11.38
Students without Disabilities	6433	82	0.90	10.12	2986	81	0.92	10.17
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	5167	82	0.92	10.17	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	1842	82	0.77	12.35	150	80	0.65	14.23

Table 8.18 Summary of Test Reliability Estimates for Subgroups: Algebra I

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	6058	96	0.91	9.67	3810	97	0.91	9.80
Gender								
Male	3113	96	0.91	9.63	1889	97	0.91	9.81
Female	2945	96	0.90	9.69	1921	97	0.90	9.76
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	3015	96	0.90	9.51	2041	97	0.90	9.65
African American	1165	96	0.84	10.56	640	97	0.84	10.54
Asian/Pacific Islander	305	96	0.94	8.82	144	97	0.94	9.26
American Indian/Alaska Native	175	97	0.84	10.36	-	-	-	-
Hispanic	1211	96	0.85	10.52	301	97	0.88	10.25
Multiple	357	97	0.91	9.82	110	97	0.88	10.12
Special Instructional Needs								
Economically Disadvantaged	2461	96	0.85	10.48	1119	97	0.87	10.33
Not-economically Disadvantaged	3135	96	0.91	9.42	2050	97	0.91	9.54
English Learner (EL)	327	96	0.80	10.86	268	97	0.83	10.91
Non English Learner	5519	96	0.91	9.60	3241	97	0.91	9.75
Students with Disabilities (SWD)	563	96	0.82	10.57	300	97	0.83	10.78
Students without Disabilities	4259	96	0.91	9.48	2308	97	0.91	9.52
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	2908	96	0.89	10.05	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	2043	96	0.58	13.40	-	-	-	-

Table 8.19 Summary of Test Reliability Estimates for Subgroups: Geometry

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	2493	96	0.93	6.79	1083	96	0.93	6.82
Gender								
Male	1266	96	0.94	6.74	512	96	0.94	6.83
Female	1227	96	0.92	6.82	571	96	0.93	6.79
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	1312	96	0.93	6.75	612	96	0.92	6.61
African American	300	96	0.86	7.81	141	96	0.88	7.47
Asian/Pacific Islander	170	96	0.95	6.33	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	584	96	0.88	7.55	130	96	0.90	7.70
Multiple	115	95	0.92	6.70	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	927	96	0.88	7.45	277	96	0.91	7.30
Not-economically Disadvantaged	1516	96	0.93	6.71	694	96	0.93	6.60
English Learner (EL)	226	96	0.84	8.38	-	-	-	-
Non English Learner	2324	96	0.93	6.75	933	96	0.93	6.79
Students with Disabilities (SWD)	223	96	0.85	7.78	152	96	0.86	8.21
Students without Disabilities	2120	96	0.93	6.69	885	96	0.93	6.72
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	865	96	0.92	7.21	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	806	96	0.74	10.23	-	-	-	-

Table 8.20 Summary of Test Reliability Estimates for Subgroups: Algebra II

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	4606	104	0.92	10.30	539	104	0.91	10.98
Gender								
Male	2293	104	0.93	10.18	283	105	0.92	10.91
Female	2313	104	0.91	10.41	280	104	0.90	11.05
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	2351	104	0.92	10.35	332	104	0.91	10.87
African American	682	104	0.85	11.43	106	105	0.77	13.05
Asian/Pacific Islander	361	104	0.94	9.59	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	1001	104	0.88	11.03	112	105	0.88	11.20
Multiple	169	104	0.92	10.16	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	1574	104	0.87	11.19	178	105	0.85	11.57
Not-economically Disadvantaged	2970	104	0.93	10.18	352	104	0.91	10.88
English Learner (EL)	166	104	0.88	10.89	-	-	-	-
Non English Learner	4409	104	0.92	10.27	491	104	0.91	10.97
Students with Disabilities (SWD)	262	104	0.86	10.89	112	107	0.87	11.08
Students without Disabilities	3442	104	0.92	10.17	299	104	0.91	10.79
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	1370	104	0.91	10.77	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	518	101	0.74	13.15	-	-	-	-

Table 8.21 Summary of Test Reliability Estimates for Subgroups: Integrated Mathematics I

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	6507	93	0.91	10.53	678	90	0.92	9.81
Gender								
Male	3356	93	0.91	10.49	351	90	0.93	9.69
Female	3152	93	0.90	10.53	327	90	0.92	9.90
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	2867	93	0.91	10.20	316	90	0.92	9.78
African American	730	93	0.84	11.78	435	91	0.80	11.01
Asian/Pacific Islander	194	93	0.93	9.59	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	1838	93	0.86	11.26	-	-	-	-
Multiple	164	92	0.91	10.58	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	3005	93	0.86	11.34	329	90	0.87	10.20
Not-economically Disadvantaged	2888	93	0.92	10.07	306	90	0.92	9.64
English Learner (EL)	597	93	0.74	12.26	-	-	-	-
Non English Learner	5295	93	0.91	10.39	608	90	0.92	9.76
Students with Disabilities (SWD)	640	93	0.82	11.62	307	91	0.67	11.97
Students without Disabilities	3800	93	0.91	10.18	525	90	0.92	9.53
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	2107	92	0.85	11.75	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	185	92	0.65	13.30	-	-	-	-

Table 8.22 Summary of Test Reliability Estimates for Subgroups: Integrated Mathematics II

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	2775	95	0.90	9.42	205	96	0.95	8.07
Gender								
Male	1420	95	0.91	9.28	116	96	0.95	8.05
Female	1355	95	0.90	9.58	121	96	0.94	8.26
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	1226	95	0.91	9.42	103	96	0.95	8.08
African American	175	95	0.84	10.34	-	-	-	-
Asian/Pacific Islander	104	95	0.94	8.45	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	724	95	0.83	10.14	-	-	-	-
Multiple	-	-	-	-	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	903	95	0.84	10.09	111	97	0.83	8.98
Not-economically Disadvantaged	1437	95	0.91	9.33	106	94	0.92	7.96
English Learner (EL)	165	95	0.57	12.64	-	-	-	-
Non English Learner	2176	95	0.91	9.33	167	96	0.95	8.05
Students with Disabilities (SWD)	213	95	0.78	10.92	-	-	-	-
Students without Disabilities	1889	95	0.91	9.29	165	96	0.94	7.92
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	701	95	0.90	9.65	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	-	-	-	-	-	-	-	-

Table 8.23 Summary of Test Reliability Estimates for Subgroups: Integrated Mathematics III

	CBT				PBT			
	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM	Average Sample Size	Average Maximum Possible Raw Scores	Average Reliability	Average Scale Score SEM
Total Group	1743	98	0.91	11.28	356	103	0.85	12.49
Gender								
Male	884	98	0.92	11.06	156	103	0.88	12.53
Female	859	98	0.91	11.50	200	103	0.83	12.49
Unknown/Missing	-	-	-	-	-	-	-	-
Ethnicity								
White	733	98	0.92	11.08	176	103	0.87	12.11
African American	-	-	-	-	-	-	-	-
Asian/Pacific Islander	-	-	-	-	-	-	-	-
American Indian/Alaska Native	-	-	-	-	-	-	-	-
Hispanic	448	98	0.88	12.11	-	-	-	-
Multiple	-	-	-	-	-	-	-	-
Special Instructional Needs								
Economically Disadvantaged	528	98	0.89	11.90	201	103	0.82	12.88
Not-economically Disadvantaged	852	98	0.93	11.10	146	103	0.87	12.21
English Learner (EL)	-	-	-	-	-	-	-	-
Non English Learner	1303	98	0.92	11.20	335	103	0.85	12.44
Students with Disabilities (SWD)	149	100	0.83	11.11	-	-	-	-
Students without Disabilities	1026	98	0.92	11.09	-	-	-	-
Students Taking Accommodated Forms								
A: ASL	-	-	-	-	-	-	-	-
C: Closed-Caption	-	-	-	-	-	-	-	-
R: Screen Reader	-	-	-	-	-	-	-	-
T: Text-to-Speech	156	100	0.83	11.78	-	-	-	-
Students Taking Translated Forms								
Spanish Language Form	-	-	-	-	-	-	-	-

8.5 Reliability Results for English Language Arts/Literacy Subscores

PARCC developed subclaims in addition to major claims based on the Common Core State Standards. English Language Arts/Literacy has two Major Claims relating to Reading Complex Text and Writing. The Major Claim for Reading Complex Text is that students read and comprehend a range of sufficiently complex texts independently. The Major Claim for Writing is that students write effectively when using and/or analyzing sources. Refer to Table 8.24 for a summary of the English language arts/literacy claims and subclaims.

Table 8.24 Descriptions of ELA/L Claims and Subclaims

English Language Arts/Literacy		
Major Claim	Subclaim	Description
Reading	Reading Literature	Students demonstrate comprehension and draw evidence from readings of grade-level, complex literary text
Reading	Reading Information	Students demonstrate comprehension and draw evidence from readings of grade-level, complex informational text
Reading	Reading Vocabulary	Students use context to determine the meaning of words and phrases
Writing	Writing Written Expression	Students produce clear and coherent writing in which the development, organization, and style are appropriate to the task, purpose, and audience
Writing	Writing Knowledge Language and Conventions	Students demonstrate knowledge of conventions and other important elements of language

Reliability indices were calculated for each major claim and subclaim. Table 8.25 presents the average reliability estimates for all forms of the test at the specified grade and testing mode for the English Language Arts/Literacy tests. In order to assist in understanding the reliability estimates, the average maximum number of points for each major claim and subclaim are also provided.

The reliabilities for the Reading Complex Text claim, are larger for grades 6-11, which average 91 possible points, than for grades 3-5, which average 68 possible points. The reliability for grades 3-5 ranges from .87 to .90 and the average reliability for grades 6-11 ranges from .90 to .92.

The Writing claim reliabilities are lower than for the Reading claim. All the reliabilities for the Writing claim for grades 3-5 are based on 36 points and the all the average reliabilities for the grades 6-11 Writing claims are based on 45 points. The reliability for grades 3-5 ranges from .78 to .82 with a median of .80, and the average reliability for grades 6-11 ranges from .80 to .85, with a median of .84. Taking the number of points into consideration, the per-point information of the two claims are quite similar, as are the per-point information when comparing grades 3-5 with grades 6-11.

Reliability of the Reading Literature subclaim scores over testing mode and grade has a median average of .78, and the reliabilities vary from .71 to .83. For grades 3-5, the Reading Information subclaim reliabilities are based on 25 points and have a median of .73. For grades 6-11, the Reading Information subclaim is based on an average of 37 points, and the median subclaim reliability is .80. Once again, when taking the number of points into consideration, the per-point information of the claim is quite similar when comparing grades 3-5 with grades 6-11. The Reading Vocabulary subclaim is based on the fewest points, ranging from 12 to 22 points. The average subclaim reliability has a median of .67 and ranges from .58 to .74.

The Writing: Written Expression subclaim is based on 27 points for grades 3-5 and 36 points for grades 6-11. The median average reliability for grades 3-5 is .73, and the median average reliability for grade 6-11 is .82. The Writing: Knowledge of Language and Conventions subclaims are all based on nine points. The average reliabilities are consistent, varying from .76 to .83, with a median of .80.

Table 8.25 Average ELA/L Reliability Estimates for Total Test and Subscores

Grade Level	Testing Mode	Reading: Total		Reading: Literature		Reading: Information		Reading: Vocabulary		Writing: Total		Writing: Written Expression		Writing: Knowledge Language and Conventions	
		Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability
3	CBT	64	0.90	29	0.83	22	0.69	13	0.64	36	0.78	27	0.67	9	0.78
	PBT	64	0.90	31	0.83	21	0.64	12	0.69	36	0.79	27	0.70	9	0.77
4	CBT	69	0.89	28	0.77	27	0.77	15	0.65	36	0.82	27	0.77	9	0.80
	PBT	69	0.87	27	0.71	26	0.74	17	0.63	36	0.80	27	0.74	9	0.77
5	CBT	70	0.88	27	0.75	26	0.75	17	0.63	36	0.82	27	0.76	9	0.81
	PBT	69	0.87	24	0.71	27	0.77	18	0.58	36	0.80	27	0.72	9	0.76
6	CBT	92	0.90	30	0.78	42	0.80	19	0.67	45	0.84	36	0.82	9	0.83
	PBT	92	0.92	31	0.83	40	0.80	21	0.71	45	0.82	36	0.78	9	0.80
7	CBT	90	0.91	33	0.79	37	0.80	20	0.70	45	0.84	36	0.82	9	0.83
	PBT	90	0.91	35	0.80	33	0.78	22	0.71	45	0.80	36	0.75	9	0.77
8	CBT	91	0.90	35	0.75	37	0.78	19	0.67	45	0.85	36	0.83	9	0.83
	PBT	91	0.90	39	0.78	33	0.77	19	0.74	45	0.83	36	0.80	9	0.80
9	CBT	92	0.91	35	0.78	39	0.82	19	0.67	45	0.85	36	0.82	9	0.82
	PBT	92	0.91	35	0.78	37	0.80	20	0.67	45	0.83	36	0.79	9	0.80
10	CBT	92	0.92	32	0.78	40	0.84	20	0.69	45	0.85	36	0.83	9	0.83
	PBT	92	0.92	34	0.80	39	0.85	20	0.67	45	0.84	36	0.82	9	0.81
11	CBT	91	0.90	35	0.81	36	0.78	20	0.64	45	0.85	36	0.83	9	0.83
	PBT	91	0.90	33	0.79	36	0.78	22	0.70	45	0.82	36	0.78	9	0.78

8.6 Reliability Results for Mathematics Subscores

For mathematics, there are four subclaims related to the major claim that students are on track or ready for college and careers:

- Subclaim A: Students solve problems involving the *major content* for their grade level with connections to the Standards for Mathematical Practice.
- Subclaim B: Students solve problems involving the *additional and supporting content* for their grade level with connections to the Standards for Mathematical Practice.
- Subclaim C: Students express grade/course-level appropriate *mathematical reasoning* by constructing viable mathematical arguments and critiquing the reasoning of others, and/or attending to precision when making mathematical statements
- Subclaim D: Students solve real-world problems with a degree of difficulty appropriate to the grade/course by applying knowledge and skills articulated in the standards and by engaging particularly in the *modeling practice*.

Reliability estimates were calculated for each subclaim for mathematics. Table 8.26 presents the average reliability estimates for mathematics subclaims by mode (CBT and PBT) and grade/subject.

Subclaims with greater numbers of points tend to have greater internal consistency reliability estimates. The Major Content subclaim has the largest number of points for each assessment and accordingly has higher average reliabilities than the other three subclaims. For grades 3 through 8, the average reliability for the subclaim is .89. As with the total test reliabilities, the Major Content reliabilities were lower for the six EOC assessments than for grade level test assessments. The average reliability for the Major Content subclaim for the EOC tests is .82.

The reliabilities for the Additional and Supporting Content subclaim range from .51 for the 9-point sections in the grade 4 PBT test form to .83 for the 23-point section in the Integrated Mathematics II PBT test forms. Due to the number of subclaim items being more similar across grades and courses, the subclaim reliabilities for Mathematics Reasoning are less variable than for the Additional and Supporting Content subclaim. The Mathematics Reasoning subclaim reliability ranges from .52 for Integrated Mathematics II test forms to .75 for both the grade 5 PBT test forms and Integrated Mathematics I PBT test forms.

For the Modeling Practice subclaim, the average reliability is .61, with a low of .41 for the Integrated Mathematics III PBT assessments and a high of .70 for both Geometry test form modes.

Table 8.26 Average Mathematics Reliability Estimates for Total Test and Subscores

Grade Level	Testing Mode	Major Content		Additional & Supporting Content		Mathematics Reasoning		Modeling Practice	
		Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability	Average Max Possible Raw Score	Average Reliability
3	CBT	43	0.91	13	0.73	13	0.61	12	0.61
	PBT	43	0.91	13	0.75	14	0.62	12	0.62
4	CBT	47	0.92	9	0.60	14	0.68	12	0.60
	PBT	47	0.91	9	0.51	14	0.72	12	0.60
5	CBT	44	0.90	12	0.68	14	0.69	12	0.51
	PBT	44	0.90	12	0.64	14	0.75	12	0.50
6	CBT	36	0.89	20	0.82	14	0.68	12	0.63
	PBT	36	0.89	20	0.82	14	0.67	12	0.66
7	CBT	43	0.89	13	0.65	14	0.64	12	0.51
	PBT	42	0.88	12	0.68	14	0.72	12	0.47
8	CBT	43	0.85	13	0.68	14	0.57	12	0.59
	PBT	42	0.84	13	0.69	14	0.62	12	0.59
A1	CBT	35	0.77	30	0.74	13	0.56	18	0.67
	PBT	36	0.75	29	0.79	14	0.57	18	0.64
GO	CBT	39	0.87	26	0.78	14	0.61	18	0.70
	PBT	39	0.88	25	0.77	14	0.61	18	0.70
A2	CBT	37	0.82	26	0.79	17	0.63	24	0.65
	PBT	38	0.82	27	0.72	17	0.65	23	0.57
M1	CBT	39	0.82	22	0.66	14	0.61	18	0.58
	PBT	39	0.82	23	0.74	11	0.75	18	0.68
M2	CBT	38	0.82	25	0.67	14	0.52	18	0.64
	PBT	42	0.89	23	0.83	14	0.65	17	0.74
M3	CBT	35	0.80	23	0.72	17	0.69	24	0.67
	PBT	37	0.72	26	0.61	18	0.57	23	0.41

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

8.7 Reliability of Classification

The reliability of the classifications for the test takers was calculated using the computer program RELCLASS, which operationalizes a statistical method developed by Livingston and Lewis (1993, 1995). As Livingston and Lewis (1993, 1995) explain, this method uses information from the administration of one test form (i.e., distribution of scores, the minimum and maximum possible scores, the cut points used for classification, and the reliability coefficient) to estimate two kinds of statistics, “decision accuracy” and “decision consistency.” Decision accuracy refers to the extent to which the classifications of test takers based on their scores on the test form agree with the classifications made on the basis of the classifications that would be made if the test scores were perfectly reliable. Decision consistency refers to the agreement between these classifications based on two non-overlapping, equally difficult forms of the test.

Decision consistency values are always lower than the corresponding decision accuracy values, because in decision consistency, both of the classifications of the student are based on scores that depend on which form of the test the student took. In decision accuracy, only one of the classifications is based on a score that can vary in this way. It is not possible to know which students were accurately classified, but it is possible to estimate the proportion of the students who were accurately classified. Similarly, it is not possible to know which students would be consistently classified if they were retested with another form, but it is possible to estimate the proportion of the students who would be consistently classified.

English Language Arts/Literacy

Table 8.27 provides information about the accuracy and the consistency of two classifications made on the basis of the scores on the grades 3-11 English Language Arts/Literacy assessments. The columns labeled “Exact level” provide the classification of the student into one of five achievement levels. The columns labeled “Level 4 or higher vs. 3 or lower” provide the classification of the student as being either in one of the upper two levels (Levels 4 and 5) or in one of the lower three levels (Levels 1, 2, and 3).

The table shows that for classifying each student into one of the five achievement levels, the proportion accurately classified ranges from .71 to .77; the proportion who would be consistently classified on two different test forms ranges from .62 to .68. For classifying each student simply as being at Level 4 or higher vs. being at Level 3 or lower, the proportion accurately classified ranges from .89 to .92; the proportion who would be consistently classified on two different test forms ranges from .85 to .89.

Table 8.27 Reliability of Classification: Summary for ELA/L

Grade Level	Testing Mode	Decision Accuracy: Proportion Accurately Classified		Decision Consistency: Proportion Consistently Classified	
		Exact level	Level 4 or higher vs. 3 or lower	Exact level	Level 4 or higher vs. 3 or lower
3	CBT	0.73	0.91	0.64	0.87
	PBT	0.72	0.90	0.63	0.87
4	CBT	0.74	0.90	0.64	0.87
	PBT	0.72	0.89	0.62	0.85
5	CBT	0.76	0.91	0.67	0.87
	PBT	0.75	0.89	0.65	0.85
6	CBT	0.77	0.91	0.68	0.88
	PBT	0.76	0.90	0.67	0.87
7	CBT	0.75	0.91	0.66	0.88
	PBT	0.73	0.90	0.63	0.86
8	CBT	0.75	0.91	0.65	0.88
	PBT	0.74	0.90	0.65	0.87
9	CBT	0.74	0.91	0.65	0.88
	PBT	0.73	0.90	0.64	0.86
10	CBT	0.74	0.92	0.65	0.89
	PBT	0.72	0.91	0.63	0.87
11	CBT	0.73	0.91	0.64	0.88
	PBT	0.71	0.91	0.62	0.87

Table 8.28 provides more detailed information about the accuracy and the consistency of the classification of students into proficiency levels for ELA/L grade 3. Each cell in the 5-by-5 table shows the estimated proportion of students who would be classified into a particular combination of proficiency levels. The sum of the five bold italicized values on the diagonal should equal the exact level of decision accuracy or consistency presented in Table 8.27. For “Level 4 and higher vs. 3 and lower” found in Table 8.27, the sum of the shaded values in the Table 8.28 should equal the level of decision accuracy or consistency presented in Table 8.27. Note that the sums based on values in Table 8.28 may not match exactly to the values in Table 8.27 due to truncation and rounding.

Detailed information for all ELA/L Spring results are provided in Appendix 8 Tables A.8.1 to A.8.9. Fall block results for ELA/L grades 9-11 are provided in the addendum to Section 8. The structure of these tables is the same as that of Table 8.28 and the values in the tables should be interpreted in the same manner as Table 8.28.

Table 8.28 Reliability of Classification: Grade 3 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category Total
		Summative Scale Score						
CBT	Decision Accuracy	650 – 699	0.16	0.04	0.00	0.00	0.00	0.20
		700 – 724	0.03	0.13	0.04	0.00	0.00	0.20
		725 – 749	0.00	0.05	0.14	0.05	0.00	0.24
		750 – 809	0.00	0.00	0.05	0.28	0.00	0.33
		810 – 850	0.00	0.00	0.00	0.02	0.01	0.03
	Decision Consistency	650 – 699	0.16	0.04	0.00	0.00	0.00	0.20
		700 – 724	0.04	0.10	0.05	0.00	0.00	0.20
		725 – 749	0.01	0.06	0.11	0.06	0.00	0.24
		750 – 809	0.00	0.01	0.06	0.26	0.01	0.33
		810 – 850	0.00	0.00	0.00	0.01	0.01	0.03
PBT	Decision Accuracy	650 – 699	0.14	0.04	0.00	0.00	0.00	0.18
		700 – 724	0.03	0.12	0.05	0.00	0.00	0.19
		725 – 749	0.00	0.04	0.13	0.04	0.00	0.22
		750 – 809	0.00	0.00	0.05	0.30	0.01	0.36
		810 – 850	0.00	0.00	0.00	0.02	0.03	0.05
	Decision Consistency	650 – 699	0.13	0.04	0.00	0.00	0.00	0.18
		700 – 724	0.04	0.09	0.05	0.01	0.00	0.19
		725 – 749	0.01	0.05	0.10	0.06	0.00	0.22
		750 – 809	0.00	0.01	0.06	0.27	0.02	0.36
		810 – 850	0.00	0.00	0.00	0.02	0.03	0.05

Note: This table includes the same information as Table A.8.1. The sum of the five bold italicized values on the diagonal should equal the exact level of decision accuracy or consistency presented in Table 8.27. For “Level 4 and higher vs. 3 and lower” presented in Table 8.27, the sum of the shaded values in Table 8.28 should equal the level of decision accuracy or consistency presented in Table 8.27. Any differences between the sums based on values in Table 8.28 and the values in Table 8.27 are due to truncation and rounding.

Mathematics

Table 8.29 provides information about the accuracy and the consistency of two classifications made on the basis of the scores on the mathematics assessments. The decision accuracy and decision consistency proportions are generally higher and more variable for the mathematics tests than for the English Language/Literacy tests. For the grades 3-8 EOY tests, the table shows that for classifying each student into one of the five achievement levels, the proportion accurately classified ranges from .74 to .80; the proportion who would be consistently classified on two different test forms ranges from .65 to .72. For the six EOC courses, the table shows that for classifying each student into one of the five achievement

levels, the proportion accurately classified ranges from .67 to .79; the proportion who would be consistently classified on two different test forms ranges from .60 to .71.

For classifying each student simply as being at Level 4 or higher vs. being at Level 3 or lower, for grades 3-8 tests, the proportion accurately classified ranges from .92 to .93; the proportion who would be consistently classified on two different test forms ranges from .89 to .91. For the six EOC courses, the proportion accurately classified as being at Level 4 or higher vs. being at Level 3 or lower ranges from .88 to .94; the proportion who would be consistently classified on two different test forms ranges from .85 to .91.

Appendix 8 tables A.8.10 to A.8.21 provide more detailed information about the accuracy and the consistency of the classification of students into proficiency levels on the basis of the mathematics. Each cell in the 5-by-5 table shows the estimated proportion of students who would be classified into a particular combination of proficiency levels. Fall block results for Algebra I, Geometry, and Algebra II are provided in the addendum to Section 8.

Table 8.29 Reliability of Classification: Summary for Mathematics

Grade Level	Testing Mode	Decision Accuracy: Proportion Accurately Classified		Decision Consistency: Proportion Consistently Classified	
		Exact Level	Level 4 or higher vs. 3 or lower	Exact Level	Level 4 or higher vs. 3 or lower
3	CBT	0.79	0.92	0.71	0.89
	PBT	0.78	0.92	0.70	0.90
4	CBT	0.80	0.93	0.72	0.90
	PBT	0.80	0.93	0.72	0.91
5	CBT	0.78	0.92	0.70	0.90
	PBT	0.79	0.93	0.71	0.91
6	CBT	0.80	0.93	0.72	0.90
	PBT	0.79	0.93	0.71	0.90
7	CBT	0.80	0.93	0.72	0.90
	PBT	0.79	0.93	0.71	0.90
8	CBT	0.74	0.92	0.65	0.89
	PBT	0.75	0.92	0.66	0.89
A1	CBT	0.76	0.92	0.66	0.89
	PBT	0.76	0.91	0.67	0.87
GO	CBT	0.79	0.93	0.71	0.90
	PBT	0.79	0.91	0.71	0.88
A2	CBT	0.76	0.94	0.68	0.91
	PBT	0.74	0.93	0.66	0.90
M1	CBT	0.75	0.92	0.65	0.89
	PBT	0.75	0.92	0.65	0.89
M2	CBT	0.73	0.92	0.63	0.89
	PBT	0.78	0.93	0.70	0.91
M3	CBT	0.75	0.93	0.68	0.90
	PBT	0.67	0.88	0.60	0.85

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

8.8 Inter-rater Agreement

Inter-rater Agreement is the agreement between the first and second scores assigned to student responses. Inter-rater agreement measurements include exact, adjacent, and nonadjacent agreement. Pearson scoring staff used these statistics as one factor in determining the needs for continuing training and intervention on both individual and group levels. Table 8.30 displays both PARCC's expectations and the actual Spring 2015 agreement percentages for perfect agreement and perfect plus adjacent agreement.

Table 8.30 Inter-rater Agreement Expectations and Results

Subject	Score Point Range	Perfect Agreement Expectation	Perfect Agreement Result	Within One Point Expectation	Within One Point Result
Mathematics	0-1	90%	95%	96%*	100%
Mathematics	0-2	80%	93%	96%	99%
Mathematics	0-3	70%	91%	96%	99%
Mathematics	0-4	65%	88%	95%	98%
Mathematics	0-6	65%	83%	95%	95%
ELA/L	Multi-trait	65%	65%	96%	98%

Note: * A 0 or 1 score compared to a blank score will have a disagreement greater than 1 point.

Pearson's ePEN2 scoring system included comprehensive inter-rater agreement reports that allowed supervisory personnel to monitor both individual and group performance. Based on reviews of these reports, scoring experts targeted individuals for increased backreading and feedback and, if necessary, retraining. Table 8.29 shows that the actual percentages for both exact reader agreement and the percentages of agreement within one-point were higher than the inter-rater agreement expectations. Refer to Section 4 for more information on handscoring.

Section 9: Validity

9.1 Overview

The Standards for Educational and Psychological Testing, issued jointly by the American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME] (2014) reports:

Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests. The process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretations (p. 11).

The purpose of test validation is not to validate the test itself but to validate interpretations of the test scores for particular uses. Test validation is not a quantifiable property but an ongoing process, beginning at initial conceptualization and continuing throughout the lifetime of an assessment. Every aspect of an assessment provides evidence in support of its validity (or evidence of lack of validity), including design, content specifications, item development, and psychometric characteristics. The 2015 operational assessment provided an opportunity to gather evidence of validity based on both test content and on the internal structure of the tests.

9.2 Evidence Based on Test Content

Evidence based on content of achievement tests is supported by the degree of correspondence between test content and instructional content. The degree to which the test measures what it claims to measure is known as construct validity. The PARCC Assessments adhere to the principles of evidence-centered design, in which the standards to be measured (the Common Core State Standards) are identified, and the performance a student needs to achieve to meet those standards is delineated in the PARCC evidence statements.

In addition to the PARCC evidence statements, content is aligned through the articulation of performance in the performance level descriptors. At the policy level, the performance level descriptors include policy claims about the educational achievement of students who attain a particular performance level, and a broad description of the grade-level knowledge, skills and practices students performing at a particular achievement level are able to demonstrate. Those policy-level descriptors are the foundation for the subject- and grade-specific performance level descriptors which, along with the PARCC Evidence frameworks, guide the development of the items and tasks.

The PARCC College- and Career-Ready determinations (CCRD) in English Language Arts/literacy and mathematics describe the academic knowledge, skills and practices students must demonstrate to show readiness for success in entry-level, credit-bearing college courses and relevant technical courses. The PARCC states determined that this level means graduating from high school and having at least a 75% likelihood of earning a grade of “C” or better in credit-bearing courses without the need for remedial coursework. After reviewing the standards and assessment design, the PARCC Governing Board (made

up of the K-12 education chiefs in PARCC states) in conjunction with the PARCC Advisory Committee on College Readiness (composed of higher education chiefs in the PARCC states), determined that students who achieve at levels 4 and 5 on the final PARCC high school assessments are likely to have acquired the skills and knowledge to meet the definition of college- and career-readiness. To validate the determinations, PARCC conducted a Postsecondary Educator Judgment Study and a Benchmark study of the SAT, ACT, National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), Programme of International Student Assessment (PISA), and Progress in International Reading Literacy Study (PIRLS) tests (McClarty, Korbin, Moyer, Griffin, Huth, Carey, and Medberry, 2015).

Gathering construct validity evidence for PARCC is embedded in the process by which the PARCC assessment content is developed and validated. At each step in the assessment development process, PARCC states involved hundreds of educators, assessment experts, and bias and sensitivity experts in review of text, items and tasks for accuracy, appropriateness, alignment to the instructional standards, and freedom from bias. See Chapter 2 for an overview of the content development process. In the early stages of development Pearson conducted research studies to validate the PARCC item and task development approach. One such study was a Student Task Interaction Study designed to collect data on the student's experience with the assessment tasks and technological functionalities, as well as the amount of time needed for answering each task. Pearson also conducted a Rubric Choice Study that compared the functioning of two rubrics developed to score the Prose Constructed Response (PCR) tasks in ELA. Quantitative and qualitative evidence was collected to support the use of a condensed or expanded trait scoring rubric in scoring student responses.

PARCC items and tasks were field tested prior to their use on an assessment. During the initial field test administration in 2014, PARCC states collected feedback from students, test administrators, test coordinators, and classroom teachers on their experience with the PARCC assessments, including the quality of test items and student experience. A summary of the feedback can be found at: <http://parconline.org/news-and-video/317-lessons-learned-part-2-digging-into-the-survey-results>. The feedback from that survey was used to inform test directions, test timing, and the function of online task interactions. Performance data from the field test also informed the future development of additional items and tasks.

An important consideration when constructing test forms is recognition of items that may introduce construct-irrelevant variance. Such items should not be included on test forms to help ensure fairness to all subgroups of test takers. PARCC convened bias and sensitivity committees to review all items. Additionally, content experts facilitated reviews of all items. All reviewers were trained using PARCC Bias and Sensitivity Guidelines, and the Guidelines were used to review items and ELA/L passages. Accommodations were made available based on individual need documented in the student's approved IEP, 504 Plan, or if required by the PARCC member state, an English learner (EL) Plan (refer to Section 3.4). An accessibility specialist worked in consultation with the PARCC accessibility specialist to review forms and determine which forms should be used for students with accommodations.

The ELA/L and mathematics operational test forms, as described in Section 2, were carefully constructed to align with the test blueprints and specifications that are based on the Common Core State Standards (CCSS). During the fall of 2014, content experts representing Parcc, Inc. and various PARCC states, along with content experts, held a series of meetings to review the operational forms for ELA/L and mathematics. These meetings provided opportunity to evaluate tests forms in their entirety and recommend changes. Requested item replacements were accommodated to the extent possible while striving to maintain the integrity of the various linking designs required for the operational test analyses. Psychometricians were available throughout this process to provide guidance with regard to implications of item replacements for the linking and statistical requirements.

Further information regarding the PARCC assessment college- and career-ready content standards, performance level descriptors, and accessibility features and accommodations is provided at the following URL: <http://www.parcconline.org/policies-and-guidance>.

9.3 Evidence Based on Internal Structure

Analyses of the internal structure of a test typically involve studies of the relationships among test items and/or test components (i.e., subclaims) in the interest of establishing the degree to which the items or components appear to reflect the construct on which a test score interpretation is based (AERA, APA, & NCME, 2014, p. 16). The term construct is used here to refer to the characteristics that a test is intended to measure; in the case of the PARCC operational tests, the characteristics of interest are the knowledge and skills defined by the test blueprint for ELA/L and for mathematics.

The PARCC assessments provide a full summative test score, Reading claim score, and Writing claim score as well as ELA/L subclaims and mathematics subclaim scores. The goal of reporting at this level is to provide criterion-referenced data to assess the strengths and weaknesses of a student's achievement in specific components of each content area. This information can then be used by teachers to plan for further instruction, to plan for curriculum development, and to report progress to parents. The results can also be used as one factor in making administrative decisions about program effectiveness, teacher effectiveness, class grouping, and needs assessment.

9.3.1 Intercorrelations

The ELA/L full summative tests comprise two claim scores: Reading (RD) and Writing (WR) and five subclaim scores: Reading Literature (RL), Reading Information (RI), Reading Vocabulary (RV), Writing Written Expression (WE), and Writing Knowledge Language and Conventions (WKL). The RD claim score is a composite of RL, RI, and RV. The writing claim score, a composite of WE and WKL, comprises only PCR items and the same PCR items are in each subclaim. The ELA/L operational test analyses were performed by evaluating the separate trait scores of WE and WKL, and for some PCR items also RL or RI, therefore the trait scores were used for the intercorrelations.

The mathematics full summative tests have four subclaim scores: Major Content (MC), Mathematical Reasoning (MR), Modeling Practice (MP), and Additional and Supporting Content (ASC).

High total group internal consistencies as well as similar reliabilities across subgroups provide additional evidence of validity. High reliability of test scores implies that the test items within a domain are measuring a single construct, which is a necessary condition for validity when the intention is to measure a single construct. Refer to Section 8 for reliability estimates for the overall population, subgroups of interest, as well as for subscores for ELA/L and subclaims for mathematics.

Another way to assess the internal structure of a test is through the evaluation of correlations among subscores. These analyses were conducted between the ELA/L reading and writing claim scores and the ELA/L subclaims (RL, RI, RV, WE, and WKL) and between the mathematics subclaims. If these components within a content area are strongly related to each other, this is evidence of unidimensionality.

A series of tables are provided to summarize the results for the Spring 2015 administration.¹¹ Tables 9.1 through 9.9 present the Pearson correlations observed between the ELA/L reading and writing claim scores and subclaim scores for each grade; correlations are reported separately for online (CBT) and paper (PBT) versions of the tests. The tables provide the average intercorrelations by averaging the intercorrelations computed for all the individual forms of the test, separately for the CBT and PBT tests within each grade level. The average sample size is provided in the upper triangle portion of the tables. The sample sizes may vary amongst the correlations as some students may not have a valid subclaim score. The subclaim reliabilities (from Section 8) are reported along the diagonal. The WR, WE, and WKL scores tended to be highly correlated; this is expected given that these three intercorrelations are based on the same Writing items. RL, RI, and RV, all subclaims of Reading, are moderately to highly correlated. Additionally, the WR subclaims, WE and WKL, are moderately correlated with RD subclaims (of RL, RI, and RV). These moderate to high ELA/L intercorrelations amongst the subclaims are sufficiently high enough to provide evidence that the ELA/L tests are unidimensional.

The intercorrelations and reliability estimates for mathematics are provided in Tables 9.10 to 9.21. The mathematics intercorrelations are moderate. The only observable pattern in the mathematics intercorrelations is that the MC subclaim has consistently slightly higher correlations with the ASC, MR, and MP subclaims; the intercorrelations amongst the ASC, MR, and MP subclaims are slightly lower. The mathematics intercorrelations are sufficiently high enough to suggest that the mathematics tests are likely to be unidimensional with some minor secondary dimensions.

Additionally, the ELA/L and mathematics correlations for the two modes, PBT and CBT, displayed similar patterns of intercorrelations suggesting that the structure of the PBT assessments and CBT assessments are similar.

¹¹ Addendum 9 provides a summary of results for the Fall 2014 administration.

Table 9.1 Average Intercorrelations and Reliability between Grade 3 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.90	10,544	10,544	10,544	10,544	10,544	10,544	RD	0.90	8,638	8,638	8,638	8,638	8,638	8,638
RL	0.94	0.83	10,544	10,544	105,44	10,544	10,544	RL	0.95	0.83	8,638	8,638	8,638	8,638	8,638
RI	0.87	0.71	0.69	10,544	105,44	10,544	10,544	RI	0.86	0.71	0.64	8,638	8,638	8,638	8,638
RV	0.85	0.71	0.64	0.64	105,44	10,544	10,544	RV	0.85	0.73	0.64	0.69	8,638	8,638	8,638
WR	0.72	0.67	0.65	0.57	0.78	10,544	10,544	WR	0.74	0.70	0.66	0.60	0.79	8,638	8,638
WE	0.69	0.64	0.63	0.54	0.99	0.67	10,544	WE	0.72	0.68	0.65	0.58	0.99	0.70	8,638
WKL	0.70	0.66	0.62	0.57	0.90	0.82	0.78	WKL	0.71	0.67	0.63	0.58	0.92	0.85	0.77

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.2 Average Intercorrelations and Reliability between Grade 4 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.89	13,322	13,322	133,22	13,322	13,322	13,322	RD	0.87	9,460	9,460	9,460	9,460	9,460	9,460
RL	0.92	0.77	13,322	13,322	13,322	13,322	13,322	RL	0.90	0.71	9,460	9,460	9,460	9,460	9,460
RI	0.92	0.74	0.77	13,322	13,322	13,322	13,322	RI	0.89	0.67	0.74	9,460	9,460	9,460	9,460
RV	0.85	0.70	0.69	0.65	13,322	13,322	13,322	RV	0.84	0.66	0.65	0.63	9,460	9,460	9,460
WR	0.76	0.71	0.70	0.61	0.82	13,321	13,321	WR	0.75	0.70	0.67	0.60	0.80	9,460	9,460
WE	0.75	0.70	0.69	0.60	0.99	0.77	13,321	WE	0.74	0.69	0.67	0.59	0.99	0.74	9,460
WKL	0.73	0.68	0.67	0.59	0.94	0.90	0.80	WKL	0.71	0.66	0.64	0.57	0.94	0.89	0.77

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.3 Average Intercorrelations and Reliability between Grade 5 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.88	13,791	13,791	13,791	13,791	13,791	13,791	RD	0.87	8,779	8,779	8,779	8,779	8,779	8,779
RL	0.91	0.75	13,791	13,791	13,791	13,791	13,791	RL	0.89	0.71	8,779	8,779	8,779	8,779	8,779
RI	0.90	0.72	0.75	13,791	13,791	13,791	13,791	RI	0.91	0.71	0.77	8,779	8,779	8,779	8,779
RV	0.85	0.67	0.65	0.63	13,791	13,791	13,791	RV	0.84	0.64	0.64	0.58	8,779	8,779	8,779
WR	0.74	0.69	0.69	0.59	0.82	13,791	13,791	WR	0.74	0.68	0.69	0.56	0.80	8,778	8,778
WE	0.73	0.68	0.67	0.57	0.99	0.76	13,791	WE	0.72	0.67	0.68	0.55	0.99	0.72	8,778
WKL	0.72	0.67	0.66	0.57	0.93	0.88	0.81	WKL	0.70	0.64	0.66	0.53	0.93	0.88	0.76

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.4 Average Intercorrelations and Reliability between Grade 6 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.90	14,228	14,228	14,228	14,228	14,228	14,228	RD	0.92	7,425	7,425	7,425	7,425	7,425	7,425
RL	0.91	0.78	14,228	14,228	14,228	14,228	14,228	RL	0.93	0.83	7,425	7,425	7,425	7,425	7,425
RI	0.93	0.75	0.80	14,228	14,228	14,228	14,228	RI	0.93	0.78	0.80	7,425	7,425	7,425	7,425
RV	0.86	0.71	0.71	0.67	14,228	14,228	14,228	RV	0.88	0.75	0.73	0.71	7,425	7,425	7,425
WR	0.75	0.71	0.71	0.58	0.84	14,228	14,228	WR	0.75	0.72	0.72	0.59	0.82	7,424	7,424
WE	0.74	0.70	0.70	0.57	1.00	0.82	14,228	WE	0.74	0.71	0.71	0.59	1.00	0.78	7,424
WKL	0.74	0.70	0.70	0.58	0.96	0.93	0.83	WKL	0.74	0.70	0.71	0.59	0.96	0.93	0.80

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.5 Average Intercorrelations and Reliability between Grade 7 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.91	14,272	14,272	14,272	14,272	14,272	14,272	RD	0.91	6,807	6,807	6,807	6,807	6,807	6,807
RL	0.93	0.79	14,272	14,272	14,272	14,272	14,272	RL	0.93	0.80	6,807	6,807	6,807	6,807	6,807
RI	0.93	0.78	0.80	14,271	14,271	14,271	14,271	RI	0.92	0.76	0.78	6,807	6,807	6,807	6,807
RV	0.87	0.73	0.73	0.70	14,272	14,272	14,272	RV	0.88	0.74	0.72	0.71	6,807	6,807	6,807
WR	0.76	0.71	0.73	0.61	0.84	14,271	14,271	WR	0.75	0.70	0.73	0.61	0.80	6,806	6,806
WE	0.75	0.70	0.72	0.60	1.00	0.82	14,271	WE	0.74	0.69	0.72	0.60	1.00	0.75	6,806
WKL	0.75	0.70	0.73	0.61	0.97	0.95	0.83	WKL	0.73	0.68	0.70	0.60	0.96	0.93	0.77

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.6 Average Intercorrelations and Reliability between Grade 8 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.90	14,023	14,023	14,023	14,023	14,023	14,023	RD	0.90	6,970	6,970	6,970	6,970	6,970	6,970
RL	0.92	0.75	14,023	14,023	14,023	14,023	14,023	RL	0.93	0.78	6,970	6,970	6,970	6,970	6,970
RI	0.93	0.75	0.78	14,023	14,023	14,023	14,023	RI	0.92	0.76	0.77	6,969	6,969	6,969	6,969
RV	0.86	0.71	0.71	0.67	14,023	14,023	14,023	RV	0.89	0.76	0.73	0.74	6,970	6,970	6,970
WR	0.77	0.72	0.73	0.61	0.85	14,022	14,022	WR	0.76	0.71	0.72	0.62	0.83	6,969	6,969
WE	0.76	0.71	0.72	0.60	1.00	0.83	14,022	WE	0.75	0.70	0.72	0.61	1.00	0.80	6,969
WKL	0.77	0.72	0.73	0.62	0.97	0.94	0.83	WKL	0.75	0.70	0.71	0.62	0.97	0.94	0.80

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.7 Average Intercorrelations and Reliability between Grade 9 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.91	5,364	5,364	5,364	5,364	5,364	5,364	RD	0.91	4,128	4,128	4,128	4,128	4,128	4,128
RL	0.92	0.78	5,364	5,364	5,364	5,364	5,364	RL	0.92	0.78	4,128	4,128	4,128	4,128	4,128
RI	0.94	0.78	0.82	5,363	5,363	5,363	5,363	RI	0.94	0.77	0.80	4,127	4,127	4,127	4,127
RV	0.86	0.71	0.73	0.67	5,364	5,364	5,364	RV	0.87	0.73	0.74	0.67	4,128	4,128	4,128
WR	0.77	0.73	0.74	0.61	0.85	5,362	5,362	WR	0.76	0.72	0.73	0.61	0.83	4,127	4,127
WE	0.76	0.72	0.73	0.60	1.00	0.82	5,362	WE	0.75	0.71	0.72	0.61	1.00	0.79	4,127
WKL	0.77	0.73	0.73	0.61	0.97	0.95	0.82	WKL	0.75	0.71	0.72	0.61	0.97	0.95	0.80

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.8 Average Intercorrelations and Reliability between Grade 10 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.92	3,813	3,813	3,813	3,813	3,813	3,813	RD	0.92	1,372	1,372	1,372	1,372	1,372	1,372
RL	0.92	0.78	3,813	3,813	3,813	3,813	3,813	RL	0.93	0.80	1,372	1,372	1,372	1,372	1,372
RI	0.95	0.78	0.84	3,813	3,813	3,813	3,813	RI	0.95	0.80	0.85	1,372	1,372	1,372	1,372
RV	0.87	0.72	0.74	0.69	3,813	3,813	3,813	RV	0.86	0.72	0.76	0.67	1,372	1,372	1,372
WR	0.78	0.74	0.75	0.63	0.85	3,812	3,812	WR	0.79	0.75	0.77	0.63	0.84	1,371	1,371
WE	0.77	0.73	0.74	0.62	1.00	0.83	3,812	WE	0.79	0.75	0.76	0.62	1.00	0.82	1,371
WKL	0.78	0.73	0.75	0.63	0.97	0.94	0.83	WKL	0.78	0.74	0.76	0.63	0.97	0.95	0.81

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.9 Average Intercorrelations and Reliability between Grade 11 ELA/L Subclaims

	CBT								PBT						
	RD	RL	RI	RV	WR	WE	WKL		RD	RL	RI	RV	WR	WE	WKL
RD	0.90	4,452	4,452	4,452	4,452	4,452	4,452	RD	0.90	587	587	587	587	587	587
RL	0.93	0.81	4,452	4,452	4,452	4,452	4,452	RL	0.92	0.79	587	587	587	587	587
RI	0.91	0.75	0.78	4,451	4,451	4,451	4,451	RI	0.90	0.72	0.78	586	586	586	586
RV	0.85	0.71	0.68	0.64	4,452	4,452	4,452	RV	0.87	0.73	0.68	0.70	587	587	587
WR	0.78	0.75	0.73	0.59	0.85	4,450	4,450	WR	0.75	0.74	0.68	0.59	0.82	586	586
WE	0.77	0.74	0.73	0.58	1.00	0.83	4,450	WE	0.75	0.73	0.68	0.58	1.00	0.78	586
WKL	0.76	0.74	0.71	0.58	0.97	0.94	0.83	WKL	0.73	0.72	0.65	0.58	0.97	0.95	0.78

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.10 Average Intercorrelations and Reliability between Grade 3 Mathematics Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.91	12,555	12,555	12,555	MC	0.91	12,236	12,236	12,236
ASC	0.80	0.73	12,555	12,555	ASC	0.81	0.75	12,236	12,236
MR	0.70	0.60	0.61	12,554	MR	0.71	0.60	0.62	12,235
MP	0.76	0.64	0.68	0.61	MP	0.74	0.62	0.65	0.62

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.11 Average Intercorrelations and Reliability between Grade 4 Mathematics Subclaims

	CBT				PBT				
	MC	ASC	MR	MP	MC	ASC	MR	MP	
MC	0.92	13,192	13,192	13,192	MC	0.91	9,851	9,851	9,851
ASC	0.70	0.60	13,192	13,192	ASC	0.65	0.51	9,851	9,851
MR	0.73	0.60	0.68	13,192	MR	0.76	0.57	0.72	9,850
MP	0.69	0.56	0.65	0.60	MP	0.72	0.52	0.68	0.60

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.12 Average Intercorrelations and Reliability between Grade 5 Mathematics Subclaims

	CBT				PBT				
	MC	ASC	MR	MP	MC	ASC	MR	MP	
MC	0.90	13,272	13,272	13,272	MC	0.90	9,160	9,160	9,160
ASC	0.71	0.68	13,272	13,272	ASC	0.71	0.64	9,160	9,160
MR	0.76	0.60	0.69	13,271	MR	0.79	0.62	0.75	9,160
MP	0.70	0.56	0.65	0.51	MP	0.70	0.57	0.67	0.50

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.13 Average Intercorrelations and Reliability between Grade 6 Mathematics Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.89	13,850	13,850	13,850	MC	0.89	7,304	7,304	7,304
ASC	0.80	0.82	13,850	13,850	ASC	0.81	0.82	7,304	7,304
MR	0.73	0.70	0.68	13,847	MR	0.76	0.70	0.67	7,297
MP	0.71	0.68	0.68	0.63	MP	0.74	0.68	0.67	0.66

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.14 Average Intercorrelations and Reliability between Grade 7 Mathematics Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.89	13,992	13,992	13,992	MC	0.88	6,437	6,437	6,437
ASC	0.75	0.65	13,985	13,985	ASC	0.76	0.68	6,433	6,433
MR	0.76	0.63	0.64	13,988	MR	0.78	0.68	0.72	6,426
MP	0.74	0.62	0.66	0.51	MP	0.74	0.64	0.69	0.47

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.15 Average Intercorrelations and Reliability between Grade 8 Mathematics Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.85	11,442	11,442	11,442	MC	0.84	5,644	5,644	5,644
ASC	0.76	0.68	11,442	11,442	ASC	0.77	0.69	5,644	5,644
MR	0.69	0.62	0.57	11,438	MR	0.72	0.65	0.62	5,637
MP	0.67	0.61	0.61	0.59	MP	0.72	0.68	0.66	0.59

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.16 Average Intercorrelations and Reliability between Algebra I Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.77	6,213	6,213	6,213	MC	0.75	4,543	4,543	4,543
ASC	0.78	0.74	6,213	6,213	ASC	0.78	0.79	4,543	4,543
MR	0.67	0.64	0.56	6,177	MR	0.64	0.63	0.57	4,478
MP	0.70	0.68	0.62	0.67	MP	0.69	0.69	0.61	0.64

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.17 Average Intercorrelations and Reliability between Geometry Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.87	2,559	2,559	2,559	MC	0.88	1,253	1,253	1,253
ASC	0.82	0.78	2,559	2,559	ASC	0.83	0.77	1,253	1,253
MR	0.69	0.64	0.61	2,544	MR	0.69	0.65	0.61	1,237
MP	0.76	0.70	0.65	0.70	MP	0.78	0.72	0.65	0.70

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.18 Average Intercorrelations and Reliability between Algebra II Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.82	4,807	4,807	4,807	MC	0.82	806	806	806
ASC	0.80	0.79	4,807	4,807	ASC	0.77	0.72	806	806
MR	0.72	0.70	0.63	4,805	MR	0.71	0.65	0.65	794
MP	0.69	0.68	0.68	0.65	MP	0.68	0.64	0.66	0.57

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.19 Average Intercorrelations and Reliability between Integrated Mathematics I Subclaims

	CBT				PBT				
	MC	ASC	MR	MP	MC	ASC	MR	MP	
MC	0.82	6,737	6,737	6,737	MC	0.82	792	792	792
ASC	0.76	0.66	6,737	6,737	ASC	0.80	0.74	792	792
MR	0.71	0.64	0.61	6,735	MR	0.70	0.67	0.75	784
MP	0.74	0.67	0.67	0.58	MP	0.76	0.74	0.68	0.68

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.20 Average Intercorrelations and Reliability between Integrated Mathematics II Subclaims

	CBT				PBT				
	MC	ASC	MR	MP	MC	ASC	MR	MP	
MC	0.82	2,841	2,841	2,841	MC	0.89	225	225	225
ASC	0.75	0.67	2,841	2,841	ASC	0.86	0.83	225	225
MR	0.69	0.61	0.52	2,841	MR	0.75	0.71	0.65	224
MP	0.68	0.63	0.61	0.64	MP	0.78	0.75	0.67	0.74

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table 9.21 Average Intercorrelations and Reliability between Integrated Mathematics III Subclaims

	CBT					PBT			
	MC	ASC	MR	MP		MC	ASC	MR	MP
MC	0.80	1,819	1,819	1,819	MC	0.72	564	564	564
ASC	0.77	0.72	1,819	1,819	ASC	0.69	0.61	564	564
MR	0.74	0.68	0.69	1,819	MR	0.63	0.54	0.57	553
MP	0.72	0.65	0.72	0.67	MP	0.50	0.49	0.57	0.41

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

9.3.2 Reliability

Additionally, the reliability analyses presented in Section 8 of this technical report provide information about the internal consistency of the PARCC assessments. Internal consistency is typically measured via correlations amongst the items on an assessment and provides an indication of how much the items measure the same general construct. The reliability estimates, computed using coefficient alpha (Cronbach, 1951) are presented in Tables 8.1 and 8.2 and are along the diagonals of Tables 9.1 through 9.21.¹² The average reliabilities for ELA/L PARCC assessments range from .89 to .93 and for the mathematics assessments range from .85 up to .95. Along with the subclaim intercorrelations, the reliability estimates indicate that the items within each PARCC assessment are measuring the same construct and provides further evidence of unidimensionality.

9.3.3 Local Item Dependence

In addition to the intercorrelations for ELA/L and mathematics, the local independence was evaluated. Local independence is one of the primary assumptions of IRT that states the probability of success on one item is not influenced by performance on other items, when controlling for ability level. This implies that ability or theta accounts for the associations among the observed items. *Local item dependence* (LID) essentially overstates the amount of information predicted by the IRT model and the resulting scale scores no longer reflect the joint likelihood of the item response. It can exert other undesirable psychometric effects and represents a threat to validity since other factors besides the construct of interest are present. Classical statistics are also affected when LID is present since estimates of test reliability like IRT information can be inflated (Zenisky, Hambleton, Sireci, 2003). When item parameter estimates exhibit LID, these estimates may not be stable. When the items are subsequently selected for inclusion on a test, inaccurate predictions and scale instability can potentially result.

The LID issue affects the choice of item scoring in IRT calibrations. Specifically, if evidence suggests these items indeed have local dependence, then it might be preferable to sum the item scores into clusters or testlets as a method of minimizing LID. However, if these items do not appear to have strong local item dependence, then retaining the scores as individual item scores in an IRT calibration is preferred since more information concerning item properties is retained. Two methods described below were used to investigate the presence of LID in PARCC.

First, item and testlet reliability analyses under classical test theory (Wainer & Thissen, 2001) were conducted as a way to evaluate the degree of LID. Two estimates of Cronbach's alpha (Cronbach, 1951) were compared based on individual items in a test and those clustered into testlets. Cronbach's alpha is formulated as:

$$\alpha = \frac{k}{k-1} \frac{\sum_{i \neq i'} \sigma_{ii'}}{\sigma_T^2}, \quad (9-1)$$

¹² Section 8 provides information on the computations of the reliability estimates.

where k is the total number of items, $\sigma_{ii'}$ is the covariance of items i and i' ($i \neq i'$), and σ_T^2 is the variance of total scores. To compute an alpha coefficient, sample standard deviations and variances are substituted for the $\sigma_{ii'}$ and σ_T^2 . The alpha for the total test based on individual items is compared with those that form testlets based on larger subparts. If the item-level configuration has appreciably higher levels of reliability compared with the testlets, LID may be present.

For IRT based methods, local dependence can be evaluated using statistics such as Q_3 (Yen, 1984). The item residual is the difference between observed and expected performance. The Q_3 index is the correlation between residuals of each item pair defined as

$$\begin{aligned} d_i &= (O - \hat{E}), \\ Q_3 &= r(d_i, d_{i'}) \end{aligned} \tag{9-2}$$

where O is the observed score and \hat{E} is the expected value of O under a proposed IRT model and the index is defined as the correlation between the two item residuals.

LID manifests itself as a residual correlation that is nonzero and large. For Q_3 , LID can either be positive or negative. Positive (negative) LID indicates that performance is higher (lower) than expectation. The residual Q_3 correlation matrix can be inspected to determine if there are any blocks of locally dependent items (e.g., perhaps blocks of items belonging to the same reading passage). For Q_3 , the null hypothesis is that local independence holds. The expected value of Q_3 is $-1/n-1$ where n is the number of items such that the statistic shows a small negative bias. As a rule-of-thumb, item pairs with moderate levels of LID for Q_3 are $|0.20|$ or greater. Significant levels of LID are present when the statistic is greater than $|0.40|$. An alternative is to use the Fisher r to z transformation and evaluate the resulting p values.

For the reliability comparison, the follows eight test levels were selected:

1. Grade 4 for span 3-5 in ELA/L,
2. Grade 4 for span 3-5 in mathematics,
3. Grade 7 for span 6-8 in ELA/L,
4. Grade 7 for span 6-8 in mathematics,
5. Grade 10 for span 9-11 in ELA/L,
6. Integrated Mathematics II for Integrated Mathematics I-III,
7. Algebra I, and
8. Algebra II.

One Spring 2015 CBT form for each of the eight tests was selected that was roughly at the median in terms of test difficulty. For ELA/L, reading items were summed according to passage assignment. For mathematics, items were summed according to subclaims. Reliability was computed for the entire forms using the two different approaches as described above, one involving calculations at the item level and the second utilizing scores on summed items (i.e., testlets). Further description of the data is given in Table 9.22.

To cross-validate the classical reliability analysis, the Q_3 statistic was computed from spring CBT data based on Grade 4 ELA/L and Integrated Mathematics II items. All items in the pool at that test level were included. The CBT item pool for grade 4 ELA/L contained 125 items while Integrated Mathematics Two had 77 items.

The results for the reliability analysis are shown in Figure 9.1. In every instance, the item-level reliability is higher than in the testlet configuration. The greatest difference was for Algebra II which showed a difference of 0.07. Although this was not unexpected, the magnitude of the differences in the respective alpha reliability coefficients in general do not suggest a concerning level of LID. Table 9.23 shows the summary for the Q_3 values. Figures 9.2 and 9.3 show graphs of the distribution of Q_3 values. Most of the Q_3 values were small and negative, again suggesting that LID is not at a level of concern.

For these two test levels, the difference in the reliability coefficients was 0.03 and was consistent with the low values of Q_3 .

In summary, this investigation did not find evidence for the existence of pervasive LID. The results of both the reliability and Q_3 methods support a claim of minimal LID. For a multiple-choice only test containing four reading passages with 5 to 12 items associated with a reading passage, Sireci, Thissen, and Wainer (1991) reported that testlet alpha was approximately 10 percent lower than the item-level coefficient. In comparison, PARCC tests have complex test structures and exhibited smaller differences in reliability. In addition, the median Q_3 values presented in Table 9.23 centered around the expectation of $-1/n-1$.

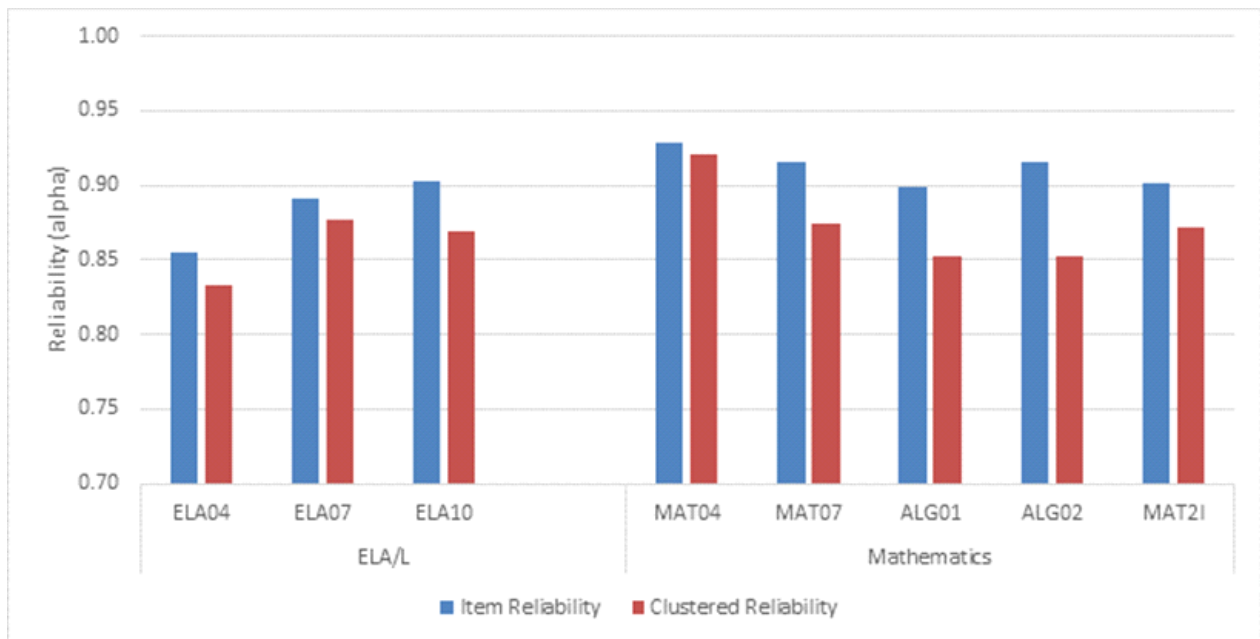


Figure 9.1 Comparison of Score Reliability by Item Reliability and Cluster (Testlet) Reliability Approaches

Table 9.22 Conditions used in LID Reliability Investigation and Results

Content	Grade	N Valid	N Complete	Percent Incomplete	No. Items	No. Tasks	Item Rel.	Task Rel.
ELA/L								
ELA/L	4	13,660	13,518	1.04	31	5	0.86	0.83
ELA/L	7	12,757	12,685	0.56	41	7	0.89	0.88
ELA/L	10	3,097	3,033	2.07	41	7	0.90	0.87
Mathematics								
Math	4	10,332	10,255	0.75	53	4	0.93	0.92
Math	7	10,295	10,188	1.04	50	6	0.92	0.87
Math	A1	5,072	4,885	3.69	52	6	0.90	0.85
Math	A2	4,982	4,769	4.28	54	6	0.92	0.85
Math	M2	2,708	2,645	2.33	51	6	0.90	0.87

Note: A1 = Algebra I, A2 = Algebra II, M2 = Integrated Mathematics II.

Table 9.23 Summary of Q₃ Values for ELA/L Grade 4 and Integrated Mathematics II

Min.	Q ₁	Median	Mean	Q ₃	Max.	SD
ELA/L, Grade 4						
-0.138	-0.047	-0.031	-0.031	-0.017	0.279	0.030
Integrated Mathematics II						
-0.160	-0.038	-0.017	-0.019	0.001	0.280	0.032

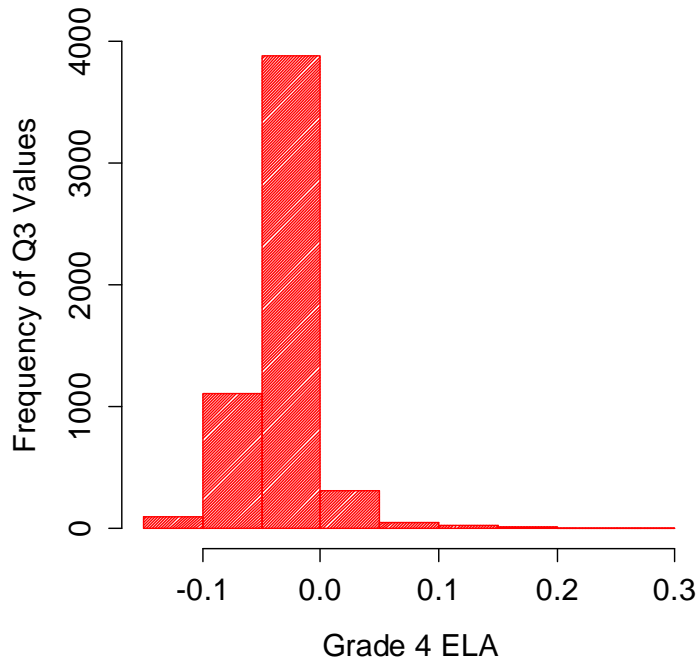


Figure 9.2 Distribution of Q_3 Values for Grade 4 ELA/L

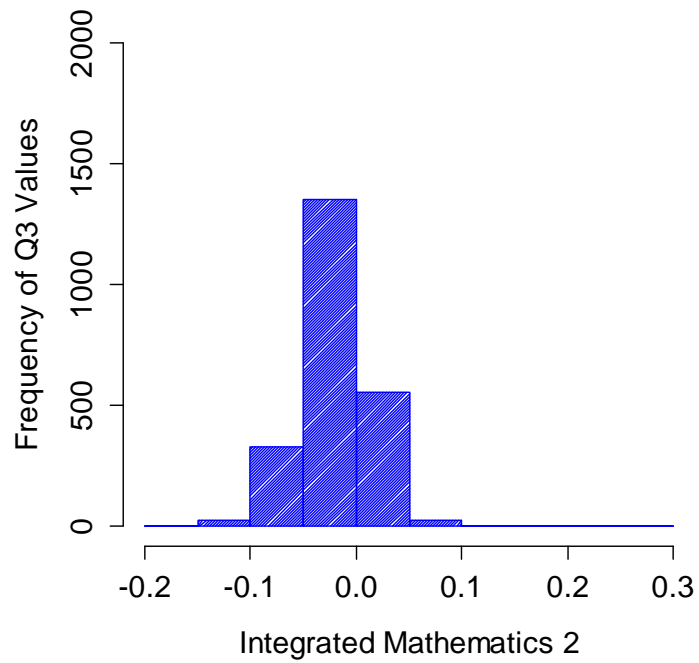


Figure 9.3 Distribution of Q_3 Values for Integrated Mathematics II

9.4 Evidence Based on Relationships to Other Variables

Empirical results concerning the relationships between score on a test and measures of other variables external to the test can also provide evidence of validity when these relationships are found to be consistent with the definition of the construct that the test is intended to measure. As indicated in the APA, AERA, and NCME Standards (2014), the variables investigated can include other tests that measure the same construct and different constructs, criterion measures that scores on the test are expected to predict, as well as demographic characteristics of test takers that are expected to be related and unrelated to test performance.

The relationship of the scores across the ELA/L and mathematics assessments was evaluated using correlational analyses. Tables 9.24 through 9.29 present the average Pearson correlations observed between the ELA/L raw scores and the mathematics raw scores for each grade; average correlations are reported separately for online (CBT) and paper (PBT) versions of the tests. For grades three through 8, students must have a valid test score for both ELA/L and mathematics at the same grade level to be included in the tables. These tables provide the average correlation in the lower triangle and the average sample size is provided in the upper triangle. Note, that the sample sizes for these tables are smaller than the within subject test. However, despite the small average sample sizes, the average correlations are relatively stable because they are based on many form combinations. In computing the correlations between a particular pair of ELA/L and mathematics test forms, test takers must have taken both tests via the same mode and form combination. For example, there are 36 core forms for grade 3 CBT ELA/L and 36 core forms for grade 3 CBT mathematics resulting in 1,296 possible combinations of core forms. ELA/L, RD, WR, are moderately to highly correlated with mathematics; the correlations range from .66 up to .78 for grades 3 through 8, and range from .44 to .74 for the high school tests. These correlations suggest that the ELA/L and mathematics tests are assessing different content.

The ELA/L and mathematics correlations for the high school tests are presented in Tables 9.30 through 9.32. Because students in high school can take the mathematics courses in different years (e.g., one student make take Algebra I in grade 9 while another student may take Algebra I in grade 10), the high school mathematics scores were correlated with several of the ELA/L grades (e.g., Algebra I correlated with both grades 9 and 10). Correlations between high school mathematics scores and corresponding ELA/L scores demonstrate low to moderate correlations.

Table 9.24 Average Correlations between ELA/L and Mathematics for Grade 3

	CBT				PBT			
	ELA/L	RD	WR	MA	ELA/L	RD	WR	MA
ELA/L		290	290	290	ELA/L		522	522
RD	0.96		290	290	RD	0.96		522
WR	0.88	0.71		290	WR	0.89	0.73	522
MA	0.78	0.76	0.67		MA	0.78	0.76	0.67

Note: ELA/L = English language arts/Literacy, RD = Reading, WR = Writing, MA = Mathematics. The average correlations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table. Students must have a valid grade 3 ELA/L score *and* a valid grade 3 mathematics score to be included in this table.

Table 9.25 Average Correlations between ELA/L and Mathematics for Grade 4

	CBT				PBT			
	ELA/L	RD	WR	MA	ELA/L	RD	WR	MA
ELA/L		362	362	362	ELA/L		577	577
RD	0.96		362	362	RD	0.95		577
WR	0.90	0.75		362	WR	0.91	0.74	576
MA	0.78	0.76	0.68		MA	0.76	0.74	0.66

Note: ELA/L = English language arts/Literacy, RD = Reading, WR = Writing, MA = Mathematics. The average correlations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table. Students must have a valid grade 4 ELA/L score *and* a valid grade 4 mathematics score to be included in this table.

Table 9.26 Average Correlations between ELA/L and Mathematics for Grade 5

	CBT				PBT			
	ELA/L	RD	WR	MA	ELA/L	RD	WR	MA
ELA/L		374	374	374	ELA/L		534	534
RD	0.96		374	374	RD	0.95		534
WR	0.90	0.74		374	WR	0.90	0.72	534
MA	0.76	0.74	0.66		MA	0.74	0.73	0.64

Note: ELA/L = English language arts/Literacy, RD = Reading, WR = Writing, MA = Mathematics. The average correlations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table. Students must have a valid grade 5 ELA/L score *and* a valid grade 5 mathematics score to be included in this table.

Table 9.27 Average Correlations between ELA/L and Mathematics for Grade 6

	CBT				PBT			
	ELA/L	RD	WR	MA	ELA/L	RD	WR	MA
ELA/L		385	385	385	ELA/L		424	424
RD	0.96		385	385	RD	0.96		424
WR	0.90	0.74		385	WR	0.89	0.74	424
MA	0.78	0.77	0.66		MA	0.77	0.76	0.65

Note: ELA/L = English language arts/Literacy, RD = Reading, WR = Writing, MA = Mathematics. The average correlations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table. Students must have a valid grade 6 ELA/L score *and* a valid grade 6 mathematics score to be included in this table.

Table 9.28 Average Correlations between ELA/L and Mathematics for Grade 7

	CBT				PBT				
	ELA/L	RD	WR	MA	ELA/L	RD	WR	MA	
ELA/L		377	377	377	ELA/L		372	372	372
RD	0.96		377	377	RD	0.96		372	372
WR	0.90	0.75		377	WR	0.90	0.73		372
MA	0.77	0.77	0.66		MA	0.75	0.75	0.63	

Note: ELA/L = English language arts/Literacy, RD = Reading, WR = Writing, MA = Mathematics. The average correlations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table. Students must have a valid grade 7 ELA/L score *and* a valid grade 7 mathematics score to be included in this table.

Table 9.29 Average Correlations between ELA/L and Mathematics for Grade 8

	CBT				PBT				
	ELA/L	RD	WR	MA	ELA/L	RD	WR	MA	
ELA/L		308	308	308	ELA/L		341	341	341
RD	0.95		308	308	RD	0.95		341	341
WR	0.91	0.75		308	WR	0.90	0.73		341
MA	0.73	0.72	0.62		MA	0.74	0.73	0.62	

Note: ELA/L = English language arts/Literacy, RD = Reading, WR = Writing, MA = Mathematics. The average correlations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table. Students must have a valid grade 8 ELA/L score *and* a valid grade 8 mathematics score to be included in this table.

Table 9.30 Average Correlations between ELA/L and Mathematics for High School

ELA/L	CBT						ELA/L	PBT					
	A1	GO	A2	M1	M2	M3		A1	GO	A2	M1	M2	M3
8	0.66 (126)						8	0.64 (111)					
9	0.69 (206)			0.74 (122)			9	0.70 (151)			0.69 (170)		
10		0.67 (107)					10	0.54 (125)					
11			0.66 (118)				11						

Note: ELA/L = English language arts/Literacy, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. The average correlations are provided with the average sample sizes, below in parentheses.

Table 9.31 Average Correlations between ELA/L Reading and Mathematics for High School

RD	CBT						RD	PBT					
	A1	GO	A2	M1	M2	M3		A1	GO	A2	M1	M2	M3
8	0.66 (126)						8	0.66 (111)					
9	0.68 (206)			0.73 (122)			9	0.69 (151)			0.68 (170)		
10		0.67 (107)					10	0.51 (125)					
11			0.66 (118)				11						

Note: RD = Reading, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. The average correlations are provided with the average sample sizes, below in parentheses.

Table 9.32 Average Correlations between ELA/L Writing and Mathematics for High School

WR	CBT						WR	PBT					
	A1	GO	A2	M1	M2	M3		A1	GO	A2	M1	M2	M3
8	0.55 (126)						8	0.53 (111)					
9	0.58 (206)			0.63 (122)			9	0.58 (151)			0.58 (170)		
10		0.58 (107)					10	0.44 (125)					
11			0.56 (118)				11						

Note: WR = Writing, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. The average correlations are provided with the average sample sizes, below in parentheses.

9.5 Evidence from the Special Studies

Several research studies were conducted to provide additional validity evidence for the PARCC's goals of assessing more rigorous academic expectations, helping to prepare students for college and careers, and providing information back to teachers and parents about their students' progress towards college and career readiness. Some of the special studies conducted include:

- benchmarking study,
- content evaluation studies,
- mode comparability study, and
- device comparability study.

The following paragraphs briefly describe each of these studies.

9.5.1 Benchmarking Study

The purpose of the PARCC Benchmarking Study (McClarty, Korbin, Moyer, Griffin, Huth, Carey, and Medberry, 2015) was to provide information that would inform the PARCC performance level setting (PLS) process. PARCC used an Evidence-Based Standard Setting approach (EBSS; McClarty, Way, Porter, Beimers, & Miles, 2013) to establish the performance levels for its assessments. In EBSS, the threshold scores for performance levels are set based on a combination of empirical research evidence and expert judgment. This benchmarking study provided one source of empirical evidence to inform the PARCC college and career readiness performance level (i.e., Level 4). The study findings were provided to PARCC's pre-policy standard-setting committee. The charge of this committee was to suggest a reasonable range for the percentage of students meeting or exceeding the PARCC Level 4 threshold score and therefore considered college- and career-ready. Section 11.3.2 of this report provides more information about the PARCC pre-policy meeting. For the PARCC Benchmarking Study, external information was analyzed to provide information about the Level 4 threshold scores for the grade 11 ELA/literacy, Algebra II, and Integrated Mathematics III assessments, the grade 8 ELA/literacy and mathematics assessments, and the grade 4 ELA/literacy and mathematics assessments. The PARCC assessments and Level 4 expectations were compared with comparable assessments and expectations for the Programme of International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), National Assessment of Educational Progress (NAEP), ACT, SAT, the Michigan Merit Exam, and the Virginia End-of-Course exams. For each external assessment, the best-matched performance level was determined and the percentage of students reaching that level across the nation and in the PARCC states was determined. Across all grades and subjects, the data indicated approximately 25 to 50 percent of students were college- and career-ready or on track to readiness based on PARCC's Level 4 expectations.

For details on how the benchmarking study was used during the standard setting process, refer to Section 11 of this technical report.

9.5.2 Content Alignment Studies

The content of the ELA/L assessments at grades 5, 8, and 11 and the Algebra II and Integrated Mathematics II assessment were evaluated to determine how well the PARCC assessments were aligned to the Common Core State Standards (CCSS; Doorey, & Polikoff, 2016, Schultz, Michaels, Dvorak, & Wiley, 2016). These content alignment studies were conducted by the Fordham Institute for grades 5 and 8 and by Human Resources Research Organization (HumRRO) for the high school assessments. Both of these studies used the same methodology by having content experts review the assessment items and answers (for the constructed response items the rubrics were reviewed). The content experts then judged how well the items aligned to the CCSS, the depth of knowledge of the items, and the accessibility of the items to all students, including English learners and students with disabilities. The authors of both studies noted that the content experts reviewing the assessments were required to be familiar with the CCSS but could not be employed by participating organizations or be the writers of the CCSS. Therefore, an effort was made to eliminate any potential conflicts of interest.

The content studies had the individual content experts review and rate each item then as a group the content experts came to a consensus on the final ratings for the content alignment, depth of knowledge, and accessibility to all students. In addition to the ratings, the content experts were asked to make comments that provided an explanation of their ratings; these comments were then used by the full group of content experts to provide narrative comments regarding the overall ratings and to provide feedback and recommendation about the assessment programs.

The PARCC assessment program was rated as *Excellent Match* for ELA/L content and depth and *Good Match* for mathematics content and depth for grades 5 and 8. However, for grade 11 ELA/L content was rated as *Excellent Match* but depth was rated as *Limited/Uneven Match*. The high school mathematics assessments were rated at *Excellent Match* for content and *Good Match* for depth.

The content studies noted some weaknesses and strengths of the PARCC assessments. For ELA/L it was noted that the assessments include complex texts, a range of cognitive demands, and have a variety of item types. Furthermore, the ELA/L “assessments require close reading, assess writing to sources, research, and inquiry, and emphasize vocabulary and language skills” (Doorey & Polikoff, 2016). The grade 11 ELA/L assessment had a smaller range of depth and included items assessing the higher-demand cognitive level. A weakness of the ELA/L assessments is the lack of a listening and speaking component. It was also suggested that the ELA/L assessments could be enhanced by the inclusion of a research task that requires the use of two or more sources of information.

The strengths of the mathematics assessments include assessments that are aligned to the major work for each grade level. While the grade 5 assessment includes a range of cognitive demand, the grade 8 assessment includes a number of higher-demand items and may not fully assess the standards at the lowest level of cognitive demand. It was suggested that the grade 5 assessment could include more focus on the major work and the grade 8 assessment could include items at the lowest cognitive demand level. Additionally, the reviewers noted that some of the mathematics items should be carefully reviewed for editorial and mathematical accuracy.

The high school report noted that the PARCC assessment program incorporates a number of accessibility features and test accommodations for students with disabilities and for English language learners. Furthermore, the PARCC assessments included items designed to accommodate the needs of students with disabilities.

9.5.3 Mode Comparability Study

The PARCC (Operational) Mode Comparability Study was conducted using the 2015 operational data to support both computer-based testing (CBT) and paper-based testing (PBT) modes of administration of the PARCC assessments (Liu, Brown, Chen, Ali, Hou, & Costanzo, 2016).

For the spring 2015 operational administration, schools and districts within each state selected the mode of test administration. The resulting CBT and PBT test-taking groups were therefore not randomly equivalent. To improve the overall comparability of the CBT and PBT groups, propensity score matching, based on test-taker demographic information, was used. Then item-level analyses (e.g., p values, and differential item functioning) and test-level analyses (e.g., test characteristic curves) were conducted.

Item-level analyses showed that there were negligible to small differences in terms of p values and average item scores across modes for the majority of items in mathematics and ELA/L. Prose Constructed Response (PCR) task traits in ELA/L had larger p value effect sizes than other items, all favoring PBT. A very small percentage of items was identified as functioning differently (with C-level DIF) in the two modes. Many items ELA/L PCR task traits were also found to have B-level (DIF), favoring PBT.

Additionally, the item response theory (IRT) difficulty and discrimination parameters estimated separately within mode were highly correlated. For grade levels with a lower correlations between modes, removing items with outlier parameter estimates provided substantial improvement in the correlation. As well, the overall the differences between common test characteristic curves (TCCs) of different modes were small and within 0.5 raw score points, except for ELA/L grade 9 and Geometry where TCC differences exceeded the differences that matter criterion in regions of the theta scale where large percentages of students were located. When comparing the performance on the common items, the effect sizes ranged from negligible to small for most of the tests evaluated. The directions of effect sizes were not consistent across subject/grade levels.

Additional analyses were conducted on students from one of the states that provided prior state assessment scores. Summary statistics of these students' prior state assessment scores suggested CBT and PBT samples from propensity score matching (PSM) were not comparable in their prior achievement. Therefore, poststratification weights based on prior state assessment score were used to calculate PBT students' PARCC scale score to minimize the impact of noncomparability of prior achievement across modes. The scale score differences were largely reduced for mathematics grade 5, 7 and Algebra I after weighting. Small effect sizes, in favor of PBT, were found for Geometry and ELA/L grade 9 and a negligible effect size was found for ELA/L grade 7 after poststratification weighting.

The PARCC (Operational) Mode Comparability Study found evidence that the score comparability was not consistent across all content domains and grade levels. As noted in the study, only one state

provided previous year's achievement data, therefore, the CBT and PBT groups were matched based on only demographic data. Furthermore, the additional analyses based on the one state that provided prior achievement data indicated that the CBT and PBT matched groups were not comparable in terms of their prior achievement. Thus, caution should be taken when interpreting the results of the Mode Comparability Study.

9.5.4 Device Comparability Study

In addition to the PARCC (Operational) Mode Comparability Study, the comparability across digital devices (e.g., tablet versus non-tablet) was evaluated using the 2015 operational data (Steedle, McBride, Johnson, & Keng, 2015).

PARCC allows students to take its assessments on a variety of digital devices, such as desktops, laptops, and tablets. It is therefore important to evaluate comparability across digital devices by investigating whether test items were of similar difficulty, whether psychometric properties of test scores were similar, and whether overall test score interpretation was similar across traditional (i.e., desktops and laptops) and non-traditional (i.e., tablet) computing devices. For the 2015 Device Comparability Study, any student who took one of the study forms on a tablet or non-tablet device were eligible for inclusion in the study. Students were matched on demographic information to create tablet and non-tablet samples that were considered randomly equivalent.

The 2015 Device Comparability Study found evidence of comparability between test scores from tablets and non-tablet devices. The item p values and IRT difficulty estimates were similar across tablets and non-tablet devices. A small number of items were flagged for device effects, and nearly all of them were part of high school mathematics assessments. The raw score and scale score distributions indicated similar overall performance on both components (PBA and EOY) of the 2015 PARCC assessments. Additionally, IRT true-score equating indicated that students who tested on non-tablet devices would be expected to score similarly had they taken the same PARCC assessment on tablets.

9.6 Evidence Based on Response Processes

As noted in the AERA, APA, and NCME Standards (2014), additional support for a particular score interpretation or use can be provided by theoretical and empirical evidence indicating that test takers are using the intended response processes when responding to the items in a test. This type of evidence may be gathered from interacting with test takers in order to understand what processes underlie their item responses. Evidence may also be derived from feedback provided by test proctors/teachers involved in the administration of the test and raters involved in the scoring of constructed response items. Evidence may also be gathered by evaluating the correct and incorrect responses to short constructed response items (e.g., items requiring a few words to respond) or by evaluating the response patterns to multi-part items.

PARCC has undertaken research investigating the quality of the items, tasks, and stimuli, focusing on whether students interact with items/tasks as intended, whether they were given enough time to

complete the assessments, and the degree to which PARCC scoring rubrics allow accurate and reliable scoring. In addition, PARCC has examined the accessibility of the test for students with disabilities and English learners. This research has included examining students' understanding of the format of the assessments and the use of technology. Although out of the purview of this technical report, several other PARCC research efforts have investigated questions relevant to response processes evidence.¹³

9.7 Interpretations of Test Scores

The PARCC ELA/L and mathematics scores are expressed as scale scores (both total scores and claim scores), along with performance levels to describe how well students met the academic standards for their grade level. Additionally, information on specific skills (the subclaims) is also provided and is reported as "Below Expectations", "Nearly Meets Expectations" and, "Meets or Exceeds". On the basis of a student's total score, an inference is drawn about how much knowledge and skill in the content area the students has acquired. The total score is also used to classify students in terms of the level of knowledge and skill in the content area as students progress in the K-12 education. These levels are called performance levels and are reported as:

- Level 5: Exceeded expectations
- Level 4: Met expectations
- Level 3: Approached expectations
- Level 2: Partially met expectations
- Level 1: Did not yet meet expectations

Students classified as either Level 4 or Level 5 are meeting or exceeding the grade level expectations. PARCC has developed performance level descriptors (PLDs) to assist with the understanding and interpretations of the ELA/L and mathematics scores (<http://www.parcconline.org/news-and-video/230-performance-level-descriptors>). Additionally, resource information is available online to educators, parents, and students (<http://www.parcconline.org/resources>), which includes information on understanding and interpreting the ELA/L and mathematics score reports (<http://www.state.nj.us/education/assessment/parcc/scores/Fall14Spring15SRIG.pdf> and <http://www.parcconline.org/assessments/score-results>).

Section 12 of this technical report provides more information on the scale scores and the subclaim scores.

9.8 Evidence Based on the Consequences to Testing

The consequence of testing should also be investigated to support the validity evidence for the use of the PARCC assessments as the Standards note that tests are usually administered "with the expectation that some benefit will be realized from the intended use of the scores" (AERA, APA, & NCME, 2014). When this is the case, evidence that the expected benefits accrue will provide support for the intended

¹³ Various PARCC research is described at: <http://www.parcconline.org/assessments/test-design/research>

use of the scores. Evidence of the consequence of testing will also accrue with the continued implementation of the CCSS and the continued administration of the PARCC assessments.

Consequences of the PARCC tests may vary by state or by school district. For example, some states may require “passing” the PARCC assessments as one of several criteria for high school graduation, while other states/districts may not require students to “pass” the PARCC assessments for high school graduation. Additionally, some school districts may use the PARCC scores along with other information such as school grades and teacher recommendations for placing students into special programs (e.g., remedial support, gifted and talented program) or for course placement (e.g., Algebra I in grade 8). Because the consequences for the PARCC assessments can vary by each state, it is suggested that each PARCC member state provide school districts, teachers, parents, and students with information on how to interpret and use the PARCC scores. Additionally, the states should monitor how PARCC scores are used to ensure that the scores are being used as intended by PARCC.

9.9 Summary

In this section of the technical report, several aspects of validity were included, such as validity evidence based on content, the internal structure of the assessments, relationships across the content assessments, and from special studies.

The PARCC item development process involved educators, assessment experts, and bias and sensitivity experts in review of text, items and tasks for accuracy, appropriateness, alignment to the instructional standards, and freedom from bias. PARCC conducted several studies during the item development process to evaluate the item development process (e.g., technological functionalities, answer time required, and student experiences). Additionally, items were field tested prior to the 2015 operational administration and data and feedback from students, test administrators, classroom teachers was used to improve the operational administration of the items and to inform future item development. The multiple item and form reviews conducted by educators and studies to evaluate item administration help to ensure the integrity of the PARCC assessments.

The intercorrelations of the subclaims, the reliability analyses, and the local item dependence analyses indicated that the ELA/L and the mathematics assessments are both essentially unidimensional. Furthermore, the correlations between ELA/L and mathematics indicated that the two assessments are measuring different content. Also, the patterns of correlations for the CBT and PBT assessments were similar indicating that the structure of the assessments were similar across the two modes.

Several studies were conducted as part of the PARCC assessment program (e.g., benchmarking study, content evaluation/alignment studies, mode and device comparability studies). The benchmarking study was conducted in support of the standard setting meeting. This study indicated students performing at or above Level 4 could be considered to be college- and career-ready or on track to readiness.

The content evaluation/alignment studies performed by the Fordham Institute and HumRRO indicate that the PARCC assessments are good to excellent matches to the CCSS in terms of content and depth of knowledge. Thus, the PARCC assessments are assessing the college- and career-readiness standards.

However, the reports noted that the PARCC program could improve by adding a wider range of depth of knowledge to some of the assessments. The reports also suggested enhancing the ELA/L assessments by including a research task that requires the use of two or more sources of information.

The mode comparability study indicated that the comparability across modes was inconsistent across content domains and grade levels. The mode comparability study indicated that outliers should be removed from the mode anchor set and that the PCR items should be considered for exclusion from the anchor set. Furthermore, the scoring of the PCR items should be carefully reviewed.

The device comparability study indicated that there were some, but small, effects of testing device when comparing tablet to non-tablet devices. While a small number of mathematics tasks were flagged for device effects, the raw and scale score distributions were similar across the testing devices. The equating analyses indicated that students could expect to receive a similar score regardless of the testing device.

In addition to the validity information presented in this Section of the technical report, other information in support of the uses and interpretations of the PARCC scores appear in the following sections:

Section 5 presents information regarding student characteristics for the spring administration of the ELA/L and mathematics administration.

Section 6 provides information concerning the test characteristics based on classical test theory.

Section 7 provides information regarding the differential item functioning analyses (DIF).

Section 8 provides information on the test reliability (total test score and for subclaims) and includes information on the interrater reliability/agreement.

Section 12 provides detailed information concerning the scores that were reported and the cut scores for ELA/L and mathematics.

The technical report addendum provides the test taker characteristics, the classical test theory characteristics, and DIF information for the 2014 Fall block administration.

Section 10: IRT Calibration and Scaling in Operational Year One

10.1 Overview

Multiple operational core forms were administered for each grade in English Language Arts/Literacy (ELA/L) and mathematics. The purpose of the item response theory (IRT) calibration and scaling was to place all operational items for a single grade/subject onto a common scale. This section addresses procedures used to calibrate and scale the PARCC operational item response data using IRT. Based on the results of the 2014 field test dimensionality and mode comparability studies, the operational data were calibrated concurrently across the Performance Based Assessment (PBA) and the End-of-Year (EOY) assessment, and separately by mode (computer-based tests, or CBT, and, paper-based tests, or PBT) using IRT models consistent with mixed format data. The PBT IRT parameter estimates were then transformed onto the CBT scale using the Stocking and Lord (1983) procedure.

In this section¹⁴ of the technical report we discuss the following topics related to IRT calibration and scaling:

Calibration:

- 10.2 IRT Sparse Data File Preparation
- 10.3 Description of the calibration process
- 10.4 Model fit evaluation criteria
- 10.5 Items excluded from score reporting

Scaling:

- 10.6 Description of scaling process (Paper to Online)
- 10.7 Items Excluded from Spring 2015 Paper to Online Linking Sets
- 10.8 Correlations and Plots of Parameter Estimates
- 10.9 Scaling constants
- 10.10 Summary Statistics and Distributions from IRT Analyses
- 10.11 Effect Sizes of Linking Average Item Score Differences versus Scale Score Differences

10.2 IRT Data Preparation

10.2.1 Overview

The first step, before IRT calibration, was to use the scored item response block data to create IRT data matrices. For operational items only, IRT data matrices were prepared by combining students' matched PBA and EOY test components (i.e., full summative) for each grade/subject and content area. In this response data, each row represented one student record while each column represented the scored responses for one item. All unique CBT items for a grade (e.g., 149 items in ELA/L grade 6) and all unique PBT items (e.g., 100 in ELA/L grade 6) were included in analysis files and calibrated separately. Valid PBA

¹⁴ This section focuses on the Spring 2015 administration, and Addendum 10 provides information on results from the Fall 2014 administration.

and valid EOY data for students were matched and merged. For students who only took only one component (either PBA or EOY, but not both), the item response data for the component they took were included in the appropriate IRT calibration data file. Also, for students who took mixed test modes, that is: one component online and the other on paper, their PBA and EOY records were not matched. Instead, their item response data were included in the appropriate CBT or PBT IRT calibration data file separately. When duplicate records for a single student existed for PBA and/or EOY, the record with the largest raw score was used in the IRT calibration data file (and the other record was excluded). No student was included more than one time in the CBT and PBT IRT calibration data file.

10.2.2 Student Inclusion/Exclusion Rules

The following are the IRT valid case criteria. These criteria are the same as the student inclusion/exclusion rules used to evaluate and filter data prior to conducting the operational item analysis (IA) and differential item functioning (DIF) analyses (steps 1-5). The rules were agreed upon with PARCC and applied to the scored IRT data. The first 5 steps were applied at the component level (PBA/EOY). Then the matching of PBA and EOY records was done for IRT analysis only at step 6.

1. All component records with an invalid form number were excluded.
2. All component records flagged as “void” based on the student file layout were excluded.
3. Records in which the student attempted fewer than 25% of the PBA items and/or attempted fewer than 25% of the EOY items were excluded. An item was deemed “not attempted” if it had a value of “M” (item omitted) or “Z” (item ‘spoiled’, do not score) in the scored item response block. For example, if there were 25 items on a form and two were flawed (“Z”), those two items were not included in the numerator or denominator of the percentage attempted calculation.
4. If a student had duplicate valid records for either PBA or EOY, the record with the higher raw score (e.g., $PBA_1 = 7$ and $PBA_2 = 29$; chose PBA_2) was chosen.
5. Individual student records were excluded based on a list. The list contained the PARCC student identifiers with administration issues or anomalies.

Test components were matched using the variables “TestCode” and “PARCCStudentIdentifier” fields from the student data files; all pairs of valid components (PBA and EOY) for students were matched in the IRT sparse matrices. (Note: All records not matched were included in the appropriate CBT or PBT data file.)

All accommodated students were included in the IRT calibrations except for students taking certain forms including: a) Spanish forms (mathematics only), b) American Sign Language (ASL) forms online, and c) AT/Screen Reader forms online. The assumption was made that mathematics items translated into Spanish were equivalent to the same items in English. The results of Spanish versus English differential item functioning (DIF) analyses supported this assumption. Also, ASL and AT/Screen Reader forms were delivered online, but were constructed from PBT items. As a result, these students could not be combined with the CBT and/or PBT IRT data files in a psychometrically defensible way.

10.2.3 Items Excluded from IRT Sparse Matrices

Before delivering data files to ETS, Pearson conducted an initial scoring and key check. Items identified by Pearson as “spoiled” (also referred to as “do not use (DNU)”) were listed and excluded from the test maps. When the IRT sparse data matrices were created, all items were included in the files unless they were marked as “spoiled” by Pearson.

10.2.4 Omitted, Not Reached and Not Presented Items

In the Pearson data files, ‘Z’ was used to represent “spoiled” or “not presented” items and ‘M’ was used to represent omitted items. For IA and IRT, *omitted* and *not reached* items were treated differently. Item response scores for omits were recoded as ‘O’ in the IRT sparse matrix files (i.e., *unless* the omitted item was a “not reached” item). Not reached items are omitted items at the end of the test – items that the student probably did not reach or try to answer. Not Reached items were recoded from ‘M’ in the Pearson SIRB to ‘N’ (i.e., not presented) in the IRT sparse matrix files, if all items from that point, to the end of the form, are ‘M’ or ‘Z’. Not reached items were counted as *missing* or *no response*, and therefore did not contribute to the item statistics for IA and IRT calibration.

10.2.5 Quality Control of the IRT Sparse Matrix Data Files

The IRT sparse data matrices created by ETS were checked for quality and accuracy by comparing the number of students (*N*-counts), item category frequencies, and item statistics (e.g., AIS values) of the IRT data files to those obtained from the IA results. Since the same inclusion rules for students were used for both IA and IRT, nearly all *N*-counts, category frequencies, and statistics for all items matched. All discrepancies in *N*-counts were resolved. The programs used to create the IA statistics and the IRT statistics were independent, so the QC procedure involved parallel computing. Tables 10.1 and 10.2 show the *N*-Counts (N), percentage of students (Percent), and number of items (*n*-Items) in the CBT and PBT IRT sparse data matrices for each grade in ELA/L and mathematics.

Table 10.1 N-Counts, Percent of Students, and Number of Items in the ELA/L IRT Calibration Files

Grade	N			Percent		n-Items	
	ALL	CBT	PBT	CBT	PBT	CBT	PBT
3	528,076	386,078	141,998	73.1	26.9	111	75
4	645,178	488,016	157,162	75.6	24.4	125	87
5	651,594	505,551	146,043	77.6	22.4	138	87
6	649,508	525,073	124,435	80.8	19.2	157	100
7	645,051	530,131	114,920	82.2	17.8	150	100
8	641,937	524,188	117,749	81.7	18.3	150	100
9	452,278	378,209	74,069	83.6	16.4	215	100
10	299,897	274,287	25,610	91.5	8.5	194	105
11	202,468	189,323	13,145	93.5	6.5	156	120

Table 10.2 N-Counts, Percent of Students, and Number of Items in the Mathematics IRT Calibration Files

Grade/ Subject	N			Percent		n-Items	
	ALL	CBT	PBT	CBT	PBT	CBT	PBT
3	657,888	456,327	201,561	69.4	30.6	269	158
4	644,122	480,810	163,312	74.6	25.4	199	135
5	650,202	498,129	152,073	76.6	23.4	206	135
6	646,836	523,925	122,911	81.0	19.0	192	132
7	627,478	518,545	108,933	82.6	17.4	238	131
8	523,074	426,743	96,331	81.6	18.4	196	117
A1	520,084	438,802	81,282	84.4	15.6	275	136
GO	227,413	204,515	22,898	89.9	10.1	288	136
A2	215,502	199,489	16,013	92.6	7.4	200	135
M1	33,844	30,164	3,680	89.1	10.9	73	69
M2	15,260	13,878	1,382	90.9	9.1	78	75
M3	11,237	9,797	1,440	87.2	12.8	81	70

Note: A1 = Algebra I, A2 = Algebra II, GO = Geometry, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, and M3 = Integrated Mathematics III.

10.3 Description of the Calibration Process

10.3.1 Special Studies to Inform the Operational Calibration and Scaling

The purposes of the 2014 field test (FT) Item Response Theory (IRT) calibration and scaling analyses were to provide data for operational test forms construction and to inform the model selection for the operational administration. Using sparse IRT data matrices, within-grade concurrent calibrations of PBA and EOY items were performed separately on CBT and PBT data using three model combinations: one-parameter/partial credit model (1PL/PC), two-parameter logistic/generalized partial credit model (2PL/GPC), and three-parameter logistic/generalized partial credit model (3PL/GPC). In addition, dimensionality and mode comparability studies were completed on the 2014 field-test data.

Dimensionality Study

The results of the 2014 FT dimensionality study were presented in the 2014 field test technical report. One of the primary questions explored in the dimensionality study was: Is a unidimensional IRT model appropriate for the calibration of PARCC item response data? On the whole, the results appear to support the use of a unidimensional scaling model. The results of the *exploratory* factor analyses indicated that more than one factor was reflected in item performance. A follow-up investigation of the secondary factor(s) for a subset of ELA/L grades suggested that items loading on secondary factors might be described as “problematic” due to extreme point-biserials or percentage correct statistics. The findings of the *confirmatory* factor analyses support this, because multidimensional models that might be expected (e.g., separate factors for Reading and Writing items) did not fit better than the one factor model.

Another critical question explored in the dimensionality study was: Is it appropriate to calibrate item response data for both components (i.e., PBA and EOY) concurrently? The results of the *confirmatory* factor analyses indicate that a two-factor model (i.e., PBA and EOY as separate factors) did not fit better than the one-factor model. As a result, it was concluded that the field test data supported a combined (PBA and EOY) unidimensional calibration approach.

Mode Comparability Study

The results of the 2014 field-test mode comparability study were presented as a special report. The study evaluated to what extent scores from the online (CBT) and paper (PBT) form versions could be considered comparable in their psychometric characteristics. The findings indicated that scores were not comparable across modes in a strict sense, particularly for PBA. However, there was substantial evidence indicating that the differences in comparability across modes were relatively minor. When comparing the performance on the common items, there were small effect sizes in favor of PBT for the mathematics and ELA/Literacy PBA components and negligible effect sizes for EOY and full summative assessments.

Specifically, the differential item functioning (DIF) results indicated that a small number of items for English Language Arts/Literacy (i.e., 0 to 7 items per grade) and a slightly higher number of items for mathematics (i.e., 2 to 17 items per grade) possessed either positive or negative C-level DIF across modes. This would indicate that these items were functioning differently across modes.

There were several implications for the operational calibration and scaling plan based on these findings. First, since DIF clearly existed for some items in the 2014 field test study, it was appropriate to calibrate operational CBT and PBT items separately for each grade/subject. Second, when scaling PBT item parameter estimates to the CBT scales, the exclusion rules used for linking 2014 field test items were appropriate and should also be used for the 2015 operational administration (i.e., items flagged for positive or negative C-DIF should be removed from the linking sets). Common items that behaved differently across modes should be treated as separate and unique items and different CBT and PBT item parameter estimates should be used for generating operational scoring tables.

The mode comparability study was repeated using the 2015 operational data and are summarized in a special report (Liu, Brown, Chen, Ali, Hou, & Costanzo, 2016). Like the 2014 field-test study, the spring 2015 findings indicated that scores across the two modes were not comparable in a strict sense. When comparing the item difficulties by evaluating the p value effect sizes, the ELA/L Prose Constructed Response (PCR) traits had larger p value effect sizes than the other items, all favoring PBT. Additionally, a very small percentage of items was identified as functioning differently (with C-level DIF) across the two modes. Many items with B-level (DIF) favoring PBT were PCR traits. The analysis of IRT parameter estimates revealed that IRT difficulties and discriminations estimated separately within mode were highly correlated. For grade levels with lower correlations between modes, removing items with outlier parameter estimates provided substantial improvement in the correlations of the common item parameters.

The implications of the 2015 operational mode comparability study on operational item calibration and scaling are similar to the implications from the 2014 field-test study. First, DIF continues to exist for some items. Thus, the operational CBT and PBT items should be calibrated separately for each grade/subject. Second, when scaling PBT item parameter estimates to the CBT scales, the exclusion rules used for the PBT-to-CBT common-item linking sets should be the same (i.e., items flagged for positive or negative C-DIF should be removed from the linking sets). Finally, common items that behaved differently across modes should be treated as separate and unique items and different CBT and PBT item parameter estimates should be used for generating operational scoring tables

Device Study

The 2014 field-test device comparability study investigated the effects of taking the CBT field test using a traditional computing device, such as a desktop or laptop, versus a nontraditional digital device, such as a tablet. Students participating in the study were either randomly assigned to groups for Grade 8 ELA and mathematics, Grade 10 ELA and geometry, or matched based on their performance on the state assessment test for Grade 4 ELA and mathematics. Although the size and scope of the study was limited, the study found no consistent or significant device effects in the selected tests. Given the study limitations, however, the results were considered preliminary and the study was repeated on a larger sample using the 2015 operational data.

The 2015 operational device comparability study (Steedle, McBride, Johnson, & Keng, 2015) found evidence of the comparability between testing on traditional and nontraditional computing devices. While no specific task type was found to consistently exhibit differences across devices, a small number of items for high school mathematics was flagged for device effects. Additionally, IRT true-score equating analyses indicated that test takers would be expected to score similarly on the PARCC assessments regardless of the testing device. The results of the device study indicate that data from traditional and nontraditional computing devices can be calibrated together and that the same CBT item parameters can be used for creating the operational scoring tables.

Choosing the 2PL/GPC IRT Model Combination

Based on the analyses conducted in 2014 using field test data, the 2PL/GPC model combination was recommended for operational use for both ELA/L and mathematics. For ELA/L, the field test results clearly indicated that the item fit was better for the 2PL/GPC than the 1PL/PC model combination. The results were consistent for the PARSCALE (Muraki & Bock, 2003) chi-square item fit statistics, the empirical and observed item plots, and the final log likelihood calibration values. Also, the 1PL/PC model assumption, that item discrimination was constant across items, did not hold true based on the field test data.

For mathematics, the 2PL/GPC model was also recommended for operational use, although consideration was made for the 3PL/GPC model combination. Based on the PARSCALE chi-square item fit statistics for most grades (e.g., 3 through 8), the empirical and observed item plots, and the final log likelihood calibration values, the fit was generally better for the 2PL/GPC than the 1PL/PC model

combination. Relatively few CBT items were single select multiple choice items (approximately 15% to 23%). Consequently, the 3PL/GPC model did not markedly improve fit when compared to the fit of the 2PL/GPC model.

A concern about using the 3PL model for mathematics SSMC items was based on Holland's (1990) findings that estimation using unconstrained versions of 3PL models might be expected to have stability problems, and that further intervention would be needed to address *c*-parameter estimation issues (by manipulating priors and holding them constant at logical values). Therefore, to minimize the need for human judgments and intervention, the 2PL model was recommended. Ultimately, PARCC made the decision to use the 2PL/GPC model for operational assessments.

Treatment of Prose Constructed Response (PCR) tasks

PCRs are writing tasks only administered on the PBA portion of the PARCC ELA/L assessment. The student receives a prompt and writes a response which is then scored using a multi-trait rubric. An *aggregated* PCR item score is determined by adding together the multiple scores the student received on the two or three traits. PCRs consist of at least two writing traits (Written Expression and Writing Knowledge and Convention) and, in some cases, a reading trait (Reading Comprehension) as well. One of the writing traits (Written Expression) is weighted by 3 to give it more emphasis in the total score. The aggregated PCR scores had total maximum points possible range from 12 to 19 depending on the item and the grade. In the 2014 field test analysis, *aggregated* PCR scores were delivered to ETS. All the 2014 FT PCR items, had to be collapsed due to few or no item responses at various categories. A special data set was performed to handle the collapsing of 2014 FT categories.

The 2014 FT PCR data revealed that it was critical to determine how best to handle the PCR items operationally and that an *aggregation* approach was not feasible for calibration or reporting claim level scores. As a result, a special study was conducted by ETS. The findings indicated that PCRs should be calibrated at the *trait* level (rather than the *aggregated* level). The one concern about calibrating PCR items as traits, was that it could potentially violate the IRT assumption of local independence.

To address the issue of local independence related to PCR items, a single-calibration "model" approach was used. When sample sizes were large (i.e., greater than 10,000 test takers), the data were manipulated using random assignment, by selecting one of the two or three traits for each PCR item for each student. Then one calibration was run so that all trait parameters were independently estimated. When sample sizes were smaller (i.e., for Integrated Mathematics and Fall block samples), a multiple-calibration "model" approach was used. In this alternative approach, the same data set was calibrated three times, each trait represented in one of the three data sets for all students. Then the PCR traits were scaled onto the base (Reading trait) scale using non-PCR items as anchor items. These two trait calibration approaches addressed the issue of local dependence, and also allowed for the accurate calculation of claim scores, and the proper weighting of traits in the summative scale scores.

10.3.2 IRT Item Exclusion Rules (Before Calibration)

In addition to checking IRT data for accuracy, ETS conducted item analyses (IA) to identify items that were not performing as expected and should be considered for removal from calibration and score reporting. The following are the criteria ETS used to flag extremely problematic items to be dropped from calibration. All “non-spoiled” items were included in the IRT data matrices, however, the **PARSCALE** control files were used to exclude from calibration items flagged for the following reasons:

1. A weighted polyserial correlation less than 0.0
2. An average item score of 0.0
3. 100% of the students have the same item score, such as:
 - a. 100% omitted the item,
 - b. 100% received the same score,
 - c. 100% of the responses were at the same score after collapsing score categories due to low frequencies, or
 - d. 100% of the responses were not presented or not reached
4. Insufficient sample sizes for the selected IRT model combinations (i.e., 300 for the 2PL/GPC).
5. High omit rates (i.e., greater than 50%) on one or more forms (usually an indication that an item may not be functioning correctly on all forms).

A master list of all problematic items before and after calibration was maintained and all flagged and potentially flawed items were brought to the PARCC Priority Alert Task Force (consisting of Parcc Inc. and participating State Leads) for content and statistical reviews. Ultimately the decisions about whether to keep or exclude an item from score reporting was made by the Task Force.

10.3.3 PARSCALE Calibration Procedures and Convergence Criteria

As stated in Section 10.1, based on the results of 2014 field test dimensionality and mode comparability studies, the data were calibrated concurrently across components (PBA and EOY) and separately by mode (CBT and PBT) using the 2PL/GPC model combination. The PBT parameter estimates were then transformed onto the CBT scale using the Stocking and Lord (1983) procedure. The primary goal was to place the operational item data within each content area and grade/subject on a common difficulty scale. The following are the steps used to calibrate the Spring 2015 operational PBA and EOY item response data:

1. Using the IRT sparse data matrices, concurrent calibrations were conducted using commercially available **PARSCALE** for Windows (version 4.1.28303.1) on CBT data (PBA and EOY items together), and separately on PBT data (PBA and EOY items together) within each grade/subject.
2. The 2PL/GPC model combination was used for all grades and subjects for each content area (with a few exceptions described later). Thus, two calibrations were completed for each grade/subject.
3. **PARSCALE** Calibration Settings: The logistic partial credit model was specified using the scale constant of 1.7. The prior distributions for latent traits were set to a mean of 0.00 and a SD of 1.00.

The number of quadrature points used in the estimation was set to 41. And, the slope starting value was set/updated before each run.

4. Calibration runs:

- 1) One goal, when calibrating the operational item response data, was to keep as many items as possible (i.e., avoid deleting items from the calibration pool). Items without IRT parameter estimates could not be included in the generation of conversion tables.
- 2) A critical step used to produce replicable results was to set/update the PARSCALE starting slope value equal to the final mean slope of the previous PARSCALE “run” so that the final mean slope (e.g., 0.659) was equal to (or within + or - .001 of) the starting slope value (i.e., 0.659). Multiple **PARSCALE** “runs” were often needed to obtain convergence and a final mean slope value equal to the starting slope value.
- 3) Convergence occurred when the “criterion of largest parameter change” was less than 0.005 and the pattern of log likelihood values decreased smoothly from the first to the final cycle.
- 4) The number of EM cycles was set to 100 for most runs. **PARSCALE** usually converged in 100 or fewer EM cycles. If the program converged in less than 100 cycles, and the final mean slope was equal to the starting slope, then the calibration was completed.
- 5) If **PARSCALE** *nearly* converged in 100 EM cycles, then the number of EM cycles was increased to slightly higher (e.g., to 120), the starting slope value was updated (if necessary), and **PARSCALE** was run again. Sometimes PARSCALE needed only a few additional cycles to converge.
- 6) If the calibration software did *not* converge the following steps were taken:
 - a. The most “problematic” item was identified and examined by reviewing Phase 1 and Phase 2 program outputs.
 - b. Apply the keyword “**category**” in the control file to problematic item so that alternative starting values, generated by PARSCALE, were applied to the item difficulty and the item category parameter starting values (i.e., the *d*-parameter estimates). Then, run the software again. If applying the keyword “category” solved the estimation problem and the run converged, then the calibration was completed.
 - c. If the item parameter estimation was still problematic, then the possibility of **collapsing** the item categories was considered. This was used in particular when there were no students in the highest score category for polytomous items. When feasible, collapsing was done by modifying the PARSCALE control file.
 - d. If steps 6b and 6c did not resolve the estimation issues, there was no choice but to **exclude** the problematic item from the calibration runs.
- 7) **PARSCALE** was run repeatedly for each data set, updating the starting slope value each time, until no problematic items remained and the software program converged as specified.

10.3.4 Calibration Quality Control

To ensure IRT calibrations and conversion tables were produced accurately, Pearson replicated the IRT calibrations and the generation of the score conversion tables. While ETS used **PARSCALE**, Pearson did the same calibrations using **IRTPRO** (Cai, Thissen, & du Toit, 2011) calibration software. Daily meetings were held so that Pearson and ETS could provide status reports and discuss issues related to the IRT

work. Measured Progress performed independent quality control comparisons between the ETS and Pearson item parameter estimates to identify any differences.

Specifically, Measured Progress completed the following quality control analyses/comparisons:

1. Made sure all items were treated the same way (i.e., if ETS collapsed a category, made sure Pearson collapsed the category in the same way for the item);
2. Compared IRT item parameter estimates by ETS and Pearson (i.e., IRT a -, b -, and d -parameter estimates);
3. Made sure the same items were excluded from the CBT/PBT common item linking sets;
4. Compared transformed PBT parameter estimates generated by ETS and Pearson;
5. Compared all conversion tables (for all PBA and EOY combinations) produced by ETS and Pearson to make sure they were accurate.

Measured Progress prepared reports documenting their findings. Exact matches were found between all ETS and Pearson conversion tables before scores were reported.

10.4 Model Fit Evaluation Criteria

The usefulness of IRT models is dependent on the extent to which they effectively reflect the data. As discussed by Hambleton, Swaminathan, and Rogers (1991), “The advantages of item response models can be obtained only when the fit between the model and the test data of interest is satisfactory. A poorly fitting IRT model will not yield invariant item and ability parameters” (p. 53).

After convergence was achieved for each IRT data set, the IRT model fit was evaluated by doing the following:

1. Reviewing item chi-square values from **PARSCALE**
2. Calculating *adjusted fit* values and flagging them if >0.45
3. Reviewing graphical **PARPLOT** (Educational Testing Service, 2009) output for all items

Since chi-square values are sensitive to sample size, these statistics are not easily compared when n -counts vary across items. As a result, *adjusted fit* values were calculated by dividing the chi-square fit statistic by the sample size using the following formula:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} \quad (10-1)$$

One limitation of the PARSCALE output is that when a chi-square value was greater than 9,999.99, PARSCALE does not print the value in the phase 2 output. Instead it prints asterisk (*****). When that happened, no chi-square or *adjusted fit* was available. In those cases, only PARPLOT graphical item response curve plots were reviewed.

The following are two examples of PARPLOT output. The first example is an ELA/L grade 3 item with 3-categories (0, 1, 2). The lines represent the empirical item characteristic curve (ICC) based on the item parameter estimates, and the blue triangles represent the observed item response data. The fit for the item in Figure 10.1 is excellent. The observed data matches the empirical ICCs. The chi-square, and adjusted fit values (305, and 0.05 respectively) are small, which indicate good model fit.

The second example is a mathematics Algebra I item that is dichotomous (i.e., scored 0, 1). Again, the line represents the empirical ICC based on the item parameter estimates, and the blue triangles represent the observed item response data. The fit for the item in Figure 10.2 is extremely poor. The observed data (the blue triangles) do not touch or follow the model predictions of the empirical ICC. The chi-square, and adjusted fit values (5168 and 0.329) are large for a sample of over 42 thousand test takers. This item was very difficult and the few students who answered the item correctly are represented by the blue triangles that rise sharply towards the high end of the ability scale. There is some guessing involved in this item and a 3PL model would likely fit better than the 2PL model.

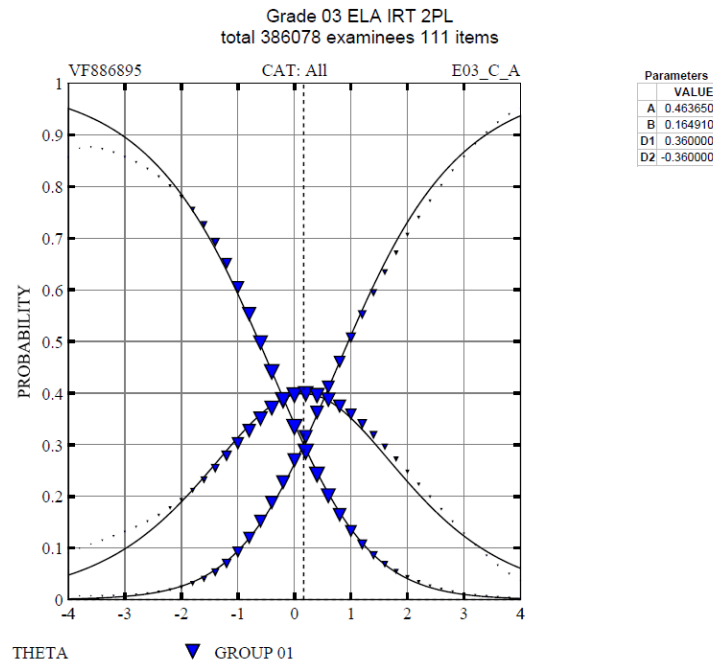


Figure 10.1 Grade 3 ELA/L 3-Category Item, 2 PL/GPC Model, N-count 124,090, chi-square 305.0, adjusted fit = 0.050.

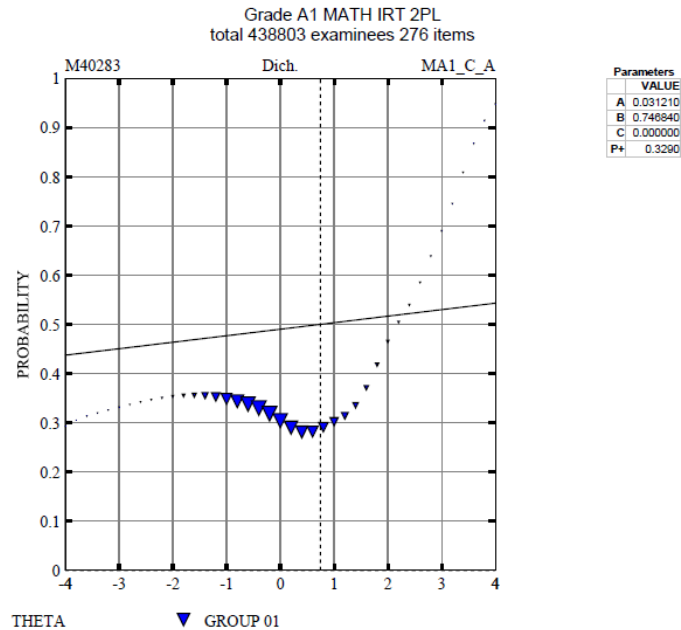


Figure 10.2 Algebra I Dichotomous (2-Category) Item, 2 PL Model, N-count 42,630, chi-square 5167.6, adjusted fit = 0.329.

10.5 Items Excluded from Score Reporting

As mentioned previously, after calibration and model fit evaluation was completed, a master list of all problematic items was compiled and potentially flawed items were brought to the PARCC Priority Alert Task Force. The Task Force reviewed each item, its content and the statistical properties, and made decisions about whether to include the item in the operational scores. Sometimes, an item was rejected because it appeared to have content issues, and sometimes an item was excluded because it could not be calibrated or showed extremely poor IRT model fit. Ultimately the decisions about whether to keep or exclude each flagged item was made by the Task Force.

10.5.1 Item Review Process

The following are the types of problematic items that were brought to the PARCC Priority Alert Task Force for evaluation and an “include or exclude” determination was made:

1. Extremely difficult items that did not converge (e.g., an item with a p value less than 0.02),
2. Items with low a -parameter estimates (e.g., slope less than 0.100) that were difficult to estimate,

3. Items with no test takers in a category or with category mean score reversals (i.e., an item that may have needed to be collapsed in order to successfully calibrate),¹⁵ and
4. Items that were calibrated but had extremely poor IRT fit.

Again, the primary goal was to minimize the number of items dropped from the operational test forms. An equally important goal was to not advantage or disadvantage any test takers. In some cases, when items were very difficult, and there was no content reason to remove an item that did not calibrate or had poor model fit, an alternative approach was used to obtain IRT parameter estimates. In all of these cases, the choice was between excluding the item from score reporting or finding an alternative way to estimate the item. The following are the five alternative approaches used in a few extreme cases to obtain operational item parameter estimates for an item that the Task Force felt should be kept:

1. Substituted Pearson generated **IRTPRO** estimates when **PARSCALE** would not converge for an item. This substitution was done for two Algebra I and six Algebra II items, as well as for one grade 3 ELA/L, two grade 8 ELA/L, one grade 3 mathematics, two grade 4 mathematics, two grade 7 mathematics, and three mathematics grade 8 items. All **IRTPRO** parameters were transformed onto the **PARSCALE** score reporting metric using the Stocking and Lord procedure.
2. Specified the 3PL model in one **PARSCALE** run for two items in grade 4 mathematics.
3. Used a weighted least squares (WLS) approach to fit 3PL models to eight Algebra I items, one Geometry items; as well as three grade 8 mathematics items. Figure 10.3 shows the improved fit for the Algebra I item displayed in Figure 10.2, after the WLS procedure was applied.
4. Modified/Collapsed items with zero students scoring in middle (items scored 0, 1, or 2) or highest categories (items scored 0, 1, 2, 3, or 4). For example, this was done for two items in grade 8 ELA/L.
5. Used CBT parameter estimates for PBT conversion tables when PBT parameter estimates could not be calibrated (usually due to sample size issues). This was only performed if no mode C-DIF was observed between the CBT and PBT versions of the item.

¹⁵ A category mean reversal is when the mean raw score of students obtaining a higher category score (e.g., 1 or 2) is less than the mean raw score of students obtaining a lower category score (e.g., 0 or 1) for a polytomous item. In other words, one would expect mean total test raw scores, for a well performing item, to increase as score categories increase from 0 to the maximum category.

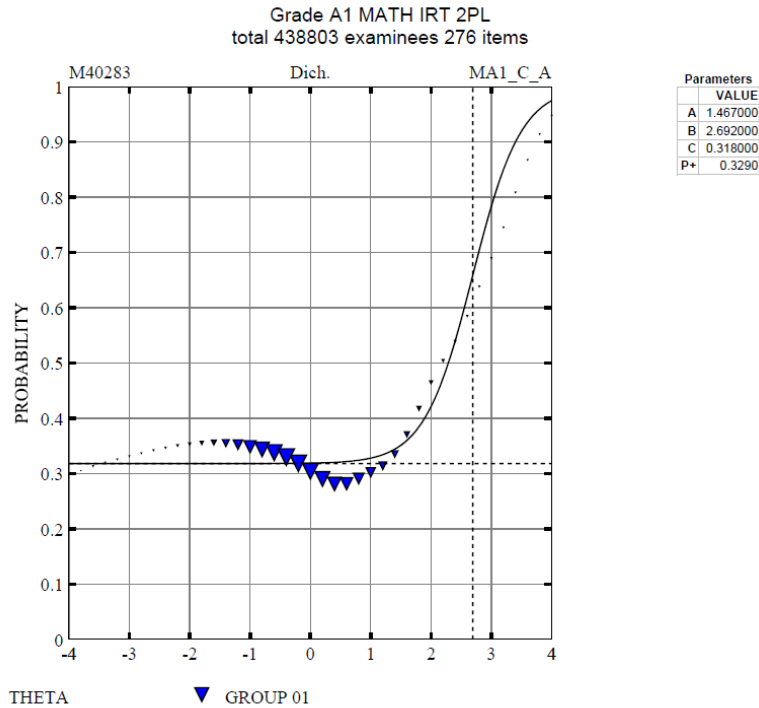


Figure 10.3 Algebra I Dichotomous Item, after the 3PL model was applied using the Weighted Least Squares approach, N-count 42,630.

Figure 10.3 demonstrates how the alternative approaches were used to obtain more appropriate item parameter estimates for a few extreme items that otherwise would have been excluded from the score reporting process. In this example, the 3PL WLS fit is far superior and more appropriate than the 2PL model fit because guessing appears to have been involved for a large number of test takers. This can be seen by the fact that the 3PL curve has a lower asymptote of about 0.31 which is around where the lower-scoring test takers’ probabilities of answering correctly is located, whereas the 2PL is forced to an asymptote at 0.00 which is far below the probability of lower-scoring test takers getting the item correct.

10.5.2 Count and Percentage of Items Excluded from Score Reporting

Tables 10.3 and 10.4 present the count and percentage of CBT and PBT items excluded from IRT calibration along with the reasons the items were excluded for ELA/L and mathematics, respectively. All items that did not have IRT item parameter estimates were excluded from the student operational scores and the conversion tables used for score reporting. For ELA/L and mathematics, at most 2% of the items were excluded from score reporting for all grades/subjects. As shown in Table 10.4, Integrated Mathematics I and III were exceptions and required a larger percentage of items to be excluded from score reporting. Integrated Mathematics I included some extremely difficult items that would not converge and Integrated Mathematics III had a number of items for which very few students scored at some of the score categories.

Table 10.3 Number and Percentage of ELA/L Items Excluded from IRT Calibration

Grade	Total <i>n</i> of CBT Items	<i>n</i> of CBT Items Excluded	Percent Excluded	Reason Excluded				Total <i>n</i> of PBT Items	<i>n</i> of PBT Items Excluded	Percent Excluded	Reason Excluded			
				Small Sample Size	Poor IA Stats	Did Not Calibrate	Other				Small Sample Size	Poor IA Stats.	Did Not Calibrate	Other
3	111	0	0.0					75	0	0.0				
4	125	1	0.8				1	87	1	1.1				1
5	138	0	0.0					87	0	0.0				
6	157	0	0.0					100	0	0.0				
7	150	3	2.0			1	2	100	2	2.0			1	1
8	150	2	1.3			2		100	1	1.0			1	
9	215	1	0.5	1				100	0	0.0				
10	194	1	0.5			1		105	0	0.0				
11	156	2	1.3		1		1	120	1	0.8		1		

Note: Grade 4 Other: Poor IRT fit for both modes. Dropped the same item in both modes. Grade 7 Other: One item identified as having a content problem. The other item had poor fit in both modes. Grade 11 CBT Other: Extremely hard, only 90 students (0.2%) out of 50,774 obtained a score of '2'.

Table 10.4 Number and Percentage of Mathematics Items Excluded from IRT Calibration

Grade/ Subject	Total n of CBT Items	n of CBT Items Excluded	Percent Excluded	Reason Excluded				Total n of PBT Items	n of PBT Items Excluded	Percent Excluded	Reason Excluded			
				Small Sample Size	Poor IA Stats	Did Not Calibrate	Other				Small Sample Size	Poor IA Stats.	Did Not Calibrate	Other
3	269	1	0.4		1			158	0	0.0				
4	199	0	0.0					135	0	0.0				
5	206	0	0.0					135	0	0.0				
6	192	0	0.0					132	0	0.0				
7	238	0	0.0					131	1	0.8		1		
8	196	1	0.5				1	117	1	0.9				1
A1	275	2	0.7			2		136	0	0.0				
GO	288	2	0.7			2		136	1	0.7		1		
A2	200	0	0.0					135	0	0.0				
M1	73	2	2.7			2		69	4	5.8		1		3
M2	78	1	1.3			1		75	1	1.3		1		
M3	81	7	8.6			7		70	4	5.7				4

Note: Grade 8 Other: Item had poor fit in CBT and needed to be collapsed for PBT. A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

10.6 Scaling Parameter Estimates from Paper to Online

10.6.1 Scaling Procedures (PBT to CBT)

Once the CBT and PBT item response data were calibrated for all grades/subjects for each content area, all available item parameter estimates of common items across modes (CBT and PBT), were used to transform the PBT item parameter estimates onto the CBT scales. The software program **STUIRT** (Kim & Kolen, 2004) was used to obtain Stocking and Lord (1983) transformation values to link the PBT scales.

Linking is an iterative process. **SAS** code was developed to calculate weighted root mean square difference (WRMSD) values to compare the item characteristic curves (ICCs) across modes and to identify items for possible removal from the linking sets. In addition, based on the 2014 field test analysis (which found some common items performed differently across modes), differential item functioning (DIF) analyses were completed on the 2015 operational common items to exclude linking items that appeared to show systematic differences across CBT and PBT administrations. Mantel-Haenszel D-DIF procedures were used for dichotomous items and standardized mean difference (SMD) were calculated for polytomous items. The following are the rules used to identify items for exclusion from the linking sets. The first four rules were automatic exclusions. Closer inspection was warranted for items flagged for large WRMSD values.

1. Exclude the item from linking set if the CBT and/or PBT weighted polyserial correlation, based on the item analysis, was less than 0.10.
2. Exclude polytomous items from linking set if categories were collapsed differently across modes.
3. Exclude items dropped by the PARCC Priority Alert Task Force (i.e., due to content or parameter estimation issues).
4. Exclude items from linking set if the item that were flagged for positive or negative C-level mode DIF.
5. Round 1: Run **STUIRT** and flag items for further inspection if the weighted root mean square difference (WRMSD) was greater than the values in Table 10.5. (*Note: these values were developed and used for the 2014 field test Analysis.*)

Table 10.5 WRMSD Flagging Criteria for Inspection and Possible Removal of Linking Items

Categories	Points	WRMSD/ points	WRMSD
2	1	0.100	0.100
3	2	0.075	0.150
4	3	0.075	0.225
5	4	0.075	0.300
6	5	0.075	0.375
7	6	0.075	0.450
>=8	>= 7	0.090	0.999

When inspecting items with large WRMSD values for exclusion from the linking sets, content representation was also considered to avoid removing large numbers of items from the same subclaim. Tables 10.6 and 10.7 (in Section 10.7) present the numbers of items excluded from the CBT/PBT linking sets for each grade/subject by content area. Few items were dropped from the linking sets based on WRMSD.

After calculating WRMSD and excluding items with large WRMSD values from the linking sets in Round 1, transformation and scaling software (STUIRT) were run a second time to calculate the scaling parameters and place the PBT item and ability parameter estimates onto the CBT scales. WRMSDs of the linking items were reviewed again. It was not necessary to run a third round of STUIRT.

10.6.2 Comparability across Spanish and English Versions

All items on one CBT and one PBT form of the mathematics test at each grade/subject was translated from English into Spanish. However, data from the Spanish forms were not included in the calibration with the English data. As is often done with braille forms, the item parameter estimates based on data from the English forms was used to generate conversion tables for the Spanish forms. To check that the Spanish and English items were performing similarly across language versions, when sample size was large enough, Mantel-Haenszel (MH) and the Standardized Mean Difference (SMD) DIF procedures were run and items showing C-DIF were dropped from the Spanish forms. The DIF analyses¹⁶ required at least 100 students for the smaller group (either reference or focal group) and 400 student for the combined group (reference and focal groups). If either of these sample size requirements were not met, then the DIF analyses were not performed. It is appropriate to drop an item from the Spanish form if: a) it appears the item was poorly translated, and b) if it provided either an advantage or a disadvantage to those students taking the Spanish forms. Spanish items flagged for C-DIF were reviewed by Pearson content specialists to decide if the translations were an issue. These items were also reviewed by the Priority Alert Task Force for determination of whether to exclude these items from score reporting. The number and content representation of the flagged items (to be dropped) was monitored closely to avoid dropping large number of items and points for a single form, and to avoid dropping too many items from

¹⁶ Refer to Section 7 for more information on DIF analysis.

a single subclaim. Items excluded from score reporting due to Spanish DIF were: mathematics grades 5, 6, 8, and Integrated Mathematics II one item each; Algebra II two items; Integrated Mathematics I three items; and Integrated Mathematics III six items.

10.6.3 Scaling Quality Control

Pearson not only conducted independent calibrations of item response data using **IRTPRO** scaling software, they also used **STUIRT** (Kim & Kolen, 2004) software to transform their **IRTPRO** PBT item parameter estimates onto the **IRTPRO** CBT scales for each grade/subject. Pearson's scaling constants were compared to those generated by ETS and found to be consistent. As described in Section 10.3.4, Measured Progress independently made certain that the same items were excluded from the CBT/PBT linking sets, and compared transformed PBT parameter estimates by ETS and Pearson. If items had large differences across modes, the items were discussed and any remaining issues resolved.

10.7 Items Excluded from Spring 2015 Paper to Online Linking Sets

Tables 10.6 and 10.7 present: the total number of common items, items excluded from the CBT/PBT linking sets, and items kept in the linking sets for each grade/subject by content area. For ELA/L the numbers of linking items ranged from 50 (in grade 3) to 78 (in grade 7). For mathematics, the numbers of linking items ranged from 29 (in Integrated Mathematics III) to 104 (in grade 3). Algebra II had the largest number of items removed from the linking sets due to mode C-DIF (i.e., 12 items).

Table 10.6 Number of ELA/L Items Excluded from the CBT/PBT Linking Sets

Grade	Total <i>n</i> of Common Items	Number Excluded	Final Number in Linking Set	Number of Excluded Items by Reason for Exclusion				
				Low Polyserial	Diff. No. of Cat.	Mode C-DIF	Other*	High WRMSD
3	52	2	50					2
4	57	3	54			1	1	1
5	62	0	62					
6	56	1	55					1
7	79	1	78				1	
8	81	4	77			1	2	1
9	80	3	77			2		1
10	64	1	63					1
11	77	0	77					

Note: *Grade 4 Other: item had poor fit in both modes. Grade 7 Other: item excluded due to content issue. Grade 8 Other: two items with IRTPRO parameter estimates.

Table 10.7 Number of Mathematics Items Excluded from the CBT/PBT Linking Sets

Grade /Subj.	Total <i>n</i> of Common Items	Number Excluded	Final Number in Linking Set	Number of Excluded Items by Reason for Exclusion				
				Low Polyserial	Diff. No. of Cat.	Mode C-DIF	Other	High WRMSD
3	111	7	104			4	1	2
4	96	7	89			3	3	1
5	75	2	73			1		1
6	82	4	78			4		
7	91	4	97	1		3		
8	72	4	68			1		3
A1	101	8	93	2		2	3	1
GO	115	12	103			8	1	3
A2	75	14	61			12	1	1
M1	35	4	31			3		1
M2	49	7	42			4		3
M3	35	6	29		1	5		

Note: Grade 3 other items had IRTPRO parameters. Grade 4 Other: two items had IRTPRO parameters and one used 3PL model. HS Items in "Other" column were excluded because items had WLS or IRTPRO parameter estimates. A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

10.8 Correlations and Plots of Scaling Item Parameter Estimates

Once the final group of items for each linking set was determined, and the PBT item parameter estimates were transformed onto the CBT scales, the a - and b -parameter estimates across modes were plotted and the correlation between the a -parameter estimates and the b -parameter estimates were calculated. Tables 10.8 and 10.9 present the number of: linking items, score points of the linking items, and the correlation of the a - and b -parameter estimates across modes.

Table 10.8 Number of Items, Number of Points and Correlations for ELA/L Linking Items

Grade	Number		Parameter Correlations	
	Items	Points	a -	b -
3	50	143	0.96	0.95
4	54	150	0.98	0.97
5	62	192	0.97	0.98
6	55	147	0.97	0.98
7	78	219	0.95	0.98
8	77	227	0.96	0.98
9	77	228	0.97	0.99
10	63	176	0.98	0.98
11	77	206	0.92	0.97

Table 10.9 Number of Items, Number of Points and Correlations for Mathematics Linking Items

Grade/ Subject	Number		Parameter Correlations	
	Items	Points	a -	b -
3	104	157	0.95	0.99
4	89	137	0.96	0.99
5	73	116	0.97	0.99
6	78	125	0.96	0.98
7	87	149	0.96	0.98
8	68	117	0.97	0.97
A1	93	183	0.93	0.95
GE	103	198	0.93	0.97
A2	61	116	0.95	0.97
M1	31	58	0.88	0.89
M2	42	75	0.87	0.89
M3	29	46	0.69	0.96

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

The following are plots of a - and b -parameter estimates for linking items and the TCCs for the common item sets for two ELA/L grades (i.e., 3 and 11) and two mathematics grades (i.e., 8 and Integrated Mathematics III). Figures 10.4 to 10.6 and 10.7 to 10.9 relate to ELA/L grades 3 and 11, respectively. Figures 10.10 to 10.12 and 10.13 to 10.15 relate to mathematics grades 8 and Integrated Mathematics III, respectively.

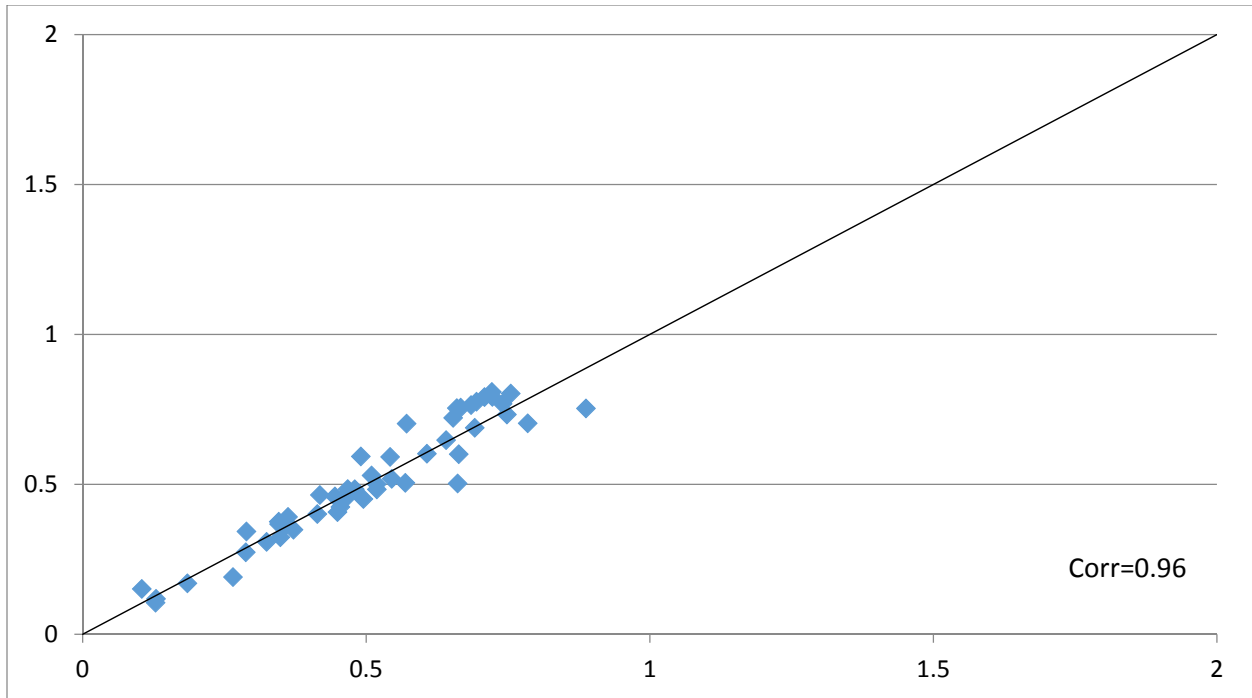


Figure 10.4 ELA/L Grade 3 Transformed New α - vs. Reference α -Parameter Estimates

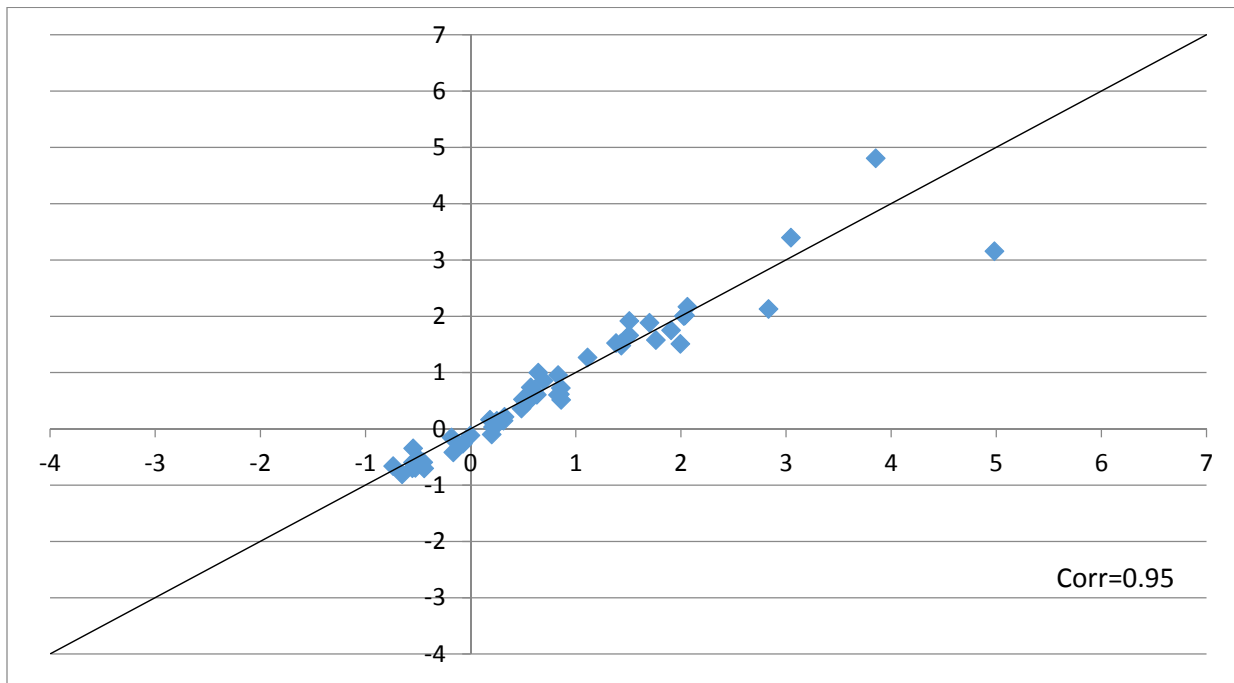


Figure 10.5 ELA/L Grade 3 Transformed New b - vs. Reference b -Parameter Estimates¹⁷

¹⁷ Note: One item (0382_A): Ref $b=3.1578$, Transformed New $b=4.9814$, measures Reading (RI) and had a WRMSD=0.038 which is far below the threshold for exclusion. This item was included in the linking set.

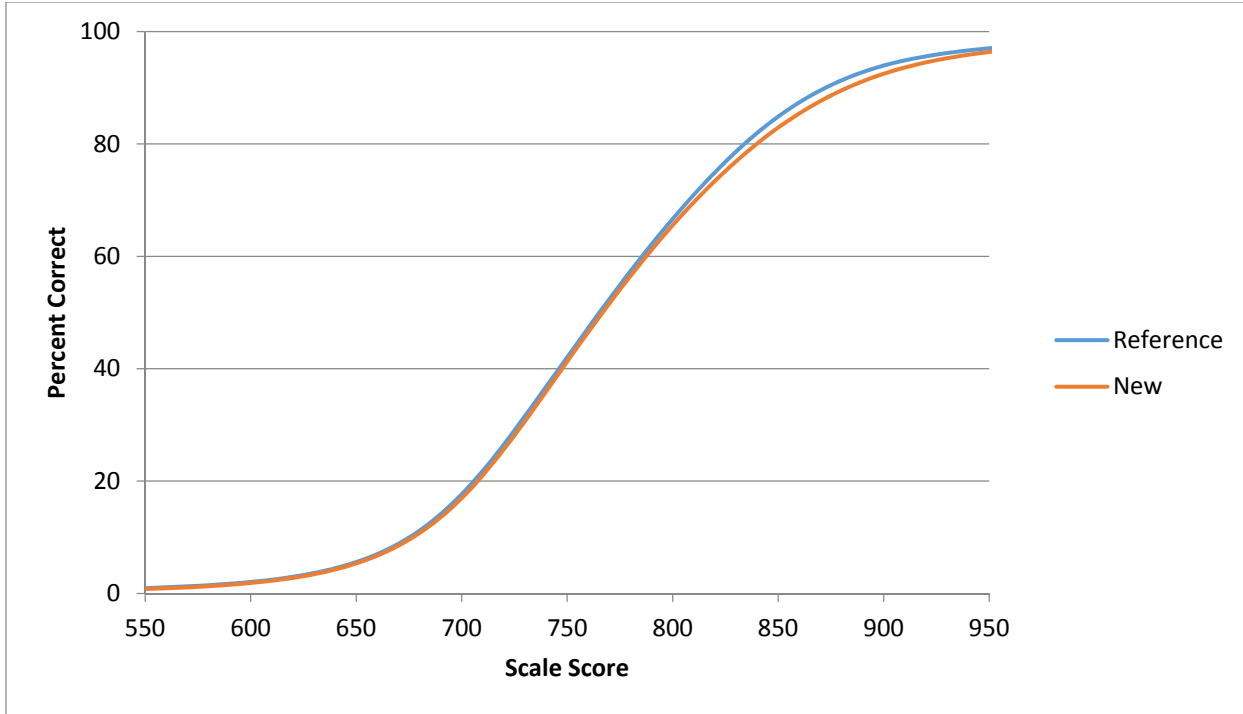


Figure 10.6 ELA/L Grade 3 Transformed New and Reference TCC for Linking Items

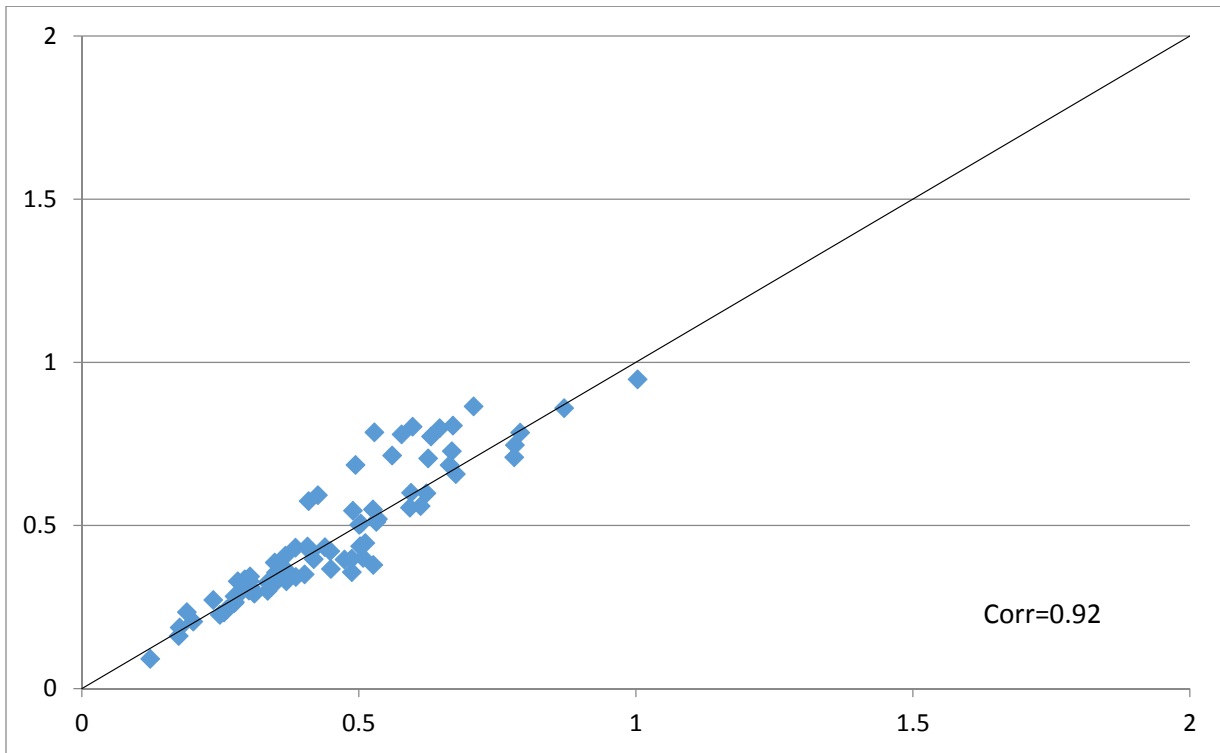


Figure 10.7 ELA/L Grade 11 Transformed New α - vs. Reference α -Parameter Estimates

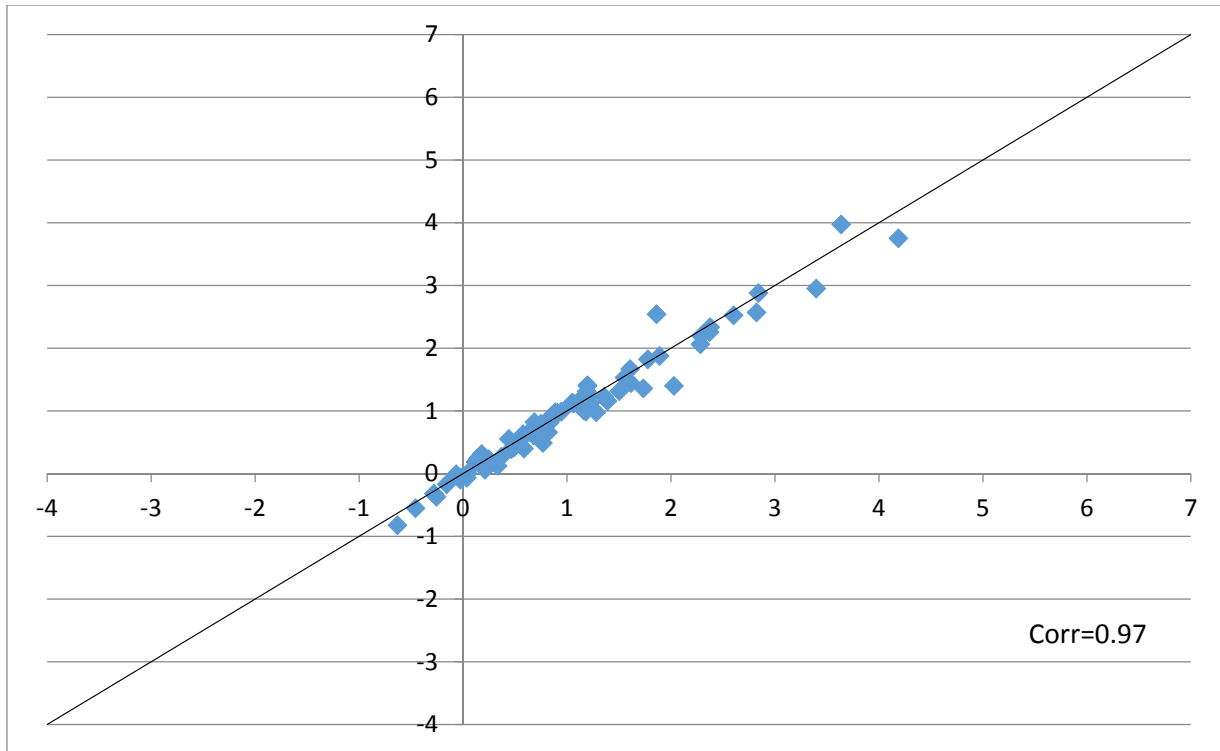


Figure 10.8 ELA/L Grade 11 Transformed New b - vs. Reference b -Parameter Estimates

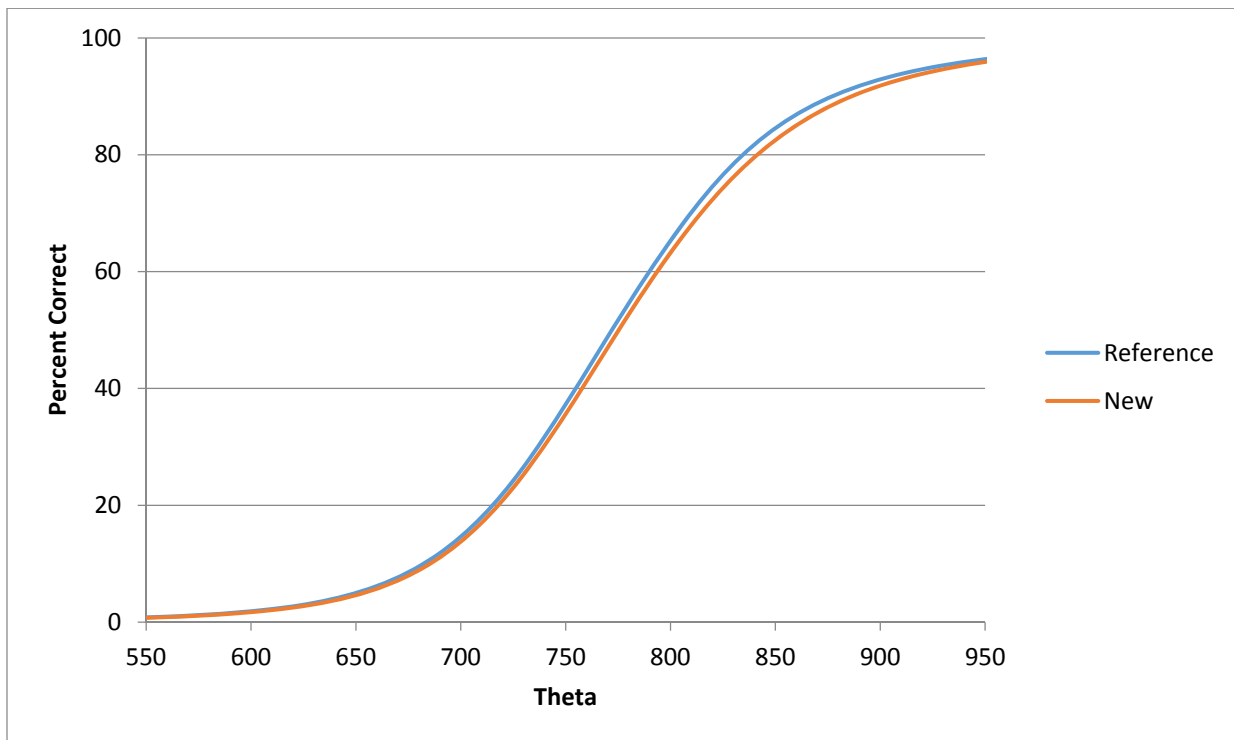


Figure 10.9 ELA/L Grade 11 Transformed New and Reference TCC for Linking Items

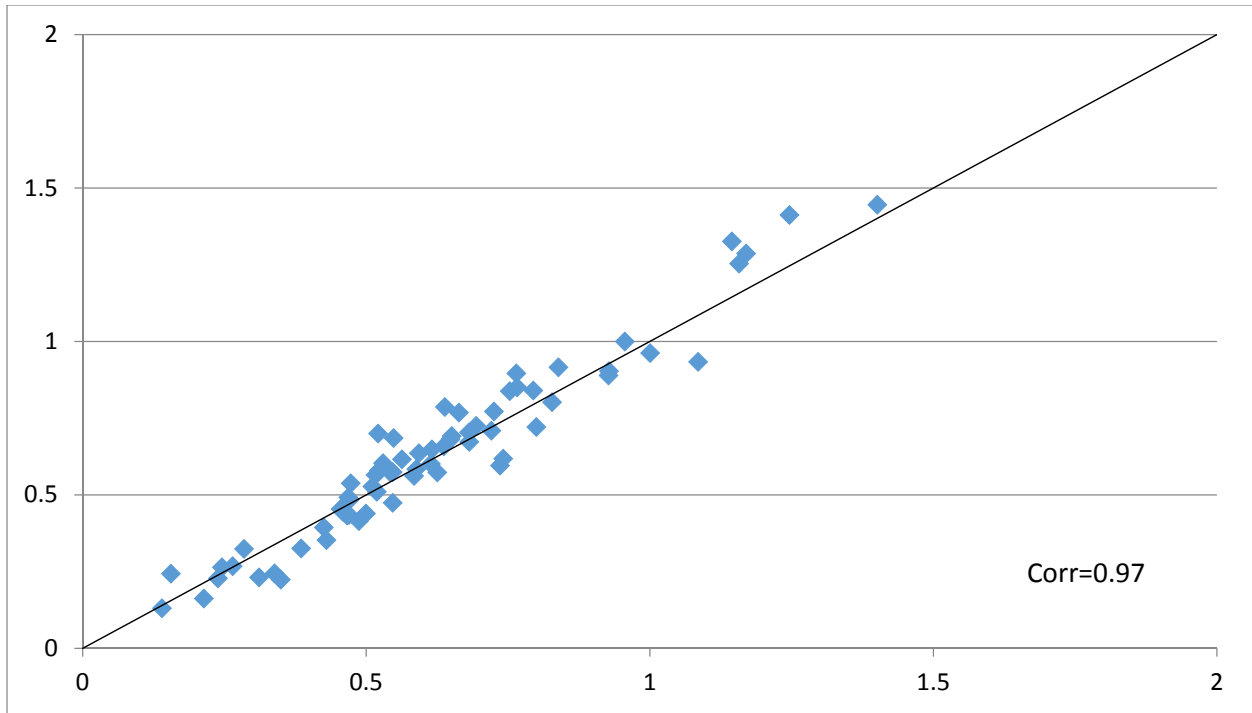


Figure 10.10 Mathematics Grade 8 Transformed New a - vs. Reference a -Parameter Estimates

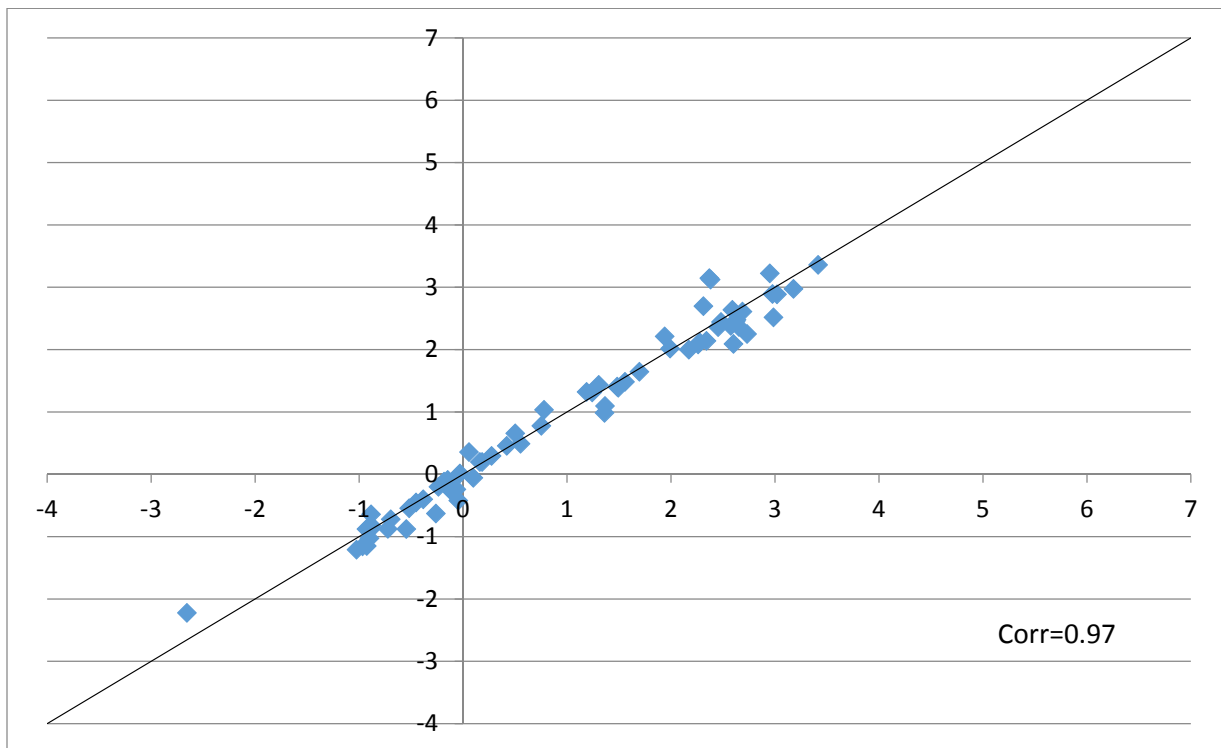


Figure 10.11 Mathematics Grade 8 Transformed New b - vs. Reference b -Parameter Estimates

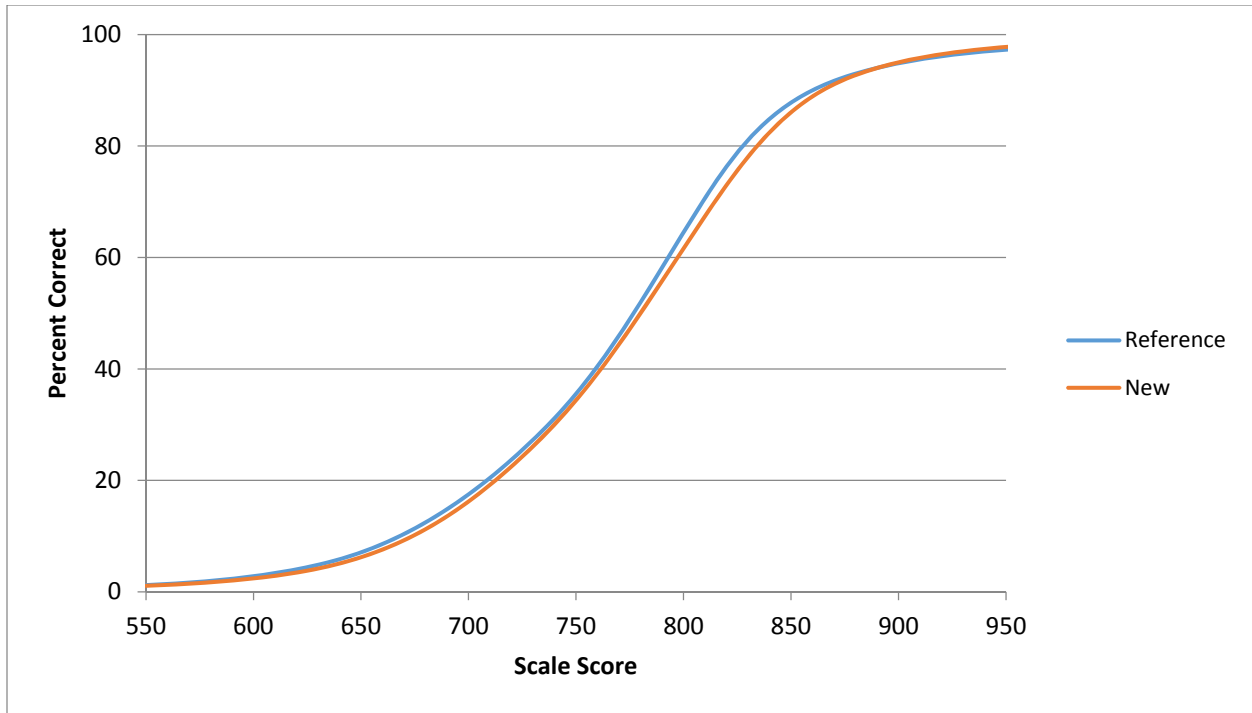


Figure 10.12 Mathematics Grade 8 Transformed New and Reference TCC for Linking Items

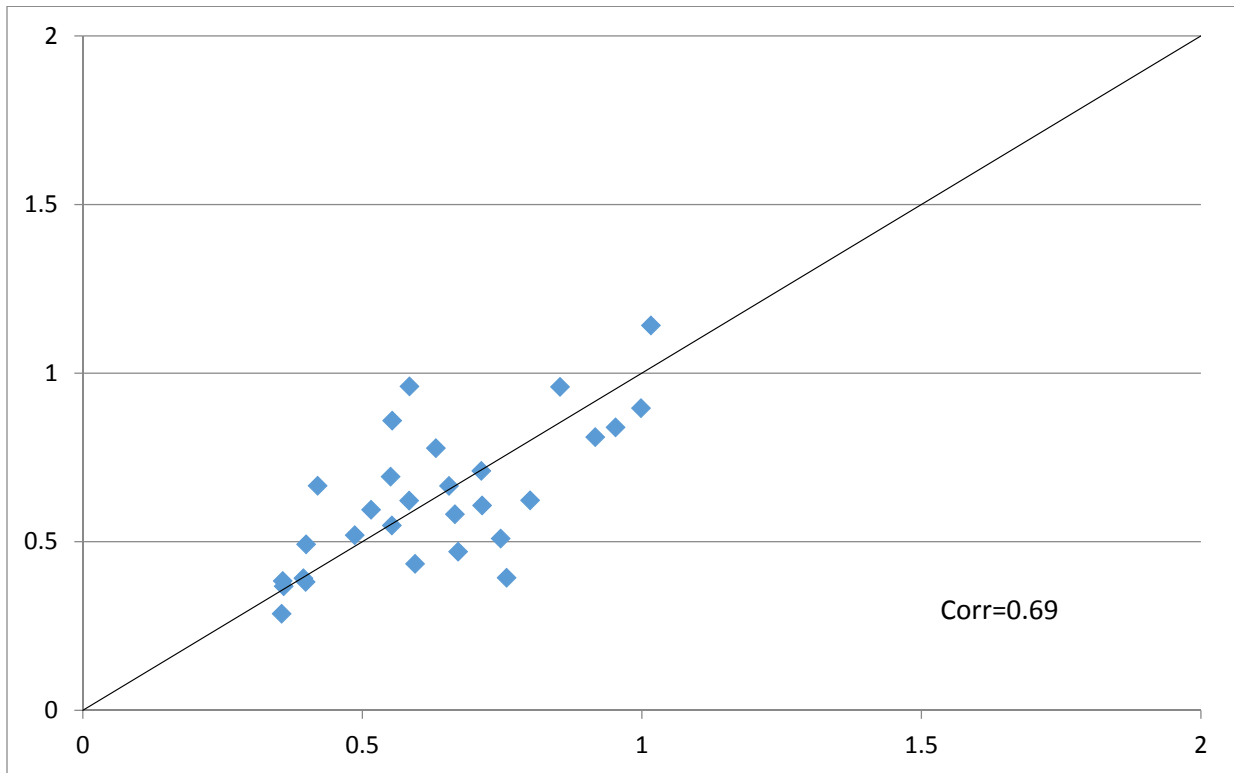


Figure 10.13 Integrated Mathematics III Transformed New α - vs. Reference α -Parameter Estimates

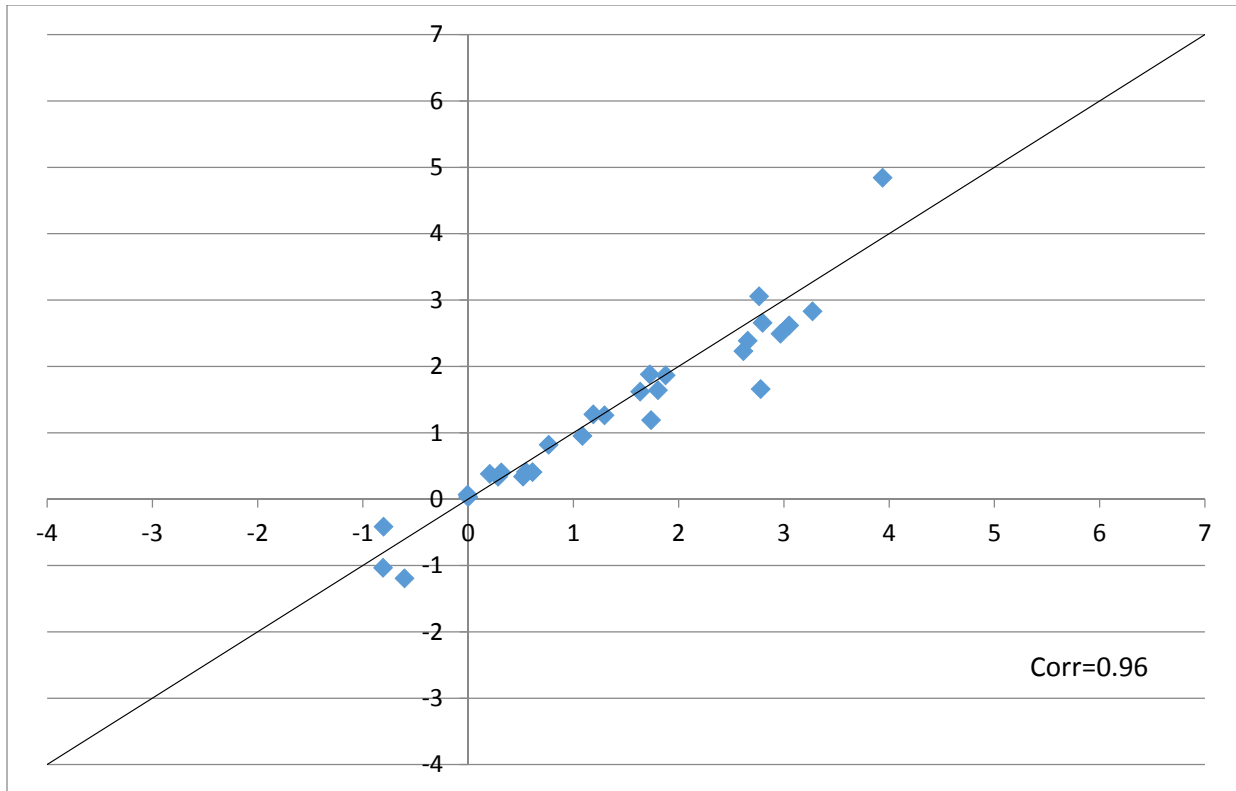


Figure 10.14 Integrated Mathematics III Transformed New b - vs. Reference b -Parameter Estimates

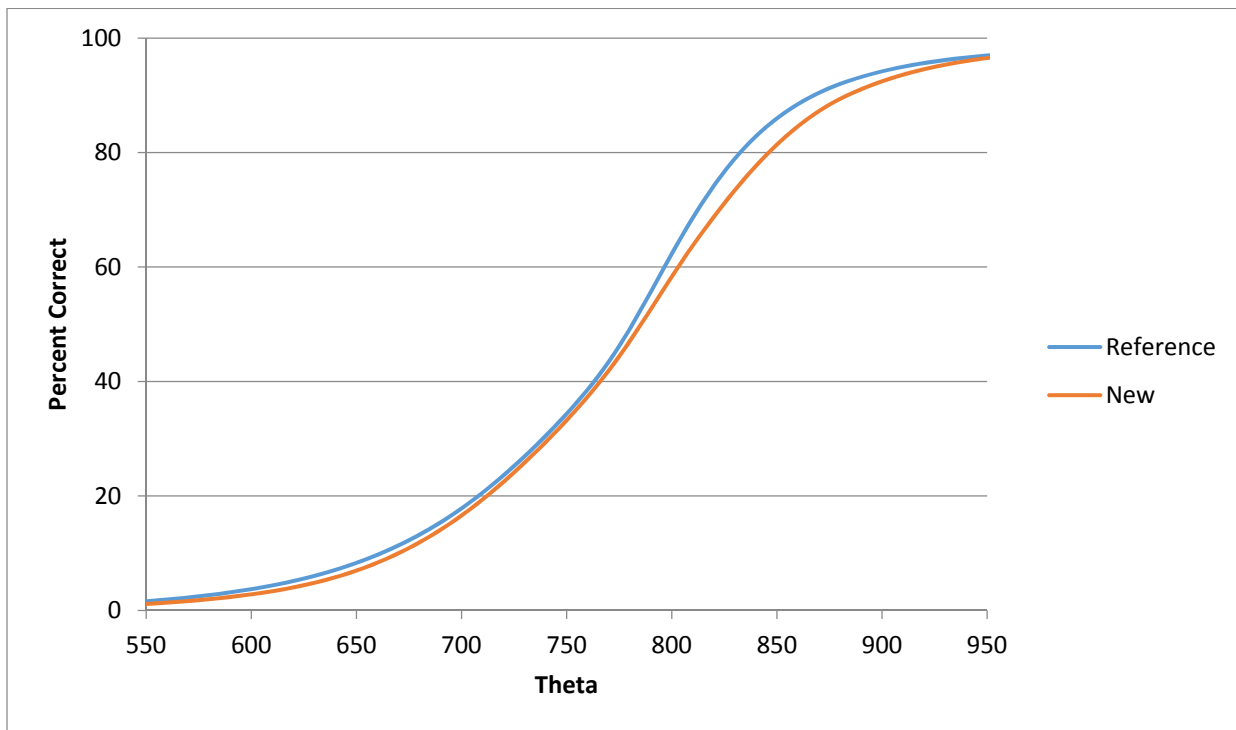


Figure 10.15 Integrated Mathematics III Transformed New and Reference TCC for Linking Items

10.9 Scaling Constants

Tables 10.10 and 10.11 present the slope and intercept scaling constants for ELA/L and mathematics, derived from **STUIRT** (Kim & Kolen, 2004) using the Stocking and Lord (1983) test characteristic curve procedure. Of interest are the intercept values which provide an indication of the difference in ability/performance between the students in the CBT and PBT groups on the common items across modes. If the two groups are similar in ability, the intercept will be close to 0.00 (as it is for mathematics grades 3 through 5).

Table 10.10 shows that, for ELA/L, the intercept is fairly consistent, ranging between 0.10 and 0.18 for grades 3 through 8, with a much larger difference at grade 9 (0.366) and nearly no difference (-0.022) at grade 11. This consistent difference in the intercept for ELA/L indicates that, on the common items across modes, students taking the items on paper obtained slightly higher mean scores than students taking the common items online at most grades.

Table 10.11 shows that, for mathematics, there were similar performance between CBT and PBT at the lower grades (e.g., grades 3 through 6) with the differences in the intercepts becoming larger, in favor of PBT test takers in high school. The largest intercepts are for Geometry and Integrated Mathematics II. These differences may be due to test takers taking tests on paper being more able than students taking the tests online. It is also possible that there are mode effects influencing the performance of CBT and PBT test takers on the common items. A mode comparability study, investigating the possibility of mode effects contributing to the performance differences across modes, was completed by ETS as part of the Year 1 PARCC operational research studies. The results of the mode comparability study are available in a separate report.

Table 10.10 Scaling Constants Spring 2015 PBT to CBT for ELA/L

Grade/Subject	Spring 2015 PBT to CBT	
	Slope	Intercept
3	1.044781	0.101355
4	0.971881	0.156461
5	0.970331	0.128022
6	0.985520	0.177576
7	0.967701	0.162102
8	0.981375	0.180024
9	0.982731	0.365813
10	1.008393	0.146266
11	1.002837	-0.022134

Table 10.11 Scaling Constants Spring 2015 PBT to CBT for Mathematics

Grade/ Subject	Spring 2015 PBT to CBT	
	Slope	Intercept
3	1.025761	0.003053
4	1.014104	-0.007029
5	1.024348	-0.013060
6	1.043489	0.066416
7	1.015457	0.104058
8	1.052324	0.232566
A1	1.017697	0.290914
GO	1.058918	0.593719
A2	0.974191	0.118806
M1	1.024626	0.333109
M2	1.162125	0.625184
M3	0.875079	0.266588

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

10.10 Summary Statistics and Distributions from IRT Analyses

Tables 10.12 through 10.23 present the b - and a -parameter estimates, the Standard Errors (SEs) of the b - and a -parameter estimates, and the IRT model fit values (chi-square and adjusted fit) for ELA/L and mathematics, for CBT and PBT for all items at each grade/subject. The mean, standard deviation (SD) and minimum and maximum values are presented for each statistic in addition to the number of items and points for all CBT and PBT items at each grade/subject. Chi-square, represented as G^2 , was calculated by **PARSCALE** and was then converted to the adjusted fit value (between 0.00 and 1.00) using the formula presented in Section 10.4. Again, when chi-square was greater than 9,999.99 the value was not provided as part of the **PARSCALE** phase 2 output and the adjusted fit value could not be calculated.

10.10.1 IRT Summary Statistics for English Language Arts/Literacy

Tables 10.12 and 10.13 show the b - and a -parameter estimates for the ELA/L assessments. Since there were no single select multiple choice (SSMC) items in ELA/L, all item responses were estimated using the 2PL/GPC model combination. For CBT items, the mean b -value ranged from 0.44 in grade 4 to 1.14 in grade 11. These parameter estimates are not on the same scale across grades, but are linked across CBT and PBT. For PBT items, the mean b -value ranged from 0.47 in grade 7 to 1.07 in grade 3. Notice that the standard deviation of the b -values in Tables 10.12 and 10.13 is largest at grade 3. The mean a -values across both CBT and PBT administrations were consistent, ranging from 0.44 to 0.50. Tables 10.14 and 10.15 present the standard errors of estimate for CBT and PBT ELA/L assessments, and Tables 10.16 and 10.17 provide model fit information. IRT summary statistics are provided in the **Appendix 10** for ELA/L for all items, reading-only, writing-only, PBA-only, and EOY-only. The addendum to this section presents similar information for the Fall block data.

Since chi-square values are sensitive to sample size, and the number of test takers per item varied across forms and modes, we focused on the adjusted fit values. In general, fit was good for the 2PL/GPC model combination. Only when items exceeded two standard deviations above the mean did fit begin to become problematic (greater than 0.23 to 0.25). Only a few items exceeded these levels. A few items with extremely poor IRT fit were removed by the Priority Alert Task Force from score reporting.

Table 10.12 CBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade

Grade	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	300	111	0.84	1.45	-1.45	8.21	0.50	0.20	0.10	1.03
4	326	124	0.44	0.96	-1.30	3.36	0.49	0.21	0.12	0.94
5	414	138	0.78	0.96	-1.35	4.80	0.54	0.25	0.07	1.04
6	425	157	0.56	1.07	-2.17	6.16	0.45	0.19	0.11	0.90
7	405	147	0.47	0.97	-1.63	3.03	0.46	0.20	0.12	0.91
8	407	148	0.64	1.31	-1.56	6.40	0.45	0.22	0.06	1.21
9	576	214	0.88	1.46	-1.29	9.87	0.47	0.22	0.06	0.98
10	543	193	0.66	0.90	-2.02	3.78	0.48	0.18	0.15	0.86
11	419	154	1.14	1.31	-1.02	7.76	0.48	0.20	0.09	0.95

Table 10.13 PBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade

Grade	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	202	75	1.07	1.74	-0.74	10.44	0.47	0.19	0.08	0.89
4	232	86	0.63	1.09	-1.30	4.51	0.46	0.21	0.12	0.86
5	251	87	0.83	1.05	-1.35	5.31	0.49	0.22	0.07	0.89
6	274	100	0.60	1.03	-1.50	5.58	0.45	0.17	0.14	0.80
7	270	98	0.47	0.98	-1.68	4.36	0.44	0.17	0.11	0.87
8	272	99	0.63	1.67	-1.26	12.48	0.45	0.20	0.04	1.01
9	274	100	0.86	1.47	-1.03	8.42	0.46	0.20	0.09	0.93
10	284	105	0.71	0.87	-1.03	3.57	0.49	0.18	0.15	0.95
11	325	119	1.03	1.03	-0.76	5.71	0.45	0.18	0.09	1.00

Table 10.14 CBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade

Grade	No. of Score Points	No. of Items	<i>SE of b Estimates</i>				<i>SE of a Estimates</i>			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	300	111	0.012	0.026	0.003	0.206	0.003	0.002	0.000	0.008
4	326	124	0.007	0.006	0.003	0.054	0.003	0.001	0.001	0.007
5	414	138	0.008	0.006	0.003	0.038	0.004	0.003	0.000	0.012
6	425	157	0.007	0.005	0.003	0.032	0.003	0.001	0.001	0.009
7	405	147	0.007	0.004	0.003	0.026	0.002	0.001	0.001	0.006
8	407	148	0.010	0.017	0.003	0.171	0.003	0.001	0.001	0.007
9	576	214	0.018	0.058	0.004	0.753	0.004	0.003	0.001	0.014
10	543	193	0.011	0.009	0.005	0.078	0.004	0.002	0.001	0.010
11	419	154	0.018	0.032	0.005	0.285	0.005	0.002	0.001	0.010

Table 10.15 PBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade

Grade	No. of Score Points	No. of Items	<i>SE of b Estimates</i>				<i>SE of a Estimates</i>			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	202	75	0.016	0.028	0.004	0.178	0.004	0.002	0.001	0.009
4	232	86	0.013	0.011	0.005	0.064	0.004	0.002	0.001	0.010
5	251	87	0.012	0.010	0.005	0.071	0.005	0.003	0.001	0.014
6	274	100	0.012	0.009	0.005	0.052	0.004	0.002	0.001	0.009
7	270	98	0.011	0.009	0.005	0.071	0.004	0.002	0.001	0.009
8	272	99	0.014	0.018	0.005	0.117	0.004	0.002	0.001	0.010
9	274	100	0.021	0.051	0.007	0.414	0.005	0.003	0.001	0.013
10	284	105	0.025	0.017	0.011	0.100	0.010	0.004	0.003	0.021
11	325	119	0.055	0.068	0.017	0.663	0.016	0.007	0.006	0.042

Table 10.16 CBT IRT Model Fit for All Items for ELA/L by Grade

Grade	No. of Score Points	No. of Items	G^2				Adjusted Fit			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	292	107	2394	1951	267	8589	0.13	0.05	0.05	0.25
4	312	117	2600	2395	234	9876	0.13	0.05	0.04	0.24
5	404	133	2352	2371	114	9245	0.13	0.04	0.03	0.23
6	409	149	2647	2321	200	9782	0.12	0.05	0.04	0.23
7	387	138	2891	2448	305	9939	0.12	0.05	0.04	0.23
8	393	141	2674	2152	242	9795	0.12	0.04	0.04	0.24
9	576	214	1823	1769	105	9426	0.14	0.05	0.05	0.31
10	543	193	1609	1565	223	9725	0.15	0.06	0.06	0.36
11	419	154	1176	1087	144	7407	0.14	0.05	0.05	0.33

Table 10.17 PBT IRT Model Fit for All Items for ELA/L by Grade

Grade	No. of Score Points	No. of Items	G^2				Adjusted Fit			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	202	75	1539	1411	169	5173	0.15	0.06	0.06	0.27
4	232	86	1494	1338	172	5967	0.14	0.05	0.05	0.26
5	251	87	1284	1188	196	6777	0.14	0.04	0.05	0.31
6	274	100	1270	1033	154	4859	0.14	0.05	0.05	0.27
7	270	98	1202	958	186	3661	0.14	0.05	0.06	0.24
8	270	98	1118	937	160	4105	0.14	0.05	0.05	0.27
9	274	100	759	627	123	2799	0.14	0.05	0.06	0.27
10	284	105	313	254	68	1544	0.17	0.06	0.08	0.37
11	325	119	186	273	62	2984	0.20	0.06	0.10	0.59

10.10.2 IRT Summary Statistics for Mathematics

Tables 10.18 and 10.19 show the b - and a -parameter estimates for mathematics. Since there were single select multiple choice (SSMC) items in mathematics and some did not fit well using the 2PL/GPC model combination, the 3PL model was used for a few items as described in Section 10.5.1. For CBT items, the mean b -value ranged from 0.10 in Grade 4 to 2.04 for Algebra I. These parameter estimates are not on the same scale across grades, but are linked across CBT and PBT. For PBT items, the mean b -value ranged from 0.12 in Grade 4 to 2.19 for Algebra II. The mean a -values across both CBT and PBT administrations were consistent, ranging from 0.52 to 0.80. Tables 10.20 and 10.21 present the standard errors of estimate for CBT and PBT mathematics assessments, and Tables 10.22 and 10.23 provide model fit information. IRT summary statistics are provided in the **Appendix 10** for mathematics for all items, single select multiple choice items, constructed response items, PBA-only, EOY-only, and subclaims. The addendum to this section presents similar information for the fall block mathematics data.

Since chi-square values are sensitive to sample size and the number of test takers per item varied across forms and modes, we focused on the adjusted fit values. In general fit was good for the 2PL/GPC model combination. Only when items exceeded two standard deviations above the mean did fit begin to become problematic. In general, mean adjusted fit values increased from the lower grades (e.g., 3 and 4) as grade increased (e.g., high school). Only a few items with extremely poor IRT fit were removed by the Priority Alert Task Force from score reporting.

Table 10.18 CBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject

Grade/ Subject	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	380	268	0.11	1.43	-3.99	4.18	0.80	0.24	0.13	1.41
4	307	199	0.10	1.27	-2.43	4.10	0.77	0.24	0.13	1.45
5	317	206	0.38	1.29	-2.88	4.72	0.69	0.25	0.21	1.54
6	305	192	0.62	1.28	-2.20	4.73	0.77	0.27	0.10	1.47
7	386	238	1.13	1.58	-3.26	9.04	0.78	0.32	0.08	1.72
8	303	195	1.33	1.44	-2.22	7.57	0.71	0.33	0.13	1.76
A1	482	273	2.04	1.66	-0.92	12.83	0.64	0.36	0.06	2.62
GO	506	286	1.35	1.16	-1.67	5.87	0.75	0.34	0.09	1.97
A2	393	200	1.98	1.67	-1.26	10.46	0.69	0.33	0.04	1.79
M1	134	71	1.68	1.24	-1.35	4.40	0.67	0.36	0.07	1.62
M2	141	77	1.86	1.39	-0.99	6.62	0.72	0.41	0.06	1.99
M3	144	74	1.67	1.33	-1.19	6.31	0.68	0.28	0.19	1.39

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table 10.19 PBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject

Grade/ Subject	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	235	158	0.16	1.44	-4.01	5.33	0.74	0.23	0.17	1.41
4	215	135	0.12	1.27	-2.39	3.34	0.73	0.27	0.15	2.14
5	207	135	0.25	1.22	-2.49	5.29	0.65	0.23	0.11	1.34
6	211	132	0.63	1.21	-3.43	4.34	0.68	0.26	0.09	1.46
7	208	130	1.09	1.39	-4.13	5.64	0.64	0.28	0.05	1.43
8	188	116	1.23	1.42	-2.66	5.16	0.63	0.28	0.14	1.49
A1	242	136	2.05	1.77	-0.83	12.09	0.52	0.24	0.02	1.22
GO	253	135	1.34	1.28	-1.41	5.87	0.67	0.29	0.13	1.97
A2	266	135	2.19	1.67	-1.05	11.65	0.55	0.27	0.00	1.17
M1	118	65	1.83	1.26	-1.05	5.36	0.56	0.29	0.00	1.42
M2	131	74	1.59	1.22	-1.39	3.95	0.61	0.24	0.13	1.34
M3	120	66	2.04	1.62	-0.81	6.72	0.56	0.26	0.13	1.31

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table 10.20 CBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject

Grade/ Subject	No. of Score Points	No. of Items	<i>SE of b</i> Estimates				<i>SE of a</i> Estimates			
			Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
3	380	268	0.011	0.010	0.002	0.081	0.007	0.003	0.001	0.018
4	307	199	0.009	0.010	0.003	0.128	0.006	0.002	0.002	0.013
5	317	206	0.010	0.012	0.003	0.136	0.006	0.002	0.002	0.015
6	305	192	0.010	0.019	0.002	0.190	0.006	0.003	0.001	0.021
7	386	238	0.018	0.043	0.002	0.383	0.007	0.004	0.001	0.028
8	303	195	0.018	0.029	0.000	0.274	0.007	0.004	0.000	0.033
A1	482	273	0.039	0.092	0.000	1.097	0.008	0.005	0.000	0.033
GO	506	286	0.024	0.033	0.000	0.306	0.012	0.007	0.000	0.063
A2	393	200	0.046	0.122	0.005	1.086	0.011	0.008	0.001	0.057
M1	134	71	0.050	0.080	0.006	0.527	0.017	0.013	0.003	0.062
M2	141	77	0.068	0.101	0.013	0.674	0.028	0.025	0.003	0.182
M3	144	74	0.070	0.101	0.013	0.821	0.030	0.020	0.007	0.099

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table 10.21 PBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject

Grade/ Subject	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
			Mean	SD	Min	Max	Mean	SD	Min	Max
3	235	158	0.013	0.013	0.004	0.091	0.008	0.003	0.002	0.019
4	215	135	0.012	0.009	0.004	0.071	0.008	0.006	0.002	0.061
5	207	135	0.015	0.023	0.003	0.256	0.007	0.003	0.002	0.017
6	211	132	0.019	0.027	0.004	0.173	0.008	0.004	0.003	0.026
7	208	130	0.030	0.071	0.004	0.769	0.010	0.005	0.003	0.033
8	188	116	0.022	0.022	0.000	0.156	0.009	0.005	0.000	0.030
A1	242	136	0.062	0.124	0.000	1.117	0.010	0.005	0.000	0.031
GO	253	135	0.040	0.053	0.000	0.512	0.020	0.011	0.000	0.055
A2	266	135	0.144	0.295	0.013	1.866	0.024	0.015	0.000	0.077
M1	118	65	0.154	0.257	0.016	1.775	0.033	0.020	0.000	0.116
M2	131	74	0.129	0.096	0.042	0.598	0.058	0.034	0.014	0.190
M3	120	66	0.217	0.211	0.026	0.940	0.062	0.037	0.010	0.208

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table 10.22 CBT IRT Model Fit for All Items for Mathematics by Grade/Subject

Grade/ Subject	No. of Score Points	No. of Items	G^2				Adjusted Fit			
			Mean	SD	Min	Max	Mean	SD	Min	Max
3	378	267	693	1125	41	9550	0.07	0.04	0.03	0.27
4	294	193	1242	1440	77	9073	0.09	0.04	0.03	0.26
5	312	204	1302	1611	68	9390	0.09	0.04	0.03	0.24
6	303	191	1230	1469	85	8343	0.08	0.04	0.03	0.25
7	382	236	1105	1356	35	7471	0.09	0.05	0.02	0.26
8	296	193	907	1199	49	7813	0.08	0.04	0.03	0.24
A1	482	273	805	1055	30	6887	0.09	0.04	0.02	0.23
GO	505	285	421	487	46	3434	0.10	0.05	0.03	0.25
A2	373	194	646	834	50	4541	0.10	0.05	0.03	0.31
M1	134	71	255	254	38	1243	0.11	0.05	0.05	0.31
M2	141	77	143	128	20	541	0.12	0.05	0.04	0.27
M3	144	74	130	111	24	535	0.14	0.05	0.07	0.32

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table 10.23 PBT IRT Model Fit for All Items for Mathematics by Grade/Subject

Grade/ Subject	No. of Score Points	No. of Items	G^2				Adjusted Fit			
			Mean	SD	Min	Max	Mean	SD	Min	Max
3	233	157	560	722	67	5886	0.08	0.04	0.03	0.28
4	213	134	764	1029	66	8137	0.10	0.05	0.03	0.32
5	207	135	721	862	39	6293	0.10	0.04	0.04	0.22
6	210	131	590	812	65	5419	0.10	0.05	0.04	0.29
7	208	130	553	581	39	2512	0.10	0.05	0.04	0.24
8	186	114	480	629	47	4449	0.10	0.04	0.03	0.27
A1	237	131	358	574	42	5843	0.10	0.05	0.04	0.34
GO	252	134	151	131	37	822	0.13	0.05	0.05	0.30
A2	265	134	159	296	17	2358	0.15	0.08	0.06	0.70
M1	118	65	90	82	20	483	0.18	0.06	0.08	0.35
M2	131	74	43	25	4	111	0.23	0.05	0.10	0.37
M3	119	65	53	36	14	228	0.22	0.05	0.14	0.38

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

10.11 Mean Linking Item AIS Effect Size Compared to Overall Scale Score Effect Size

Linking sets across modes (CBT and PBT) made use of all available common items administered both online and on paper. The final quality control step done to review the scaling results was to calculate a mean difference and effect size across modes for the linking item sets (i.e., average item score) and compare them to the mean difference and effect size across modes for the overall reported scale scores. If the linking set is representative of the total test, the effect size for the linking set should be similar in magnitude and direction to the overall reported scale scores. If the patterns are not similar for the linking sets and overall scale scores, there may be issues with the linking. Again, because students were not randomly assigned to take online and paper tests, there was no reason to expect that the ability of CBT and PBT students would be equivalent. Table 10.25 and Table 10.26 show the AIS mean difference and effect size and the overall scale score mean difference and effect size for CBT and PBT test takers by grade/subject for ELA/L and mathematics, respectively.

For ELA/L, students taking the test on paper performed better on the linking items and the overall test than did online test takers at all grades except at Grade 11. The effect sizes of the AIS were fairly consistent across grades 3 through 10, ranging from 0.10 at Grade 8 to 0.30 at Grade 9. The effect sizes for the overall scale scores were similar in magnitude and direction for grades 3 through 10, ranging from 0.12 at grades 3 and 5, to 0.39 at Grade 9. The results suggest test takers on paper may have been more able than students taking the test online, or that a mode effect may have existed.

For mathematics, the results were different than ELA/L. Students taking the test on paper performed about the same as students taking the test online at the lower grades (e.g., grades 3 through 7). The effect sizes were larger on the high school subjects, favoring students taking the tests on paper. The

effect sizes of the AIS were fairly consistent across grades 3 through 7, ranging from -0.02 at Grade 3 to 0.05 at Grade 9. The effect sizes for the overall scale scores were similar in magnitude and direction for grades 3 through 7, ranging from -0.03 at Grade 3 to 0.09 at Grade 7. The results for these grades suggest that test takers on paper and online were similar in ability on both the linking sets and the overall tests, or that mode effects were not observed. For the high school end-of-course tests, however, AIS effect sizes ranged from 0.17 for Algebra I to 0.59 for Integrated Mathematics II. For the overall scale scores, effect sizes ranged from 0.09 for Algebra II to 0.59 for Integrated Mathematics II. In all high school mathematics subjects test takers performed better on paper than online.

It should be noted that a mode comparability study was also conducted to determine if there was any type of mode effect contributing to the differences in performance. Since the CBT and PBT groups were not randomly assigned, the results do not provide evidence as to whether students taking the test on paper were simply more able than students taking the tests online or if there was some type of mode effect influencing the results. The consistency in magnitude and direction of the linking sets and the overall scale scores was a positive sign that the scaling was functioning as expected.

Table 10.24 Average Item Score Effect Size Compared to Overall Scale Score Effect Size for ELA/L

Grade	Items (Points)	Common Linking Items							Overall Scale Scores			
		CBT N-Count	PBT N-Count	Stat	CBT AIS	PBT AIS	Dif. AIS	ES AIS	CBT	PBT	Diff.	ES
3	50 (143)	379,595	138,123	Mean	1.00	1.08	0.08	0.16	735.0	739.8	4.9	0.12
				SD	0.48	0.60			39.3	40.7		
4	54 (150)	479,579	151,280	Mean	1.14	1.23	0.09	0.11	740.7	746.3	5.6	0.17
				SD	0.79	0.96			33.3	32.8		
5	62 (192)	496,486	140,342	Mean	1.11	1.20	0.09	0.13	740.3	744.1	3.9	0.12
				SD	0.64	0.77			31.9	31.5		
6	55 (147)	512,211	118,654	Mean	1.02	1.12	0.09	0.12	738.9	744.7	5.8	0.19
				SD	0.70	0.85			31.0	30.2		
7	78 (219)	513,778	108,757	Mean	1.12	1.23	0.11	0.17	739.4	746.3	6.9	0.19
				SD	0.63	0.80			36.3	34.8		
8	77 (227)	504,831	111,380	Mean	1.27	1.38	0.11	0.10	739.3	747.0	7.8	0.21
				SD	1.07	1.26			37.0	35.8		
9	77 (228)	343,279	67,032	Mean	1.08	1.3	0.22	0.30	737.3	751.4	14.17	0.39
				SD	0.7	0.98			36.3	34.9		
10	63 (176)	244,041	22,248	Mean	1.03	1.15	0.11	0.14	734.6	743.2	8.55	0.19
				SD	0.75	0.93			44.5	44.6		
11	77 (206)	160,270	9,707	Mean	0.9	0.88	-0.01	-0.02	739.2	737.4	-1.8	-0.05
				SD	0.65	0.67			38.9	39.5		

Table 10.25 Average Item Score Effect Size Compared to Overall Scale Score Effect Size for Mathematics

Grade	Items (Points)	Common Linking Items							Overall Scale Scores			
		CBT N-Count	PBT N-Count	Stat	CBT AIS	PBT AIS	Dif. AIS	Effect Size	CBT	PBT	Diff.	Effect Size
3	104 (157)	447,753	193,942	Mean	0.63	0.62	-0.01	-0.02	739.0	737.9	-1.10	-0.03
				SD	0.43	0.43			32.8	33.5		
4	89 (137)	472,440	156,865	Mean	0.62	0.62	0.00	0.00	735.9	736.0	0.10	0.00
				SD	0.30	0.30			30.3	30.8		
5	73 (116)	489,021	146,057	Mean	0.67	0.66	-0.01	-0.02	735.6	735.4	-0.10	0.00
				SD	0.47	0.46			29.9	30.9		
6	78 (125)	510,692	116,575	Mean	0.56	0.58	0.02	0.05	735.0	736.4	1.40	0.05
				SD	0.40	0.42			29.4	30.9		
7	87 (149)	501,624	102,602	Mean	0.50	0.52	0.02	0.05	733.7	736.2	2.50	0.09
				SD	0.41	0.41			27.0	27.5		
8	68 (117)	409,867	89,949	Mean	0.49	0.55	0.06	0.15	727.0	735.0	8.00	0.23
				SD	0.40	0.44			34.7	36.8		
A1	93 (183)	401,289	73,319	Mean	0.48	0.55	0.07	0.17	733.2	740.3	7.11	0.22
				SD	0.40	0.48			31.8	33.0		
GO	103 (198)	183,006	20,173	Mean	0.48	0.65	0.17	0.41	731.0	746.0	14.67	0.56
				SD	0.40	0.54			26.0	28.0		
A2	61 (116)	172,334	12,858	Mean	0.40	0.45	0.05	0.14	719.2	722.5	3.29	0.09
				SD	0.35	0.43			36.6	37.8		
M1	31 (58)	26,727	3,236	Mean	0.37	0.47	0.10	0.28	729.9	741.7	11.77	0.34
				SD	0.34	0.46			34.6	36.5		
M2	42 (75)	11,363	907	Mean	0.33	0.47	0.14	0.59	727.9	746.4	18.51	0.59
				SD	0.23	0.32			30.8	35.5		
M3	29 (46)	7,265	1,122	Mean	0.39	0.45	0.06	0.19	716.2	727.3	11.14	0.29
				SD	0.31	0.38			39.0	35.3		

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Section 11: Performance Level Setting

11.1 Performance Standards

Performance standards relate levels of performance on an assessment directly to what students are expected to learn. This is done by establishing threshold scores that distinguish between performance levels. Performance level setting (PLS) is the process of establishing these threshold scores that define the performance levels for an assessment.

11.2 Performance Levels and Policy Definitions

For the PARCC assessments, the performance levels are

- Level 5: Exceeded expectations
- Level 4: Met expectations
- Level 3: Approached expectations
- Level 2: Partially met expectations
- Level 1: Did not yet meet expectations

More detailed descriptions of each performance level, known as policy definitions, are:

Level 5: Exceeded expectations

Students performing at this level **exceed academic expectations** for the knowledge, skills, and practices contained in the standards assessed at their grade level or course.

Grades 3-10: Students performing at this level **exceed academic expectations** for the knowledge, skills, and practices contained in the standards for English language arts/literacy or mathematics assessed at their grade level. They are **academically well prepared** to engage successfully in further studies in this content area.

Algebra II, Integrated Mathematics III, and ELA/L 11: Students performing at this level **exceed academic expectations** for the knowledge, skills, and practices contained in the Mathematics and ELA/L standards assessed at grade 11. They are very likely to engage successfully in entry-level, credit-bearing courses in Mathematics and ELA/L, as well as technical courses requiring an equivalent command of the content area. Students performing at this level are exempt from having to take and pass placement tests in two- and four-year public institutions of higher education designed to determine whether they are academically prepared for such courses without need for remediation.

Level 4: Met expectations

Students performing at this level **meet academic expectations** for the knowledge, skills, and practices contained in the standards assessed at their grade level or course.

Grades 3-10: Students performing at this level **meet academic expectations** for the knowledge, skills, and practices contained in the standards for English language arts/literacy or Mathematics assessed at

their grade level. They are **academically prepared** to engage successfully in further studies in this content area.

Algebra II, Integrated Mathematics III, and ELA/L 11: Students performing at this level **meet academic expectations** for the knowledge, skills and practices contained in Mathematics and ELA/L at grade 11. They are very likely to engage successfully in entry-level, credit bearing courses in mathematics and ELA/L, as well as technical courses requiring an equivalent command of the content area. Students performing at this level are exempt from having to take and pass placement tests in two- and four-year public institutions of higher education designed to determine whether they are academically prepared for such courses without need for remediation.

Level 3: Approached expectations

Students performing at this level **approach academic expectations** for the knowledge, skills, and practices contained in the standards assessed at their grade level or course

Grades 3-10: Students performing at this level **approach academic expectations** for the knowledge, skills, and practices contained in the standards for English language arts/literacy or Mathematics assessed at their grade level. They are likely prepared to engage successfully in further studies in this content area.

Algebra II, Integrated Mathematics III, and ELA/L 11: Students performing at this level **approach academic expectations** for the knowledge, skills, and practices contained in the ELA/L and mathematics standards assessed at grade 11. They are likely to engage successfully in entry-level, credit-bearing courses in mathematics and ELA/L, as well as technical courses requiring an equivalent command of the content area. **Students performing at Level 3 are strongly encouraged to continue to take challenging high school coursework in English and mathematics through graduation.** Postsecondary institutions are encouraged to use additional information about students performing at Level 3, such as course completion, course grades and scores on other assessments to determine whether to place them directly into entry level courses.

Level 2: Partially met expectations

Students performing at this level **partially meet academic expectations** for the knowledge, skills, and practices contained in the standards assessed at their grade level or course.

Grades 3-10: Students performing at this level **partially meet academic expectations** for the knowledge, skills, and practices contained in the standards for English language arts/literacy or Mathematics assessed at their grade level. They will likely need academic support to engage successfully in further studies in this content area.

Algebra II, Integrated Mathematics III, and ELA/L 11: Students performing at this level **partially meet academic expectations** for the knowledge, skills, and practices contained in the ELA/L and mathematics standards assessed at grade 11. They will likely need academic support to engage successfully in entry-level, credit-bearing courses, and technical courses requiring an equivalent command of the content area. Students performing at this level are not exempt from having to take and pass placement tests

designed to determine whether they are academically prepared for such courses without the need for remediation in two- and four-year public institutions of higher education.

Level 1: Did not meet expectations

Students performing at this level **do not yet meet academic expectations** for the knowledge, skills, and practices contained in the standards assessed at their grade level or course.

Grades 3-10: Students performing at this level **do not yet meet academic expectations** for the knowledge, skills, and practices contained in the standards for English language arts/literacy or Mathematics assessed at their grade level. They will need academic support to engage successfully in further studies in this content area.

Algebra II, Integrated Mathematics III, and ELA/L 11: Students performing at this level **do not yet meet academic expectations** for the knowledge, skills, and practices contained in the ELA/L and mathematics standards assessed at grade 11. They will need academic support to engage successfully in entry-level, credit-bearing courses in College Algebra, Introductory College Statistics, and technical courses requiring an equivalent level of mathematics. Students performing at this level are not exempt from having to take and pass placement tests in two- and four-year public institutions of higher education designed to determine whether they are academically prepared for such courses without need for remediation.

11.3 Performance Level Setting Process for the PARCC Assessment System

One of the main objectives of the PARCC assessment system is to provide information to students, parents, educators, and administrators as to whether students are on track in their learning for success after high school, defined as college- and career-readiness. To set performance levels associated with this objective, PARCC used the Evidence-Based Standard Setting (EBSS) method (Beimers, Way, McClarty, & Miles, 2012) for the PARCC Performance Level Setting (PLS) process. The EBSS method is a systematic method for combining various considerations into the process for setting performance levels, including policy considerations, content standards, educator judgment about what student should know and be able to demonstrate, and research to support PARCC's policy goals related to college- and career- readiness. A defined multistep process was used to allow a diverse set of stakeholders to consider the interaction of these elements in recommending performance level threshold scores for each PARCC assessment.

The seven steps of the EBSS process that were followed in order to establish performance standards for the PARCC assessments are:

- Step 1: Define outcomes of interest and policy goals
- Step 2: Develop research, data collection, and analysis plans
- Step 3: Synthesize the research results
- Step 4: Conduct pre-policy meeting
- Step 5: Conduct performance level setting (PLS) meetings with panels
- Step 6: Conduct reasonableness review with post-policy panel
- Step 7: Continue to gather evidence in support of standards

A summary of key components within these steps is provided below. Additional detail about each step in the PARCC Performance Level Setting (PLS) process is provided in the “PARCC Performance Level Setting Technical Report”.

11.3.1 PARCC Research Studies

PARCC conducted two research studies in support of their policy goals—the Benchmarking study and the Postsecondary Educators’ Judgment (PEJ) study. The Benchmarking study included a review of the literature relative to college- and career- readiness as well as consideration of the percentage of students obtaining a level equivalent to college- and career- readiness on a set of external assessments (e.g., ACT, SAT, NAEP). The PEJ study involved a group of nearly 200 college faculty reviewing items on the Algebra II and ELA/L 11 PARCC assessments and making judgments about the level of performance needed on each item to be academically ready for an entry-level college-credit bearing course in mathematics or ELA/L. Additional detail¹⁸ about the Benchmarking study can be found in the “PARCC Performance Level Setting Technical Report” as well as in the “PARCC Benchmarking Study Report.” Additional detail about the PEJ study can be found in the in the “PARCC Performance Level Setting Technical Report” as well as in the “Postsecondary Educators’ Judgment Study Final Report.”

11.3.2 PARCC Pre-Policy Meeting

Prior to the PLS meetings, a pre-policy meeting was convened to determine reasonable ranges which would be shown to panelists during the high school PLS meetings. Pre-policy meeting participants included representatives from both K-12 and higher education who served in roles such as Commissioner/Superintendent, Deputy/Assistant Commissioner, State Board Member, Director of Assessment, Director of Academic Affairs, Senior Policy Associate, and so on. The reasonable ranges recommended by the pre-policy meeting defined the minimum and maximum percentage of students that would be expected to be classified as college- and career-ready. The pre-policy meeting participants reviewed the test purpose, how the performance standards will be used, and the results of the research studies to provide the recommendations for the reasonable ranges without viewing any student performance data.

11.3.3 Performance Level Setting Meetings

The task of the performance level setting committee was to recommend four threshold scores that would define the five performance levels for each PARCC assessment. PARCC solicited nominations from all states that had administered the PARCC assessments in 2014-2015 for panelists to serve on the performance level setting committees. Nominations were solicited both from state departments of public education (K-12) and higher education (primarily for participation on the high school panels). When selecting panelists, PARCC placed an emphasis on those educators who had content knowledge as

¹⁸ More information can be requested online from <http://www.parcconline.org/assessments/test-design/research>.

well as experience with a variety of student groups and attempted to balance the panels in terms of state representation.

PARCC used an Extended Modified Angoff (Yes/No) method to collect educator judgments on the PARCC items. This method asked panelists to review each item on a reference form of the PARCC assessment and to make the following judgment:

How many points would a borderline student at each performance level likely earn if they answered the question?

This extension to the Yes/No standard setting method (Plake, Ferdous, Impara, & Buckendahl, 2005) allowed for incorporation of the multipoint PARCC items by asking educators to evaluate (Yes or No) whether a borderline student would earn the maximum number of points on an item, a lesser number of points on an item, or no points on the item. In the case of a single point or multiple-choice item, this task simplifies to the standard Yes/No method.

After receiving training on the PLS procedure, panelists participated in three rounds of judgments for each assessment. Within each round, panelists were asked to consider the items in the test form, starting with the PBA component and then the EOY component. Each panelist made a judgment for the Level 2 performance level, followed by judgments for the Level 3 performance level, the Level 4 performance level, and the Level 5 performance level, in this order. The panelists entered their item judgments for each round by completing an online item judgment survey. Educator judgments were summed across items to create an estimated total score on the reference form for each performance level threshold. Feedback data relative to panelist agreement, student performance on the items, and student performance on the test as a whole were provided in between each of the three rounds of judgment. Panelists were shown the pre-policy reasonable ranges prior to making their Round 1 judgments and again as feedback data following each round of judgment.

A dry-run of the PARCC PLS meeting process was held for Grade 11 English language arts/literacy (ELA/L) and Algebra II in order to evaluate the implementation of the performance level setting method with the innovative characteristics of the PARCC assessments. These content areas were selected because they combined all of the various aspects of the PARCC assessments, including the various types of items, scoring rules, and performance level decisions. The dry-run PLS meetings provided the opportunity to implement and evaluate multiple aspects of the operational plan for the actual PLS meeting, including pre-work, meeting materials, data analysis and feedback, and staff and panelist functions. The results of the dry-run PLS meeting were used to implement improvements in the process for the operational PLS meetings. Additional information about the methods and results of the dry-run PLS meeting is available in the full report for the “PARCC Performance Level Setting Dry-Run Meeting Report.”

The PLS meetings for the PARCC assessments were conducted during three one-week sessions. The dates of the twelve PLS committee meetings that were conducted are shown in Table 11.1.

Table 11.1 PARCC PLS Committee Meetings and Dates

Dates	Committees by Subjects and Grades
July 27 - 31, 2015	Algebra I/Integrated Mathematics I
	Geometry/Integrated Mathematics II
	Algebra II/Integrated Mathematics III
	Grade 9 English Language Arts/Literacy
	Grade 10 English Language Arts/Literacy
	Grade 11 English Language Arts/Literacy
August 17 - 21, 2015	Grades 7 & 8 Mathematics
	Grades 7 & 8 English Language Arts/Literacy
August 24 - 28, 2015	Grades 3 & 4 Mathematics
	Grades 5 & 6 Mathematics
	Grades 3 & 4 English Language Arts/Literacy
	Grades 5 & 6 English Language Arts/Literacy

Additional information about the methods and results of the PLS meetings is available in the “Performance Level Setting Technical Report.”

11.3.4 PARCC Post-Policy Reasonableness Review

Performance standards for all PARCC assessments were recommended by PLS committees and reviewed by the PARCC Governing Board and (for the Algebra II, Integrated Mathematics III, and ELA/L 11 assessments) the Advisory Committee on College Readiness as part of a post-policy reasonableness review. This group reviewed both the median threshold score recommendations from each committee and the variability in the threshold scores as represented by the Standard Error of Judgment (SEJ) of the committee. Adjustments to the median threshold scores that were within 2 SEJ were considered to be consistent with the PLS panels’ recommendation.

In addition to voting to adopt the performance standards based on the committees’ recommendations, this group also voted to conduct a shift in the PARCC performance levels to better meet the intended inferences about student performance. Holding the college- and career- ready (or on track) expectations (i.e., the current level 4) constant, performance levels above this expectation were combined and performance levels below this expectation were expanded to create the final system of performance levels with three below and two above the college- and career- ready (or on track) expectation. The shift in performance levels was accomplished using a scale anchoring process which involved two primary steps. In the first step, the top two performance levels, above college- and career-ready (or on-track), were combined into a single performance level and an additional performance level below college- and career- ready (or on track) was created by empirically determining the mid-point between the existing two levels. In the second step, the performance level descriptors (PLDs) were updated using items which discriminated student performance well at this level to create a PLD aligned with the new empirically determined performance level. At this same time, PLDs for all performance levels were reviewed for consistency and continuity. Members of the original PLS committees were recruited to participate in this process. Additional information about this process can be found in the Performance Level Setting Technical Report.

Section 12: Scale Scores

PARCC assessments are designed to measure and report results in categories called master claims and subclaims. Master claims (or simply “claims”) are at a higher level than subclaims with content representing multiple subclaims contributing to each claim outcome.

Four scale scores were reported for PARCC assessments.¹⁹ A full summative (FS) claim score was reported for each mathematics assessment. A FS claim score, and separate claim scores for Reading and Writing were reported for each English Language Arts/Literacy (ELA/L) assessment. PARCC reports results according to five performance levels that delineate the knowledge, skills, and practices students are able to demonstrate:

- Level 5: Exceeded expectations
- Level 4: Met expectations
- Level 3: Approached expectations
- Level 2: Partially met expectations
- Level 1: Did not yet meet expectations

Subclaim outcomes describe student performance for content-specific subsets of the item scores contributing to a particular claim. For example, Written Expression and Knowledge of Conventions subclaim outcomes are reported along with Writing claim scores. Subclaim outcomes are reported as *Below Expectations*, *Nearly Meets Expectations* or *Meets or Exceeds Expectations*.

12.1 Operational Test Content (Claims and Subclaims)

A claim is a statement about student performance based on how students respond to test questions. PARCC tests are designed to elicit evidence from students that support valid and reliable claims about the extent to which they are college and career ready or on track toward that goal and are making expected academic gains based on the Common Core State Standards (CCSS).

The number of items associated with each claim and subclaim outcome vary depending on test subject and grade. The item types vary in terms of the number of points associated with them, so that both the number of items and the number of points are important in evaluating the quality of a claim or subclaim score.

12.1.1 English Language Arts/Literacy

Table 12.1²⁰ includes the number of items and the number of points by subclaim and claim for ELA/L Grade 3. Corresponding information is provided in **Appendix 12.1** for all ELA/L grades.

¹⁹ Addendum 12 presents a summary of results on scale scores for the Fall 2014 administration.

²⁰ Table A.12.1 in Appendix 12.1 is identical to Table 12.1.

Table 12.1 Form Composition for ELA/L Grade 3

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	13 - 16	27 - 33
	Reading Informational Text	9 - 11	19 - 23
	Vocabulary	5 - 9	10 - 18
	Claim Total	31	64
Writing			
	Written Expression	3	27
	Knowledge of Conventions	3	9
	Claim Total	6	36
SUMMATIVE TOTAL		37	100

Note: Prose constructed responses (PCRs) consist of at least two writing traits (Written Expression and Writing Knowledge and Conventions) and, in some cases, a reading trait as well. An *aggregated* PCR item score is determined by summing the multiple scores the student received on two or three traits depending on the item. Therefore, each PCR trait is identified as a separate item in this table for the two writing subclaims and, in some cases, either the Reading Literary Text or the Reading Informational Text subclaim.

Each ELA/L form contains items of varying types. The prose constructed response (PCR) traits contribute to different claims and the aggregate of the traits contributes to the summative scale score. The following details the number of possible points and the associated subclaims for the three PCR tasks:

- Literary Analysis Task;
- Research Simulation Task;
- Narrative Writing Task.

The Literary Analysis Task and the Research Simulation Task are scored for three traits/subclaims: Reading, Written Expression, and Knowledge of Conventions. The Narrative Writing Task is scored for two traits/subclaims: Written Expression and Knowledge of Conventions. All traits/subclaims are initially scored as either 0-3 or 0-4; the Written Expression subclaims are multiplied by 3 (or weighted) to increase their contribution to the total score, making possible subclaim scores 0, 3, 6, and 9, or 0, 3, 6, 9, and 12. The maximum possible points for ELA/L PCR items are provided in Table 12.2.

Table 12.2 Contribution of Prose Constructed Response Items to ELA/L

Grade	Score	Possible Points		
		Literary Analysis Task	Research Simulation Task	Narrative Writing Task
3	Reading	3	3	0
	Written Expression	9	9	9
	Knowledge of Conventions	3	3	3
	Total	15	15	12
4-5	Reading	3	3	0
	Written Expression	9	9	9
	Knowledge of Conventions	3	3	3
	Total	15	15	12
6-11	Reading	4	4	0
	Written Expression	12	12	12
	Knowledge of Conventions	3	3	3
	Total	19	19	15

12.1.2 Mathematics

Table 12.3²¹ includes the numbers of items and points associated with subclaim scores for mathematics grade 3, as an example of the composition of the mathematics tests.

Table 12.3 Mathematics Form Composition for Grade 3

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	36	43
	Additional & Supporting Content	12 - 13	12 - 13
	Expressing Mathematical Reasoning	3 - 4	11 - 14
	Modeling and Applications	3	12
TOTAL		54 - 56	78 - 82

Because there is substantial variation in the composition of the tests, corresponding information is provided in **Appendix 12.1** tables for all mathematics grades/courses.

12.2 Establishing the Reporting Scales

PARCC reporting scales designate student performance into one of five Performance Levels²² with Level 1 indicating the lowest level of performance and Level 5 indicating the highest level of performance.

²¹ Table A.12.10 in Appendix 12.1 is identical to Table 12.3.

²² Section 11 provides an overview of the Performance Level Setting process, and detailed information can be found in the Performance Level Setting Technical Report.

Threshold or cut scores associated with performance levels were initially expressed as raw scores on the performance level setting (PLS) forms approved by the PARCC Governing Board. A number of adjustments were made by the PARCC Governing Board following the performance levels recommended by committees in the performance level setting.

- Level 2 threshold scores from the performance level setting were retained as the operational Level 2 threshold scores.
- Level 3 threshold scores from the performance level setting were redefined to be the operational Level 4 threshold scores.
- Level 4 threshold scores from the performance level setting were redefined to be the operational Level 5 threshold scores.
- Operational Level 3 threshold scores were mathematically defined by choosing a scale score value halfway between the Level 2 and Level 4 threshold scores.

A scale score task force was assembled by PARCC, which made recommendations about how threshold levels would be represented on the reporting scale.

12.2.1 Full Summative Score Scale and Performance Levels

There are 201 defined full summative scale score points for both ELA/L and mathematics, ranging from 650 to 850. A scale score of 700 is associated with minimum Level 2 performance, a scale score of 750 is associated with minimum Level 4 performance. The threshold for summative performance levels on the scale score metric recommended by the scale score task force are described in Table 12.4.

Table 12.4 Defined Summative Scale Scores

	Lowest Obtainable Scale Score	Level 2	Level 4	Highest Obtainable Scale Score
Full Summative	650	700	750	850

Scale scores were defined for each test as a linear transformation of the IRT theta (θ) scale. The theta values associated with the Level 2 and Level 4 performance levels were identified using the test characteristic curve associated with the performance level setting form.²³ With Level 2 and 4 scale scores defined to be 700 and 750, respectively, the relationship between theta and scale scores was established as

$$Scale_{FS} = A * \theta + B \quad (12-1)$$

where

²³ The Level 3 cut was defined to be the performance associated with a scale score of 725 rather than being determined in the performance level setting process. Caution is advised when making inferences from Level 3 cuts because there is not always a raw score associated with 725 on the performance level setting form. Therefore, Level 3 cuts should be regarded as approximations.

$$A = \frac{750 - 700}{\theta_{Level\ 4} - \theta_{Level\ 2}} \tag{12-2}$$

and

$$B = 750 - A * \theta_{Level\ 4} \tag{12-3}$$

A and B values resulting from these calculations as well as the performance level setting form raw scores and the theta values associated with the threshold performance levels are included in **Appendix 12.2**; **Appendix 12.3** includes raw to scale score conversion tables for the performance level setting forms.

12.2.2 ELA/L Reading and Writing Claim Scale

There are 81 defined scale score points for Reading, ranging from 10 to 90. A scale score of 30 is associated with minimum Level 2 performance, a scale score of 50 is associated with minimum Level 4 performance. There are 51 defined scale score points for Writing, ranging from 10 to 60. A scale score of 25 is associated with minimum Level 2 performance, a scale score of 35 is associated with minimum Level 4 performance. The threshold Reading and Writing performance levels on the scale score metric recommended by the scale score task force are described in Table 12.5.

Table 12.5 Defined Scaled Scores for Reading and Writing Claim Scores

	Lowest Obtainable Scale Score	Level 2	Level 4	Highest Obtainable Scale Score
Reading	10	30	50	90
Writing	10	25	35	60

As with the full summative scores, scale scores for Reading and Writing were defined for each test as a linear transformation of the IRT theta (θ) scale. The same IRT theta scale was used for Reading and Writing as was used for the ELA/L full summative scores. The theta values associated with the Level 2 and Level 4 performance levels were identified using the test characteristic curve associated with the performance level setting form. As with the full summative scores, the relationship between theta and scale scores was established with Level 2 and 4 theta scores and the corresponding predefined scale scores. The formulas used for this are provided in Table 12.6.

Table 12.6 Calculating Scaling Constants for Reading and Writing Claim Scores

Reading	Writing
$Scale = A_R * \theta + B_R$	$Scale = A_W * \theta + B_W$
$A_R = \frac{50 - 30}{\theta_{Level\ 4} - \theta_{Level\ 2}}$	$A_W = \frac{35 - 25}{\theta_{Level\ 4} - \theta_{Level\ 2}}$
$B_R = 50 - A * \theta_{Level\ 4}$	$B_W = 35 - A * \theta_{Level\ 4}$

A and B values resulting from these calculations are included in **Appendix 12.2**.

12.2.3 Subclaims Scale

The Level 4 cut is defined as *Meets or Exceeds Expectations* because high school students at Level 4 or above are likely to have the skills and knowledge to meet the definition of career and college readiness. Subclaim outcomes center on that performance level and are reported at three levels:

- Below Expectations;
- Nearly Meets Expectations; or
- Meets or Exceeds Expectations.

To move these cut scores from the full summative scale where they were originally defined to the subclaim raw score scale, average raw subclaim scores for students at the Level 3 and 4 thresholds were calculated. Students earning a raw subclaim score equal to or greater than the average raw subclaim score earned by students at the Level 4 threshold were designated as *Meets or Exceeds Expectations*. Students not earning a raw subclaim score equal to or greater than the average raw subclaim score earned by students at the Level 3 threshold were designated as *Below Expectations*. Other students, whose raw subclaim score fell between the average raw subclaim score earned by students at the Level 3 and 4 thresholds, were designated as *Nearly Meets Expectations*.

12.3 Creating Conversion Tables

A PARCC conversion table relates the number of points earned by a student on the ELA/L full summative score, the mathematics full summative score, the Reading claim score, or the Writing claim score to the corresponding scale score for the test form administered to that student. An IRT inverse test characteristic curve (TCC) approach is used to develop the relationship between point scores and IRT ability estimates (θ s). In carrying out the calculations, estimates of item parameters and thetas are substituted for parameters in the formulas in each step.

Step 1: Calculate the expected item score (i.e., estimated item true score) for every scale score in the selected range (determined by LOSS, HOSS, and scale score increment) based on the generalized partial credit model for both dichotomous and polytomous items:

$$s_i(\theta_j) = \sum_{m=0}^{M_i-1} m p_{im}(\theta_j), \quad (12-4)$$

$$p_{im}(\theta_j) = \frac{\exp\left[\sum_{k=0}^m D a_i(\theta_j - b_i + d_{ik})\right]}{\sum_{v=0}^{M_i-1} \exp\left[\sum_{k=0}^v D a_i(\theta_j - b_i + d_{iv})\right]}, \quad (12-5)$$

where $a_i(\theta_j - b_i + d_{i0}) \equiv 0$; $s_i(\theta_j)$ is the expected item score for item i on a scale score, θ_j ; $p_{im}(\theta_j)$ is the probability of a test taker with θ_j getting score m on item i ; M_i is the number of score categories of item i with possible item scores as consecutive integers from 0 to $M_i - 1$; D is the IRT scale constant

(1.7); a_i is a slope parameter; b_i is a location parameter reflecting overall item difficulty; d_{ik} is a location parameter incrementing the overall item difficulty to reflect the difficulty of earning score category k ; v is the number of score categories.

Step 2: Calculate the expected (weighted) test score for every scale score in the selected range:

$$T_j = \sum_{i=1}^I w_i s_i(\theta_j), \quad (12-6)$$

where T_j is the expected (weighted) test score on a scale score, θ_j ; w_i is the item weight for item i (e.g., with $w_i = 2$, a dichotomous item is scored as 0 or 2, and a three-category item is scored as 0, 2, or 4); I is the total number of items in a test form.

Step 3: Calculate the estimated conditional standard error of measurement (CSEM) for each scale score in the selected range:

$$\text{CSEM}_j = \sqrt{\frac{1}{\sum_{i=1}^I L_i(\theta_j)}}, \quad (12-7)$$

$$L_i(\theta_j) = (Da_i)^2 [s_{i2}(\theta_j) - s_i^2(\theta_j)], \quad (12-8)$$

$$s_{i2}(\theta_j) = \sum_{m=0}^{M_i-1} m^2 p_{im}(\theta_j), \quad (12-9)$$

where $L_i(\theta_j)$ is the estimated item information function for item i on scale score θ_j .

Step 4: Match every raw score with a scale score. θ_j is the scale score for a raw score r_h if $T_j - r_h$ is minimum across all T_j s.

Figure 12.1 contains TCCs, Figure 12.2 contains estimated information (INF) curves, and Figure 12.3 contains estimated conditional standard error of measurement (CSEM) curves for ELA/L grade 3.²⁴ The curves in each figure are for the average difficulty form, the easiest form, and the most difficult form for both CBT and PBT delivery modes. The form used in the standard setting (or performance level setting) process is also included. Maximum CSEM values were set to 15 for ELA/L and 20 for mathematics, as recommended by the scale score task force previously mentioned. Graphs were extended slightly

²⁴ Grade 3 TCC, INF, and CSEM curves are also included in Appendix 12.4 as Figures A.12.1, A.12.22, and A.12.43.

beyond the 650 LOSS and 850 HOSS in an effort to fully capture the full range of student performance, particularly in percent correct for the TCCs.

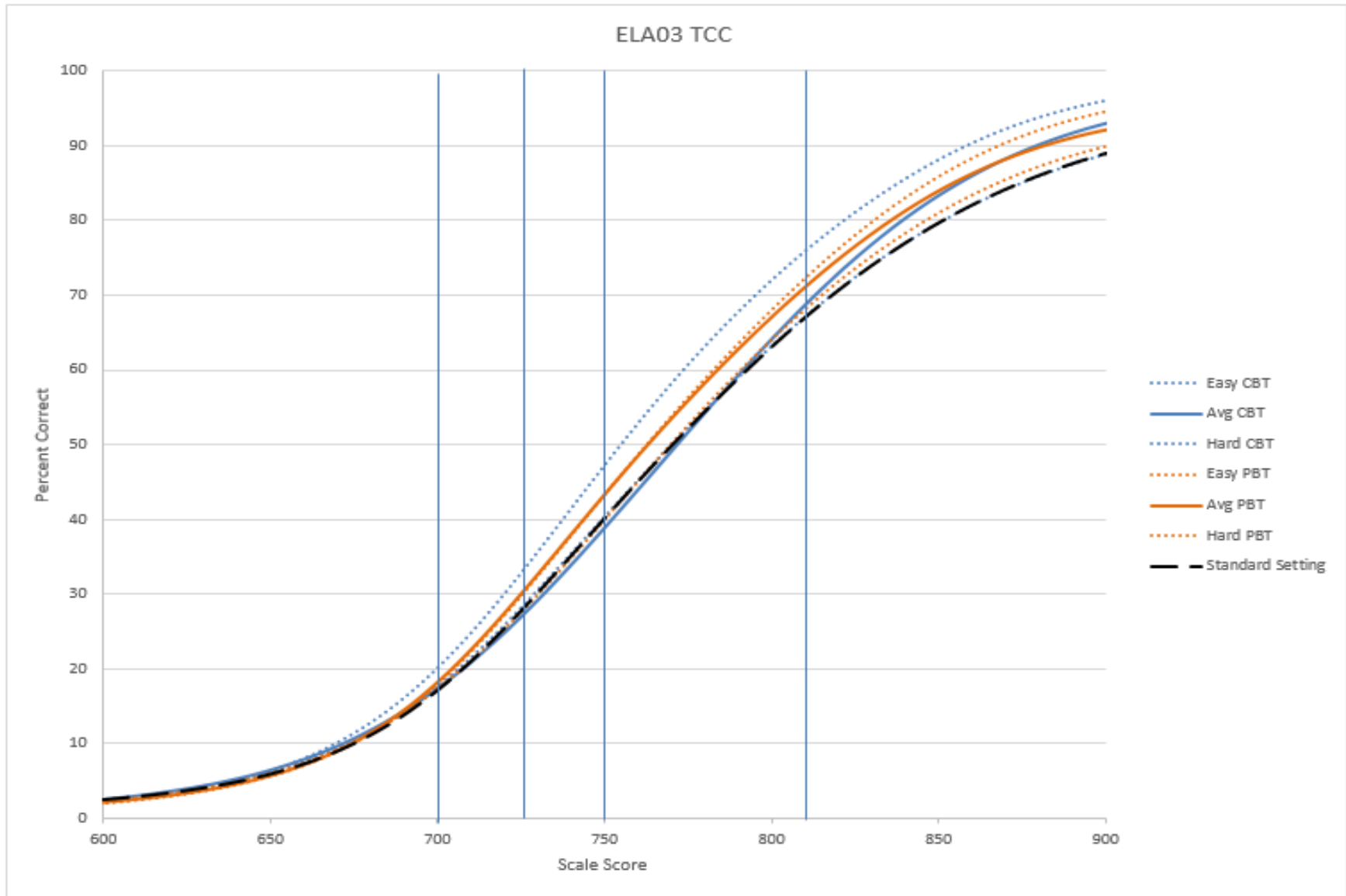


Figure 12.1 Test Characteristic Curves for ELA/L Grade 3

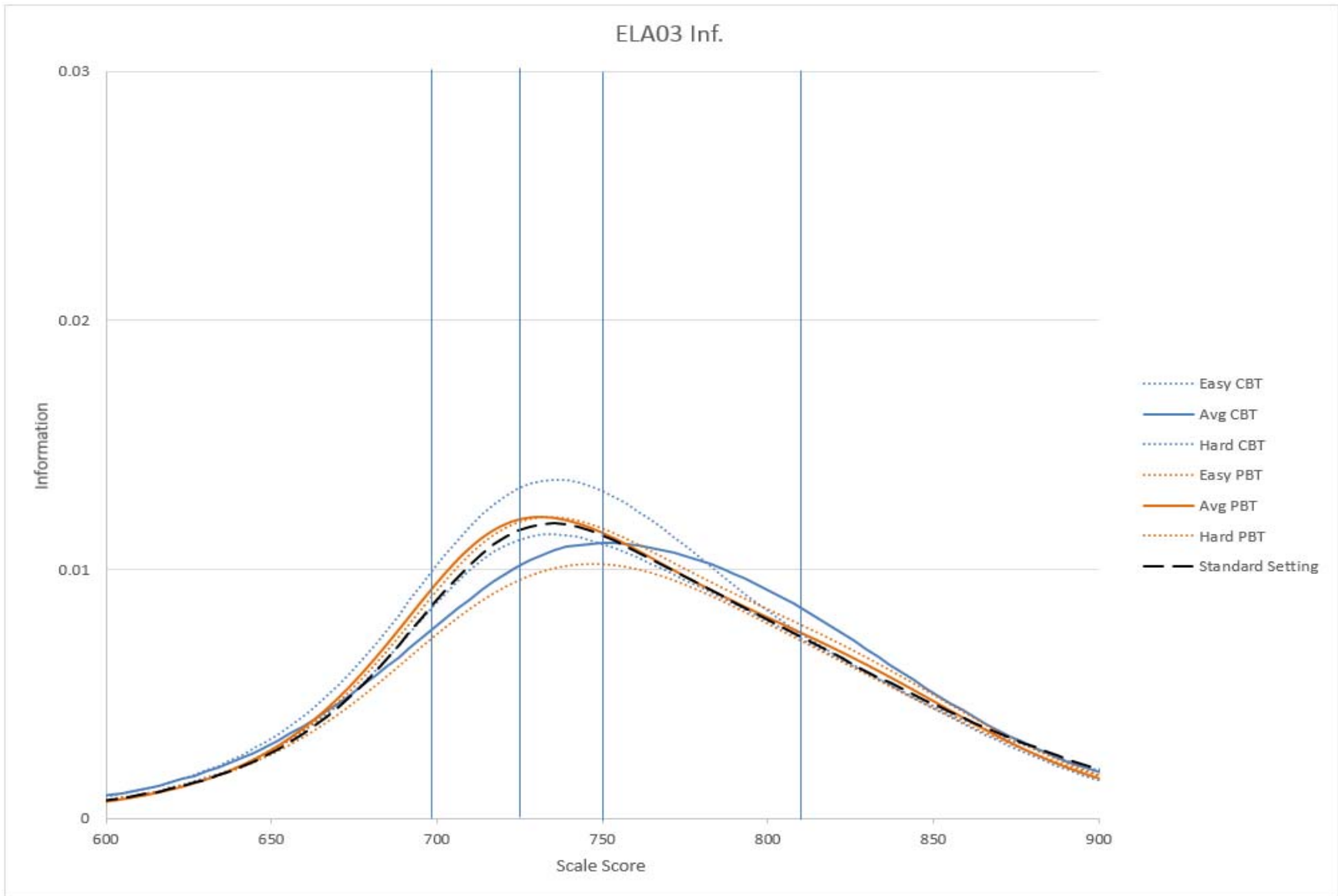


Figure 12.2 Information Curves for ELA/L Grade 3

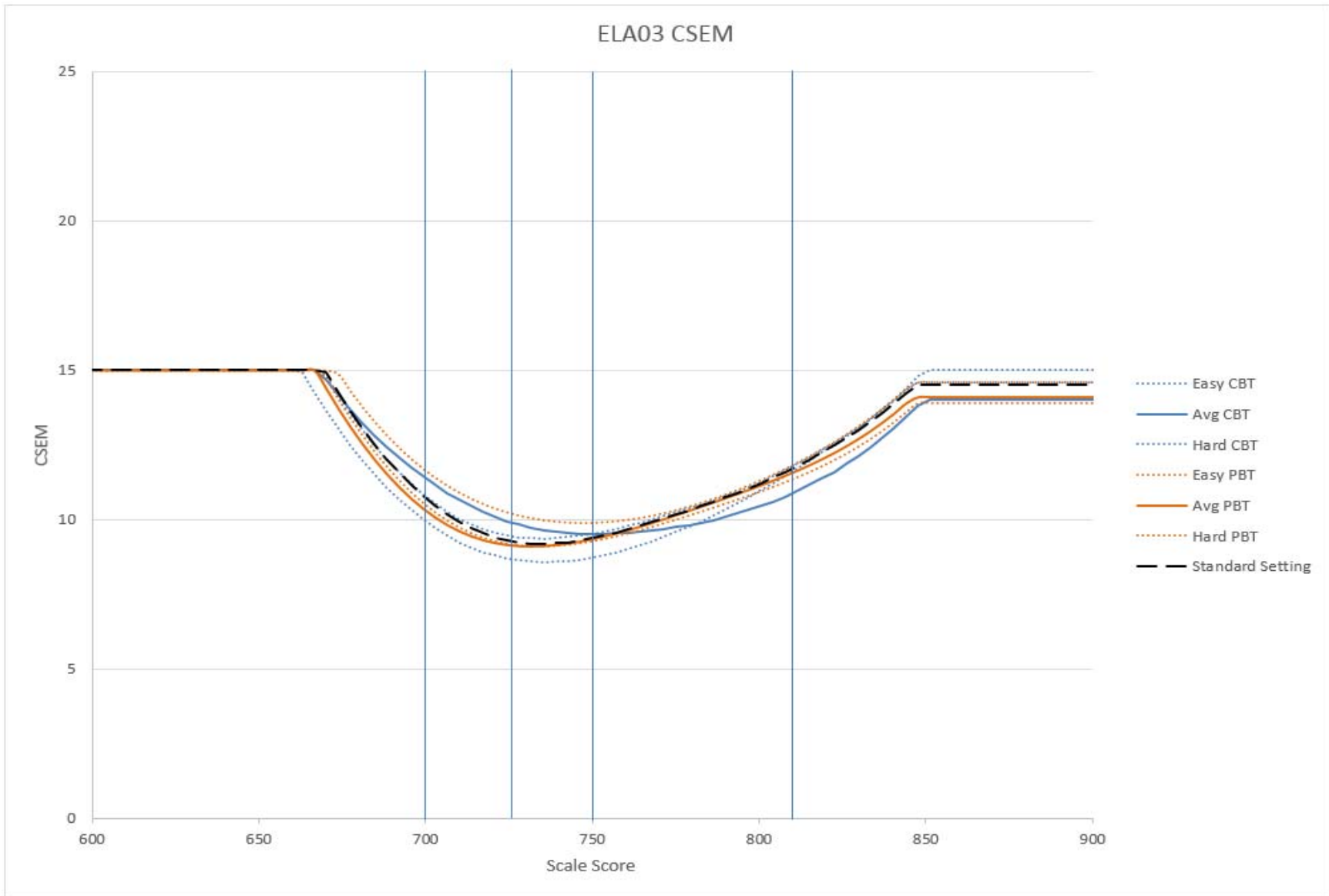


Figure 12.3 Conditional Standard Error of Measurement Curve for ELA/L Grade 3

For ELA/L grade 3, CBT and PBT forms had very similar TCCs. The standard setting form was a bit more difficult than the average CBT or PBT forms, as reflected by a higher scale score being associated with a particular percentage correct. Information and CSEM curves were similar for CBT forms, PBT forms, and for the standard setting form. **Appendix 12.4** contains TCC, INF, and CSEM curves for all ELA/L grades and all mathematics grades/courses.

12.4 Score Distributions

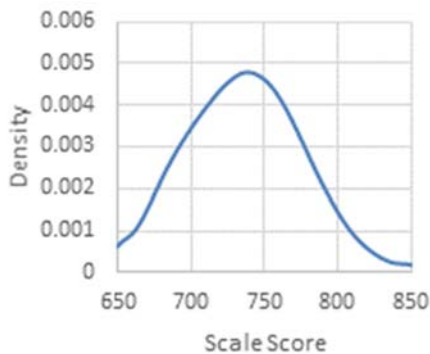
12.4.1 Score Distributions for ELA/L

All Students

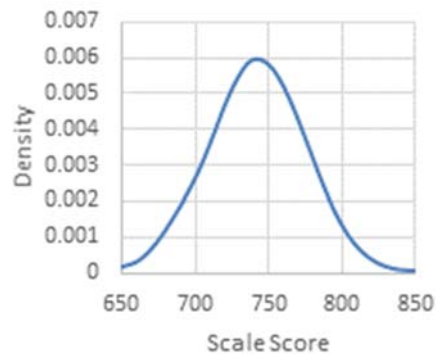
Figures 12.4 through 12.6 graphically represent the distributions of scale scores for grades 3 through 11 ELA/L FS, Reading, and Writing, respectively. The vertical axis of each graph, labeled “Density”, represents the proportion of students earning the scale score point indicated along the horizontal axis.

ELA/L scale scores that were a bit below the Level 4 cut score (i.e., 750) were most commonly observed for grades 3 to 11.

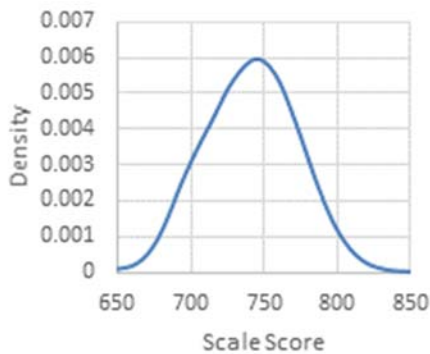
Grade 3



Grade 4



Grade 5



Grade 6

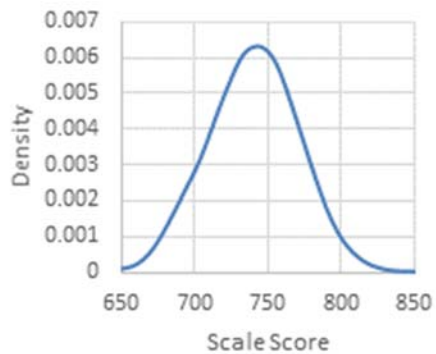
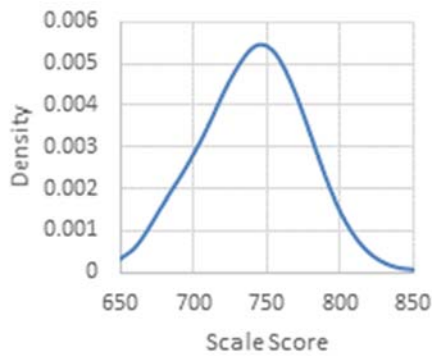
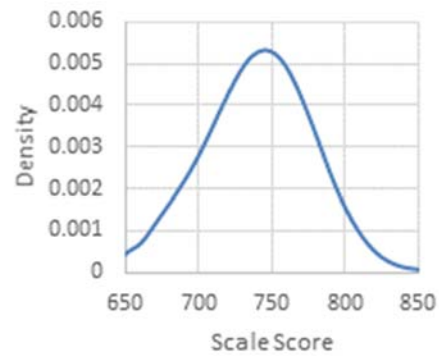


Figure 12.4 Distributions of ELA/L Scale Scores: Grades 3 to 11

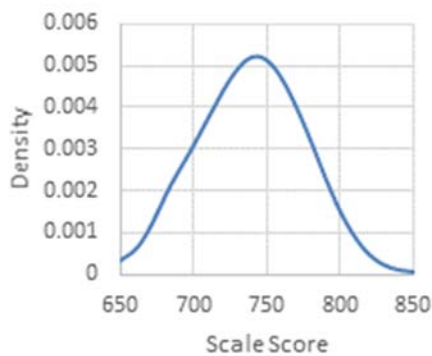
Grade 7



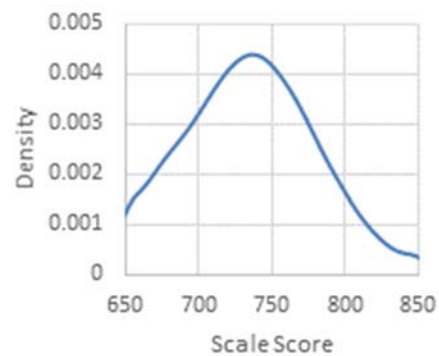
Grade 8



Grade 9



Grade 10



Grade 11

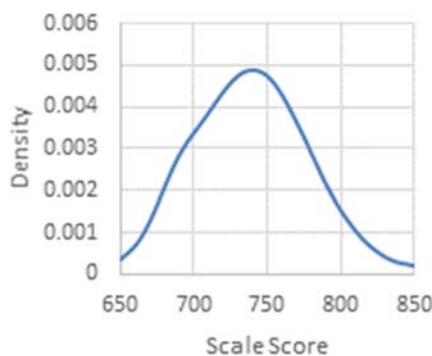
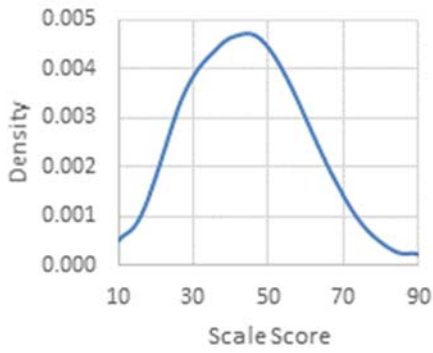


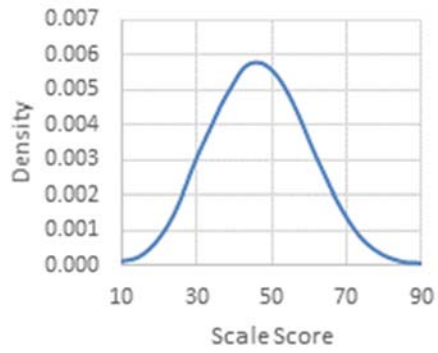
Figure 12.4 (continued) Distributions of ELA/L Scale Scores: Grades 3 to 11

Reading scale scores that were a bit below the Level 4 cut score of 50 were most often observed. Distributions were fairly symmetric, with scores below the Level 4 cut score being a bit more common than higher scores. A portion of this is due to larger numbers of students earning near zero raw scores than near perfect raw scores. Near zero raw scores can occur for a variety of reasons, such as student illness, and do not always indicate weak student skills.

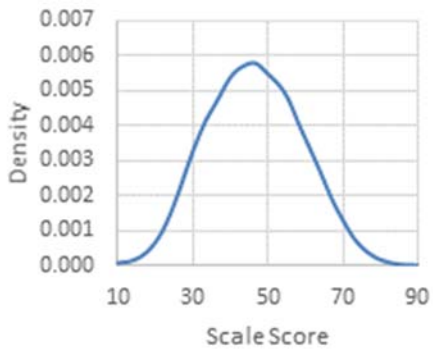
Grade 3



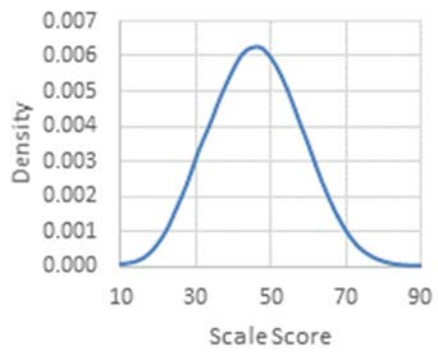
Grade 4



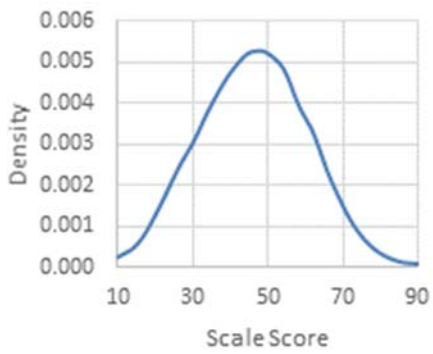
Grade 5



Grade 6



Grade 7



Grade 8

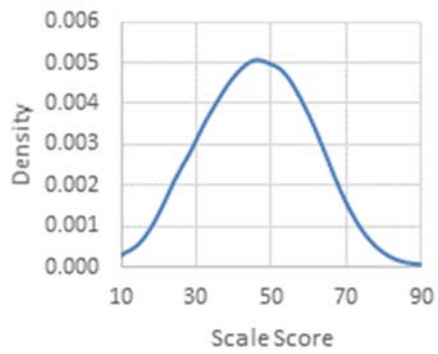
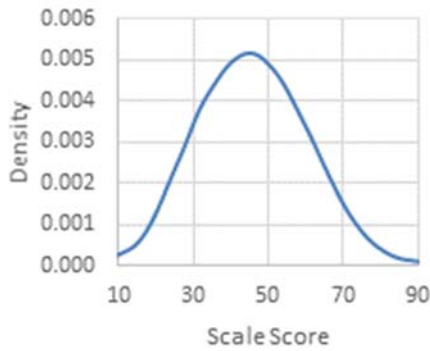
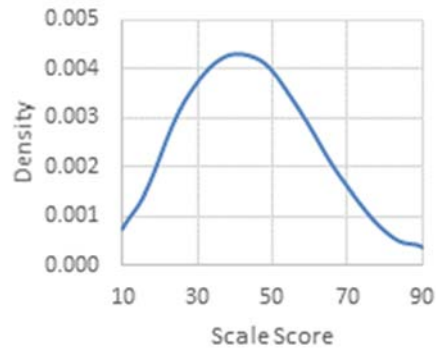


Figure 12.5 Distributions of Reading Scale Scores: Grades 3 to 11

Grade 9



Grade 10



Grade 11

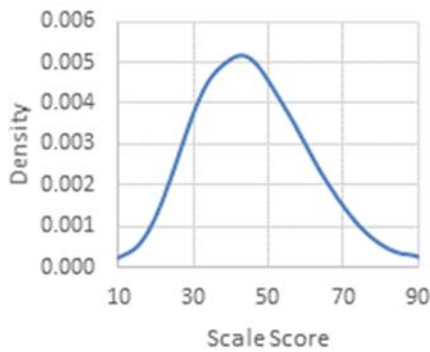
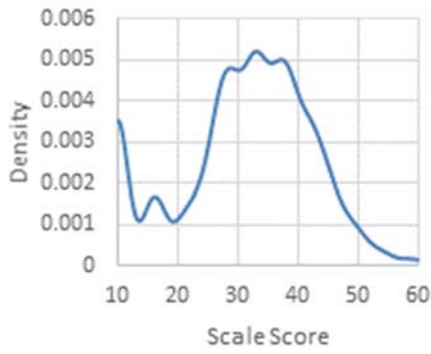


Figure 12.5 (continued) Distributions of Reading Scale Scores: Grades 3 to 11

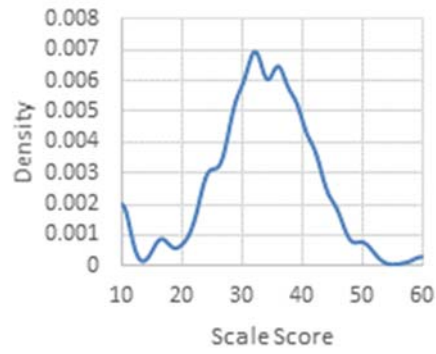
Writing scale score distributions are noticeably less smooth than Reading or ELA/L FS distributions due to peaks related to the weighting of the Written Expression portion of the PCR tasks.

The proportion of students earning the lowest obtainable scale score is fairly high for Writing. This occurred even though a score point of zero is typically necessary to obtain the minimum scale score. Writing items are embedded exclusively in PCR tasks, which tended to be difficult. Written Expression trait/subclaim also tended to be the most difficult of the PCR traits.

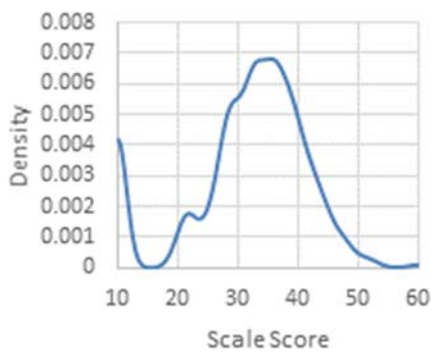
Grade 3



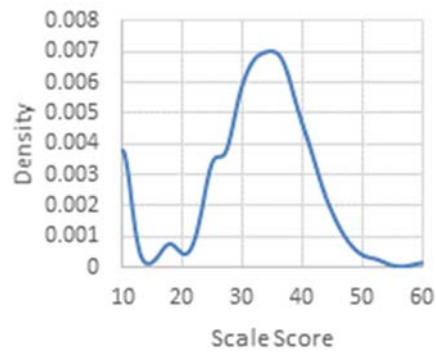
Grade 4



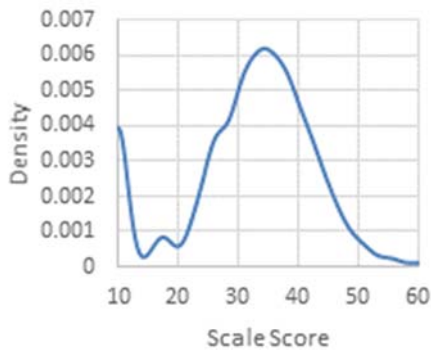
Grade 5



Grade 6



Grade 7



Grade 8

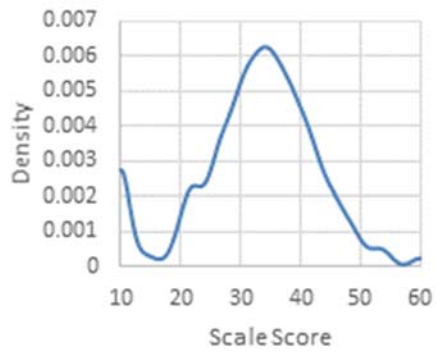
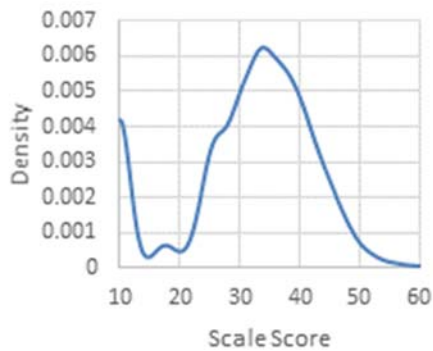
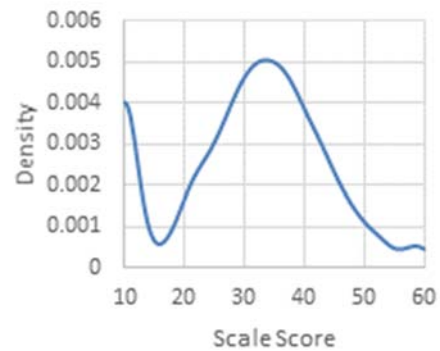


Figure 12.6 Distributions of Writing Scale Scores: Grades 3 to 11

Grade 9



Grade 10



Grade 11

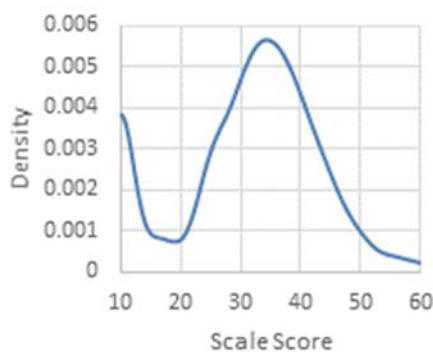


Figure 12.6 (continued) Distributions of Writing Scale Scores: Grades 3 to 11

Groups

Grade 3 group statistics for ELA/L FS, Reading, and Writing scale scores are presented in Table 12.7.²⁵ Mean scores were higher for female students relative to male students. Mean scores were highest for Asian students followed by “Multiple Race Selected” students and were lowest for American Indian/Alaska Native students. Economically disadvantaged students performed less well than students who are not economically disadvantaged. English learners (EL) performed less well than Non English learner students. Students with disabilities (SWD) performed less well than students without disabilities.

Patterns of mean scale scores were extremely similar in other grades; corresponding tables for all grades are presented in **Appendix 12.5**.

²⁵ Table A.12.48 in Appendix 12.5 is identical to Table 12.7.

Table 12.7 Subgroup Performance for ELA/L: Grade 3

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		518,857	736.20	39.74	650	850
Gender	Female	254,214	740.84	39.98	650	850
	Male	264,643	731.73	38.99	650	850
Ethnicity	American Indian/Alaska Native	5,088	715.25	33.61	650	850
	Asian	24,035	762.06	39.66	650	850
	Black or African American	95,098	719.84	35.79	650	850
	Hispanic/Latino	126,309	722.69	36.19	650	850
	Native Hawaiian or Pacific Islander	997	732.97	41.78	650	850
	Multiple Race Selected	43,004	750.47	39.45	650	850
	White	221,943	745.95	37.96	650	850
Economic Status*	Economically Disadvantaged	249,056	721.61	35.77	650	850
	Not Economically Disadvantaged	236,384	752.60	37.95	650	850
English Learner Status	English Learner (EL)	67,638	713.33	33.16	650	850
	Non English Learner	416,752	740.51	39.67	650	850
Disabilities	Students with Disabilities (SWD)	46,945	707.95	37.10	650	850
	Students without Disabilities	319,190	741.47	38.65	650	850
Reading Score		518,857	44.35	15.87	10	90
Gender	Female	254,214	45.71	15.89	10	90
	Male	264,643	43.05	15.74	10	90
Ethnicity	American Indian/Alaska Native	5,088	35.87	13.12	10	90
	Asian	24,035	53.62	15.91	10	90
	Black or African American	95,098	37.81	13.91	10	90
	Hispanic/Latino	126,309	38.55	14.02	10	90
	Native Hawaiian or Pacific Islander	997	42.78	16.25	10	90
	Multiple Race Selected	43,004	50.02	16.06	10	90
	White	221,943	48.60	15.38	10	90
Economic Status*	Economically Disadvantaged	249,056	38.39	13.95	10	90
	Not Economically Disadvantaged	236,384	50.90	15.41	10	90
English Learner Status	English Learner (EL)	67,638	34.67	12.48	10	90
	Non English Learner	416,752	46.08	15.88	10	90
Disabilities	Students with Disabilities (SWD)	46,945	34.02	14.62	10	90
	Students without Disabilities	319,190	46.40	15.52	10	90
Writing Score		518,857	31.10	11.24	10	60
Gender	Female	254,214	32.76	10.99	10	60
	Male	264,643	29.50	11.24	10	60
Ethnicity	American Indian/Alaska Native	5,088	26.52	10.63	10	60

Group Type	Group	N	Mean	SD	Min	Max
	Asian	24,035	37.97	10.22	10	60
	Black or African American	95,098	27.34	11.03	10	60
	Hispanic/Latino	126,309	28.56	11.05	10	60
	Native Hawaiian or Pacific Islander	997	30.62	11.77	10	60
	Multiple Race Selected	43,004	34.43	10.65	10	60
	White	221,943	32.92	10.72	10	60
Economic Status*	Economically Disadvantaged	249,056	27.94	11.00	10	60
	Not Economically Disadvantaged	236,384	34.86	10.39	10	60
English Learner Status	English Learner (EL)	67,638	26.64	10.90	10	60
	Non English Learner	416,752	32.07	11.12	10	60
Disabilities	Students with Disabilities (SWD)	46,945	23.09	11.44	10	60
	Students without Disabilities	319,190	32.40	10.82	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Grade 9 group statistics for ELA/L, Reading, and Writing scale scores are presented in Table 12.8.²⁶ Mean scores were very similar to what was observed for grades 3 to 8. Mean scores were higher for female students than for male students. Mean scores were highest for Asian students followed by White students; scores were lowest for Black or African American students. Economically disadvantaged students performed less well than students who are not economically disadvantaged. English learners (EL) performed less well than Non English Learner students. Students with disabilities (SWD) performed less well than students without disabilities.

Very similar patterns are observed in other grades, and corresponding tables for all grades are presented in **Appendix 12.5**.

²⁶ Table A.12.54 in Appendix 12.5 is identical to Table 12.8.

Table 12.8 Subgroup Performance for ELA/L: Grade 9

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		413,098	739.45	36.55	650	850
Gender	Female	201,280	746.57	35.57	650	850
	Male	211,818	732.68	36.18	650	850
Ethnicity	American Indian/Alaska Native	3,982	721.02	31.41	650	836
	Asian	17,616	762.36	37.86	650	850
	Black or African American	56,553	720.83	32.84	650	850
	Hispanic/Latino	85,733	725.20	33.80	650	850
	Native Hawaiian or Pacific Islander	702	734.33	37.76	650	850
	Multiple Race Selected	15,568	741.13	36.50	650	850
	White	211,054	747.33	34.58	650	850
Economic Status*	Economically Disadvantaged	175,922	725.87	33.29	650	850
	Not Economically Disadvantaged	228,674	749.50	35.49	650	850
English Learner Status	English Learner (EL)	20,828	703.19	28.74	650	843
	Non English Learner	377,194	741.60	35.85	650	850
Disabilities	Students with Disabilities (SWD)	40,963	705.72	29.15	650	850
	Students without Disabilities	278,126	745.43	34.73	650	850
Reading Score		413,098	45.78	14.58	10	90
Gender	Female	201,280	47.80	14.26	10	90
	Male	211,818	43.86	14.62	10	90
Ethnicity	American Indian/Alaska Native	3,982	38.32	12.45	10	90
	Asian	17,616	53.95	15.25	10	90
	Black or African American	56,553	38.58	12.80	10	90
	Hispanic/Latino	85,733	39.93	13.12	10	90
	Native Hawaiian or Pacific Islander	702	43.06	14.60	10	89
	Multiple Race Selected	15,568	46.47	14.55	10	90
	White	211,054	49.04	13.98	10	90
Economic Status*	Economically Disadvantaged	175,922	40.45	13.11	10	90
	Not Economically Disadvantaged	228,674	49.75	14.30	10	90
English Learner Status	English Learner (EL)	20,828	31.41	10.77	10	90
	Non English Learner	377,194	46.64	14.33	10	90
Disabilities	Students with Disabilities (SWD)	40,963	33.04	11.52	10	90
	Students without Disabilities	278,126	48.07	13.99	10	90
Writing Score		413,098	31.44	10.83	10	60
Gender	Female	201,280	34.19	9.89	10	60
	Male	211,818	28.83	11.04	10	60
Ethnicity	American Indian/Alaska Native	3,982	27.33	10.26	10	57

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
	Asian	17,616	37.77	10.01	10	60
	Black or African American	56,553	26.50	10.96	10	60
	Hispanic/Latino	85,733	28.18	10.93	10	60
	Native Hawaiian or Pacific Islander	702	30.95	11.31	10	60
	Multiple Race Selected	15,568	31.77	10.89	10	60
	White	211,054	33.27	10.08	10	60
Economic Status*	Economically Disadvantaged	175,922	28.05	10.76	10	60
	Not Economically Disadvantaged	228,674	33.92	10.18	10	60
English Learner Status	English Learner (EL)	20,828	22.36	10.62	10	60
	Non English Learner	377,194	31.96	10.63	10	60
Disabilities	Students with Disabilities (SWD)	40,963	21.84	10.53	10	60
	Students without Disabilities	278,126	33.08	10.09	10	60

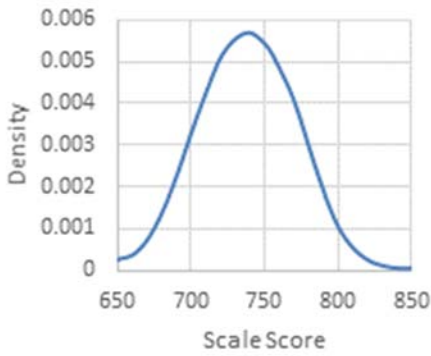
Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

12.4.2 Score Distributions for Mathematics

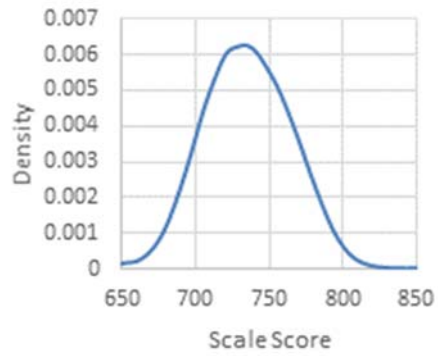
All Students

Figure 12.7 graphically represents the distributions of scale scores for grades 3 through 8 mathematics. Scale score distributions peaked midway between 700 and the Level 4 performance level cut of 750.

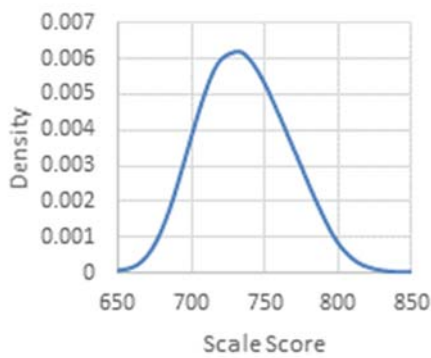
Grade 3



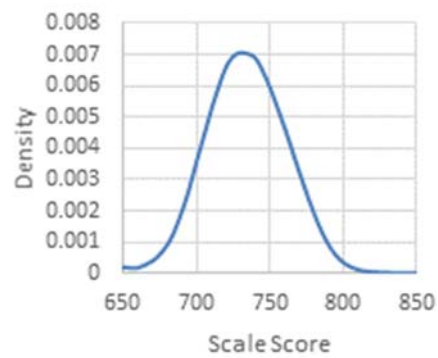
Grade 4



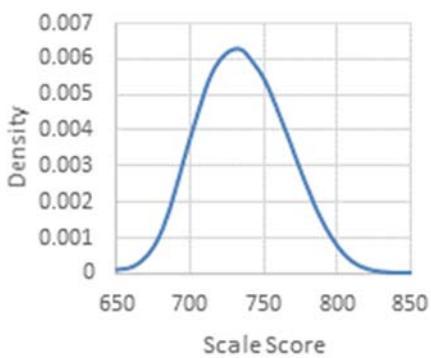
Grade 5



Grade 6



Grade 7



Grade 8

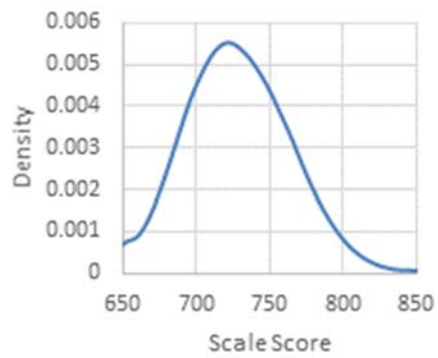
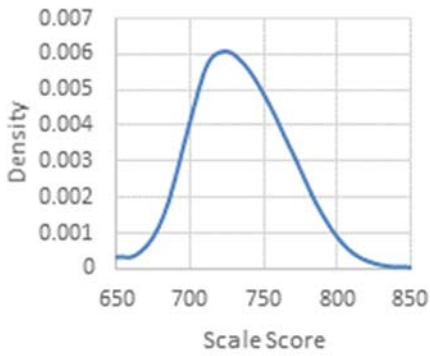


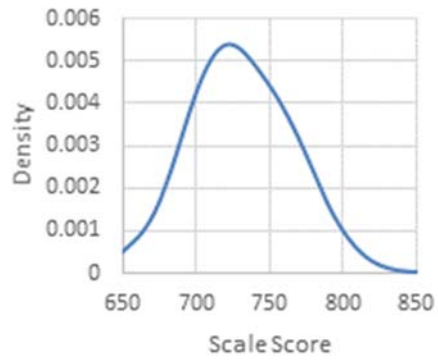
Figure 12.7 Distributions of Mathematics Scale Scores: Grades 3 to 8

Figure 12.8 graphically represents the distributions of scale scores for Algebra I, Geometry, Algebra II, and Integrated Mathematics I through III. Scale score distributions peaked between 700 and the 750 Level 4 performance level cut score. Distributions were similar for Algebra I and Integrated Mathematics I, for Geometry and Integrated Mathematics II, and for Algebra II and Integrated Mathematics III.

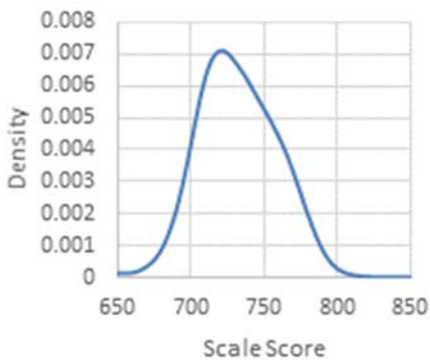
Algebra I



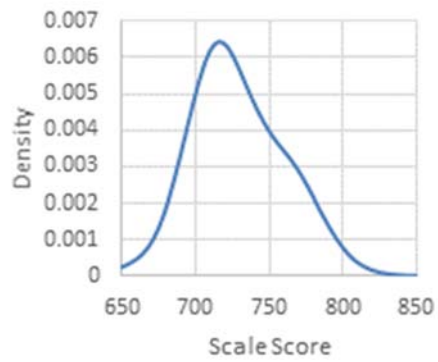
Integrated Mathematics I



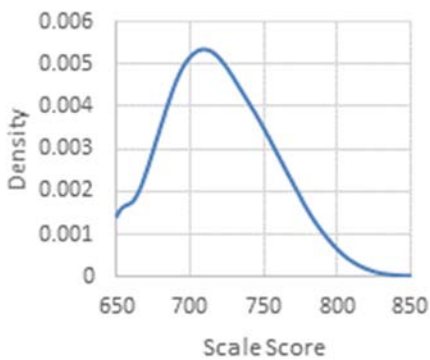
Geometry



Integrated Mathematics II



Algebra II



Integrated Mathematics III

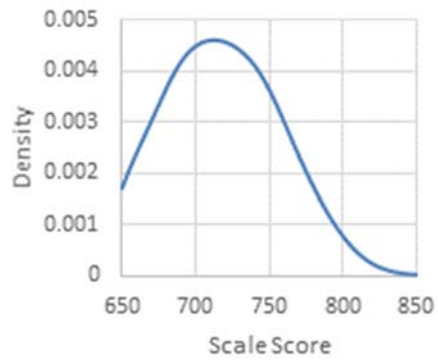


Figure 12.8 Distributions of Mathematics Scale Scores: High School

Groups

Grade 3 group statistics for mathematics scale scores are presented in Table 12.9.²⁷ Mean scores were higher for female students relative to male students. Mean scores were highest for Asian students

²⁷ Table A.12.57 in Appendix 12.5 is identical to Table 12.9.

followed by “Multiple Race Selected” students and were lowest for Black or African American students. Economically disadvantaged students performed less well than students who are not economically disadvantaged. English learners (EL) performed less well than Non English learner students. Students with disabilities (SWD) performed less well than students without disabilities.

Table 12.9 Subgroup Performance for Mathematics Scale Scores: Grade 3

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		650,262	738.42	33.12	650	850
Gender	Female	318,517	738.86	32.04	650	850
	Male	331,745	737.99	34.11	650	850
Ethnicity	American Indian/Alaska Native	5,270	722.18	28.62	650	836
	Asian	26,866	765.78	33.45	650	850
	Black or African American	114,909	721.97	29.90	650	850
	Hispanic/Latino	136,010	727.57	30.03	650	850
	Native Hawaiian or Pacific Islander	1,149	736.09	35.00	650	845
	Multiple Race Selected	48,249	748.37	33.56	650	850
	White	307,191	745.89	31.05	650	850
Economic Status*	Economically Disadvantaged	311,763	726.10	30.05	650	850
	Not Economically Disadvantaged	296,843	751.70	31.42	650	850
English Learner Status	English Learner (EL)	77,582	723.18	29.75	650	850
	Non English Learner	528,210	741.04	33.13	650	850
Disabilities	Students with Disabilities (SWD)	61,364	716.60	32.90	650	850
	Students without Disabilities	424,037	742.84	31.73	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Similar patterns were observed in other grades; corresponding tables for all grades are presented in **Appendix 12.5**.

Algebra I scale score statistics are presented in Table 12.10.²⁸ Mean scores were higher for female students relative to male students. Mean scores were highest for Asian students followed by White students and were lowest for Black or African American students. Economically disadvantaged students performed less well than students who are not economically disadvantaged. English learners (EL) performed less well than Non English learner students. Students with disabilities (SWD) performed less well than students without disabilities.

²⁸ Table A.12.63 in Appendix 12.5 is identical to Table 12.10.

Table 12.10 Subgroup Performance for Mathematics Scale Scores: Algebra I

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		480,604	733.96	32.23	650	850
Gender	Female	233,330	735.72	31.13	650	850
	Male	247,274	732.29	33.15	650	850
Ethnicity	American Indian/Alaska Native	4,024	718.89	25.82	650	822
	Asian	20,209	761.40	35.88	650	850
	Black or African American	93,290	718.84	26.79	650	850
	Hispanic/Latino	91,047	722.22	28.10	650	850
	Native Hawaiian or Pacific Islander	775	730.50	32.40	650	839
	Multiple Race Selected	18,497	740.96	35.21	650	850
	White	230,890	741.54	30.97	650	850
Economic Status*	Economically Disadvantaged	194,799	722.37	27.77	650	850
	Not Economically Disadvantaged	249,242	743.28	32.68	650	850
English Learner Status	English Learner (EL)	25,094	711.30	25.68	650	850
	Non English Learner	438,285	735.42	32.05	650	850
Disabilities	Students with Disabilities (SWD)	45,685	709.45	25.77	650	850
	Students without Disabilities	333,456	739.43	32.28	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Very similar patterns were observed in Geometry and Algebra II; corresponding tables are presented in **Appendix 12.5**.

Integrated Mathematics I scale score statistics are presented in Table 12.11.²⁹ Mean scores were higher for female students relative to male students. Mean scores were highest for Asian students followed by White students and were lowest for American Indian/Alaska Native students. Economically disadvantaged students performed less well than students who are not economically disadvantaged. English learners (EL) performed less well than non-English learner students. Students with disabilities (SWD) performed less well than students without disabilities.

²⁹ Table A.12.66 in Appendix 12.5 is identical to Table 12.11.

Table 12.11 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics I

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		30,217	730.86	35.11	650	850
Gender	Female	14,546	732.49	34.06	650	850
	Male	15,671	729.34	35.99	650	850
Ethnicity	American Indian/Alaska Native	341	712.77	29.53	650	799
	Asian	826	746.95	38.94	650	850
	Black or African American	4,147	716.88	29.20	650	825
	Hispanic/Latino	9,212	718.92	30.37	650	842
	Native Hawaiian or Pacific Islander	48	730.42	31.81	668	800
	Multiple Race Selected	1,283	738.76	38.58	650	850
	White	13,336	741.21	34.67	650	850
Economic Status*	Economically Disadvantaged	15,591	719.96	30.39	650	850
	Not Economically Disadvantaged	14,424	742.42	35.80	650	850
English Learner Status	English Learner (EL)	2,670	705.67	25.54	650	842
	Non English Learner	25,207	734.75	34.84	650	850
Disabilities	Students with Disabilities (SWD)	3,515	701.55	26.69	650	821
	Students without Disabilities	18,880	737.05	34.79	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Very similar patterns were observed in Integrated Mathematics II and Integrated Mathematics III; corresponding tables are presented in **Appendix 12.5**.

12.5 Interpreting Claim Scores and Subclaim Scores

12.5.1 Interpreting Claim Scores

PARCC ELA/L assessments provide separate claim scale scores for both Reading and Writing. The claim scale scores and the summative scale score are on different scales; therefore, the sum of the scale scores for each claim will not equal the summative scale score. PARCC Reading scale scores range from 10 to 90 and PARCC Writing scale scores range from 10 to 60.

The claim scores can be interpreted by comparing a student's claim scale score to the average performance for the school, district, and state. The PARCC Individual Student Report (ISR) provides the student scale score results and the average scale score results for the school, district, and state.

12.5.2 Interpreting Subclaim Scores

Within each reporting category are specific skill sets (subclaims) students demonstrate on the PARCC assessments. Subclaim categories are not reported using scale scores or Performance Levels. Subclaim performance for PARCC assessments is reported using graphical representations that indicate how the

student performed relative to the overall performance of students who met or nearly met expectations for the content area.

Subclaim indicators represent how well students performed in a subclaim category relative to how well the reference group performed in that same subclaim category. To determine a student's subclaim performance, the performance of all students in the PARCC consortium taking the same test form whose overall scale score identified them as having just achieved Performance Level 3 (i.e., summative scale score = 725) or having just achieved Performance Level 4 (i.e., summative scale score = 750) were considered, as described below. Once these reference groups were identified, their average performance in each subclaim was noted.

Student performance for each subclaim is marked with a subclaim performance indicator.

- An 'up' arrow for the specified subclaim for "Meets or Exceeds Expectations" indicates that the student's performance for the subclaim was equal to or better than the average performance of students who just achieved Performance Level 4 (i.e., students whose summative scale score was 750).
- A 'bidirectional' arrow for the specified subclaim for "Nearly Meets Expectations" indicates that the student's performance was below the average performance of students who just achieved Performance Level 4 (i.e., students whose summative scale score was 750) but better than or equal to the average performance of students who just achieved Performance Level 3 (i.e., students whose summative scale score was 725).
- A 'down' arrow for the specified subclaim for "Below Expectations" indicates that the student's performance for the subclaim was below the average performance for students who just achieved Performance Level 3 (i.e., students whose summative scale score was 725).

Section 13: Quality Control Procedures

Quality control in a testing program is a comprehensive and ongoing process. This section describes procedures put into place to monitor the quality of the item bank, test form, and ancillary material development. The quality checks for scanning, image editing, scoring, and data screening during psychometric analyses are also outlined. Additional quality information can be found in the PARCC Program Quality Plan document.

13.1 Quality Control of the Item Bank

The PARCC summative item bank consists of test passages and items, their associated metadata, and status (e.g., operational-ready, field-test ready, released, etc.). The items on the 2015 assessments were developed by Pearson, West Ed, and ETS, and put in the item bank once created.

In 2015, the PARCC summative bank underwent a conversion from the existing repository to the Assessment Banking for Building and Interoperability (ABBI) system. The ABBI bank houses the passages and items, art, associated metadata, rubrics, alternate text for use on accommodated forms, and text complexity documentation. It provides an item previewer that allows items to be viewed and interacted with in the same way students see and interact with items and tools, and manages versioning of items with a date/time stamp. It allows PARCC reviewers to vote on item acceptance, and to record and retain their review notes for later reconciliation and reference. As part of the bank transition, quality processes were conducted to ensure that the content of the passages and items, and the underlying QTI structure associated with those items, remained consistent from the old bank to the new bank. A validation of scoring and metadata was conducted. Mathematics rubrics were loaded and the versions validated.

In 2015, the bank transition occurred after initial development, but prior to the PARCC editorial review. PARCC editorial review committee participants conducted their review in the item banking system. As with all reviews, the committee members viewed the items as the student would, and could vote to alter the item, accept the item or reject the item and record their comments in the system. After each meeting, reports were forwarded to PARCC Inc. The reports were generated by the item banking system and summarized feedback from the committee reviewers.

All new development for the PARCC assessments is being created within the ABBI system, which employs templates to control the consistency of the underlying scoring logic and QTI creation for each item type. The ABBI system incorporates a previewer that allows the PARCC reviewers to validate the content of the item and validate the expected scoring of tasks. It supports the full range of PARCC review activities, including content review, bias and sensitivity review, expert editorial review, data review, and test construction review. It provides insight into the item edit process through versioning. A series of metadata validations at key points in the development cycle provide support for metadata consistency. The bank can be queried on the full range of PARCC metadata values to support bank analysis.

13.2 Quality Control of Test Form Development

Test Forms were built based upon targets and the established blueprints set by PARCC. The construction process started with specification and requirement capture to create the Test Specification Document. From there items were pulled into forms based on the criteria approved in the Test Specifications document. Quality control steps were conducted on the items and forms evaluating several item characteristics (e.g., content accuracy, completeness, style guides conformity, tools function). After forms composition, the forms went through a review process that involved groups from Pearson, Parcc Inc., and the PARCC states. Revisions were incorporated into the forms before final review and approval. Section 2.2 provides more details on the form development process.

The forms quality assurance was performed by Pearson's Assessment and Information Quality (AIQ) organization. AIQ completed a comprehensive review of all *online* forms for the PARCC administration cycle. This group is part of Pearson's larger Organizational Quality group and operates exclusively to validate form operability. The group validates that the functionality of every online form is working to specifications. The overall functionality and maneuverability of each form is checked, and the behavior of each item within the form is verified. (Quality processes for paper forms are described in section 13.3).

The items within each form were tested to verify that they operated as expected for test takers. As a further aspect of the testing process, AIQ confirmed that forms were loaded correctly and that the audio was correct when compared to text. Sections, seal codes, and overviews were reviewed. Technology enhanced items also were tested as an additional measure. As enumerated in the Technology Guidelines for PARCC Assessments,³⁰ user interfaces were compatible with a range of common computer devices, operating systems, and browsers.

Pearson also performed QC tests to verify that a standard set of responses was outputted to the XML as expected after PARCC had approved the final version of the form. These responses were based on the keys provided in the test map or a standard open-ended (OE) responses string that contained a valid range of characters. The test maps also were validated against the form layout and item types for correctness as part of these tests.

Pearson conducted a multifaceted validation of all item layout, rendering, and functionality. Reviewers conducted comparisons between the approved item and the item as it appeared in the field test form, validated that tools and functions in the test delivery system, TestNav, were accurately applied, and verified that the style and layout met requirements documented in the PARCC Publishing Style Guide as

³⁰ This document is available online from:

http://www.parcconline.org/files/72/Technology%20Guidelines%20for%20PARCC%20Assessments/389/TechnologyGuidelinesPARCCAssessments-v5_0-August2015.pdf

part of the PARCC Item Development Technical Guide.³¹ In addition, all answer keys were validated through a formal key review process. More details on the Test Development procedures are provided in Section 2.

13.3 Quality Control of Test Materials

Pearson provided high quality materials in a timely and efficient manner to meet PARCC's test administration needs. Since the majority of printing work was done in-house it was possible to fully control the production environment, press schedule, and quality process for print materials. Additionally, strict security requirements were employed to protect secure materials production; Section 3 provides details on the secure handling of test materials. Materials were produced according to the PARCC Style Guide and to the detailed specifications supplied in the Materials List.

Pearson Print Service operates within the sanctions of an ISO 9001:2008 Quality Management System, and practices process improvement through Lean principles and employee involvement.

Raw materials (paper and ink) used for scannable forms production were manufactured exclusively for Pearson Print Service using specifications created by Pearson Print Service. Samples of ink and paper were tested by Pearson prior to use in production. Project Specialists were the point of contact for incoming production.

Purchase orders and other order information were assessed against manufacturing capabilities and assigned to the optimal production methodology. PARCC expectations, quality requirements, and cost considerations were foremost in these decisions. Prior to release for manufacture, order information was checked against PARCC specifications, technical requirements, and other communication that includes expected outcomes. Records of these checks were maintained.

Files for image creation flow through one of two file preparation functions: Digital pre-press (DPP) for digital print methodology, or plateroom for offset print methodology. Both the DPP and plateroom functions verify content, file naming, imposition, pagination, numbering stream, registration of technical components, color mapping, workflow, and file integrity. Records of these checks are created and saved.

Offset production requires printing that uses a lithographic process. Offline finishing activities are required to create books and package offset output. Digital output may flow through an inkjet Digital Production Line (DPL) or a sheet-fed toner application process in the Xpress Center. A battery of quality checks was performed in these areas. The checks included color match, correct file selection, content match to proof, litho-code to serial number synchronization, registration of technical components, ink density controlled by densitometry, inspection for print flaws, perforations, punching, pagination, scanning requirements, and any unique features specified for the order. Records of these checks and

³¹ The PARCC Item Development Technical Guide is available online from:
<http://parccinc.org/wp-content/uploads/2014/07/PARCCItemDevelopmentTechnicalGuidePUBLICDRAFTFORRELEASE-20130912.pdf>

samples pulled from planned production points were maintained. Offline finishing included cutting, shrink wrapping, folding, and collating. The collation process has three robust inline detection systems that inspected each book for:

- Caliper validation that detects too few or too many pages. This detector will stop the collator if an incorrect caliper reading is registered.
- An optical reader that will only accept one sheet. Two or zero sheets will result in a collator stoppage.
- The correct bar code for the signature being assembled. An incorrect or upside down signature will be rejected by the bar code scanner and will result in a collator stoppage.

Pearson's Quality Assurance (QA) department personnel inspected print output prior to collation and shipment. QA also supported process improvement, work area documentation, audited process adherence, and established training programs for employees.

13.4 Quality Control of Scanning

Establishing and maintaining the accuracy of scanning, editing, and imaging processes is a cornerstone of the Pearson scoring process. While the scanners are designed to perform with great precision, Pearson implements other quality assurance processes to confirm that the data captured from scan processing produce a complete and accurate map to the expected results.

Pearson pioneered optical mark reading (OMR) and image scanning, and continues to improve in-house scanners for this purpose. Software programs drive the capture of student demographic data and student responses from the test materials during scan processing. Routinely scheduled maintenance and adjustments to the scanner components (e.g., camera) maintain scanner calibration. Test sheets inserted into every batch test scanner accuracy and calibration.

Controlled processes for developing and testing software specifications included a series of validation and verification procedures to confirm the captured data can be mapped accurately and completely to the expected results and that editing application rules are properly applied.

13.5 Quality Control of Image Editing

The final step in producing accurate data for scoring is the editing process. Once information from the documents was captured in the scanning process, the scan program file was executed, comparing the data captured from the student documents to the project specifications. The result of the comparison was a report (or edit listing) of documents needing corrections or validation. Image Editing Services performed the tasks necessary to correct and verify the student data prior to scoring.

Using the report, editors verified that all unscanned documents were scanned, or the data were imported into the system through some other method such as flatbed scan or key entry.

Documents with missing or suspect data were pulled, verified, and corrections or additional data were entered. Standard edits included:

- Incorrect or double gridding
- Incorrect dates (including birth year)
- Mismatches between pre-ID label and gridded information
- Incomplete names

When all edits were resolved, corrections were incorporated into the document file containing student records.

Additional quality checks were also performed. These included student *N*-count checks to make certain:

- Students were placed under the correct header
- All sheets belonged to the appropriate document
- Documents were not scanned twice
- No blank documents existed

Finally, accuracy checks were performed by checking random documents against scanned data to verify the accuracy of the scanning process.

Once all corrections were made, the scan program was tested a second time to verify all data were valid. When the resulting output showed that no fields were flagged as suspect, the file was considered clean and scoring began. Once all scanning was completed, the right/wrong response data were securely handed off.

13.6 Quality Control of Answer Document Processing and Scoring

Quality control of answer document and scoring involves all aspects of the scoring procedures, including key-based and rule-based machine scoring and handscoring for constructed response items and performance tasks.

For the 2015 PARCC operational administration, Pearson's validation team prepared test plans used throughout the scoring process. Test plan preparation was organized around detailed specifications.

Based on lessons learned from previous administrations the following quality steps were implemented:

- Raw score validation (e.g., score key validation; evidence statement, field test non-score; double-grid combinations; possible correct combination, if applicable; out-of-range/negative test cases);
- Matching (e.g., validation of high-confidence criteria, low-confidence criteria, cross document, external or forced matching by customer; prior to and after data updates; extract file of matched and unmatched documents);
- Demographic update tests (e.g., verification of data extract against corresponding layout; valid values for updatable fields; invalid values for updatable/non-updatable fields; negative test for non-existing record or empty file).

The following components were added to the quality control process specifically for the PARCC program. These additional steps were introduced to address issues with item-level scoring that were identified in the 2014 PARCC field test administration:

- XML Validation: A combination of automated validation against 100% of item XMLs and human inspection of XML from selected difficult item types or composite items.
- Administration/End-to-End Data Validation: An automated generation of response data from approved test maps that have known conditions were executed against the operational scoring systems and data generation systems to verify scoring accuracy.
- Psychometric Validation: Verification of data integrity using criteria typically used in psychometric processes (e.g., statistical keychecks) and categorization of identified issues to help inform investigation by other groups
- Content Validation: An examination, by subject matter experts, of all items using a combination of automated tools to generate response and scoring data.

In addition to the steps described above, the following quality control process for answer keys and scoring that was implemented for the first PARCC operational administration were used:

1. Pearson’s psychometrics team conducted empirical analyses based on preliminary data files and flagged items based on statistical criteria;
2. Pearson content team reviewed the flagged items and provided feedback on the accuracy of content, answer keys, and scoring;
3. Items potentially requiring changes were added to the product validation (PV) log for further investigation by other Pearson teams;
4. PARCC staff was notified of items for which keys or scoring changes were recommended;
5. PARCC approved/rejected scoring changes; and
6. All approved scoring changes were implemented and validated prior to the generation of the data files used for psychometric processing.

13.7 Quality Control of Psychometric Processes

13.7.1 ETS Psychometric Quality Control Process

ETS was responsible for the psychometric analyses of the 2014-2015 PARCC operational administration and implemented measures to ensure the quality of work. The psychometric analyses were all conducted according to well-defined specifications. Data cleaning rules were clearly articulated and applied consistently throughout the process. Results from all analyses underwent comprehensive quality checks by a team of psychometricians and data analysts. Detailed checklists were used by members of the team for each statistical procedure.

Described below is an overview of the quality control steps performed at different stages of the psychometric analyses. Greater detail is provided in Sections 6 (Classical Item Analysis), 7 (Differential Item Functioning), 10 (IRT Calibration and Scaling), and 12 (Scale Scores).

Data Screening

Data screening is an important first step to ensure quality data input for meaningful analysis. Given this was the first PARCC operational administration, data challenges due to low student performance were expected that may be associated with factors such as lack of opportunity to learn the new content,

unfamiliarity with innovative item types and new testing platforms. ETS consulted with Pearson and established predefined valid case criteria, which were implemented consistently throughout the process. Refer to Section 5.3 for rules for inclusion of students in analyses and Section 10.2 for IRT calibration data preparation criteria and procedures.

Classical Item Analysis

Classical item analysis (IA) produces item level statistics (e.g., item difficulty and item-total correlations). The IA results were reviewed by ETS psychometricians. Items flagged for unusual statistical properties were sent to Pearson for content review. A subset of items identified as having key issues, scoring issues, or content issues was presented to the PARCC Priority Alert Task Force, which made decisions on whether to exclude them from the IRT calibration process and, consequently, the reported student scores. Refer to Section 6.4 for classical IA item flagging criteria.

Calibrations

Creation of item response theory (IRT) sparse data matrices is an important step before the calibrations can begin. Although both the IA data files and the IRT data matrices originated from the same scored item response data, provided by Pearson, the two sets of data were prepared independently. Processing of all data was done in parallel by two psychometricians or data analysts for IA, and the IA and IRT student *N*-counts were compared. The *N*-counts of students included in the IRT data files closely matched the *N*-counts of students included in the IA data files. The small number of discrepancies were due to test takers having multiple test records. For IA, the test record with the higher raw score was included for analysis regardless of test forms. For example, if a test taker took a regular form and a Spanish form, with the Spanish form having the higher score, then the score on the Spanish form was used. However, for IRT that test taker's response data was from the regular form because all IRT calibrations were based on the English forms only. For some high school grade levels, IRT calibrations used later data extracts that included slightly more student records than the IA data files. This verification of the data preparation was important to ensure that student exclusion rules were applied consistently across the analyses.

During the calibration process, checks were made to ensure that the correct PARSCALE options for the analyses were selected. Checks were also made on the number of items, number of test takers with valid scores, IRT item difficulties, standard errors for the item difficulties, and the match of selected statistics to the same statistics obtained during item analyses. Psychometricians also performed detailed reviews of statistics to investigate the extent to which the model fit the data. Refer to Section 10.4 for IRT model fit evaluation criteria.

Scaling

During the scaling process, checks were made on the number of linking items, the number of items dropped during the stability check of the scaling process, and the scaling constants. For the linking items the reviewers verified the number of linking items, and linking items that did not meet the anchor criteria were dropped. For example, C-DIF items flagged in the mode comparability study were dropped.

Additionally, items with large weighted root mean square difference (WRMSD) values in Round 1 of scaling were excluded as linking items in Round 2. For the scaling constants the reviewers computed the linking constants and then checked that the linking constants were correctly applied. Refer to Section 10.6 for description of scaling process.

Conversion Tables

Conversion tables must be accurate because they are used to generate reported scores for test takers. Comprehensive records were meticulously maintained on item-level decisions, and thorough checks were made to insure that the correct items were included in the final score. All conversion tables were parallel processed at ETS and completely matched. A reasonableness check was also conducted by psychometricians for each content and grade level to make sure the results were in alignment with observations during the analyses prior to conversion table creation. Refer to Section 12.3 for the procedure to create conversion tables.

Delivering Item Statistics

Item statistics based on classical item analyses and the IRT analyses were obtained during the psychometric analysis process. The statistics were compiled by two data analysts independently to ensure that the correct statistics were delivered to Pearson for the item bank.

13.7.2 Pearson Psychometric Quality Control Process

Pearson Psychometrics performed two roles for the quality control of psychometric processes for the 2014-2015 PARCC administration: data validation and replication of the IRT analyses and conversion files. The Pearson Psychometrics team validated all student data files used in the psychometric operational analyses.

The data validation for the student data files (SDF) included the following steps:

1. Validated variables in the data file for values in acceptable ranges.
2. Validated that the test form ID, unique item numbers (UINs) and item sequence on the data file were consistent with the test form values on the corresponding Test Map
3. Computed the composite raw score, claim raw score, and subclaim raw scores, given the item scores on the student data file.
4. Compared computed raw scores to the raw scores on the student data file.
5. Compared the student item response block (SIRB) to the item scores.
6. Flagged student records with inconsistencies for further investigation.

Data validation of the item response files (IRF) included the following steps:

1. Validated variables in the data file for values being in acceptable ranges.
2. Validated the test form ID, item UINs and item sequence on the data file were consistent with the test form values on the corresponding Test Map
3. Compared the maximum score possible for an item to the item scores.

4. Compared the item response to the key file (when available) and computed an expected score. Due to the complexity of the scoring logic for items with multiple parts and/or multiple cardinality, the key was not available for such items for this check.
5. Compared the item score to the expected score.
6. Flagged item records with inconsistencies for further investigation.

For early data extracts, Pearson psychometrics generated classical item statistics and response score distributions. Items were flagged if they did not meet the statistical criteria and provided to Pearson content specialists for review. The statistical criteria included:

- mean item score (i.e., p value) ≥ 0.95 or ≤ 0.10 ;
- item-to-total score correlation < 0.05 ;
- missing score categories; or
- unexpected (i.e., out of range) score(s).

In addition to conducting data validation, Pearson Psychometrics was the replicator for the IRT analyses³² and the conversion file creations. The following steps outline the replication process:

1. Calibrated online and paper data separately.
2. Scaled the paper item parameter estimates to the online scale.
3. Sent data files with the item parameter estimates to Measured Progress for comparison.
4. Reconciled differences in results with ETS and Measured Progress.
5. Generated the raw score to theta tables for the Performance Level Setting (PLS) forms.
6. Sent data files to Measured Progress for comparison and reconcile differences, if any.
7. Generated the scale score scaling constants based on the approved PLS cuts.
8. Compared scale score scaling constants to the ETS scale score scaling constants.
9. Generated the performance levels, summative, claim, and subclaim conversion tables.
10. Sent conversion tables to Measured Progress for comparison and reconcile differences, if any.

13.7.3 Measured Progress Psychometric Quality Control Process

Measured Progress (MP) served as the external evaluator for the 2014-15 PARCC operational administration. MP's main task was to compare the IRT calibration results and conversion tables created by ETS and Pearson.

IRT Calibrations Comparison

MP reviewed and compared the psychometric IRT calibrations performed primarily by ETS using PARSCALE, and replicated by Pearson using IRTPRO, for all grade levels in ELA/L and mathematics administered both online and on paper.

ETS and Pearson each provided the comparison files containing IRT item parameter estimates. ETS also provided the classic item statistics (item-level mean score, item-total correlation) of the items. For each

³² Pearson used the IRTPro (Cai, Thissen & du Toit, 2011) software for IRT analysis.

test, the comparability of IRT parameter estimates between ETS and Pearson were evaluated in the following aspects:

- Number of items and types of interventions in the IRT calibration process
- Descriptive statistics of the IRT a -, b -, and d -parameter estimates
- Scatterplot of IRT a -, b -, and d -parameter estimates
- Absolute differences in IRT a -, b -, and d -parameter estimates
- Mean absolute difference (MAD) and root mean square difference (RMSD) in IRT-model-predicted item mean scores

IRT Comparison Results

In grades 3-8 ELA/L and mathematics, MP compared the items excluded from the IRT calibration by ETS and Pearson. For polytomously-scored items, MP compared the interventions imposed by ETS and Pearson in collapsing the response categories in the IRT calibration. Overall, MP observed very comparable IRT item parameter estimates between ETS and Pearson for the grades 3 through 8 tests. The only noticeable differences were two items (one in ELA/L grade 4 and one in Mathematics grade 8) excluded from IRT calibration by Pearson but included by ETS, and one very difficult item in grade 8 mathematics that had an unusually large b -parameter estimate. After further investigation, a decision was made to move forward using the ETS parameter estimates for these three items. MP did observe some nontrivial discrepancies in the IRT parameter estimates between ETS and Pearson for a few items in the high school tests. These included Algebra I, Algebra II, and Integrated Mathematics III, and for one item each in ELA/L grade 11, Integrated Mathematics I, and Geometry. A few possible reasons may have contributed to these differences: the weighted least square (WLS) 3PL item estimates for a few poor-fitting items from ETS vs. the 2PL item estimates from Pearson; and the more extreme b -parameter estimates provide by IRTPRO for difficult items. MP also observed that a few items had unusually large d -parameter estimates in Algebra II, and Integrated Mathematics I and III. The decision was to use ETS estimates for producing final scores.

Conversion Files Comparison

Measured Progress provided comparison results for raw-to-theta (RST) lookup tables, summative and claim scale scores, summative performance levels, and subclaim performance levels.

Raw-to-Theta Lookup Tables

Raw-to-Theta lookup tables provided by ETS and Pearson were compared for the PLS forms across the grades 3-8 and High School tests administered in the Spring 2015 PARCC assessment.

Summative and Claim Scale Score Conversation Tables

For summative and claim scale score conversation tables, MP compared the identifiers of forms/form-combinations provided by each contractor. For each form/form-combination matched on both their identifiers and raw score points, MP compared the theta and scale scores associated with each raw score point.

Summative Performance Level

For the summative performance level files, MP compared and found identical lower and upper limits of scale scores across the five performance levels for each test for both grades 3-8 and high school.

Subclaim Performance Level

For each test and by each subclaim area, MP compared the identifiers of forms/form-combinations provided by each contractor, and their maximum raw score points.

For the forms/form-combinations that matched on both identifiers and maximum raw score points, MP further compared their lower and upper limits of raw scores for each subclaim performance level, and computed the mean absolute difference (MAD) of the two unrounded category cut scores.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Beimers, J. N., Way, W. D., McClarty, K. L., & Miles, J. A. (2012, January). Evidence based standard setting: Establishing cut scores by integrating research evidence with expert content judgments. Austin, TX: Pearson. Retrieved from http://researchnetwork.pearson.com/wp-content/uploads/Bulletin21_Evidence_Based_Standard_Setting.pdf
- Cai, L., Thissen, D., & du Toit, S. (2011). IRTPRO 2.1 [Computer software]. Lincolnwood, IL: Scientific Software International.
- Council of Chief State School Officers. (2015). Common Core State Standards Initiative. Retrieved from <http://www.corestandards.org/>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Doorey, N. & Polikoff, M. (2016). *Evaluating the Content and Quality of Next Generation Assessments*. Washington, DC: Thomas B. Fordham Institute.
- Dorans, N. J., & Schmitt, A. P. (1991). *Constructed response and differential item functioning: A pragmatic approach. (ETS Research Report No. 91-47)*. Princeton, NJ: Educational Testing Service.
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35-66). Hillsdale, NJ: Erlbaum.
- Dorans, N. J. (2013). *ETS Contributions to the Quantitative Assessment of Item, Test and Score Fairness (ETS R&D Science and Policy Contributions Series, ETS SPC-13-04)*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service (2009). *Parplot: A Computer Program for Producing IRT Residual Plots*. Princeton, NJ: ETS. Version 2.6.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performances and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Erlbaum.
- Holland, P. W. (1990). On the sampling theory foundations of item response theory models. *Psychometrika*, 55, 577-601.

- Kim, S., & Kolen M. J. (2004). *STUIRT: A Computer Program for Scale Transformation under Unidimensional Item Response Theory Models*. University of Iowa. Version 1.0.
- Kolen, M.J., & Brennan, R.L. (2004). *Test equating, scaling, and linking: Methods and practices*. New York: Springer.
- Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, *58*, 690-700.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, *22*, 719-748.
- McClary, K. L., Kobrin, J. L., Moyer, E, Griffin, S., Huth, K., Carey, S., & Medberry, S. (2015). *PARCC Benchmarking Study*. Pearson Educational Measurement, Iowa City, IA: Pearson.
- McClarty, K. L., Way, W. D., Porter, A. C., Beimers, J. N., & Miles, J. A. (2013). Evidence-based standard setting: establishing a validity framework for cut scores. *Educational Researcher*, *42*(2), 78-88.
- Muraki, E., & Bock R. D. (2003). *PARSCALE: IRT Item Analysis and Test Scoring for Rating-Scale Data*. Chicago, IL: Scientific Software International. Version 4.1.
- Liu, J., Brown, T., Chen, J., Ali, U., Hou, L., & Costanzo, K. (2016). *Mode Comparability Study Based on Spring 2015 Operational Test Data*. Princeton, NJ: Educational Testing Service.
- Olsson, U., Drasgow, F., & Dorans, N. J. (1982), "The Polyserial Correlation Coefficient," *Biometrika*, *47*, 337-347.
- Plake, B. S., Ferdous, A. A., Impara, J. C., & Buckendahl, C. W. (2005). Setting Multiple Performance Standards Using the Yes/No Method: An Alternative Item Mapping Method. *Meeting of the National Council on Measurement in Education*. Montreal. Canada.
- Schultz, S. R., Michaels, H. R., Dvorak, R. N., & Wiley, C. R. H. (2016). *Evaluating the Content and Quality of Next Generation High School Assessments*. (HumRRO Report 2016 No. 001). Alexandria, VA: Human Resources Research Organization.
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, *28*, 237-247.
- Steedle, J., McBride, M. Johnson, M., & Keng, L. (2015). *Spring 2015 Digital Devices Comparability Research Study*. Pearson Educational Measurement, Iowa City, IA: Pearson.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*, 201-210.
- Thacker, A., Sinclair, A., Wise, L., & Becker, S. (April 10, 2014). *PARCC validity studies including predictive and longitudinal studies memorandum* (HumRRO Report 2014 No. 020). Alexandria, VA: Human Resources Research Organization.

- Wainer, H., & Thissen, D. (2001). *Test scoring*. Mahwah, NJ: Erlbaum.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 8*, 125-145.
- Zenisky, A. L., Hambleton, R. K., & Sireci, S.C. (2003). *Effects of local dependence on the validity of IRT item test, and ability statistics*. (Technical Report). American College Admissions Test.
- Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). *Describing and Categorizing DIF in Polytomous Items (ETS Research Report RR-97-05)*. Princeton, NJ: Educational Testing Service.

Appendices

Appendix 5: Test Takers by Grade and Mode, for Each State

Table A.5.1 ELA/L Test Takers, by State, Grade, and Mode

State	Category	English Language Arts/Literacy									
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
PARCC	N of Students	4,514,914	518,857	632,312	638,239	632,506	624,008	617,564	413,098	267,159	171,171
PARCC	N of CBT	3,637,535	379,768	479,742	496,795	512,391	513,949	505,003	344,594	244,345	160,948
PARCC	% of CBT	80.6	73.2	75.9	77.8	81.0	82.4	81.8	83.4	91.5	94.0
PARCC	N of PBT	869,496	138,309	151,501	140,505	118,840	108,963	111,477	67,600	22,363	9,938
PARCC	% of PBT	19.3	26.7	24.0	22.0	18.8	17.5	18.1	16.4	8.4	5.8
PARCC	N of Mixed Modes	7,883	780	1,069	939	1,275	1,096	1,084	904	451	285
PARCC	% of Mixed Modes	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.2
AR	N of Students	291,307	34,871	34,946	35,052	34,852	35,437	35,342	35,011	33,193	12,603
AR	% of PARCC Data	6.5	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.3
AR	N of CBT	266,381	31,505	31,702	31,791	32,165	32,574	32,474	32,350	29,843	11,977
AR	% of CBT	91.4	90.3	90.7	90.7	92.3	91.9	91.9	92.4	89.9	95.0
AR	N of PBT	23,569	3,222	3,083	3,096	2,489	2,649	2,635	2,516	3,255	624
AR	% of PBT	8.1	9.2	8.8	8.8	7.1	7.5	7.5	7.2	9.8	5.0
AR	N of Mixed Modes	1,357	144	161	165	198	214	233	145	95	2
AR	% of Mixed Modes	0.5	0.4	0.5	0.5	0.6	0.6	0.7	0.4	0.3	0.0
CO	N of Students	474,015	62,712	62,352	61,996	60,953	57,373	54,567	45,938	38,386	29,738
CO	% of PARCC Data	10.5	1.4	1.4	1.4	1.4	1.3	1.2	1.0	0.9	0.7
CO	N of CBT	465,936	55,012	62,284	61,921	60,896	57,324	54,530	45,884	38,370	29,715
CO	% of CBT	98.3	87.7	99.9	99.9	99.9	99.9	99.9	99.9	100	99.9
CO	N of PBT	7,965	7,634	58	64	56	45	32	42	15	19
CO	% of PBT	1.7	12.2	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.1
CO	N of Mixed Modes	114	66	10	11	1	4	5	12	1	4
CO	% of Mixed Modes	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DC	N of Students	34,972	5,616	5,041	4,615	4,376	4,102	4,198	3,312	3,711	1
DC	% of PARCC Data	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0
DC	N of CBT	32,241	5,119	4,730	4,302	3,967	3,879	3,978	2,999	3,266	1

State	Category	English Language Arts/Literacy									
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
DC	% of CBT	92.2	91.2	93.8	93.2	90.7	94.6	94.8	90.5	88.0	100
DC	N of PBT	2,711	492	309	307	407	221	220	313	442	n/a
DC	% of PBT	7.8	8.8	6.1	6.7	9.3	5.4	5.2	9.5	11.9	n/a
DC	N of Mixed Modes	20	5	2	6	2	2	n/a	n/a	3	n/a
DC	% of Mixed Modes	0.1	0.1	0.0	0.1	0.0	0.0	n/a	n/a	0.1	n/a
IL	N of Students	992,432	146,118	144,033	146,382	145,746	142,801	143,568	79,138	5,605	39,041
IL	% of PARCC Data	22.0	3.2	3.2	3.2	3.2	3.2	3.2	1.8	0.1	0.9
IL	N of CBT	731,951	82,164	85,844	98,198	121,293	117,308	121,109	67,799	4,991	33,245
IL	% of CBT	73.8	56.2	59.6	67.1	83.2	82.1	84.4	85.7	89.0	85.2
IL	N of PBT	258,923	63,700	57,957	48,018	23,966	25,286	22,289	11,308	612	5,787
IL	% of PBT	26.1	43.6	40.2	32.8	16.4	17.7	15.5	14.3	10.9	14.8
IL	N of Mixed Modes	1,558	254	232	166	487	207	170	31	2	9
IL	% of Mixed Modes	0.2	0.2	0.2	0.1	0.3	0.1	0.1	0.0	0.0	0.0
MA	N of Students	240,440	37,129	36,650	37,312	37,621	37,651	37,660	9,377	n/a	7,040
MA	% of PARCC Data	5.3	0.8	0.8	0.8	0.8	0.8	0.8	0.2	n/a	0.2
MA	N of CBT	129,779	19,455	19,257	19,450	20,141	20,760	20,003	6,161	n/a	4,552
MA	% of CBT	54.0	52.4	52.5	52.1	53.5	55.1	53.1	65.7	n/a	64.7
MA	N of PBT	110,491	17,648	17,361	17,830	17,455	16,864	17,629	3,216	n/a	2,488
MA	% of PBT	46.0	47.5	47.4	47.8	46.4	44.8	46.8	34.3	n/a	35.3
MA	N of Mixed Modes	170	26	32	32	25	27	28	n/a	n/a	n/a
MA	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.1	n/a	n/a	n/a
MD	N of Students	427,818	65,198	63,935	63,463	62,181	61,351	59,494	14	52,167	15
MD	% of PARCC Data	9.5	1.4	1.4	1.4	1.4	1.4	1.3	0.0	1.2	0.0
MD	N of CBT	337,468	41,539	55,495	55,273	44,541	51,976	45,338	14	43,277	15
MD	% of CBT	78.9	63.7	86.8	87.1	71.6	84.7	76.2	100	83.0	100
MD	N of PBT	89,820	23,568	8,370	8,137	17,545	9,299	14,040	n/a	8,861	n/a
MD	% of PBT	21.0	36.1	13.1	12.8	28.2	15.2	23.6	n/a	17.0	n/a
MD	N of Mixed Modes	530	91	70	53	95	76	116	n/a	29	n/a
MD	% of Mixed Modes	0.1	0.1	0.1	0.1	0.2	0.1	0.2	n/a	0.1	n/a
MS	N of Students	246,194	37,695	36,352	36,239	35,362	36,009	35,896	n/a	28,641	n/a

English Language Arts/Literacy											
State	Category	Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
MS	% of PARCC Data	5.5	0.8	0.8	0.8	0.8	0.8	0.8	n/a	0.6	n/a
MS	N of CBT	202,919	28,720	28,512	30,157	30,170	31,445	31,173	n/a	22,742	n/a
MS	% of CBT	82.4	76.2	78.4	83.2	85.3	87.3	86.8	n/a	79.4	n/a
MS	N of PBT	43,074	8,928	7,804	6,046	5,167	4,533	4,708	n/a	5,888	n/a
MS	% of PBT	17.5	23.7	21.5	16.7	14.6	12.6	13.1	n/a	20.6	n/a
MS	N of Mixed Modes	201	47	36	36	25	31	15	n/a	11	n/a
MS	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.0	n/a	0.0	n/a
NJ	N of Students	773,710	95,276	93,852	94,655	92,635	90,331	88,494	82,467	73,235	62,765
NJ	% of PARCC Data	17.1	2.1	2.1	2.1	2.1	2.0	2.0	1.8	1.6	1.4
NJ	N of CBT	768,968	94,697	93,368	94,196	92,340	90,044	88,161	81,629	72,441	62,092
NJ	% of CBT	99.4	99.4	99.5	99.5	99.7	99.7	99.6	99.0	98.9	98.9
NJ	N of PBT	3,450	474	413	379	236	246	288	513	495	406
NJ	% of PBT	0.4	0.5	0.4	0.4	0.3	0.3	0.3	0.6	0.7	0.6
NJ	N of Mixed Modes	1,292	105	71	80	59	41	45	325	299	267
NJ	% of Mixed Modes	0.2	0.1	0.1	0.1	0.1	0.0	0.1	0.4	0.4	0.4
NM	N of Students	207,100	23,514	23,944	23,844	23,770	23,539	22,949	23,365	22,222	19,953
NM	% of PARCC Data	4.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.4
NM	N of CBT	189,485	14,050	21,511	22,687	22,471	22,415	21,969	23,141	21,904	19,337
NM	% of CBT	91.5	59.8	89.8	95.1	94.5	95.2	95.7	99.0	98.6	96.9
NM	N of PBT	17,529	9,436	2,418	1,152	1,296	1,110	968	222	314	613
NM	% of PBT	8.5	40.1	10.1	4.8	5.5	4.7	4.2	1.0	1.4	3.1
NM	N of Mixed Modes	86	28	15	5	3	14	12	2	4	3
NM	% of Mixed Modes	0.0	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0
NY	N of Students	1,111	394	n/a	148	495	74	n/a	n/a	n/a	n/a
NY	% of PARCC Data	0.0	0.0	n/a	0.0	0.0	0.0	n/a	n/a	n/a	n/a
NY	N of CBT	1,111	394	n/a	148	495	74	n/a	n/a	n/a	n/a
NY	% of CBT	100	100	n/a	100	100	100	n/a	n/a	n/a	n/a
NY	N of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	% of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	N of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

State	Category	English Language Arts/Literacy										
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	
NY	% of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
OH	N of Students	750,022	76	121,068	124,208	124,644	125,582	125,815	125,956	2,661	12	
OH	% of PARCC Data	16.6	0.0	2.7	2.8	2.8	2.8	2.8	2.8	0.1	0.0	
OH	N of CBT	452,546	28	68,678	70,195	75,651	79,441	78,854	77,857	1,830	12	
OH	% of CBT	60.3	36.8	56.7	56.5	60.7	63.3	62.7	61.8	68.8	100	
OH	N of PBT	295,003	48	51,962	53,632	48,622	45,679	46,515	47,714	831	n/a	
OH	% of PBT	39.3	63.2	42.9	43.2	39.0	36.4	37.0	37.9	31.2	n/a	
OH	N of Mixed Modes	2,473	n/a	428	381	371	462	446	385	n/a	n/a	
OH	% of Mixed Modes	0.3	n/a	0.4	0.3	0.3	0.4	0.4	0.3	n/a	n/a	
RI	N of Students	75,793	10,258	10,139	10,325	9,871	9,758	9,581	8,520	7,338	3	
RI	% of PARCC Data	1.7	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.0	
RI	N of CBT	58,750	7,085	8,361	8,477	8,261	6,709	7,414	6,760	5,681	2	
RI	% of CBT	77.5	69.1	82.5	82.1	83.7	68.8	77.4	79.3	77.4	66.7	
RI	N of PBT	16,961	3,159	1,766	1,844	1,601	3,031	2,153	1,756	1,650	1	
RI	% of PBT	22.4	30.8	17.4	17.9	16.2	31.1	22.5	20.6	22.5	33.3	
RI	N of Mixed Modes	82	14	12	4	9	18	14	4	7	n/a	
RI	% of Mixed Modes	0.1	0.1	0.1	0.0	0.1	0.2	0.1	0.0	0.1	n/a	

Note: CBT = computer-based test; PBT = paper-based test; Mixed Modes = one test component (i.e., PBA or EOY) as CBT and the other component as PBT. n/a = not applicable.

Table A.5.2 Mathematics Test Takers, by State, Grade, and Mode

State	Category	Total	Mathematics											
			Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
PARCC	N of Students	4,594,725	650,262	635,075	640,102	632,127	608,990	504,561	480,604	205,219	186,890	30,217	12,282	8,396
PARCC	N of CBT	3,662,173	452,907	475,918	492,387	513,755	504,839	413,073	405,661	184,494	173,572	26,953	11,374	7,240
PARCC	% of CBT	79.7	69.6	74.9	76.9	81.3	82.9	81.9	84.4	89.9	92.9	89.2	92.6	86.2
PARCC	N of PBT	924,396	196,140	158,116	146,743	117,147	103,054	90,517	73,955	20,310	13,099	3,251	908	1,156
PARCC	% of PBT	20.1	30.2	24.9	22.9	18.5	16.9	17.9	15.4	9.9	7.0	10.8	7.4	13.8
PARCC	N of Mixed Modes	8,156	1,215	1,041	972	1,225	1,097	971	988	415	219	13	n/a	n/a
PARCC	% of Mixed Modes	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.0	n/a	n/a
AR	N of Students	280,856	34,931	34,987	35,104	34,876	35,250	28,978	34,181	32,110	10,439	n/a	n/a	n/a
AR	% of PARCC Data	6.1	0.8	0.8	0.8	0.8	0.8	0.6	0.7	0.7	0.2	n/a	n/a	n/a
AR	N of CBT	256,661	31,565	31,737	31,833	32,181	32,433	26,558	31,612	28,885	9,857	n/a	n/a	n/a
AR	% of CBT	91.4	90.4	90.7	90.7	92.3	92.0	91.6	92.5	90.0	94.4	n/a	n/a	n/a
AR	N of PBT	22,871	3,223	3,091	3,107	2,506	2,611	2,193	2,454	3,126	560	n/a	n/a	n/a
AR	% of PBT	8.1	9.2	8.8	8.9	7.2	7.4	7.6	7.2	9.7	5.4	n/a	n/a	n/a
AR	N of Mixed Modes	1,324	143	159	164	189	206	227	115	99	22	n/a	n/a	n/a
AR	% of Mixed Modes	0.5	0.4	0.5	0.5	0.5	0.6	0.8	0.3	0.3	0.2	n/a	n/a	n/a
CO	N of Students	465,421	63,799	62,352	61,930	60,895	55,404	40,649	41,810	29,614	23,462	11,000	8,635	5,871
CO	% of PARCC Data	10.1	1.4	1.4	1.3	1.3	1.2	0.9	0.9	0.6	0.5	0.2	0.2	0.1
CO	N of CBT	418,761	55,633	54,890	54,678	55,309	50,478	37,399	37,661	26,607	20,917	10,836	8,516	5,837
CO	% of CBT	90.0	87.2	88.0	88.3	90.8	91.1	92.0	90.1	89.8	89.2	98.5	98.6	99.4
CO	N of PBT	46,176	8,081	7,397	7,197	5,552	4,896	3,209	4,062	2,964	2,503	162	119	34
CO	% of PBT	9.9	12.7	11.9	11.6	9.1	8.8	7.9	9.7	10.0	10.7	1.5	1.4	0.6
CO	N of Mixed Modes	484	85	65	55	34	30	41	87	43	42	2	n/a	n/a
CO	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.0	n/a	n/a
DC	N of Students	34,462	5,654	5,088	4,661	4,402	3,899	3,248	3,780	3,569	150	n/a	11	n/a
DC	% of PARCC Data	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	n/a	0.0	n/a
DC	N of CBT	31,752	5,156	4,778	4,340	3,993	3,679	3,029	3,481	3,146	150	n/a	n/a	n/a
DC	% of CBT	92.1	91.2	93.9	93.1	90.7	94.4	93.3	92.1	88.1	100	n/a	n/a	n/a
DC	N of PBT	2,682	493	309	313	406	215	219	296	420	n/a	n/a	11	n/a
DC	% of PBT	7.8	8.7	6.1	6.7	9.2	5.5	6.7	7.8	11.8	n/a	n/a	100	n/a

State	Category	Total	Mathematics											
			Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
DC	N of Mixed Modes	28	5	1	8	3	5	n/a	3	3	n/a	n/a	n/a	n/a
DC	% of Mixed Modes	0.1	0.1	0	0.2	0.1	0.1	n/a	0.1	0.1	n/a	n/a	n/a	n/a
IL	N of Students	973,452	146,422	144,580	146,873	146,007	142,998	143,473	54,488	4,891	33,586	7,691	660	1,783
IL	% of PARCC Data	21.2	3.2	3.1	3.2	3.2	3.1	3.1	1.2	0.1	0.7	0.2	0.0	0.0
IL	N of CBT	714,246	82,293	86,083	98,413	121,625	117,496	121,044	44,963	4,362	28,846	7,654	593	874
IL	% of CBT	73.4	56.2	59.5	67.0	83.3	82.2	84.4	82.5	89.2	85.9	99.5	89.8	49.0
IL	N of PBT	257,725	63,873	58,269	48,292	23,914	25,311	22,277	9,510	529	4,738	36	67	909
IL	% of PBT	26.5	43.6	40.3	32.9	16.4	17.7	15.5	17.5	10.8	14.1	0.5	10.2	51.0
IL	N of Mixed Modes	1,481	256	228	168	468	191	152	15	n/a	2	1	n/a	n/a
IL	% of Mixed Modes	0.2	0.2	0.2	0.1	0.3	0.1	0.1	0.0	n/a	0.0	0.0	n/a	n/a
MA	N of Students	239,598	37,376	36,885	37,484	37,742	37,545	33,620	11,391	1,945	4,731	593	27	259
MA	% of PARCC Data	5.2	0.8	0.8	0.8	0.8	0.8	0.7	0.2	0.0	0.1	0.0	0.0	0.0
MA	N of CBT	129,240	19,578	19,387	19,527	20,185	20,819	17,153	7,599	1,090	3,464	359	25	54
MA	% of CBT	53.9	52.4	52.6	52.1	53.5	55.5	51.0	66.7	56.0	73.2	60.5	92.6	20.8
MA	N of PBT	110,190	17,770	17,471	17,933	17,531	16,699	16,438	3,791	855	1,261	234	2	205
MA	% of PBT	46.0	47.5	47.4	47.8	46.4	44.5	48.9	33.3	44.0	26.7	39.5	7.4	79.2
MA	N of Mixed Modes	168	28	27	24	26	27	29	1	n/a	6	n/a	n/a	n/a
MA	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	n/a	0.1	n/a	n/a	n/a
MD	N of Students	448,793	65,596	64,290	63,828	62,195	55,012	41,167	58,858	6	37,841	n/a	n/a	n/a
MD	% of PARCC Data	9.8	1.4	1.4	1.4	1.4	1.2	0.9	1.3	0.0	0.8	n/a	n/a	n/a
MD	N of CBT	376,121	43,057	56,608	56,385	48,222	50,718	34,970	51,252	6	34,903	n/a	n/a	n/a
MD	% of CBT	83.8	65.6	88.1	88.3	77.5	92.2	84.9	87.1	100	92.2	n/a	n/a	n/a
MD	N of PBT	72,252	22,447	7,622	7,392	13,909	4,256	6,138	7,560	n/a	2,928	n/a	n/a	n/a
MD	% of PBT	16.1	34.2	11.9	11.6	22.4	7.7	14.9	12.8	n/a	7.7	n/a	n/a	n/a
MD	N of Mixed Modes	420	92	60	51	64	38	59	46	n/a	10	n/a	n/a	n/a
MD	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	n/a	0	n/a	n/a	n/a
MS	N of Students	255,857	37,730	36,376	36,225	35,347	35,975	35,856	38,348	n/a	n/a	n/a	n/a	n/a
MS	% of PARCC Data	5.6	0.8	0.8	0.8	0.8	0.8	0.8	0.8	n/a	n/a	n/a	n/a	n/a
MS	N of CBT	210,460	28,687	28,514	30,155	30,151	31,406	31,162	30,385	n/a	n/a	n/a	n/a	n/a
MS	N of CBT	210,460	28,687	28,514	30,155	30,151	31,406	31,162	30,385	n/a	n/a	n/a	n/a	n/a

State	Category	Total	Mathematics											
			Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
MS	% of CBT	82.3	76.0	78.4	83.2	85.3	87.3	86.9	79.2	n/a	n/a	n/a	n/a	n/a
MS	N of PBT	45,203	9,002	7,829	6,031	5,172	4,544	4,677	7,948	n/a	n/a	n/a	n/a	n/a
MS	% of PBT	17.7	23.9	21.5	16.6	14.6	12.6	13.0	20.7	n/a	n/a	n/a	n/a	n/a
MS	N of Mixed Modes	194	41	33	39	24	25	17	15	n/a	n/a	n/a	n/a	n/a
MS	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	n/a	n/a	n/a	n/a	n/a
NJ	N of Students	745,646	95,988	94,572	95,354	93,201	87,430	58,158	91,740	71,137	58,026	29	8	3
NJ	% of PARCC Data	16.2	2.1	2.1	2.1	2.0	1.9	1.3	2.0	1.5	1.3	0.0	0.0	0.0
NJ	N of CBT	741,041	95,356	94,051	94,857	92,860	87,094	57,808	90,877	70,506	57,596	26	8	2
NJ	% of CBT	99.4	99.3	99.4	99.5	99.6	99.6	99.4	99.1	99.1	99.3	89.7	100	66.7
NJ	N of PBT	3,530	519	443	411	284	285	312	579	392	301	3	n/a	1
NJ	% of PBT	0.5	0.5	0.5	0.4	0.3	0.3	0.5	0.6	0.6	0.5	10.3	n/a	33.3
NJ	N of Mixed Modes	1,075	113	78	86	57	51	38	284	239	129	n/a	n/a	n/a
NJ	% of Mixed Modes	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.3	0.2	n/a	n/a	n/a
NM	N of Students	206,645	24,974	24,717	24,430	23,911	23,553	19,061	24,435	20,888	18,575	691	930	480
NM	% of PARCC Data	4.5	0.5	0.5	0.5	0.5	0.5	0.4	0.5	0.5	0.4	0.0	0.0	0.0
NM	N of CBT	188,106	15,887	22,325	23,245	22,595	22,424	18,106	23,842	20,149	17,760	638	662	473
NM	% of CBT	91.0	63.6	90.3	95.1	94.5	95.2	95.0	97.6	96.5	95.6	92.3	71.2	98.5
NM	N of PBT	18,443	9,066	2,375	1,180	1,312	1,114	941	584	736	807	53	268	7
NM	% of PBT	8.9	36.3	9.6	4.8	5.5	4.7	4.9	2.4	3.5	4.3	7.7	28.8	1.5
NM	N of Mixed Modes	96	21	17	5	4	15	14	9	3	8	n/a	n/a	n/a
NM	% of Mixed Modes	0.0	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	n/a	n/a	n/a
NY	N of Students	826	419	n/a	189	104	114	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	% of PARCC Data	0.0	0.0	n/a	0.0	0.0	0.0	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	N of CBT	826	419	n/a	189	104	114	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	% of CBT	100	100	n/a	100	100	100	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	N of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	% of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	N of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NY	% of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
OH	N of Students	867,817	126,997	120,957	123,576	123,462	121,914	93,151	112,183	33,282	77	10,207	2,011	n/a

State	Category	Total	Mathematics											
			Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
OH	% of PARCC Data	18.9	2.8	2.6	2.7	2.7	2.7	2.0	2.4	0.7	0.0	0.2	0.0	n/a
OH	N of CBT	535,250	68,109	69,079	70,195	78,181	81,389	60,474	75,949	22,794	76	7,434	1,570	n/a
OH	% of CBT	61.7	53.6	57.1	56.8	63.3	66.8	64.9	67.7	68.5	98.7	72.8	78.1	n/a
OH	N of PBT	329,760	58,470	51,515	53,013	44,935	40,033	32,297	35,826	10,466	1	2,763	441	n/a
OH	% of PBT	38.0	46.0	42.6	42.9	36.4	32.8	34.7	31.9	31.4	1.3	27.1	21.9	n/a
OH	N of Mixed Modes	2,807	418	363	368	346	492	380	408	22	n/a	10	n/a	n/a
OH	% of Mixed Modes	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.1	n/a	0.1	n/a	n/a
RI	N of Students	75,352	10,376	10,271	10,448	9,985	9,896	7,200	9,390	7,777	3	6	n/a	n/a
RI	% of PARCC Data	1.6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.0	0.0	n/a	n/a
RI	N of CBT	59,709	7,167	8,466	8,570	8,349	6,789	5,370	8,040	6,949	3	6	n/a	n/a
RI	% of CBT	79.2	69.1	82.4	82.0	83.6	68.6	74.6	85.6	89.4	100	100	n/a	n/a
RI	N of PBT	15,564	3,196	1,795	1,874	1,626	3,090	1,816	1,345	822	n/a	n/a	n/a	n/a
RI	% of PBT	20.7	30.8	17.5	17.9	16.3	31.2	25.2	14.3	10.6	n/a	n/a	n/a	n/a
RI	N of Mixed Modes	79	13	10	4	10	17	14	5	6	n/a	n/a	n/a	n/a
RI	% of Mixed Modes	0.1	0.1	0.1	0.0	0.1	0.2	0.2	0.1	0.1	n/a	n/a	n/a	n/a

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. CBT = computer-based test; PBT = paper-based test; Mixed Modes = one test component (i.e., PBA or EOY) as CBT and the other component as PBT. n/a = not applicable.

Table A.5.3 Spanish-Language Mathematics Test Takers, by State, Grade, and Mode

State*	Category	Mathematics (Spanish-Language forms)												
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
PARCC	N of Students	20,077	650,262	635,075	640,102	632,127	608,990	504,561	480,604	205,219	186,890	30,217	12,282	8,396
PARCC	N of CBT	16,854	452,907	475,918	492,387	513,755	504,839	413,073	405,661	184,494	173,572	26,953	11,374	7,240
PARCC	% of CBT	83.9	69.6	74.9	76.9	81.3	82.9	81.9	84.4	89.9	92.9	89.2	92.6	86.2
PARCC	N of PBT	3,164	196,140	158,116	146,743	117,147	103,054	90,517	73,955	20,310	13,099	3,251	908	1,156
PARCC	% of PBT	15.8	30.2	24.9	22.9	18.5	16.9	17.9	15.4	9.9	7.0	10.8	7.4	13.8
PARCC	N of Mixed Modes	59	1,215	1,041	972	1,225	1,097	971	988	415	219	13	n/a	n/a
PARCC	% of Mixed Modes	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.0	n/a	n/a
AR	N of Students	28	n/a	5	4	5	2	2	2	8	n/a	n/a	n/a	n/a
AR	% of PARCC Data	0.1	n/a	0.0	0.0	0.0	0.0	0.0	0.0	0.0	n/a	n/a	n/a	n/a
AR	N of CBT	28	n/a	5	4	5	2	2	2	8	n/a	n/a	n/a	n/a
AR	% of CBT	100	n/a	100	100	100	100	100	100	100	n/a	n/a	n/a	n/a
AR	N of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
AR	% of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
AR	N of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
AR	% of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
CO	N of Students	2,416	965	241	198	197	194	171	173	95	28	108	36	10
CO	% of PARCC Data	12.0	4.8	1.2	1.0	1.0	1.0	0.9	0.9	0.5	0.1	0.5	0.2	0.0
CO	N of CBT	2,229	894	195	168	183	176	163	173	95	28	108	36	10
CO	% of CBT	92.3	92.6	80.9	84.8	92.9	90.7	95.3	100	100	100	100	100	100
CO	N of PBT	186	70	46	30	14	18	8	n/a	n/a	n/a	n/a	n/a	n/a
CO	% of PBT	7.7	7.3	19.1	15.2	7.1	9.3	4.7	n/a	n/a	n/a	n/a	n/a	n/a
CO	N of Mixed Modes	1	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
CO	% of Mixed Modes	0.0	0.1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
DC	N of Students	204	15	17	22	26	21	13	85	5	n/a	n/a	n/a	n/a
DC	% of PARCC Data	1.0	0.1	0.1	0.1	0.1	0.1	0.1	0.4	0.0	n/a	n/a	n/a	n/a
DC	N of CBT	189	15	17	22	25	18	13	79	n/a	n/a	n/a	n/a	n/a
DC	% of CBT	92.6	100	100	100	96.2	85.7	100	92.9	n/a	n/a	n/a	n/a	n/a
DC	N of PBT	8	n/a	n/a	n/a	n/a	n/a	n/a	3	5	n/a	n/a	n/a	n/a

State*	Category	Mathematics (Spanish-Language forms)												
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
DC	% of PBT	3.9	n/a	n/a	n/a	n/a	n/a	n/a	3.5	100	n/a	n/a	n/a	n/a
DC	N of Mixed Modes	7	n/a	n/a	n/a	1	3	n/a	3	n/a	n/a	n/a	n/a	n/a
DC	% of Mixed Modes	3.4	n/a	n/a	n/a	3.8	14.3	n/a	3.5	n/a	n/a	n/a	n/a	n/a
IL	N of Students	4,494	1,878	875	541	323	347	363	104	n/a	18	45	n/a	n/a
IL	% of PARCC Data	22.4	9.4	4.4	2.7	1.6	1.7	1.8	0.5	n/a	0.1	0.2	n/a	n/a
IL	N of CBT	2,678	850	422	358	293	276	320	98	n/a	16	45	n/a	n/a
IL	% of CBT	59.6	45.3	48.2	66.2	90.7	79.5	88.2	94.2	n/a	88.9	100	n/a	n/a
IL	N of PBT	1,800	1,026	452	182	26	66	40	6	n/a	2	n/a	n/a	n/a
IL	% of PBT	40.1	54.6	51.7	33.6	8.0	19.0	11.0	5.8	n/a	11.1	n/a	n/a	n/a
IL	N of Mixed Modes	16	2	1	1	4	5	3	n/a	n/a	n/a	n/a	n/a	n/a
IL	% of Mixed Modes	0.4	0.1	0.1	0.2	1.2	1.4	0.8	n/a	n/a	n/a	n/a	n/a	n/a
MA	N of Students	84	15	19	15	5	4	9	10	1	6	n/a	n/a	n/a
MA	% of PARCC Data	0.4	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	n/a	n/a	n/a
MA	N of CBT	81	15	19	15	4	4	9	8	1	6	n/a	n/a	n/a
MA	% of CBT	96.4	100	100	100	80	100	100	80.0	100	100	n/a	n/a	n/a
MA	N of PBT	3	n/a	n/a	n/a	1	n/a	n/a	2	n/a	n/a	n/a	n/a	n/a
MA	% of PBT	3.6	n/a	n/a	n/a	20.0	n/a	n/a	20.0	n/a	n/a	n/a	n/a	n/a
MA	N of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MA	% of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MS	N of Students	60	9	8	5	10	4	9	15	n/a	n/a	n/a	n/a	n/a
MS	% of PARCC Data	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.1	n/a	n/a	n/a	n/a	n/a
MS	N of CBT	60	9	8	5	10	4	9	15	n/a	n/a	n/a	n/a	n/a
MS	% of CBT	100	100	100	100	100	100	100	100	n/a	n/a	n/a	n/a	n/a
MS	N of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MS	% of PBT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MS	N of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MS	% of Mixed Modes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NJ	N of Students	7,515	966	769	745	819	1,038	1,019	1,267	519	373	n/a	n/a	n/a
NJ	% of PARCC Data	37.4	4.8	3.8	3.7	4.1	5.2	5.1	6.3	2.6	1.9	n/a	n/a	n/a

State*	Category	Mathematics (Spanish-Language forms)												
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
NJ	N of CBT	7,236	903	734	704	784	994	990	1,248	511	368	n/a	n/a	n/a
NJ	% of CBT	96.3	93.5	95.4	94.5	95.7	95.8	97.2	98.5	98.5	98.7	n/a	n/a	n/a
NJ	N of PBT	257	58	30	34	35	41	29	17	8	5	n/a	n/a	n/a
NJ	% of PBT	3.4	6.0	3.9	4.6	4.3	3.9	2.8	1.3	1.5	1.3	n/a	n/a	n/a
NJ	N of Mixed Modes	22	5	5	7	n/a	3	n/a	2	n/a	n/a	n/a	n/a	n/a
NJ	% of Mixed Modes	0.3	0.5	0.7	0.9	n/a	0.3	n/a	0.2	n/a	n/a	n/a	n/a	n/a
NM	N of Students	3,672	1,311	775	532	175	225	219	204	141	81	3	5	1
NM	% of PARCC Data	18.3	6.5	3.9	2.6	0.9	1.1	1.1	1.0	0.7	0.4	0.0	0.0	0.0
NM	N of CBT	3,243	958	750	511	171	221	205	201	137	80	3	5	1
NM	% of CBT	88.3	73.1	96.8	96.1	97.7	98.2	93.6	98.5	97.2	98.8	100	100	100
NM	N of PBT	423	353	23	20	3	4	12	3	4	1	n/a	n/a	n/a
NM	% of PBT	11.5	26.9	3.0	3.8	1.7	1.8	5.5	1.5	2.8	1.2	n/a	n/a	n/a
NM	N of Mixed Modes	6	n/a	2	1	1	n/a	2	n/a	n/a	n/a	n/a	n/a	n/a
NM	% of Mixed Modes	0.2	n/a	0.3	0.2	0.6	n/a	0.9	n/a	n/a	n/a	n/a	n/a	n/a
OH	N of Students	757	64	78	93	95	132	95	185	n/a	n/a	15	n/a	n/a
OH	% of PARCC Data	3.8	0.3	0.4	0.5	0.5	0.7	0.5	0.9	n/a	n/a	0.1	n/a	n/a
OH	N of CBT	504	40	41	49	62	114	72	115	n/a	n/a	11	n/a	n/a
OH	% of CBT	66.6	62.5	52.6	52.7	65.3	86.4	75.8	62.2	n/a	n/a	73.3	n/a	n/a
OH	N of PBT	249	24	37	44	33	18	22	67	n/a	n/a	4	n/a	n/a
OH	% of PBT	32.9	37.5	47.4	47.3	34.7	13.6	23.2	36.2	n/a	n/a	26.7	n/a	n/a
OH	N of Mixed Modes	4	n/a	n/a	n/a	n/a	n/a	1	3	n/a	n/a	n/a	n/a	n/a
OH	% of Mixed Modes	0.5	n/a	n/a	n/a	n/a	n/a	1.1	1.6	n/a	n/a	n/a	n/a	n/a
RI	N of Students	846	115	75	90	60	81	96	211	118	n/a	n/a	n/a	n/a
RI	% of PARCC Data	4.2	0.6	0.4	0.4	0.3	0.4	0.5	1.1	0.6	n/a	n/a	n/a	n/a
RI	N of CBT	605	65	60	61	43	38	61	181	96	n/a	n/a	n/a	n/a
RI	% of CBT	71.5	56.5	80.0	67.8	71.7	46.9	63.5	85.8	81.4	n/a	n/a	n/a	n/a
RI	N of PBT	238	50	15	29	17	41	34	30	22	n/a	n/a	n/a	n/a
RI	% of PBT	28.1	43.5	20.0	32.2	28.3	50.6	35.4	14.2	18.6	n/a	n/a	n/a	n/a
RI	N of Mixed Modes	3	n/a	n/a	n/a	n/a	2	1	n/a	n/a	n/a	n/a	n/a	n/a

State*	Category	Mathematics (Spanish-Language forms)												
		Total	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	A1	GO	A2	M1	M2	M3
RI	% of Mixed Modes	0.4	n/a	n/a	n/a	n/a	2.5	1	n/a	n/a	n/a	n/a	n/a	n/a

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. CBT = computer-based test; PBT = paper-based test; Mixed Modes = one test component (i.e., PBA or EOY) as CBT and the other component as PBT. n/a = not applicable.

* No students in MD tested in mathematics using Spanish-language forms; One student in NY (grade 6 CBT) tested in mathematics using a Spanish-language form.

Table A.5.4 All States Combined: ELA/L Test Takers by Grade, Mode, and Gender

Grade	Mode	Valid Cases	Gender			
			Female		Male	
			N	%	N	%
3	All	518,857	254,214	49.0	264,643	51.0
3	CBT	379,768	186,012	49.0	193,756	51.0
3	PBT	138,309	67,845	49.1	70,464	50.9
3	Mixed Modes	780	357	45.8	423	54.2
4	All	632,312	309,674	49.0	322,638	51.0
4	CBT	479,742	234,759	48.9	244,983	51.1
4	PBT	151,501	74,423	49.1	77,078	50.9
4	Mixed Modes	1,069	492	46.0	577	54.0
5	All	638,239	312,329	48.9	325,910	51.1
5	CBT	496,795	242,977	48.9	253,818	51.1
5	PBT	140,505	68,927	49.1	71,578	50.9
5	Mixed Modes	939	425	45.3	514	54.7
6	All	632,506	308,806	48.8	323,700	51.2
6	CBT	512,391	250,399	48.9	261,992	51.1
6	PBT	118,840	57,806	48.6	61,034	51.4
6	Mixed Modes	1,275	601	47.1	674	52.9
7	All	624,008	304,602	48.8	319,406	51.2
7	CBT	513,949	250,921	48.8	263,028	51.2
7	PBT	108,963	53,163	48.8	55,800	51.2
7	Mixed Modes	1,096	518	47.3	578	52.7
8	All	617,564	300,713	48.7	316,851	51.3
8	CBT	505,003	245,752	48.7	259,251	51.3
8	PBT	111,477	54,427	48.8	57,050	51.2
8	Mixed Modes	1,084	534	49.3	550	50.7
9	All	413,098	201,280	48.7	211,818	51.3
9	CBT	344,594	167,964	48.7	176,630	51.3
9	PBT	67,600	32,947	48.7	34,653	51.3
9	Mixed Modes	904	369	40.8	535	59.2
10	All	267,159	130,374	48.8	136,785	51.2
10	CBT	244,345	118,912	48.7	125,433	51.3
10	PBT	22,363	11,290	50.5	11,073	49.5
10	Mixed Modes	451	172	38.1	279	61.9
11	All	171,171	83,114	48.6	88,057	51.4
11	CBT	160,948	78,046	48.5	82,902	51.5
11	PBT	9,938	4,954	49.8	4,984	50.2
11	Mixed Modes	285	114	40.0	171	60.0

Note: CBT = computer-based tests; PBT = paper-based tests; Mixed Modes = one test component (i.e., PBA or EOY) as CBT and the other component as PBT.

Table A.5.5 All States Combined: All Mathematics Test Takers by Grade, Mode, and Gender

Grade	Mode	Valid Cases	Gender			
			Female		Male	
			N	%	N	%
3	All	650,262	49.0	331,745	51.0	
3	CBT	452,907	49.0	230,985	51.0	
3	PBT	196,140	49.0	100,107	51.0	
3	Mixed Modes	1,215	46.3	653	53.7	
4	All	635,075	49.0	324,131	51.0	
4	CBT	475,918	48.9	243,145	51.1	
4	PBT	158,116	49.1	80,431	50.9	
4	Mixed Modes	1,041	46.7	555	53.3	
5	All	640,102	48.9	326,880	51.1	
5	CBT	492,387	48.9	251,672	51.1	
5	PBT	146,743	49.1	74,684	50.9	
5	Mixed Modes	972	46.1	524	53.9	
6	All	632,127	48.8	323,345	51.2	
6	CBT	513,755	48.9	262,549	51.1	
6	PBT	117,147	48.7	60,151	51.3	
6	Mixed Modes	1,225	47.3	645	52.7	
7	All	608,990	48.8	311,598	51.2	
7	CBT	504,839	48.9	258,208	51.1	
7	PBT	103,054	48.7	52,827	51.3	
7	Mixed Modes	1,097	48.7	563	51.3	
8	All	504,561	48.2	261,515	51.8	
8	CBT	413,073	48.1	214,294	51.9	
8	PBT	90,517	48.4	46,729	51.6	
8	Mixed Modes	971	49.3	492	50.7	
A1	All	480,604	48.5	247,274	51.5	
A1	CBT	405,661	48.4	209,131	51.6	
A1	PBT	73,955	49.2	37,574	50.8	
A1	Mixed Modes	988	42.4	569	57.6	
GO	All	205,219	49.2	104,292	50.8	
GO	CBT	184,494	49.0	94,021	51.0	
GO	PBT	20,310	50.6	10,034	49.4	
GO	Mixed Modes	415	42.9	237	57.1	
A2	All	186,890	50.1	93,327	49.9	
A2	CBT	173,572	50.1	86,627	49.9	
A2	PBT	13,099	49.7	6,586	50.3	
A2	Mixed Modes	219	47.9	114	52.1	
M1	All	30,217	48.1	15,671	51.9	
M1	CBT	26,953	48.3	13,938	51.7	
M1	PBT	3,251	47.0	1,722	53.0	
M1	Mixed Modes	13	15.4	11	84.6	
M2	All	12,282	48.9	6,278	51.1	
M2	CBT	11,374	48.7	5,838	51.3	
M2	PBT	908	51.5	440	48.5	
M2	Mixed Modes	n/a	n/a	n/a	n/a	
M3	All	8,396	49.9	4,210	50.1	
M3	CBT	7,240	49.1	3,686	50.9	
M3	PBT	1,156	54.7	524	45.3	
M3	Mixed Modes	n/a	n/a	n/a	n/a	

Note: Includes students taking English-language mathematics tests, students taking Spanish-language mathematics tests, and students taking accommodated forms. A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. CBT = computer-based test; PBT = paper-based test; Mixed Modes = one test component (i.e., PBA or EOY) as CBT and the other component as PBT. n/a = not applicable.

Table A.5.6 All States Combined: Spanish-Language Mathematics Test Takers, by Grade, Mode, and Gender

Grade	Mode	Valid Cases N	Gender			
			Female		Male	
			N	%	N	%
3	All	5,338	2,655	49.7	2,683	50.3
3	CBT	3,749	1,869	49.9	1,880	50.1
3	PBT	1,581	782	49.5	799	50.5
3	Mixed Modes	8	4	50.0	4	50.0
4	All	2,862	1,375	48.0	1,487	52.0
4	CBT	2,251	1,074	47.7	1,177	52.3
4	PBT	603	298	49.4	305	50.6
4	Mixed Modes	8	3	37.5	5	62.5
5	All	2,245	1,090	48.6	1,155	51.4
5	CBT	1,897	917	48.3	980	51.7
5	PBT	339	169	49.9	170	50.1
5	Mixed Modes	9	4	44.4	5	55.6
6	All	1,716	831	48.4	885	51.6
6	CBT	1,581	767	48.5	814	51.5
6	PBT	129	61	47.3	68	52.7
6	Mixed Modes	6	3	50.0	3	50.0
7	All	2,048	1,009	49.3	1,039	50.7
7	CBT	1,847	911	49.3	936	50.7
7	PBT	188	90	47.9	98	52.1
7	Mixed Modes	13	8	61.5	5	38.5
8	All	1,996	917	45.9	1,079	54.1
8	CBT	1,844	849	46.0	995	54.0
8	PBT	145	64	44.1	81	55.9
8	Mixed Modes	7	4	57.1	3	42.9
A1	All	2,256	1,032	45.7	1,224	54.3
A1	CBT	2,120	971	45.8	1,149	54.2
A1	PBT	128	58	45.3	70	54.7
A1	Mixed Modes	8	3	37.5	5	62.5
GO	All	887	444	50.1	443	49.9
GO	CBT	848	424	50.0	424	50.0
GO	PBT	39	20	51.3	19	48.7
GO	Mixed Modes	n/a	n/a	n/a	n/a	n/a
A2	All	506	264	52.2	242	47.8
A2	CBT	498	260	52.2	238	47.8
A2	PBT	8	4	50.0	4	50.0
A2	Mixed Modes	n/a	n/a	n/a	n/a	n/a
M1	All	171	61	35.7	110	64.3
M1	CBT	167	59	35.3	108	64.7
M1	PBT	4	2	50.0	2	50.0
M1	Mixed Modes	n/a	n/a	n/a	n/a	n/a
M2	All	41	23	56.1	18	43.9
M2	CBT	41	23	56.1	18	43.9
M2	PBT	n/a	n/a	n/a	n/a	n/a
M2	Mixed Modes	n/a	n/a	n/a	n/a	n/a
M3	All	11	6	54.5	5	45.5
M3	CBT	11	6	54.5	5	45.5
M3	PBT	n/a	n/a	n/a	n/a	n/a
M3	Mixed Modes	n/a	n/a	n/a	n/a	n/a

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III. CBT = computer-based test; PBT = paper-based test; Mixed Modes = one test component (i.e., PBA or EOY) as CBT and the other component as PBT. n/a = not applicable.

Table A.5.7 Demographic Information for Grade 3 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	48.0	65.7	44.2	76.6	57.6	36.2	49.1	6.3	41.4	75.0	3.3	31.6	49.5
SWD (%)	9.0	11.9	9.8	13.2	0.1	16.7	9.1	12.6	14.8	13.5	2.0	6.6	15.2
EL (%)	13.0	9.3	17.2	13.9	19.4	12.8	10.7	0.8	7.4	18.6	1.5	0.0	10.2
Male (%)	51.0	51.1	51.1	50.3	50.9	51.1	51.2	50.9	51.0	50.3	50.3	50.0	51.9
Female (%)	49.0	48.9	48.9	49.7	49.1	48.9	48.8	49.1	49.0	49.7	49.7	50.0	48.1
AmInd/ANat (%)	1.0	0.6	0.7	0.0	0.3	0.0	0.3	0.7	0.1	14.2	1.3	0.0	0.6
Asian (%)	4.6	1.4	3.1	1.4	4.6	0.0	6.3	1.0	10.2	1.1	7.9	0.0	3.3
Black/AA (%)	18.3	19.9	4.6	69.4	17.6	0.0	33.4	47.3	15.4	2.0	28.7	44.7	7.6
Hisp/Lat (%)	24.3	12.9	32.7	16.1	26.5	20.0	15.7	3.6	27.5	57.5	53.0	3.9	25.7
Wh/Caus (%)	42.8	62.0	54.6	11.0	47.5	0.0	39.5	42.4	44.6	24.8	7.6	30.3	58.7
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.2	0.1	0.0	0.2	0.1	0.2	0.3	0.0	0.0	0.1
Two or More (%)	8.3	2.3	4.0	1.9	3.4	79.9	4.4	0.0	1.9	0.0	1.5	6.6	4.1
Unknown (%)	0.5	0.3	0.1	0.0	0.0	0.0	0.2	5.0	0.1	0.1	0.0	14.5	0.1

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.8 Demographic Information for Grade 4 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	46.5	65.0	44.5	75.3	55.8	34.3	47.1	6.6	39.9	75.3	n/a	44.2	49.5
SWD (%)	10.0	11.9	10.5	13.2	0.1	17.3	9.5	11.8	15.9	15.1	n/a	12.2	14.1
EL (%)	7.4	8.9	14.6	10.0	9.8	9.5	5.8	0.7	4.2	14.2	n/a	3.3	9.0
Male (%)	51.0	51.2	51.2	50.5	51.2	50.7	50.9	51.2	51.2	50.6	n/a	50.8	51.0
Female (%)	49.0	48.8	48.8	49.5	48.8	49.3	49.1	48.8	48.8	49.4	n/a	49.2	49.0
AmInd/ANat (%)	0.8	0.6	0.7	0.0	0.3	0.0	0.3	0.7	0.1	13.2	n/a	0.1	0.7
Asian (%)	4.3	1.6	3.3	1.2	4.9	0.0	6.5	1.1	10.2	1.2	n/a	2.1	3.1
Black/AA (%)	17.0	20.1	4.7	70.4	16.4	0.0	33.5	46.2	15.2	2.0	n/a	13.7	8.3
Hisp/Lat (%)	20.4	12.2	34.4	14.9	26.3	18.7	14.8	3.4	26.8	58.9	n/a	4.4	24.7
Wh/Caus (%)	48.3	62.4	52.9	11.3	48.7	0.0	40.3	43.7	45.7	24.4	n/a	68.7	59.0
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.2	0.1	0.0	0.2	0.0	0.2	0.2	n/a	0.1	0.1
Two or More (%)	7.4	2.1	3.8	1.9	3.3	81.2	4.2	0.0	1.6	0.1	n/a	3.9	4.1
Unknown (%)	1.7	0.3	0.1	0.0	0.0	0.0	0.3	4.8	0.1	0.1	n/a	6.9	0.1

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table A.5.9 Demographic Information for Grade 5 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	45.5	64.0	43.9	74.7	54.6	33.6	45.7	6.3	38.5	74.6	0.0	43.6	48.6
SWD (%)	10.1	11.3	10.8	16.7	0.1	17.6	10.1	10.8	16.0	15.4	0.0	12.4	14.6
EL (%)	5.7	8.7	12.7	9.0	6.0	8.0	4.1	0.6	3.0	12.8	2.0	2.8	7.0
Male (%)	51.1	50.6	51.2	49.9	51.1	50.9	51.0	51.0	51.3	50.6	54.1	51.1	51.3
Female (%)	48.9	49.4	48.8	50.1	48.9	49.1	49.0	49.0	48.7	49.4	45.9	48.9	48.7
AmInd/ANat (%)	0.8	0.6	0.8	0.1	0.3	0.0	0.2	0.7	0.1	13.4	0.0	0.1	0.8
Asian (%)	4.3	1.6	3.2	1.4	4.8	0.0	6.4	1.0	10.6	1.3	4.7	1.9	3.5
Black/AA (%)	16.9	19.8	4.5	71.6	16.4	0.0	33.7	45.9	15.2	2.0	35.1	13.5	8.2
Hisp/Lat (%)	19.8	12.1	34.3	14.4	25.9	18.1	13.9	3.3	25.4	59.6	42.6	3.8	23.8
Wh/Caus (%)	47.6	62.7	53.2	10.3	49.3	0.0	41.1	44.5	47.1	23.4	16.9	62.8	59.8
NtvHawai/Pacific (%)	0.2	0.6	0.2	0.2	0.1	0.0	0.1	0.1	0.2	0.1	0.0	0.1	0.2
Two or More (%)	7.3	2.2	3.5	2.0	3.2	81.8	4.2	0.0	1.4	0.0	0.7	3.0	3.6
Unknown (%)	3.2	0.3	0.2	0.0	0.0	0.0	0.2	4.5	0.1	0.1	0.0	14.8	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.10 Demographic Information for Grade 6 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	44.8	62.3	43.4	71.8	54.4	32.3	45.2	6.0	37.9	72.5	0.2	42.5	47.8
SWD (%)	9.8	10.9	10.5	17.3	0.2	17.6	9.7	10.1	15.3	14.7	0.2	12.4	14.5
EL (%)	4.7	8.1	11.4	6.5	4.6	6.7	3.4	0.4	2.7	11.1	0.0	2.1	5.7
Male (%)	51.2	51.5	51.0	51.2	51.3	50.9	51.0	51.3	51.2	50.1	47.5	51.3	51.5
Female (%)	48.8	48.5	49.0	48.8	48.7	49.1	49.0	48.7	48.8	49.9	52.5	48.7	48.5
AmInd/ANat (%)	0.8	0.6	0.7	0.1	0.3	0.0	0.3	0.6	0.1	12.9	0.2	0.1	0.5
Asian (%)	4.2	1.5	3.4	1.9	4.7	0.0	6.5	1.0	10.6	1.4	22.4	2.0	3.1
Black/AA (%)	17.0	19.9	4.6	73.7	16.8	0.0	33.8	46.2	15.8	1.9	20.0	13.5	8.2
Hisp/Lat (%)	19.5	11.9	34.1	13.7	25.6	18.0	13.7	3.2	24.7	59.7	51.3	3.8	23.6
Wh/Caus (%)	47.7	62.9	53.1	8.9	49.4	0.0	41.1	44.7	47.3	23.9	4.6	62.6	61.2
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.1	0.1	0.0	0.1	0.0	0.2	0.1	0.4	0.1	0.1
Two or More (%)	7.2	2.0	3.5	1.6	3.1	81.9	4.2	0.0	1.1	0.1	0.2	2.9	3.2
Unknown (%)	3.3	0.5	0.3	0.0	0.0	0.0	0.3	4.2	0.1	0.1	0.8	15.1	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.11 Demographic Information for Grade 7 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	43.8	61.3	43.5	73.3	52.7	32.4	43.8	6.5	37.1	70.5	27.0	41.8	46.4
SWD (%)	9.6	10.2	10.1	18.1	0.2	17.0	9.5	9.5	15.1	14.3	0.0	12.5	14.3
EL (%)	4.7	7.7	11.8	5.4	4.6	6.4	3.8	0.4	2.6	11.8	27.0	1.9	5.6
Male (%)	51.2	50.6	51.6	50.0	51.1	50.8	51.2	50.6	51.7	51.3	47.3	51.0	53.2
Female (%)	48.8	49.4	48.4	50.0	48.9	49.2	48.8	49.4	48.3	48.7	52.7	49.0	46.8
AmInd/ANat (%)	0.8	0.6	0.8	0.1	0.3	0.0	0.3	0.7	0.1	13.2	1.4	0.1	0.6
Asian (%)	4.1	1.3	3.5	1.6	4.6	0.0	6.4	1.1	10.2	1.3	2.7	1.9	3.2
Black/AA (%)	17.1	20.2	4.8	73.4	16.4	0.0	34.0	47.3	15.8	2.0	33.8	13.4	7.9
Hisp/Lat (%)	19.1	11.5	34.4	14.2	25.0	17.7	13.4	2.8	24.4	59.7	50.0	3.6	23.6
Wh/Caus (%)	48.3	63.4	52.5	9.0	50.6	0.0	41.5	44.2	48.2	23.5	9.5	63.5	61.2
NtvHawai/Pacific (%)	0.1	0.6	0.2	0.2	0.1	0.0	0.1	0.0	0.2	0.1	0.0	0.1	0.1
Two or More (%)	7.2	1.8	3.4	1.4	3.0	82.2	4.0	0.0	1.0	0.0	0.0	2.8	3.1
Unknown (%)	3.3	0.5	0.4	0.1	0.0	0.0	0.3	3.9	0.1	0.1	2.7	14.6	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.12 Demographic Information for Grade 8 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	43.0	59.9	42.1	74.3	52.2	31.9	42.2	6.3	35.9	69.3	n/a	41.0	46.1
SWD (%)	9.5	9.9	10.0	18.5	0.1	17.0	9.4	9.1	15.0	13.6	n/a	12.5	14.8
EL (%)	4.6	6.8	10.5	5.9	5.0	6.5	3.9	0.3	2.7	12.3	n/a	1.7	5.8
Male (%)	51.3	51.2	51.5	49.7	51.0	50.8	51.1	51.1	52.1	51.6	n/a	51.3	51.6
Female (%)	48.7	48.8	48.5	50.3	49.0	49.2	48.9	48.9	47.9	48.4	n/a	48.7	48.4
AmInd/ANat (%)	0.7	0.7	0.8	0.2	0.3	0.0	0.2	0.6	0.1	12.5	n/a	0.1	0.6
Asian (%)	4.0	1.5	3.3	1.3	4.5	0.0	6.3	1.0	10.4	1.3	n/a	1.7	3.1
Black/AA (%)	17.3	20.1	4.9	75.9	16.8	0.0	35.0	47.5	16.1	2.0	n/a	13.3	8.3
Hisp/Lat (%)	18.6	10.7	34.8	13.7	24.7	17.2	12.6	3.0	23.5	59.5	n/a	3.4	23.9
Wh/Caus (%)	48.4	64.1	52.0	7.4	50.8	0.1	41.7	44.0	48.8	24.4	n/a	63.4	60.6
NtvHawai/Pacific (%)	0.1	0.6	0.2	0.1	0.1	0.0	0.1	0.0	0.2	0.1	n/a	0.1	0.1
Two or More (%)	7.2	1.8	3.4	1.2	2.8	82.8	3.8	0.0	0.9	0.1	n/a	2.7	3.3
Unknown (%)	3.5	0.5	0.5	0.1	0.0	0.0	0.3	3.8	0.1	0.1	n/a	15.4	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table A.5.13 Demographic Information for Grade 9 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	42.6	58.9	42.0	74.6	45.8	33.7	n/r	n/a	34.0	60.9	n/a	38.0	46.2
SWD (%)	9.9	9.4	9.4	16.1	0.3	15.5	n/r	n/a	14.3	12.6	n/a	12.0	14.8
EL (%)	5.0	6.6	10.3	8.8	4.6	5.8	n/r	n/a	4.8	13.2	n/a	1.5	5.5
Male (%)	51.3	51.0	52.1	50.0	50.8	51.9	n/r	n/a	52.1	50.9	n/a	50.8	53.3
Female (%)	48.7	49.0	47.9	50.0	49.2	48.1	n/r	n/a	47.9	49.1	n/a	49.2	46.7
AmInd/ANat (%)	1.0	0.7	0.8	0.7	0.3	0.0	n/r	n/a	0.1	12.0	n/a	0.1	0.6
Asian (%)	4.3	1.6	3.3	1.6	5.1	0.0	n/r	n/a	10.4	1.3	n/a	1.8	3.1
Black/AA (%)	13.7	20.3	5.4	75.7	19.5	0.0	n/r	n/a	15.2	2.2	n/a	12.1	8.1
Hisp/Lat (%)	20.8	10.7	36.5	13.8	27.8	21.2	n/r	n/a	25.6	58.8	n/a	3.2	23.2
Wh/Caus (%)	51.1	64.0	49.9	6.5	44.5	0.0	n/r	n/a	47.1	25.2	n/a	63.7	61.6
NtvHawai/Pacific (%)	0.2	0.6	0.3	0.1	0.1	0.0	n/r	n/a	0.2	0.1	n/a	0.1	0.1
Two or More (%)	3.8	1.7	3.0	1.4	2.6	78.8	n/r	n/a	1.0	0.1	n/a	2.3	3.2
Unknown (%)	5.3	0.3	0.8	0.1	0.0	0.0	n/r	n/a	0.3	0.2	n/a	16.8	0.0

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Table A.5.14 Demographic Information for Grade 10 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	38.8	55.3	40.6	74.4	37.0	n/a	35.9	16.9	33.6	57.6	n/a	29.5	42.9
SWD (%)	10.1	7.9	9.3	16.7	0.0	n/a	8.0	7.6	14.2	10.8	n/a	4.2	13.9
EL (%)	4.4	6.2	7.4	7.0	0.7	n/a	2.3	0.5	4.0	8.7	n/a	0.4	5.5
Male (%)	51.2	50.3	52.5	49.2	51.3	n/a	50.7	50.0	52.2	50.6	n/a	44.2	52.2
Female (%)	48.8	49.7	47.5	50.8	48.7	n/a	49.3	50.0	47.8	49.4	n/a	55.8	47.8
AmInd/ANat (%)	1.3	0.7	0.8	0.6	0.1	n/a	0.2	0.2	0.1	11.4	n/a	0.2	0.8
Asian (%)	5.2	1.7	3.2	1.3	2.5	n/a	6.8	1.0	10.2	1.6	n/a	2.8	3.1
Black/AA (%)	21.4	19.6	5.5	78.2	7.9	n/a	34.9	50.5	15.2	1.9	n/a	13.6	8.8
Hisp/Lat (%)	21.7	10.5	36.5	12.8	10.7	n/a	11.4	3.6	24.3	58.7	n/a	2.3	22.6
Wh/Caus (%)	47.6	64.9	49.8	5.9	76.1	n/a	42.9	41.3	48.7	26.0	n/a	64.8	61.6
NtvHawai/Pacific (%)	0.2	0.6	0.3	0.1	0.0	n/a	0.1	0.0	0.2	0.1	n/a	0.1	0.1
Two or More (%)	1.7	1.6	3.0	1.0	2.7	n/a	3.5	0.0	0.8	0.1	n/a	4.7	2.8
Unknown (%)	0.8	0.4	0.8	0.0	0.0	n/a	0.1	3.4	0.3	0.2	n/a	11.5	0.1

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table A.5.15 Demographic Information for Grade 11 ELA/L, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	39.9	53.3	39.3	n/r	39.2	31.5	n/r	n/a	33.7	56.6	n/a	n/r	n/r
SWD (%)	9.3	8.5	8.6	n/r	0.3	13.9	n/r	n/a	14.7	10.2	n/a	n/r	n/r
EL (%)	4.2	4.4	6.6	n/r	0.9	6.8	n/r	n/a	3.7	7.2	n/a	n/r	n/r
Male (%)	51.4	50.2	51.9	n/r	50.9	50.5	n/r	n/a	52.5	49.7	n/a	n/r	n/r
Female (%)	48.6	49.8	48.1	n/r	49.1	49.5	n/r	n/a	47.5	50.3	n/a	n/r	n/r
AmInd/ANat (%)	1.7	0.6	0.7	n/r	0.3	0.0	n/r	n/a	0.1	11.9	n/a	n/r	n/r
Asian (%)	5.3	1.2	3.5	n/r	3.7	0.0	n/r	n/a	9.7	1.6	n/a	n/r	n/r
Black/AA (%)	11.2	22.5	5.3	n/r	11.8	0.0	n/r	n/a	15.5	2.0	n/a	n/r	n/r
Hisp/Lat (%)	26.5	8.7	36.4	n/r	13.7	19.8	n/r	n/a	24.2	57.6	n/a	n/r	n/r
Wh/Caus (%)	50.1	65.3	50.4	n/r	67.5	0.0	n/r	n/a	49.1	26.6	n/a	n/r	n/r
NtvHawai/Pacific (%)	0.2	0.1	0.3	n/r	0.1	0.0	n/r	n/a	0.3	0.1	n/a	n/r	n/r
Two or More (%)	4.8	1.5	2.8	n/r	2.8	80.1	n/r	n/a	0.8	0.1	n/a	n/r	n/r
Unknown (%)	0.2	0.1	0.6	n/r	0.0	0.0	n/r	n/a	0.3	0.1	n/a	n/r	n/r

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Table A.5.16 Demographic Information for Grade 3 Mathematics, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	47.9	65.8	45.1	76.6	57.6	36.3	49.2	6.2	41.5	76.1	27.0	46.4	49.8
SWD (%)	9.4	11.9	9.9	13.1	0.1	16.6	9.1	12.6	14.7	13.1	0.2	11.1	15.1
EL (%)	11.9	9.5	18.7	14.5	19.5	13.3	11.3	0.8	8.1	23.0	7.4	4.4	11.2
Male (%)	51.0	51.1	51.1	50.4	50.9	51.1	51.3	50.8	51.0	50.3	57.3	51.0	51.8
Female (%)	49.0	48.9	48.9	49.6	49.1	48.9	48.7	49.2	49.0	49.7	42.7	49.0	48.2
AmInd/ANat (%)	0.8	0.6	0.7	0.0	0.3	0.0	0.3	0.7	0.1	13.4	1.2	0.1	0.6
Asian (%)	4.1	1.4	3.0	1.5	4.7	0.0	6.3	1.0	10.2	1.1	6.2	2.0	3.4
Black/AA (%)	17.7	19.9	4.5	69.0	17.6	0.0	33.3	47.3	15.3	1.9	28.9	15.5	7.6
Hisp/Lat (%)	20.9	13.0	33.9	16.4	26.5	20.2	16.0	3.6	27.8	59.7	48.9	4.8	26.2
Wh/Caus (%)	47.2	61.8	53.6	11.1	47.4	0.1	39.3	42.4	44.3	23.5	14.1	67.0	58.1
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.2	0.1	0.0	0.2	0.1	0.2	0.2	0.2	0.1	0.1
Two or More (%)	7.4	2.2	3.9	1.9	3.4	79.6	4.3	0.0	1.8	0.0	0.5	4.0	4.0
Unknown (%)	1.6	0.3	0.1	0.0	0.0	0.0	0.2	5.0	0.1	0.1	0.0	6.5	0.1

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.17 Demographic Information for Grade 4 Mathematics, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	46.6	65.0	44.6	75.3	55.9	34.4	47.3	6.6	40.0	75.8	n/a	44.3	49.8
SWD (%)	9.9	11.9	10.5	13.2	0.1	17.2	9.5	11.9	15.7	15.0	n/a	12.2	14.0
EL (%)	7.8	9.0	14.7	10.7	10.0	10.2	6.4	0.7	4.9	17.1	n/a	3.4	10.1
Male (%)	51.0	51.2	51.2	50.5	51.2	50.7	50.9	51.2	51.2	50.6	n/a	50.8	50.9
Female (%)	49.0	48.8	48.8	49.5	48.8	49.3	49.1	48.8	48.8	49.4	n/a	49.2	49.1
AmInd/ANat (%)	0.8	0.6	0.7	0.0	0.3	0.0	0.3	0.7	0.1	12.7	n/a	0.1	0.7
Asian (%)	4.3	1.6	3.3	1.3	4.9	0.0	6.5	1.1	10.3	1.2	n/a	2.1	3.2
Black/AA (%)	17.0	20.0	4.6	70.0	16.4	0.0	33.4	46.2	15.1	2.0	n/a	13.8	8.3
Hisp/Lat (%)	20.6	12.3	34.5	15.3	26.3	19.0	15.0	3.4	27.1	60.2	n/a	4.4	25.1
Wh/Caus (%)	48.1	62.3	52.8	11.3	48.6	0.1	40.2	43.7	45.5	23.6	n/a	68.7	58.5
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.2	0.1	0.0	0.2	0.0	0.2	0.1	n/a	0.1	0.1
Two or More (%)	7.4	2.1	3.8	1.9	3.3	80.9	4.1	0.0	1.6	0.1	n/a	3.9	4.0
Unknown (%)	1.7	0.3	0.1	0.0	0.0	0.0	0.3	4.8	0.1	0.1	n/a	6.9	0.1

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table A.5.18 Demographic Information for Grade 5 Mathematics, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	45.6	64.0	43.9	74.8	54.7	33.6	45.9	6.3	38.7	75.0	0.0	43.8	48.9
SWD (%)	10.0	11.3	10.8	16.7	0.1	17.5	10.0	10.8	15.9	15.3	0.0	12.5	14.5
EL (%)	6.0	8.9	12.7	9.4	6.2	8.6	4.7	0.6	3.7	14.8	0.5	3.0	8.0
Male (%)	51.1	50.6	51.2	49.9	51.1	51.0	51.0	51.0	51.3	50.6	47.6	51.1	51.3
Female (%)	48.9	49.4	48.8	50.1	48.9	49.0	49.0	49.0	48.7	49.4	52.4	48.9	48.7
AmInd/ANat (%)	0.8	0.6	0.8	0.1	0.3	0.0	0.2	0.7	0.1	13.1	0.0	0.1	0.8
Asian (%)	4.3	1.6	3.2	1.4	4.9	0.0	6.4	1.0	10.6	1.3	3.7	1.9	3.5
Black/AA (%)	16.9	19.7	4.5	71.3	16.4	0.0	33.6	45.9	15.1	2.0	53.4	13.6	8.2
Hisp/Lat (%)	20.0	12.2	34.3	14.7	25.9	18.4	14.2	3.3	25.7	60.5	13.2	3.8	24.2
Wh/Caus (%)	47.5	62.6	53.2	10.3	49.2	0.1	40.9	44.5	46.8	23.0	23.3	62.7	59.3
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.2	0.1	0.0	0.1	0.1	0.2	0.1	2.1	0.1	0.2
Two or More (%)	7.3	2.2	3.5	2.0	3.2	81.5	4.2	0.0	1.4	0.0	4.2	3.0	3.6
Unknown (%)	3.2	0.3	0.2	0.0	0.0	0.0	0.2	4.5	0.1	0.1	0.0	14.7	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.19 Demographic Information for Grade 6 Mathematics, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	44.9	62.3	43.4	71.8	54.4	32.3	45.6	6.0	38.1	72.6	1.0	42.7	48.2
SWD (%)	9.8	10.9	10.5	17.2	0.2	17.4	9.7	10.1	15.3	14.5	0.0	12.4	14.4
EL (%)	5.1	8.3	11.4	7.2	4.8	7.2	4.2	0.4	3.4	11.7	0.0	2.2	6.7
Male (%)	51.2	51.5	51.0	51.0	51.3	50.9	51.0	51.3	51.2	50.0	48.1	51.2	51.5
Female (%)	48.8	48.5	49.0	49.0	48.7	49.1	49.0	48.7	48.8	50.0	51.9	48.8	48.5
AmInd/ANat (%)	0.8	0.6	0.7	0.1	0.3	0.0	0.3	0.6	0.1	12.8	0.0	0.1	0.5
Asian (%)	4.2	1.5	3.4	1.9	4.7	0.0	6.4	1.0	10.6	1.3	3.8	1.9	3.2
Black/AA (%)	17.0	19.8	4.6	73.2	16.8	0.0	33.4	46.3	15.7	1.9	30.8	13.6	8.4
Hisp/Lat (%)	19.7	12.0	34.1	14.1	25.6	18.2	14.3	3.2	25.1	60.0	59.6	3.8	23.9
Wh/Caus (%)	47.6	62.8	53.0	8.9	49.4	0.1	41.0	44.7	47.1	23.7	2.9	62.6	60.6
NtvHawai/Pacific (%)	0.2	0.7	0.2	0.1	0.1	0.0	0.1	0.0	0.2	0.1	2.9	0.1	0.1
Two or More (%)	7.2	2.0	3.6	1.6	3.1	81.7	4.2	0.0	1.1	0.1	0.0	2.9	3.2
Unknown (%)	3.3	0.5	0.3	0.0	0.0	0.0	0.3	4.2	0.1	0.1	0.0	15.1	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.20 Demographic Information for Grade 7 Mathematics, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	44.7	61.6	44.7	77.0	52.7	32.4	47.5	6.5	38.4	70.7	0.0	42.4	47.0
SWD (%)	9.8	10.2	10.5	18.5	0.2	17.0	10.5	9.5	15.5	14.1	0.0	12.8	14.2
EL (%)	5.2	7.9	12.3	6.5	4.8	7.0	5.0	0.4	3.5	12.4	0.0	2.1	6.9
Male (%)	51.2	50.7	51.4	50.2	51.1	51.0	51.2	50.6	51.5	51.3	52.6	51.0	53.1
Female (%)	48.8	49.3	48.6	49.8	48.9	49.0	48.8	49.4	48.5	48.7	47.4	49.0	46.9
AmInd/ANat (%)	0.8	0.6	0.8	0.1	0.3	0.0	0.3	0.7	0.1	13.1	0.0	0.1	0.6
Asian (%)	3.7	1.3	3.3	1.4	4.6	0.0	4.6	1.1	9.0	1.3	4.4	1.8	3.2
Black/AA (%)	17.3	20.1	5.0	75.8	16.4	0.0	36.2	47.3	16.2	2.0	32.5	13.6	8.1
Hisp/Lat (%)	19.6	11.7	35.5	15.4	25.1	18.0	14.6	2.8	25.5	60.1	43.9	3.7	24.2
Wh/Caus (%)	47.9	63.4	51.6	5.9	50.5	0.1	40.1	44.2	47.8	23.2	18.4	63.4	60.4
NtvHawai/Pacific (%)	0.1	0.6	0.2	0.2	0.1	0.0	0.1	0.0	0.2	0.1	0.9	0.1	0.1
Two or More (%)	7.3	1.8	3.4	1.2	3.0	81.9	3.8	0.0	1.0	0.0	0.0	2.7	3.1
Unknown (%)	3.3	0.5	0.3	0.1	0.0	0.0	0.3	3.9	0.1	0.1	0.0	14.6	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported.

Table A.5.21 Demographic Information for Grade 8 Mathematics, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	47.4	65.1	49.5	79.6	52.2	33.4	48.9	6.3	44.5	74.7	n/a	48.2	54.4
SWD (%)	11.1	11.8	13.0	21.0	0.1	18.4	12.2	9.1	21.1	17.7	n/a	16.3	19.0
EL (%)	5.8	8.0	13.6	6.6	5.2	7.4	6.0	0.3	4.8	16.3	n/a	2.3	9.1
Male (%)	51.8	52.2	51.9	49.8	51.0	51.2	52.6	51.1	53.3	52.4	n/a	52.1	52.6
Female (%)	48.2	47.8	48.1	50.2	49.0	48.8	47.4	48.9	46.7	47.6	n/a	47.9	47.4
AmInd/ANat (%)	0.8	0.7	0.9	0.2	0.3	0.0	0.2	0.6	0.1	14.2	n/a	0.1	0.8
Asian (%)	3.0	1.1	2.5	0.8	4.6	0.0	5.4	1.0	5.4	0.9	n/a	1.2	2.6
Black/AA (%)	19.2	21.5	5.9	82.9	16.7	0.0	40.0	47.4	20.0	2.1	n/a	16.4	9.4
Hisp/Lat (%)	20.5	11.6	40.8	11.5	24.8	18.2	14.1	3.0	29.1	61.7	n/a	4.0	28.6
Wh/Caus (%)	45.9	62.2	46.0	3.7	50.8	0.1	36.4	44.1	44.2	20.8	n/a	61.6	54.8
NtvHawai/Pacific (%)	0.1	0.7	0.2	0.2	0.1	0.0	0.1	0.0	0.2	0.1	n/a	0.1	0.2
Two or More (%)	7.6	1.7	3.2	0.7	2.8	81.7	3.4	0.0	0.8	0.1	n/a	3.0	3.4
Unknown (%)	2.9	0.5	0.4	0.1	0.0	0.0	0.3	3.9	0.2	0.1	n/a	13.7	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table A.5.22 Demographic Information for Algebra I, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	40.5	58.7	40.2	75.1	53.3	31.6	41.0	17.3	34.4	61.1	n/a	36.3	46.8
SWD (%)	9.5	6.3	10.6	16.0	0.1	12.4	8.4	9.8	13.7	12.2	n/a	10.2	13.4
EL (%)	5.2	7.1	9.7	14.4	7.3	5.8	5.6	0.5	4.7	13.5	n/a	1.4	8.1
Male (%)	51.5	50.6	52.9	50.6	51.3	50.8	51.7	52.0	52.1	51.5	n/a	50.3	53.2
Female (%)	48.5	49.4	47.1	49.4	48.7	49.2	48.3	48.0	47.9	48.5	n/a	49.7	46.8
AmInd/ANat (%)	0.8	0.7	0.8	0.6	0.3	0.0	0.2	0.3	0.1	10.9	n/a	0.1	0.6
Asian (%)	4.2	1.6	3.2	1.5	3.5	0.0	6.8	0.8	10.3	1.3	n/a	1.8	3.3
Black/AA (%)	19.4	19.6	4.4	69.0	24.6	0.0	34.7	52.7	15.7	2.3	n/a	11.0	8.5
Hisp/Lat (%)	18.9	11.4	33.5	20.6	31.3	19.5	13.7	3.2	25.3	60.1	n/a	3.2	24.9
Wh/Caus (%)	48.0	64.1	53.7	7.1	37.7	0.0	40.4	39.4	47.1	25.0	n/a	64.1	59.4
NtvHawai/Pacific (%)	0.2	0.6	0.3	0.1	0.1	0.0	0.1	0.0	0.2	0.2	n/a	0.1	0.1
Two or More (%)	3.8	1.8	3.2	1.1	2.5	80.5	3.8	0.0	0.9	0.1	n/a	2.3	3.1
Unknown (%)	4.6	0.3	0.8	0.1	0.0	0.0	0.2	3.6	0.3	0.1	n/a	17.4	0.1

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table A.5.23 Demographic Information for Geometry, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	38.5	55.7	38.0	73.4	35.9	24.8	n/r	n/a	32.6	57.6	n/a	19.5	44.0
SWD (%)	8.6	4.7	7.6	15.7	0.0	8.2	n/r	n/a	13.1	10.2	n/a	1.7	13.7
EL (%)	4.8	6.6	7.2	7.4	1.6	3.7	n/r	n/a	3.7	9.1	n/a	0.3	7.5
Male (%)	50.8	49.9	51.6	49.9	49.0	48.9	n/r	n/a	52.0	51.1	n/a	48.4	52.7
Female (%)	49.2	50.1	48.4	50.1	51.0	51.1	n/r	n/a	48.0	48.9	n/a	51.6	47.3
AmInd/ANat (%)	1.4	0.7	0.7	0.7	0.1	0.0	n/r	n/a	0.1	10.8	n/a	0.1	0.8
Asian (%)	5.4	1.7	3.2	1.7	3.0	0.0	n/r	n/a	10.9	1.5	n/a	3.5	3.0
Black/AA (%)	11.6	20.4	4.7	76.3	5.3	0.0	n/r	n/a	14.5	1.9	n/a	4.6	8.7
Hisp/Lat (%)	22.9	10.7	34.0	12.5	12.8	10.9	n/r	n/a	24.2	59.7	n/a	1.6	24.4
Wh/Caus (%)	52.7	63.9	53.1	7.6	75.9	0.0	n/r	n/a	48.8	25.9	n/a	69.3	60.0
NtvHawai/Pacific (%)	0.3	0.6	0.3	0.0	0.1	0.0	n/r	n/a	0.2	0.1	n/a	0.1	0.2
Two or More (%)	2.4	1.6	3.2	1.1	2.8	89.1	n/r	n/a	0.9	0.1	n/a	1.7	2.8
Unknown (%)	3.4	0.3	0.8	0.1	0.0	0.0	n/r	n/a	0.3	0.0	n/a	19.2	0.2

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Table A.5.24 Demographic Information for Algebra II, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	35.1	51.8	32.8	51.3	35.7	32.6	29.7	n/a	30.6	52.8	n/a	29.9	n/r
SWD (%)	6.0	4.0	5.1	13.3	0.1	14.4	4.4	n/a	10.0	7.1	n/a	2.6	n/r
EL (%)	3.2	4.3	5.2	4.0	1.2	8.3	1.6	n/a	2.8	6.4	n/a	0.0	n/r
Male (%)	49.9	49.2	52.0	55.3	49.2	50.7	48.3	n/a	51.1	48.3	n/a	68.8	n/r
Female (%)	50.1	50.8	48.0	44.7	50.8	49.3	51.7	n/a	48.9	51.7	n/a	31.2	n/r
AmInd/ANat (%)	1.2	0.6	0.6	0.0	0.3	0.0	0.2	n/a	0.1	10.2	n/a	0.0	n/r
Asian (%)	7.0	1.6	3.8	2.0	4.7	0.0	8.4	n/a	12.0	2.1	n/a	9.1	n/r
Black/AA (%)	14.8	21.6	4.0	55.3	11.2	0.0	31.6	n/a	14.3	1.9	n/a	16.9	n/r
Hisp/Lat (%)	21.8	8.7	30.4	25.3	13.6	22.0	10.3	n/a	22.0	56.2	n/a	1.3	n/r
Wh/Caus (%)	50.6	65.6	56.0	14.7	67.3	0.0	45.5	n/a	50.3	29.4	n/a	64.9	n/r
NtvHawai/Pacific (%)	0.2	0.1	0.2	0.0	0.1	0.0	0.1	n/a	0.3	0.1	n/a	0.0	n/r
Two or More (%)	4.0	1.7	3.5	2.7	2.7	78.0	3.8	n/a	0.9	0.1	n/a	0.0	n/r
Unknown (%)	0.3	0.0	1.5	0.0	0.0	0.0	0.1	n/a	0.2	0.1	n/a	7.8	n/r

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Table A.5.25 Demographic Information for Integrated Mathematics I, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	51.6	n/a	50.5	n/a	53.2	28.3	n/a	n/a	51.7	87.8	n/a	50.5	n/r
SWD (%)	11.6	n/a	12.1	n/a	2.4	17.4	n/a	n/a	96.6	18.2	n/a	17.1	n/r
EL (%)	8.8	n/a	13.8	n/a	7.6	3.5	n/a	n/a	0.0	21.9	n/a	3.9	n/r
Male (%)	51.9	n/a	52.5	n/a	50.5	50.9	n/a	n/a	69.0	51.7	n/a	52.2	n/r
Female (%)	48.1	n/a	47.5	n/a	49.5	49.1	n/a	n/a	31.0	48.3	n/a	47.8	n/r
AmInd/ANat (%)	1.1	n/a	1.0	n/a	0.2	0.0	n/a	n/a	0.0	29.7	n/a	0.2	n/r
Asian (%)	2.7	n/a	4.0	n/a	2.0	0.0	n/a	n/a	0.0	0.3	n/a	2.2	n/r
Black/AA (%)	13.7	n/a	7.2	n/a	12.8	0.0	n/a	n/a	51.7	1.0	n/a	23.0	n/r
Hisp/Lat (%)	30.5	n/a	46.1	n/a	42.4	11.3	n/a	n/a	10.3	55.1	n/a	4.2	n/r
Wh/Caus (%)	44.1	n/a	38.6	n/a	40.4	0.2	n/a	n/a	37.9	13.6	n/a	57.6	n/r
NtvHawai/Pacific (%)	0.2	n/a	0.3	n/a	0.1	0.0	n/a	n/a	0.0	0.0	n/a	0.0	n/r
Two or More (%)	4.2	n/a	2.3	n/a	2.1	88.5	n/a	n/a	0.0	0.0	n/a	3.4	n/r
Unknown (%)	3.4	n/a	0.5	n/a	0.0	0.0	n/a	n/a	0.0	0.3	n/a	9.5	n/r

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Table A.5.26 Demographic Information for Integrated Mathematics II, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	44.4	n/a	43.4	n/r	57.9	29.6	n/a	n/a	n/r	87.1	n/a	25.0	n/a
SWD (%)	7.7	n/a	8.8	n/r	0.0	14.8	n/a	n/a	n/r	14.3	n/a	2.0	n/a
EL (%)	6.5	n/a	7.4	n/r	0.2	3.7	n/a	n/a	n/r	16.0	n/a	0.5	n/a
Male (%)	51.1	n/a	51.5	n/r	53.5	33.3	n/a	n/a	n/r	46.5	n/a	51.0	n/a
Female (%)	48.9	n/a	48.5	n/r	46.5	66.7	n/a	n/a	n/r	53.5	n/a	49.0	n/a
AmInd/ANat (%)	4.0	n/a	0.8	n/r	0.0	0.0	n/a	n/a	n/r	45.2	n/a	0.1	n/a
Asian (%)	3.6	n/a	4.2	n/r	0.5	0.0	n/a	n/a	n/r	0.1	n/a	3.7	n/a
Black/AA (%)	8.8	n/a	7.4	n/r	38.5	0.0	n/a	n/a	n/r	0.3	n/a	8.9	n/a
Hisp/Lat (%)	33.4	n/a	42.6	n/r	1.2	7.4	n/a	n/a	n/r	40.8	n/a	2.0	n/a
Wh/Caus (%)	46.0	n/a	42.6	n/r	58.8	0.0	n/a	n/a	n/r	8.8	n/a	74.3	n/a
NtvHawai/Pacific (%)	0.1	n/a	0.2	n/r	0.0	0.0	n/a	n/a	n/r	0.0	n/a	0.0	n/a
Two or More (%)	2.1	n/a	2.1	n/r	1.1	92.6	n/a	n/a	n/r	0.0	n/a	2.4	n/a
Unknown (%)	1.8	n/a	0.1	n/r	0.0	0.0	n/a	n/a	n/r	4.8	n/a	8.5	n/a

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Table A.5.27 Demographic Information for Integrated Mathematics III, Overall and by State

Demographic	PARCC	AR	CO	DC	IL	MA	MD	MS	NJ	NM	NY	OH	RI
Econ Dis (%)	46.2	n/a	41.8	n/a	51.6	31.7	n/a	n/a	n/r	87.9	n/a	n/a	n/a
SWD (%)	5.1	n/a	5.5	n/a	0.0	14.3	n/a	n/a	n/r	14.6	n/a	n/a	n/a
EL (%)	4.4	n/a	4.6	n/a	0.8	12.4	n/a	n/a	n/r	11.3	n/a	n/a	n/a
Male (%)	50.1	n/a	51.5	n/a	45.9	50.2	n/a	n/a	n/r	49.6	n/a	n/a	n/a
Female (%)	49.9	n/a	48.5	n/a	54.1	49.8	n/a	n/a	n/r	50.4	n/a	n/a	n/a
AmInd/ANat (%)	2.2	n/a	0.9	n/a	0.3	0.0	n/a	n/a	n/r	26.7	n/a	n/a	n/a
Asian (%)	4.7	n/a	5.1	n/a	5.0	0.0	n/a	n/a	n/r	0.8	n/a	n/a	n/a
Black/AA (%)	8.0	n/a	8.6	n/a	9.0	0.4	n/a	n/a	n/r	0.6	n/a	n/a	n/a
Hisp/Lat (%)	35.9	n/a	41.5	n/a	15.3	19.3	n/a	n/a	n/r	53.8	n/a	n/a	n/a
Wh/Caus (%)	44.4	n/a	41.4	n/a	67.6	0.4	n/a	n/a	n/r	18.1	n/a	n/a	n/a
NtvHawai/Pacific (%)	0.2	n/a	0.3	n/a	0.1	0.0	n/a	n/a	n/r	0.0	n/a	n/a	n/a
Two or More (%)	4.5	n/a	2.1	n/a	2.7	79.9	n/a	n/a	n/r	0.0	n/a	n/a	n/a
Unknown (%)	0.0	n/a	0.1	n/a	0.0	0.0	n/a	n/a	n/r	0.0	n/a	n/a	n/a

Note: PARCC = data from all participating states combined. Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/r = not reported due to n<20. n/a = not applicable.

Appendix 7: Summary of Differential Item Function (DIF) Results

Table A.7.1 Differential Item Functioning for ELA/L Grade 3

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	102	520	424	0	0	2	0	422	100	0	0	0	0
	PBT	65	152	152	0	0	2	1	150	99	0	0	0	0
White vs AmerIndian	CBT	102	520	360	2	1	11	3	347	96	0	0	0	0
	PBT	65	152	112	0	0	10	9	102	91	0	0	0	0
White vs Asian	CBT	102	520	360	0	0	5	1	350	97	5	1	0	0
	PBT	65	152	152	0	0	3	2	149	98	0	0	0	0
White vs Black	CBT	102	520	424	0	0	1	0	423	100	0	0	0	0
	PBT	65	152	152	0	0	4	3	148	97	0	0	0	0
White vs Hispanic	CBT	102	520	424	0	0	6	1	418	99	0	0	0	0
	PBT	65	152	152	1	1	4	3	147	97	0	0	0	0
White vs Pacific Islander	CBT	102	520	120	0	0	6	5	110	92	4	3	0	0
	PBT	65	152	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	102	520	360	0	0	0	0	360	100	0	0	0	0
	PBT	65	152	152	0	0	0	0	152	100	0	0	0	0
NoEcnDis vs EcnDis	CBT	102	520	424	0	0	0	0	424	100	0	0	0	0
	PBT	65	152	152	0	0	2	1	150	99	0	0	0	0
ELN vs ELY	CBT	102	520	424	0	0	7	2	417	98	0	0	0	0
	PBT	65	152	152	1	1	7	5	144	95	0	0	0	0
SWDN vs SWDY	CBT	102	520	424	0	0	0	0	424	100	0	0	0	0
	PBT	65	152	152	1	1	2	1	149	98	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.2 Differential Item Functioning for ELA/L Grade 4

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	116	589	484	0	0	5	1	479	99	0	0	0	0
	PBT	76	212	212	0	0	5	2	207	98	0	0	0	0
White vs AmerIndian	CBT	116	589	414	3	1	15	4	396	96	0	0	0	0
	PBT	76	212	166	6	4	7	4	152	92	1	1	0	0
White vs Asian	CBT	116	589	484	2	0	9	2	466	96	6	1	1	0
	PBT	76	212	212	0	0	3	1	208	98	1	0	0	0
White vs Black	CBT	116	589	484	1	0	12	2	471	97	0	0	0	0
	PBT	76	212	212	0	0	4	2	208	98	0	0	0	0
White vs Hispanic	CBT	116	589	484	0	0	9	2	475	98	0	0	0	0
	PBT	76	212	212	0	0	6	3	206	97	0	0	0	0
White vs Pacific Islander	CBT	116	589	138	2	1	11	8	122	88	3	2	0	0
	PBT	76	212	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	116	589	414	0	0	0	0	414	100	0	0	0	0
	PBT	76	212	212	0	0	2	1	210	99	0	0	0	0
NoEcnDis vs EcnDis	CBT	116	589	484	0	0	0	0	484	100	0	0	0	0
	PBT	76	212	212	0	0	0	0	212	100	0	0	0	0
ELN vs ELY	CBT	116	589	484	0	0	20	4	464	96	0	0	0	0
	PBT	76	212	212	2	1	7	3	203	96	0	0	0	0
SWDN vs SWDY	CBT	116	589	484	0	0	4	1	480	99	0	0	0	0
	PBT	76	212	212	0	0	5	2	207	98	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.3 Differential Item Functioning for ELA/L Grade 5

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	119	563	460	0	0	8	2	451	98	1	0	0	0
	PBT	73	184	184	0	0	3	2	181	98	0	0	0	0
White vs AmerIndian	CBT	119	563	390	3	1	11	3	374	96	2	1	0	0
	PBT	73	184	115	0	0	14	12	100	87	0	0	1	1
White vs Asian	CBT	119	563	460	0	0	10	2	447	97	3	1	0	0
	PBT	73	184	184	2	1	1	1	181	98	0	0	0	0
White vs Black	CBT	119	563	460	1	0	4	1	455	99	0	0	0	0
	PBT	73	184	184	3	2	4	2	177	96	0	0	0	0
White vs Hispanic	CBT	119	563	460	2	0	2	0	456	99	0	0	0	0
	PBT	73	184	184	2	1	5	3	177	96	0	0	0	0
White vs Pacific Islander	CBT	119	563	138	3	2	11	8	122	88	2	1	0	0
	PBT	73	184	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	119	563	390	0	0	0	0	390	100	0	0	0	0
	PBT	73	184	184	0	0	0	0	183	99	1	1	0	0
NoEcnDis vs EcnDis	CBT	119	563	460	0	0	0	0	460	100	0	0	0	0
	PBT	73	184	184	0	0	2	1	182	99	0	0	0	0
ELN vs ELY	CBT	119	563	460	4	1	9	2	447	97	0	0	0	0
	PBT	73	184	184	2	1	12	7	170	92	0	0	0	0
SWDN vs SWDY	CBT	119	563	460	0	0	4	1	456	99	0	0	0	0
	PBT	73	184	184	0	0	0	0	184	100	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.4 Differential Item Functioning for ELA/L Grade 6

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	158	1354	1196	0	0	56	5	1134	95	6	1	0	0
	PBT	90	422	422	10	2	10	2	393	93	9	2	0	0
White vs AmerIndian	CBT	158	1354	237	3	1	9	4	224	95	0	0	1	0
	PBT	90	422	23	5	22	4	17	14	61	0	0	0	0
White vs Asian	CBT	158	1354	1151	5	0	23	2	1113	97	10	1	0	0
	PBT	90	422	422	2	0	14	3	394	93	12	3	0	0
White vs Black	CBT	158	1354	1196	9	1	23	2	1164	97	0	0	0	0
	PBT	90	422	422	8	2	9	2	405	96	0	0	0	0
White vs Hispanic	CBT	158	1354	1196	16	1	3	0	1176	98	1	0	0	0
	PBT	90	422	422	8	2	6	1	404	96	4	1	0	0
White vs Pacific Islander	CBT	158	1354	138	8	6	8	6	121	88	1	1	0	0
	PBT	90	422	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	158	1354	1106	0	0	4	0	1097	99	5	0	0	0
	PBT	90	422	400	0	0	11	3	382	96	5	1	2	1
NoEcnDis vs EcnDis	CBT	158	1354	1196	0	0	3	0	1193	100	0	0	0	0
	PBT	90	422	422	0	0	2	0	420	100	0	0	0	0
ELN vs ELY	CBT	158	1354	1196	2	0	25	2	1169	98	0	0	0	0
	PBT	90	422	422	11	3	36	9	375	89	0	0	0	0
SWDN vs SWDY	CBT	158	1354	1196	0	0	7	1	1189	99	0	0	0	0
	PBT	90	422	422	0	0	2	0	420	100	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.5 Differential Item Functioning for ELA/L Grade 7

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	137	1209	1074	2	0	5	0	1062	99	5	0	0	0
	PBT	90	334	334	0	0	6	2	319	96	9	3	0	0
White vs AmerIndian	CBT	137	1209	734	12	2	34	5	684	93	2	0	2	0
	PBT	90	334	39	1	3	5	13	33	85	0	0	0	0
White vs Asian	CBT	137	1209	1001	0	0	10	1	979	98	11	1	1	0
	PBT	90	334	334	2	1	8	2	320	96	4	1	0	0
White vs Black	CBT	137	1209	1074	0	0	17	2	1057	98	0	0	0	0
	PBT	90	334	334	3	1	7	2	324	97	0	0	0	0
White vs Hispanic	CBT	137	1209	1074	2	0	7	1	1065	99	0	0	0	0
	PBT	90	334	334	3	1	9	3	322	96	0	0	0	0
White vs Pacific Islander	CBT	137	1209	138	0	0	13	9	124	90	0	0	1	1
	PBT	90	334	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	137	1209	984	0	0	3	0	977	99	3	0	1	0
	PBT	90	334	334	0	0	1	0	332	99	1	0	0	0
NoEcnDis vs EcnDis	CBT	137	1209	1074	0	0	0	0	1074	100	0	0	0	0
	PBT	90	334	334	0	0	1	0	333	100	0	0	0	0
ELN vs ELY	CBT	137	1209	1074	7	1	38	4	1029	96	0	0	0	0
	PBT	90	334	334	13	4	21	6	300	90	0	0	0	0
SWDN vs SWDY	CBT	137	1209	1074	0	0	1	0	1073	100	0	0	0	0
	PBT	90	334	334	0	0	0	0	334	100	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.6 Differential Item Functioning for ELA/L Grade 8

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	145	1045	910	8	1	16	2	879	97	7	1	0	0
	PBT	90	246	246	6	2	8	3	221	90	11	4	0	0
White vs AmerIndian	CBT	145	1045	820	15	2	52	6	743	91	8	1	2	0
	PBT	90	246	154	4	3	17	11	132	86	1	1	0	0
White vs Asian	CBT	145	1045	820	0	0	9	1	804	98	7	1	0	0
	PBT	90	246	246	1	0	4	2	240	98	1	0	0	0
White vs Black	CBT	145	1045	910	5	1	12	1	893	98	0	0	0	0
	PBT	90	246	246	2	1	6	2	238	97	0	0	0	0
White vs Hispanic	CBT	145	1045	910	12	1	10	1	888	98	0	0	0	0
	PBT	90	246	246	3	1	7	3	236	96	0	0	0	0
White vs Pacific Islander	CBT	145	1045	138	0	0	2	1	136	99	0	0	0	0
	PBT	90	246	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	145	1045	820	0	0	1	0	818	100	1	0	0	0
	PBT	90	246	246	0	0	0	0	245	100	1	0	0	0
NoEcnDis vs EcnDis	CBT	145	1045	910	0	0	8	1	902	99	0	0	0	0
	PBT	90	246	246	0	0	2	1	244	99	0	0	0	0
ELN vs ELY	CBT	145	1045	910	14	2	24	3	870	96	2	0	0	0
	PBT	90	246	246	12	5	25	10	209	85	0	0	0	0
SWDN vs SWDY	CBT	145	1045	910	4	0	14	2	892	98	0	0	0	0
	PBT	90	246	246	0	0	1	0	244	99	1	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.7 Differential Item Functioning for ELA/L Grade 9

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	189	1191	1056	3	0	26	2	1012	96	15	1	0	0
	PBT	90	400	400	8	2	14	4	359	90	18	5	1	0
White vs AmerIndian	CBT	189	1191	706	4	1	47	7	650	92	3	0	2	0
	PBT	90	400	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	189	1191	967	0	0	3	0	949	98	15	2	0	0
	PBT	90	400	400	2	1	14	4	369	92	13	3	2	1
White vs Black	CBT	189	1191	1056	2	0	3	0	1051	100	0	0	0	0
	PBT	90	400	400	2	1	8	2	390	98	0	0	0	0
White vs Hispanic	CBT	189	1191	1056	2	0	3	0	1051	100	0	0	0	0
	PBT	90	400	400	2	1	2	1	396	99	0	0	0	0
White vs Pacific Islander	CBT	189	1191	184	0	0	6	3	178	97	0	0	0	0
	PBT	90	400	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	189	1191	967	0	0	11	1	940	97	15	2	1	0
	PBT	90	400	114	0	0	0	0	113	99	0	0	1	1
NoEcnDis vs EcnDis	CBT	189	1191	1056	0	0	2	0	1054	100	0	0	0	0
	PBT	90	400	400	0	0	2	1	398	100	0	0	0	0
ELN vs ELY	CBT	189	1191	1056	6	1	18	2	1028	97	4	0	0	0
	PBT	90	400	311	9	3	27	9	272	87	2	1	1	0
SWDN vs SWDY	CBT	189	1191	1033	0	0	1	0	1031	100	1	0	0	0
	PBT	90	400	400	0	0	1	0	399	100	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.8 Differential Item Functioning for ELA/L Grade 10

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	186	1267	1132	5	0	15	1	1097	97	15	1	0	0
	PBT	95	268	268	1	0	6	2	248	93	12	4	1	0
White vs AmerIndian	CBT	186	1267	195	4	2	12	6	179	92	0	0	0	0
	PBT	95	268	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	186	1267	1042	0	0	12	1	1020	98	9	1	1	0
	PBT	95	268	92	3	3	4	4	81	88	4	4	0	0
White vs Black	CBT	186	1267	1132	0	0	15	1	1117	99	0	0	0	0
	PBT	95	268	268	3	1	24	9	241	90	0	0	0	0
White vs Hispanic	CBT	186	1267	1132	0	0	4	0	1128	100	0	0	0	0
	PBT	95	268	268	0	0	3	1	265	99	0	0	0	0
White vs Pacific Islander	CBT	186	1267	184	0	0	9	5	175	95	0	0	0	0
	PBT	95	268	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	186	1267	762	1	0	11	1	728	96	20	3	2	0
	PBT	95	268	22	0	0	0	0	22	100	0	0	0	0
NoEcnDis vs EcnDis	CBT	186	1267	1132	0	0	0	0	1131	100	1	0	0	0
	PBT	95	268	268	0	0	2	1	266	99	0	0	0	0
ELN vs ELY	CBT	186	1267	1042	4	0	31	3	1006	97	1	0	0	0
	PBT	95	268	89	0	0	9	10	80	90	0	0	0	0
SWDN vs SWDY	CBT	186	1267	1042	0	0	2	0	1040	100	0	0	0	0
	PBT	95	268	268	4	1	1	0	263	98	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.9 Differential Item Functioning for ELA/L Grade 11

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	161	1177	1042	6	1	39	4	975	94	15	1	7	1
	PBT	108	378	378	9	2	15	4	335	89	16	4	3	1
White vs AmerIndian	CBT	161	1177	138	11	8	18	13	109	79	0	0	0	0
	PBT	108	378	23	7	30	7	30	8	35	0	0	1	4
White vs Asian	CBT	161	1177	952	0	0	6	1	929	98	17	2	0	0
	PBT	108	378	92	0	0	4	4	83	90	5	5	0	0
White vs Black	CBT	161	1177	1042	2	0	4	0	1033	99	3	0	0	0
	PBT	108	378	378	7	2	24	6	341	90	5	1	1	0
White vs Hispanic	CBT	161	1177	1042	0	0	19	2	1022	98	0	0	1	0
	PBT	108	378	377	3	1	14	4	355	94	4	1	1	0
White vs Pacific Islander	CBT	161	1177	138	1	1	8	6	128	93	1	1	0	0
	PBT	108	378	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	161	1177	138	0	0	0	0	138	100	0	0	0	0
	PBT	108	378	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NoEcnDis vs EcnDis	CBT	161	1177	1042	0	0	1	0	1041	100	0	0	0	0
	PBT	108	378	378	0	0	11	3	366	97	1	0	0	0
ELN vs ELY	CBT	161	1177	952	23	2	43	5	879	92	7	1	0	0
	PBT	108	378	62	7	11	15	24	39	63	1	2	0	0
SWDN vs SWDY	CBT	161	1177	952	0	0	6	1	944	99	2	0	0	0
	PBT	108	378	92	2	2	1	1	87	95	2	2	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.10 Differential Item Functioning for Mathematics Grade 3

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	340	2632	2522	5	0	60	2	2414	96	36	1	7	0
	PBT	213	598	598	0	0	15	3	575	96	8	1	0	0
White vs AmerIndian	CBT	340	2632	112	0	0	4	4	108	96	0	0	0	0
	PBT	213	598	543	4	1	59	11	476	88	3	1	1	0
White vs Asian	CBT	340	2632	2410	3	0	1	0	2141	89	236	10	29	1
	PBT	213	598	543	3	1	0	0	500	92	36	7	4	1
White vs Black	CBT	340	2632	2410	40	2	228	9	2139	89	3	0	0	0
	PBT	213	598	543	9	2	62	11	472	87	0	0	0	0
White vs Hispanic	CBT	340	2632	2410	9	0	49	2	2351	98	1	0	0	0
	PBT	213	598	543	2	0	8	1	533	98	0	0	0	0
White vs Pacific Islander	CBT	340	2632	112	1	1	14	13	95	85	1	1	1	1
	PBT	213	598	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	340	2632	2410	1	0	34	1	2333	97	39	2	3	0
	PBT	213	598	543	0	0	6	1	533	98	4	1	0	0
NoEcnDis vs EcnDis	CBT	340	2632	2522	0	0	10	0	2510	100	2	0	0	0
	PBT	213	598	581	0	0	4	1	576	99	1	0	0	0
ELN vs ELY	CBT	340	2632	2410	8	0	76	3	2316	96	10	0	0	0
	PBT	213	598	543	0	0	10	2	532	98	1	0	0	0
SWDN vs SWDY	CBT	340	2632	2522	7	0	72	3	2435	97	5	0	3	0
	PBT	213	598	543	1	0	16	3	526	97	0	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.11 Differential Item Functioning for Mathematics Grade 4

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	266	2610	2504	3	0	56	2	2398	96	47	2	0	0
	PBT	188	532	532	0	0	2	0	520	98	10	2	0	0
White vs AmerIndian	CBT	266	2610	106	0	0	5	5	101	95	0	0	0	0
	PBT	188	532	479	9	2	50	10	412	86	5	1	3	1
White vs Asian	CBT	266	2610	2398	0	0	0	0	2251	94	139	6	8	0
	PBT	188	532	479	0	0	2	0	454	95	22	5	1	0
White vs Black	CBT	266	2610	2398	16	1	101	4	2279	95	2	0	0	0
	PBT	188	532	479	2	0	24	5	453	95	0	0	0	0
White vs Hispanic	CBT	266	2610	2398	2	0	38	2	2357	98	1	0	0	0
	PBT	188	532	479	0	0	6	1	473	99	0	0	0	0
White vs Pacific Islander	CBT	266	2610	106	0	0	9	8	96	91	0	0	1	1
	PBT	188	532	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	266	2610	2398	1	0	36	2	2332	97	27	1	2	0
	PBT	188	532	479	0	0	5	1	466	97	8	2	0	0
NoEcnDis vs EcnDis	CBT	266	2610	2504	0	0	12	0	2490	99	1	0	1	0
	PBT	188	532	479	0	0	2	0	477	100	0	0	0	0
ELN vs ELY	CBT	266	2610	2398	14	1	136	6	2245	94	3	0	0	0
	PBT	188	532	479	3	1	13	3	463	97	0	0	0	0
SWDN vs SWDY	CBT	266	2610	2398	3	0	62	3	2329	97	3	0	1	0
	PBT	188	532	479	0	0	8	2	468	98	3	1	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English Learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.12 Differential Item Functioning for Mathematics Grade 5

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	282	2688	2584	11	0	63	2	2503	97	7	0	0	0
	PBT	187	524	472	0	0	6	1	466	99	0	0	0	0
White vs AmerIndian	CBT	282	2688	104	0	0	3	3	101	97	0	0	0	0
	PBT	187	524	434	2	0	33	8	392	90	4	1	3	1
White vs Asian	CBT	282	2688	2480	1	0	6	0	2341	94	125	5	7	0
	PBT	187	524	472	0	0	0	0	455	96	15	3	2	0
White vs Black	CBT	282	2688	2480	41	2	127	5	2304	93	8	0	0	0
	PBT	187	524	472	1	0	18	4	453	96	0	0	0	0
White vs Hispanic	CBT	282	2688	2480	1	0	29	1	2450	99	0	0	0	0
	PBT	187	524	472	0	0	1	0	471	100	0	0	0	0
White vs Pacific Islander	CBT	282	2688	104	1	1	5	5	96	92	1	1	1	1
	PBT	187	524	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	282	2688	2447	0	0	33	1	2382	97	32	1	0	0
	PBT	187	524	472	0	0	7	1	461	98	4	1	0	0
NoEcnDis vs EcnDis	CBT	282	2688	2584	0	0	10	0	2570	99	4	0	0	0
	PBT	187	524	472	0	0	0	0	472	100	0	0	0	0
ELN vs ELY	CBT	282	2688	2480	14	1	105	4	2353	95	7	0	1	0
	PBT	187	524	472	0	0	12	3	460	97	0	0	0	0
SWDN vs SWDY	CBT	282	2688	2480	42	2	154	6	2249	91	5	0	30	1
	PBT	187	524	472	6	1	25	5	438	93	0	0	3	1

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.13 Differential Item Functioning for Mathematics Grade 6

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	264	2550	2448	20	1	63	3	2352	96	13	1	0	0
	PBT	183	510	459	0	0	10	2	447	97	2	0	0	0
White vs AmerIndian	CBT	264	2550	102	0	0	2	2	100	98	0	0	0	0
	PBT	183	510	214	6	3	19	9	188	88	0	0	1	0
White vs Asian	CBT	264	2550	2346	2	0	0	0	2204	94	122	5	18	1
	PBT	183	510	459	1	0	0	0	434	95	24	5	0	0
White vs Black	CBT	264	2550	2346	4	0	95	4	2247	96	0	0	0	0
	PBT	183	510	459	0	0	18	4	441	96	0	0	0	0
White vs Hispanic	CBT	264	2550	2346	0	0	9	0	2337	100	0	0	0	0
	PBT	183	510	459	0	0	3	1	456	99	0	0	0	0
White vs Pacific Islander	CBT	264	2550	102	0	0	9	9	92	90	1	1	0	0
	PBT	183	510	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	264	2550	2346	1	0	27	1	2288	98	30	1	0	0
	PBT	183	510	459	0	0	6	1	450	98	3	1	0	0
NoEcnDis vs EcnDis	CBT	264	2550	2448	0	0	5	0	2442	100	1	0	0	0
	PBT	183	510	459	0	0	0	0	459	100	0	0	0	0
ELN vs ELY	CBT	264	2550	2346	17	1	134	6	2190	93	3	0	2	0
	PBT	183	510	459	5	1	20	4	431	94	3	1	0	0
SWDN vs SWDY	CBT	264	2550	2346	67	3	170	7	2029	86	7	0	73	3
	PBT	183	510	459	22	5	35	8	393	86	3	1	6	1

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.14 Differential Item Functioning for Mathematics Grade 7

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	302	2495	2395	10	0	96	4	2265	95	23	1	1	0
	PBT	181	462	412	2	0	11	3	399	97	0	0	0	0
White vs AmerIndian	CBT	302	2495	100	1	1	5	5	94	94	0	0	0	0
	PBT	181	462	131	0	0	4	3	126	96	1	1	0	0
White vs Asian	CBT	302	2495	2295	5	0	4	0	2015	88	226	10	45	2
	PBT	181	462	412	0	0	4	1	360	87	37	9	11	3
White vs Black	CBT	302	2495	2295	12	1	134	6	2143	93	2	0	4	0
	PBT	181	462	412	2	0	26	6	382	93	1	0	1	0
White vs Hispanic	CBT	302	2495	2295	1	0	37	2	2256	98	1	0	0	0
	PBT	181	462	412	0	0	9	2	403	98	0	0	0	0
White vs Pacific Islander	CBT	302	2495	100	0	0	5	5	94	94	0	0	1	1
	PBT	181	462	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	302	2495	2275	1	0	21	1	2211	97	40	2	2	0
	PBT	181	462	412	0	0	4	1	406	99	2	0	0	0
NoEcnDis vs EcnDis	CBT	302	2495	2395	1	0	18	1	2373	99	3	0	0	0
	PBT	181	462	412	0	0	1	0	411	100	0	0	0	0
ELN vs ELY	CBT	302	2495	2295	25	1	112	5	2146	94	7	0	5	0
	PBT	181	462	412	6	1	22	5	382	93	1	0	1	0
SWDN vs SWDY	CBT	302	2495	2295	7	0	79	3	2201	96	8	0	0	0
	PBT	181	462	412	0	0	17	4	394	96	1	0	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.15 Differential Item Functioning for Mathematics Grade 8

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	265	2556	2454	20	1	28	1	2402	98	4	0	0	0
	PBT	168	470	419	5	1	11	3	403	96	0	0	0	0
White vs AmerIndian	CBT	265	2556	102	0	0	3	3	99	97	0	0	0	0
	PBT	168	470	82	0	0	6	7	74	90	2	2	0	0
White vs Asian	CBT	265	2556	2352	1	0	2	0	2071	88	220	9	58	2
	PBT	168	470	419	0	0	0	0	380	91	34	8	5	1
White vs Black	CBT	265	2556	2352	14	1	108	5	2212	94	16	1	2	0
	PBT	168	470	419	3	1	32	8	384	92	0	0	0	0
White vs Hispanic	CBT	265	2556	2352	0	0	28	1	2318	99	5	0	1	0
	PBT	168	470	419	0	0	1	0	417	100	1	0	0	0
White vs Pacific Islander	CBT	265	2556	102	0	0	4	4	96	94	2	2	0	0
	PBT	168	470	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	265	2556	438	1	0	8	2	421	96	7	2	1	0
	PBT	168	470	419	0	0	2	0	413	99	3	1	1	0
NoEcnDis vs EcnDis	CBT	265	2556	2454	2	0	27	1	2425	99	0	0	0	0
	PBT	168	470	419	0	0	2	0	417	100	0	0	0	0
ELN vs ELY	CBT	265	2556	2352	19	1	112	5	2189	93	27	1	5	0
	PBT	168	470	419	4	1	29	7	381	91	2	0	3	1
SWDN vs SWDY	CBT	265	2556	2352	36	2	99	4	2163	92	34	1	20	1
	PBT	168	470	419	6	1	18	4	389	93	3	1	3	1

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.16 Differential Item Functioning for Algebra I

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	343	3070	2964	5	0	68	2	2856	96	34	1	1	0
	PBT	186	547	494	0	0	3	1	485	98	6	1	0	0
White vs AmerIndian	CBT	343	3070	106	0	0	4	4	100	94	2	2	0	0
	PBT	186	547	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	343	3070	2860	2	0	8	0	2497	87	282	10	71	2
	PBT	186	547	494	1	0	3	1	454	92	34	7	2	0
White vs Black	CBT	343	3070	2860	9	0	94	3	2733	96	22	1	2	0
	PBT	186	547	494	5	1	25	5	464	94	0	0	0	0
White vs Hispanic	CBT	343	3070	2860	2	0	25	1	2824	99	7	0	2	0
	PBT	186	547	494	0	0	5	1	486	98	3	1	0	0
White vs Pacific Islander	CBT	343	3070	106	1	1	10	9	91	86	3	3	1	1
	PBT	186	547	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	343	3070	1553	0	0	16	1	1509	97	26	2	2	0
	PBT	186	547	492	0	0	13	3	468	95	8	2	3	1
NoEcnDis vs EcnDis	CBT	343	3070	2964	4	0	31	1	2928	99	1	0	0	0
	PBT	186	547	494	0	0	1	0	493	100	0	0	0	0
ELN vs ELY	CBT	343	3070	2912	22	1	76	3	2770	95	27	1	17	1
	PBT	186	547	494	0	0	26	5	457	93	9	2	2	0
SWDN vs SWDY	CBT	343	3070	2860	14	0	68	2	2749	96	25	1	4	0
	PBT	186	547	494	2	0	23	5	465	94	4	1	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.17 Differential Item Functioning for Geometry

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	346	3072	2913	21	1	86	3	2771	95	34	1	1	0
	PBT	189	529	476	3	1	22	5	443	93	8	2	0	0
White vs AmerIndian	CBT	346	3072	106	2	2	4	4	99	93	1	1	0	0
	PBT	189	529	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	346	3072	2860	2	0	12	0	2541	89	238	8	67	2
	PBT	189	529	195	0	0	2	1	176	90	11	6	6	3
White vs Black	CBT	346	3072	2860	23	1	128	4	2689	94	19	1	1	0
	PBT	189	529	476	4	1	29	6	440	92	2	0	1	0
White vs Hispanic	CBT	346	3072	2860	5	0	55	2	2796	98	4	0	0	0
	PBT	189	529	476	2	0	19	4	455	96	0	0	0	0
White vs Pacific Islander	CBT	346	3072	51	2	4	2	4	45	88	1	2	1	2
	PBT	189	529	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	346	3072	106	0	0	1	1	105	99	0	0	0	0
	PBT	189	529	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NoEcnDis vs EcnDis	CBT	346	3072	2913	0	0	67	2	2845	98	1	0	0	0
	PBT	189	529	476	0	0	8	2	468	98	0	0	0	0
ELN vs ELY	CBT	346	3072	2540	27	1	132	5	2340	92	32	1	9	0
	PBT	189	529	15	0	0	2	13	12	80	1	7	0	0
SWDN vs SWDY	CBT	346	3072	2860	26	1	147	5	2658	93	25	1	4	0
	PBT	189	529	161	0	0	11	7	147	91	3	2	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.18 Differential Item Functioning for Algebra II

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	275	2870	2710	10	0	42	2	2616	97	38	1	4	0
	PBT	189	528	474	0	0	0	0	462	97	11	2	1	0
White vs AmerIndian	CBT	275	2870	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	189	528	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	275	2870	2658	4	0	13	0	2113	79	423	16	105	4
	PBT	189	528	118	1	1	3	3	106	90	6	5	2	2
White vs Black	CBT	275	2870	2658	6	0	70	3	2574	97	3	0	5	0
	PBT	189	528	474	2	0	12	3	449	95	10	2	1	0
White vs Hispanic	CBT	275	2870	2658	3	0	67	3	2580	97	7	0	1	0
	PBT	189	528	474	0	0	12	3	455	96	7	1	0	0
White vs Pacific Islander	CBT	275	2870	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	189	528	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	275	2870	106	1	1	1	1	103	97	1	1	0	0
	PBT	189	528	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NoEcnDis vs EcnDis	CBT	275	2870	2710	1	0	52	2	2653	98	4	0	0	0
	PBT	189	528	474	0	0	2	0	471	99	1	0	0	0
ELN vs ELY	CBT	275	2870	431	4	1	23	5	393	91	7	2	4	1
	PBT	189	528	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SWDN vs SWDY	CBT	275	2870	2658	15	1	85	3	2536	95	18	1	4	0
	PBT	189	528	98	2	2	3	3	91	93	2	2	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.19 Differential Item Functioning for Integrated Mathematics I

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	144	508	310	2	1	5	2	298	96	5	2	0	0
	PBT	119	278	104	0	0	2	2	98	94	3	3	1	1
White vs AmerIndian	CBT	144	508	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	119	278	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	144	508	216	1	0	1	0	202	94	10	5	2	1
	PBT	119	278	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Black	CBT	144	508	310	1	0	12	4	291	94	6	2	0	0
	PBT	119	278	50	0	0	0	0	48	96	2	4	0	0
White vs Hispanic	CBT	144	508	310	1	0	13	4	295	95	1	0	0	0
	PBT	119	278	36	0	0	1	3	35	97	0	0	0	0
White vs Pacific Islander	CBT	144	508	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	119	278	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	144	508	46	0	0	1	2	44	96	1	2	0	0
	PBT	119	278	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NoEcnDis vs EcnDis	CBT	144	508	310	3	1	3	1	304	98	0	0	0	0
	PBT	119	278	68	0	0	2	3	66	97	0	0	0	0
ELN vs ELY	CBT	144	508	310	4	1	12	4	293	95	0	0	1	0
	PBT	119	278	16	0	0	0	0	15	94	1	6	0	0
SWDN vs SWDY	CBT	144	508	310	1	0	12	4	295	95	2	1	0	0
	PBT	119	278	68	3	4	9	13	55	81	1	1	0	0

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.20 Differential Item Functioning for Integrated Mathematics II

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	148	570	364	0	0	1	0	356	98	6	2	1	0
	PBT	126	325	36	0	0	0	0	35	97	1	3	0	0
White vs AmerIndian	CBT	148	570	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	148	570	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Black	CBT	148	570	253	2	1	13	5	235	93	2	1	1	0
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Hispanic	CBT	148	570	364	2	1	6	2	355	98	1	0	0	0
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Pacific Islander	CBT	148	570	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	148	570	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NoEcnDis vs EcnDis	CBT	148	570	364	0	0	5	1	358	98	1	0	0	0
	PBT	126	325	36	1	3	1	3	33	92	1	3	0	0
ELN vs ELY	CBT	148	570	262	4	2	14	5	240	92	4	2	0	0
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SWDN vs SWDY	CBT	148	570	284	6	2	20	7	257	90	1	0	0	0
	PBT	126	325	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.21 Differential Item Functioning for Integrated Mathematics III

DIF Comparisons	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
Male vs Female	CBT	163	453	178	1	1	3	2	172	97	2	1	0	0
	PBT	124	250	20	0	0	1	5	18	90	1	5	0	0
White vs AmerIndian	CBT	163	453	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	124	250	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Asian	CBT	163	453	40	0	0	3	8	33	83	2	5	2	5
	PBT	124	250	19	1	5	1	5	15	79	2	11	0	0
White vs Black	CBT	163	453	40	0	0	1	3	38	95	0	0	1	3
	PBT	124	250	19	0	0	3	16	16	84	0	0	0	0
White vs Hispanic	CBT	163	453	178	0	0	1	1	177	99	0	0	0	0
	PBT	124	250	20	0	0	2	10	18	90	0	0	0	0
White vs Pacific Islander	CBT	163	453	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	124	250	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
White vs Multiracial	CBT	163	453	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	PBT	124	250	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
NoEcnDis vs EcnDis	CBT	163	453	178	0	0	1	1	177	99	0	0	0	0
	PBT	124	250	20	0	0	0	0	20	100	0	0	0	0
ELN vs ELY	CBT	163	453	40	0	0	2	5	38	95	0	0	0	0
	PBT	124	250	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SWDN vs SWDY	CBT	163	453	40	0	0	3	8	36	90	0	0	1	3
	PBT	124	250	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability.

Table A.7.22 Differential Item Functioning for Mathematics: Spanish-Language vs. English-Language Forms

Grade	Mode	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
3	CBT	56	112	112	4	4	9	8	99	88	0	0	0	0
	PBT	55	55	55	1	2	4	7	49	89	1	2	0	0
4	CBT	53	106	106	9	8	17	16	80	75	0	0	0	0
	PBT	53	53	53	1	2	2	4	50	94	0	0	0	0
5	CBT	52	104	104	8	8	11	11	83	80	0	0	2	2
	PBT	52	52	52	2	4	2	4	47	90	0	0	1	2
6	CBT	51	102	102	11	11	13	13	78	76	0	0	0	0
	PBT	51	51	51	3	6	7	14	41	80	0	0	0	0
7	CBT	50	100	100	11	11	16	16	70	70	2	2	1	1
	PBT	50	50	50	3	6	2	4	43	86	1	2	1	2
8	CBT	51	102	102	3	3	4	4	95	93	0	0	0	0
	PBT	51	51	51	0	0	2	4	48	94	0	0	1	2
A1	CBT	52	104	104	1	1	8	8	95	91	0	0	0	0
	PBT	53	53	53	2	4	8	15	40	75	3	6	0	0
GO	CBT	53	106	106	3	3	7	7	96	91	0	0	0	0
	PBT	0	0	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A2	CBT	52	104	104	8	8	10	10	85	82	0	0	1	1
	PBT	0	0	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M1	CBT	49	98	67	0	0	5	7	58	87	3	4	1	1
	PBT	0	0	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I.

Appendix 8: Reliability of Classification by Content and Grade Level

Table A.8.1 Reliability of Classification: Grade 3 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category Total
		Summative Scale Score						
CBT		650 – 699	0.16	0.04	0.00	0.00	0.00	0.20
	Decision	700 – 724	0.03	0.13	0.04	0.00	0.00	0.20
	Accuracy	725 – 749	0.00	0.05	0.14	0.05	0.00	0.24
		750 – 809	0.00	0.00	0.05	0.28	0.00	0.33
		810 – 850	0.00	0.00	0.00	0.02	0.01	0.03
		650 – 699	0.16	0.04	0.00	0.00	0.00	0.20
	Decision	700 – 724	0.04	0.10	0.05	0.00	0.00	0.20
	Consistency	725 – 749	0.01	0.06	0.11	0.06	0.00	0.24
		750 – 809	0.00	0.01	0.06	0.26	0.01	0.33
		810 – 850	0.00	0.00	0.00	0.01	0.01	0.03
PBT		650 – 699	0.14	0.04	0.00	0.00	0.00	0.18
	Decision	700 – 724	0.03	0.12	0.05	0.00	0.00	0.19
	Accuracy	725 – 749	0.00	0.04	0.13	0.04	0.00	0.22
		750 – 809	0.00	0.00	0.05	0.30	0.01	0.36
		810 – 850	0.00	0.00	0.00	0.02	0.03	0.05
		650 – 699	0.13	0.04	0.00	0.00	0.00	0.18
	Decision	700 – 724	0.04	0.09	0.05	0.01	0.00	0.19
	Consistency	725 – 749	0.01	0.05	0.10	0.06	0.00	0.22
		750 – 809	0.00	0.01	0.06	0.27	0.02	0.36
		810 – 850	0.00	0.00	0.00	0.02	0.03	0.05

Table A.8.2 Reliability of Classification: Grade 4 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.09	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.01	0.14	0.04	0.00	0.00	0.19
	Accuracy	725 – 749	0.00	0.04	0.19	0.05	0.00	0.29
		750 – 789	0.00	0.00	0.05	0.27	0.01	0.33
		790 – 850	0.00	0.00	0.00	0.02	0.05	0.07
		650 – 699	0.09	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.03	0.12	0.05	0.00	0.00	0.19
	Consistency	725 – 749	0.01	0.06	0.16	0.07	0.00	0.29
		750 – 789	0.00	0.00	0.06	0.24	0.03	0.33
		790 – 850	0.00	0.00	0.00	0.02	0.05	0.07
PBT		650 – 699	0.05	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.01	0.12	0.04	0.00	0.00	0.17
	Accuracy	725 – 749	0.00	0.04	0.19	0.05	0.00	0.29
		750 – 789	0.00	0.00	0.05	0.29	0.02	0.37
		790 – 850	0.00	0.00	0.00	0.03	0.07	0.09
		650 – 699	0.05	0.03	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.02	0.10	0.05	0.00	0.00	0.17
	Consistency	725 – 749	0.01	0.06	0.15	0.07	0.00	0.29
		750 – 789	0.00	0.00	0.07	0.26	0.04	0.37
		790 – 850	0.00	0.00	0.00	0.03	0.06	0.09

Table A.8.3 Reliability of Classification: Grade 5 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.08	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.01	0.15	0.04	0.00	0.00	0.21
	Accuracy	725 – 749	0.00	0.04	0.20	0.04	0.00	0.29
		750 – 798	0.00	0.00	0.05	0.31	0.01	0.36
		799 – 850	0.00	0.00	0.00	0.01	0.02	0.03
		650 – 699	0.08	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.03	0.13	0.05	0.00	0.00	0.21
	Consistency	725 – 749	0.00	0.06	0.16	0.06	0.00	0.29
		750 – 798	0.00	0.00	0.06	0.29	0.01	0.36
		799 – 850	0.00	0.00	0.00	0.01	0.02	0.03
PBT		650 – 699	0.06	0.03	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.01	0.13	0.05	0.00	0.00	0.19
	Accuracy	725 – 749	0.00	0.04	0.20	0.05	0.00	0.29
		750 – 798	0.00	0.00	0.05	0.34	0.01	0.40
		799 – 850	0.00	0.00	0.00	0.02	0.02	0.04
		650 – 699	0.06	0.03	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.03	0.11	0.05	0.00	0.00	0.19
	Consistency	725 – 749	0.00	0.06	0.16	0.07	0.00	0.29
		750 – 798	0.00	0.00	0.07	0.31	0.02	0.40
		799 – 850	0.00	0.00	0.00	0.02	0.02	0.04

Table A.8.4 Reliability of Classification: Grade 6 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.09	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.01	0.16	0.04	0.00	0.00	0.20
	Accuracy	725 – 749	0.00	0.04	0.22	0.04	0.00	0.31
		750 – 789	0.00	0.00	0.04	0.28	0.01	0.33
		790 – 850	0.00	0.00	0.00	0.02	0.03	0.05
		650 – 699	0.08	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.02	0.13	0.05	0.00	0.00	0.20
	Consistency	725 – 749	0.00	0.06	0.18	0.06	0.00	0.31
		750 – 789	0.00	0.00	0.06	0.25	0.02	0.33
		790 – 850	0.00	0.00	0.00	0.02	0.03	0.05
PBT		650 – 699	0.05	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.01	0.14	0.04	0.00	0.00	0.18
	Accuracy	725 – 749	0.00	0.03	0.22	0.05	0.00	0.30
		750 – 789	0.00	0.00	0.05	0.31	0.01	0.38
		790 – 850	0.00	0.00	0.00	0.02	0.05	0.07
		650 – 699	0.05	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.02	0.12	0.05	0.00	0.00	0.18
	Consistency	725 – 749	0.00	0.05	0.18	0.06	0.00	0.30
		750 – 789	0.00	0.00	0.07	0.28	0.03	0.38
		790 – 850	0.00	0.00	0.00	0.02	0.05	0.07

Table A.8.5 Reliability of Classification: Grade 7 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.12	0.02	0.00	0.00	0.00	0.15
	Decision	700 – 724	0.02	0.14	0.03	0.00	0.00	0.19
	Accuracy	725 – 749	0.00	0.04	0.17	0.04	0.00	0.26
		750 – 784	0.00	0.00	0.04	0.23	0.02	0.30
		785 – 850	0.00	0.00	0.00	0.02	0.08	0.10
		650 – 699	0.12	0.03	0.00	0.00	0.00	0.15
	Decision	700 – 724	0.03	0.12	0.04	0.00	0.00	0.19
	Consistency	725 – 749	0.01	0.05	0.14	0.06	0.00	0.26
		750 – 784	0.00	0.00	0.05	0.20	0.04	0.30
		785 – 850	0.00	0.00	0.00	0.03	0.08	0.10
PBT		650 – 699	0.08	0.02	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.01	0.11	0.03	0.00	0.00	0.16
	Accuracy	725 – 749	0.00	0.04	0.17	0.05	0.00	0.26
		750 – 784	0.00	0.00	0.05	0.27	0.03	0.35
		785 – 850	0.00	0.00	0.00	0.03	0.11	0.13
		650 – 699	0.08	0.03	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.02	0.09	0.04	0.00	0.00	0.16
	Consistency	725 – 749	0.01	0.05	0.13	0.07	0.00	0.26
		750 – 784	0.00	0.00	0.06	0.23	0.05	0.35
		785 – 850	0.00	0.00	0.00	0.03	0.10	0.13

Table A.8.6 Reliability of Classification: Grade 8 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.13	0.02	0.00	0.00	0.00	0.15
	Decision	700 – 724	0.02	0.13	0.04	0.00	0.00	0.19
	Accuracy	725 – 749	0.01	0.04	0.16	0.05	0.00	0.26
		750 – 793	0.00	0.00	0.04	0.28	0.01	0.34
		794 – 850	0.00	0.00	0.00	0.02	0.05	0.07
		650 – 699	0.12	0.03	0.00	0.00	0.00	0.15
	Decision	700 – 724	0.03	0.11	0.05	0.00	0.00	0.19
	Consistency	725 – 749	0.01	0.06	0.13	0.06	0.00	0.26
		750 – 793	0.00	0.00	0.06	0.25	0.03	0.34
		794 – 850	0.00	0.00	0.00	0.03	0.04	0.07
PBT		650 – 699	0.08	0.02	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.01	0.11	0.03	0.00	0.00	0.16
	Accuracy	725 – 749	0.00	0.04	0.16	0.05	0.00	0.26
		750 – 793	0.00	0.00	0.05	0.32	0.02	0.39
		794 – 850	0.00	0.00	0.00	0.03	0.07	0.09
		650 – 699	0.08	0.02	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.03	0.09	0.04	0.00	0.00	0.16
	Consistency	725 – 749	0.01	0.05	0.13	0.07	0.00	0.26
		750 – 793	0.00	0.00	0.06	0.29	0.04	0.39
		794 – 850	0.00	0.00	0.00	0.03	0.06	0.09

Table A.8.7 Reliability of Classification: Grade 9 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.13	0.03	0.00	0.00	0.00	0.16
	Decision	700 – 724	0.02	0.14	0.04	0.00	0.00	0.20
	Accuracy	725 – 749	0.00	0.04	0.17	0.04	0.00	0.26
		750 – 790	0.00	0.00	0.04	0.25	0.01	0.31
		791 – 850	0.00	0.00	0.00	0.02	0.05	0.07
		650 – 699	0.13	0.03	0.00	0.00	0.00	0.16
	Decision	700 – 724	0.04	0.12	0.05	0.00	0.00	0.20
	Consistency	725 – 749	0.01	0.06	0.13	0.06	0.00	0.26
		750 – 790	0.00	0.00	0.06	0.22	0.03	0.31
		791 – 850	0.00	0.00	0.00	0.02	0.05	0.07
PBT		650 – 699	0.06	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.01	0.10	0.03	0.00	0.00	0.14
	Accuracy	725 – 749	0.00	0.03	0.15	0.05	0.00	0.24
		750 – 790	0.00	0.00	0.05	0.32	0.03	0.41
		791 – 850	0.00	0.00	0.00	0.03	0.10	0.13
		650 – 699	0.05	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.02	0.08	0.04	0.00	0.00	0.14
	Consistency	725 – 749	0.00	0.05	0.12	0.07	0.00	0.24
		750 – 790	0.00	0.01	0.07	0.29	0.05	0.41
		791 – 850	0.00	0.00	0.00	0.04	0.09	0.13

Table A.8.8 Reliability of Classification: Grade 10 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.21	0.03	0.00	0.00	0.00	0.23
	Decision	700 – 724	0.03	0.12	0.03	0.00	0.00	0.18
	Accuracy	725 – 749	0.01	0.04	0.12	0.04	0.00	0.21
		750 – 793	0.00	0.00	0.04	0.21	0.02	0.27
		794 – 850	0.00	0.00	0.00	0.02	0.08	0.10
		650 – 699	0.20	0.03	0.00	0.00	0.00	0.23
	Decision	700 – 724	0.04	0.10	0.04	0.00	0.00	0.18
	Consistency	725 – 749	0.01	0.05	0.09	0.05	0.00	0.21
		750 – 793	0.00	0.01	0.05	0.18	0.03	0.27
		794 – 850	0.00	0.00	0.00	0.02	0.08	0.10
PBT		650 – 699	0.15	0.03	0.00	0.00	0.00	0.17
	Decision	700 – 724	0.03	0.11	0.04	0.00	0.00	0.17
	Accuracy	725 – 749	0.01	0.04	0.12	0.05	0.00	0.21
		750 – 793	0.00	0.00	0.04	0.23	0.03	0.30
		794 – 850	0.00	0.00	0.00	0.02	0.12	0.14
		650 – 699	0.14	0.03	0.00	0.00	0.00	0.17
	Decision	700 – 724	0.04	0.08	0.04	0.00	0.00	0.17
	Consistency	725 – 749	0.01	0.05	0.09	0.06	0.00	0.21
		750 – 793	0.00	0.01	0.05	0.20	0.04	0.30
		794 – 850	0.00	0.00	0.00	0.03	0.11	0.14

Table A.8.9 Reliability of Classification: Grade 11 ELA/L

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.14	0.03	0.00	0.00	0.00	0.17
	Decision	700 – 724	0.02	0.14	0.04	0.00	0.00	0.19
	Accuracy	725 – 749	0.00	0.04	0.15	0.04	0.00	0.24
		750 – 791	0.00	0.00	0.05	0.23	0.02	0.30
		792 – 850	0.00	0.00	0.00	0.02	0.08	0.09
		650 – 699	0.13	0.04	0.00	0.00	0.00	0.17
	Decision	700 – 724	0.04	0.11	0.05	0.00	0.00	0.19
	Consistency	725 – 749	0.01	0.06	0.12	0.06	0.00	0.24
		750 – 791	0.00	0.01	0.06	0.21	0.03	0.30
		792 – 850	0.00	0.00	0.00	0.02	0.07	0.09
PBT		650 – 699	0.15	0.04	0.00	0.00	0.00	0.19
	Decision	700 – 724	0.03	0.13	0.04	0.00	0.00	0.20
	Accuracy	725 – 749	0.01	0.05	0.14	0.04	0.00	0.24
		750 – 791	0.00	0.00	0.05	0.21	0.02	0.28
		792 – 850	0.00	0.00	0.00	0.02	0.07	0.09
		650 – 699	0.14	0.04	0.00	0.00	0.00	0.19
	Decision	700 – 724	0.04	0.11	0.05	0.00	0.00	0.20
	Consistency	725 – 749	0.01	0.06	0.11	0.06	0.00	0.24
		750 – 791	0.00	0.01	0.06	0.19	0.03	0.28
		792 – 850	0.00	0.00	0.00	0.02	0.07	0.09

Table A.8.10 Reliability of Classification: Grade 3 Mathematics

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.10	0.02	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.02	0.17	0.03	0.00	0.00	0.22
	Accuracy	725 – 749	0.00	0.04	0.20	0.04	0.00	0.28
		750 – 789	0.00	0.00	0.04	0.28	0.01	0.32
		790 – 850	0.00	0.00	0.00	0.01	0.04	0.06
		650 – 699	0.10	0.02	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.03	0.14	0.04	0.00	0.00	0.22
	Consistency	725 – 749	0.00	0.05	0.17	0.05	0.00	0.28
		750 – 789	0.00	0.00	0.05	0.25	0.02	0.32
		790 – 850	0.00	0.00	0.00	0.02	0.04	0.06
PBT		650 – 699	0.11	0.02	0.00	0.00	0.00	0.14
	Decision	700 – 724	0.02	0.17	0.03	0.00	0.00	0.22
	Accuracy	725 – 749	0.00	0.04	0.20	0.04	0.00	0.27
		750 – 789	0.00	0.00	0.04	0.25	0.01	0.30
		790 – 850	0.00	0.00	0.00	0.01	0.05	0.06
		650 – 699	0.11	0.03	0.00	0.00	0.00	0.14
	Decision	700 – 724	0.03	0.15	0.05	0.00	0.00	0.22
	Consistency	725 – 749	0.00	0.05	0.17	0.05	0.00	0.27
		750 – 789	0.00	0.00	0.05	0.23	0.02	0.30
		790 – 850	0.00	0.00	0.00	0.02	0.05	0.06

Table A.8.11 Reliability of Classification: Grade 4 Mathematics

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.09	0.02	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.02	0.20	0.04	0.00	0.00	0.26
	Accuracy	725 – 749	0.00	0.04	0.23	0.03	0.00	0.30
		750 – 795	0.00	0.00	0.04	0.26	0.00	0.30
		796 – 850	0.00	0.00	0.00	0.01	0.02	0.03
		650 – 699	0.09	0.03	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.03	0.18	0.05	0.00	0.00	0.26
	Consistency	725 – 749	0.00	0.05	0.20	0.05	0.00	0.30
		750 – 795	0.00	0.00	0.05	0.24	0.01	0.30
		796 – 850	0.00	0.00	0.00	0.01	0.02	0.03
PBT		650 – 699	0.09	0.02	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.02	0.21	0.04	0.00	0.00	0.27
	Accuracy	725 – 749	0.00	0.04	0.22	0.03	0.00	0.30
		750 – 795	0.00	0.00	0.03	0.25	0.00	0.29
		796 – 850	0.00	0.00	0.00	0.01	0.02	0.03
		650 – 699	0.09	0.03	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.03	0.18	0.05	0.00	0.00	0.27
	Consistency	725 – 749	0.00	0.05	0.19	0.05	0.00	0.30
		750 – 795	0.00	0.00	0.04	0.23	0.01	0.29
		796 – 850	0.00	0.00	0.00	0.01	0.02	0.03

Table A.8.12 Reliability of Classification: Grade 5 Mathematics

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.09	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.02	0.20	0.04	0.00	0.00	0.26
	Accuracy	725 – 749	0.00	0.04	0.23	0.04	0.00	0.31
		750 – 789	0.00	0.00	0.04	0.24	0.01	0.28
		790 – 850	0.00	0.00	0.00	0.01	0.03	0.04
		650 – 699	0.09	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.03	0.17	0.06	0.00	0.00	0.26
	Consistency	725 – 749	0.00	0.06	0.20	0.05	0.00	0.31
		750 – 789	0.00	0.00	0.05	0.22	0.01	0.28
		790 – 850	0.00	0.00	0.00	0.01	0.03	0.04
PBT		650 – 699	0.10	0.03	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.02	0.21	0.04	0.00	0.00	0.28
	Accuracy	725 – 749	0.00	0.04	0.22	0.03	0.00	0.29
		750 – 789	0.00	0.00	0.03	0.22	0.01	0.27
		790 – 850	0.00	0.00	0.00	0.01	0.04	0.05
		650 – 699	0.09	0.03	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.04	0.19	0.05	0.00	0.00	0.28
	Consistency	725 – 749	0.00	0.05	0.19	0.05	0.00	0.29
		750 – 789	0.00	0.00	0.04	0.21	0.01	0.27
		790 – 850	0.00	0.00	0.00	0.01	0.04	0.05

Table A.8.13 Reliability of Classification: Grade 6 Mathematics

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.10	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.02	0.20	0.04	0.00	0.00	0.26
	Accuracy	725 – 749	0.00	0.04	0.24	0.03	0.00	0.31
		750 – 787	0.00	0.00	0.04	0.24	0.01	0.28
		788 – 850	0.00	0.00	0.00	0.01	0.02	0.03
		650 – 699	0.09	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.03	0.18	0.05	0.00	0.00	0.26
	Consistency	725 – 749	0.00	0.05	0.21	0.05	0.00	0.31
		750 – 787	0.00	0.00	0.05	0.22	0.01	0.28
		788 – 850	0.00	0.00	0.00	0.01	0.02	0.03
PBT		650 – 699	0.09	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.02	0.20	0.04	0.00	0.00	0.26
	Accuracy	725 – 749	0.00	0.04	0.22	0.03	0.00	0.29
		750 – 787	0.00	0.00	0.04	0.23	0.01	0.28
		788 – 850	0.00	0.00	0.00	0.01	0.04	0.05
		650 – 699	0.09	0.02	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.03	0.17	0.05	0.00	0.00	0.26
	Consistency	725 – 749	0.00	0.05	0.19	0.05	0.00	0.29
		750 – 787	0.00	0.00	0.05	0.22	0.02	0.28
		788 – 850	0.00	0.00	0.00	0.01	0.04	0.05

Table A.8.14 Reliability of Classification: Grade 7 Mathematics

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.08	0.01	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.02	0.21	0.04	0.00	0.00	0.28
	Accuracy	725 – 749	0.00	0.04	0.26	0.04	0.00	0.34
		750 – 785	0.00	0.00	0.03	0.23	0.00	0.26
		786 – 850	0.00	0.00	0.00	0.01	0.01	0.02
		650 – 699	0.08	0.02	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.03	0.19	0.06	0.00	0.00	0.28
	Consistency	725 – 749	0.00	0.06	0.23	0.05	0.00	0.34
		750 – 785	0.00	0.00	0.05	0.21	0.01	0.26
		786 – 850	0.00	0.00	0.00	0.01	0.01	0.02
PBT		650 – 699	0.06	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.02	0.22	0.04	0.00	0.00	0.27
	Accuracy	725 – 749	0.00	0.04	0.26	0.04	0.00	0.34
		750 – 785	0.00	0.00	0.04	0.22	0.01	0.26
		786 – 850	0.00	0.00	0.00	0.01	0.03	0.04
		650 – 699	0.06	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.03	0.19	0.06	0.00	0.00	0.27
	Consistency	725 – 749	0.00	0.06	0.23	0.05	0.00	0.34
		750 – 785	0.00	0.00	0.05	0.20	0.01	0.26
		786 – 850	0.00	0.00	0.00	0.01	0.03	0.04

Table A.8.15 Reliability of Classification: Grade 8 Mathematics

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.19	0.03	0.00	0.00	0.00	0.23
	Decision	700 – 724	0.04	0.18	0.05	0.00	0.00	0.26
	Accuracy	725 – 749	0.00	0.05	0.16	0.04	0.00	0.25
		750 – 800	0.00	0.00	0.04	0.20	0.00	0.24
		801 – 850	0.00	0.00	0.00	0.01	0.01	0.02
		650 – 699	0.18	0.04	0.00	0.00	0.00	0.23
	Decision	700 – 724	0.05	0.15	0.06	0.00	0.00	0.26
	Consistency	725 – 749	0.00	0.06	0.13	0.05	0.00	0.25
		750 – 800	0.00	0.00	0.05	0.18	0.01	0.24
		801 – 850	0.00	0.00	0.00	0.01	0.01	0.02
PBT		650 – 699	0.15	0.03	0.00	0.00	0.00	0.18
	Decision	700 – 724	0.03	0.16	0.04	0.00	0.00	0.23
	Accuracy	725 – 749	0.00	0.04	0.16	0.04	0.00	0.25
		750 – 800	0.00	0.00	0.04	0.25	0.01	0.29
		801 – 850	0.00	0.00	0.00	0.01	0.03	0.04
		650 – 699	0.14	0.04	0.00	0.00	0.00	0.18
	Decision	700 – 724	0.04	0.13	0.06	0.00	0.00	0.23
	Consistency	725 – 749	0.00	0.06	0.13	0.05	0.00	0.25
		750 – 800	0.00	0.00	0.05	0.23	0.01	0.29
		801 – 850	0.00	0.00	0.00	0.01	0.03	0.04

Table A.8.16 Reliability of Classification: Algebra I

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.11	0.02	0.00	0.00	0.00	0.13
	Decision	700 – 724	0.03	0.20	0.05	0.00	0.00	0.29
	Accuracy	725 – 749	0.00	0.05	0.20	0.04	0.00	0.28
		750 – 804	0.00	0.00	0.04	0.24	0.00	0.28
		805 – 850	0.00	0.00	0.00	0.01	0.01	0.02
		650 – 699	0.10	0.03	0.00	0.00	0.00	0.13
	Decision	700 – 724	0.05	0.17	0.07	0.01	0.00	0.29
	Consistency	725 – 749	0.00	0.06	0.16	0.05	0.00	0.28
		750 – 804	0.00	0.00	0.05	0.22	0.00	0.28
		805 – 850	0.00	0.00	0.00	0.01	0.01	0.02
PBT		650 – 699	0.08	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.02	0.16	0.05	0.00	0.00	0.23
	Accuracy	725 – 749	0.00	0.04	0.19	0.04	0.00	0.28
		750 – 804	0.00	0.00	0.05	0.31	0.00	0.36
		805 – 850	0.00	0.00	0.00	0.01	0.02	0.03
		650 – 699	0.08	0.03	0.00	0.00	0.00	0.11
	Decision	700 – 724	0.04	0.13	0.06	0.00	0.00	0.23
	Consistency	725 – 749	0.00	0.06	0.16	0.06	0.00	0.28
		750 – 804	0.00	0.00	0.06	0.28	0.01	0.36
		805 – 850	0.00	0.00	0.00	0.01	0.02	0.03

Table A.8.17 Reliability of Classification: Geometry

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.08	0.02	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.03	0.25	0.05	0.01	0.00	0.33
	Accuracy	725 – 749	0.00	0.03	0.26	0.02	0.00	0.32
		750 – 782	0.00	0.00	0.04	0.18	0.00	0.23
		783 – 850	0.00	0.00	0.00	0.01	0.02	0.02
		650 – 699	0.08	0.02	0.00	0.00	0.00	0.10
	Decision	700 – 724	0.04	0.23	0.06	0.01	0.00	0.33
	Consistency	725 – 749	0.00	0.05	0.23	0.04	0.00	0.32
		750 – 782	0.00	0.00	0.05	0.17	0.01	0.23
		783 – 850	0.00	0.00	0.00	0.01	0.02	0.02
PBT		650 – 699	0.04	0.01	0.00	0.00	0.00	0.05
	Decision	700 – 724	0.01	0.14	0.04	0.00	0.00	0.19
	Accuracy	725 – 749	0.00	0.02	0.23	0.04	0.00	0.29
		750 – 782	0.00	0.00	0.04	0.33	0.02	0.40
		783 – 850	0.00	0.00	0.00	0.02	0.06	0.08
		650 – 699	0.04	0.01	0.00	0.00	0.00	0.05
	Decision	700 – 724	0.02	0.12	0.04	0.00	0.00	0.19
	Consistency	725 – 749	0.00	0.04	0.20	0.05	0.00	0.29
		750 – 782	0.00	0.00	0.06	0.30	0.03	0.40
		783 – 850	0.00	0.00	0.00	0.03	0.05	0.08

Table A.8.18 Reliability of Classification: Algebra II

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.29	0.03	0.00	0.00	0.00	0.32
	Decision	700 – 724	0.04	0.17	0.05	0.00	0.00	0.27
	Accuracy	725 – 749	0.00	0.04	0.13	0.03	0.00	0.21
		750 – 807	0.00	0.00	0.03	0.17	0.00	0.20
		808 – 850	0.00	0.00	0.00	0.01	0.00	0.01
		650 – 699	0.27	0.04	0.00	0.00	0.00	0.32
	Decision	700 – 724	0.06	0.14	0.06	0.01	0.00	0.26
	Consistency	725 – 749	0.00	0.05	0.11	0.04	0.00	0.21
		750 – 807	0.00	0.00	0.04	0.16	0.00	0.20
		808 – 850	0.00	0.00	0.00	0.01	0.00	0.01
PBT		650 – 699	0.25	0.03	0.00	0.00	0.00	0.28
	Decision	700 – 724	0.05	0.17	0.05	0.00	0.00	0.27
	Accuracy	725 – 749	0.00	0.05	0.13	0.03	0.00	0.22
		750 – 807	0.00	0.00	0.03	0.18	0.00	0.21
		808 – 850	0.00	0.00	0.00	0.01	0.01	0.02
		650 – 699	0.24	0.04	0.00	0.00	0.00	0.28
	Decision	700 – 724	0.07	0.14	0.06	0.01	0.00	0.27
	Consistency	725 – 749	0.01	0.06	0.10	0.05	0.00	0.22
		750 – 807	0.00	0.00	0.04	0.17	0.00	0.21
		808 – 850	0.00	0.00	0.00	0.01	0.01	0.02

Table A.8.19 Reliability of Classification: Integrated Mathematics I

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.17	0.02	0.00	0.00	0.00	0.19
	Decision	700 – 724	0.04	0.18	0.05	0.00	0.00	0.27
	Accuracy	725 – 749	0.00	0.05	0.17	0.04	0.00	0.25
		750 – 798	0.00	0.00	0.04	0.23	0.00	0.27
		799 – 850	0.00	0.00	0.00	0.01	0.01	0.02
		650 – 699	0.16	0.03	0.00	0.00	0.00	0.19
	Decision	700 – 724	0.05	0.15	0.06	0.01	0.00	0.27
	Consistency	725 – 749	0.00	0.06	0.13	0.05	0.00	0.25
		750 – 798	0.00	0.00	0.05	0.21	0.01	0.27
		799 – 850	0.00	0.00	0.00	0.01	0.01	0.02
PBT		650 – 699	0.09	0.03	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.03	0.16	0.05	0.00	0.00	0.23
	Accuracy	725 – 749	0.00	0.04	0.17	0.04	0.00	0.25
		750 – 798	0.00	0.00	0.04	0.27	0.01	0.32
		799 – 850	0.00	0.00	0.00	0.02	0.06	0.08
		650 – 699	0.09	0.04	0.00	0.00	0.00	0.12
	Decision	700 – 724	0.04	0.13	0.06	0.01	0.00	0.23
	Consistency	725 – 749	0.00	0.06	0.13	0.06	0.00	0.25
		750 – 798	0.00	0.00	0.05	0.25	0.02	0.32
		799 – 850	0.00	0.00	0.00	0.02	0.06	0.08

Table A.8.20 Reliability of Classification: Integrated Mathematics II

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative Scale Score						Total
CBT		650 – 699	0.14	0.03	0.00	0.00	0.00	0.17
	Decision	700 – 724	0.04	0.23	0.06	0.01	0.00	0.34
	Accuracy	725 – 749	0.00	0.04	0.18	0.02	0.00	0.25
		750 – 784	0.00	0.00	0.04	0.15	0.01	0.20
		785 – 850	0.00	0.00	0.00	0.01	0.03	0.04
		650 – 699	0.13	0.04	0.00	0.00	0.00	0.17
	Decision	700 – 724	0.06	0.19	0.08	0.01	0.00	0.34
	Consistency	725 – 749	0.00	0.06	0.15	0.04	0.00	0.25
		750 – 784	0.00	0.00	0.05	0.13	0.01	0.20
		785 – 850	0.00	0.00	0.00	0.01	0.02	0.04
PBT		650 – 699	0.06	0.02	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.03	0.19	0.04	0.01	0.00	0.27
	Accuracy	725 – 749	0.00	0.02	0.17	0.02	0.00	0.21
		750 – 784	0.00	0.00	0.03	0.22	0.01	0.27
		785 – 850	0.00	0.00	0.00	0.03	0.15	0.18
		650 – 699	0.05	0.03	0.00	0.00	0.00	0.08
	Decision	700 – 724	0.04	0.16	0.05	0.01	0.00	0.27
	Consistency	725 – 749	0.00	0.03	0.14	0.03	0.00	0.21
		750 – 784	0.00	0.00	0.04	0.20	0.02	0.27
		785 – 850	0.00	0.00	0.00	0.03	0.14	0.18

Table A.8.21 Reliability of Classification: Integrated Mathematics III

		Full	Level 1	Level 2	Level 3	Level 4	Level 5	Category
		Summative						Total
		Scale Score						
CBT		650 – 699	0.33	0.03	0.00	0.00	0.00	0.36
	Decision	700 – 724	0.04	0.14	0.04	0.00	0.00	0.23
	Accuracy	725 – 749	0.00	0.04	0.12	0.04	0.00	0.20
		750 – 803	0.00	0.00	0.03	0.17	0.00	0.20
		804 – 850	0.00	0.00	0.00	0.01	0.00	0.01
		650 – 699	0.32	0.05	0.00	0.00	0.00	0.36
	Decision	700 – 724	0.06	0.11	0.05	0.01	0.00	0.23
	Consistency	725 – 749	0.01	0.05	0.09	0.05	0.00	0.20
		750 – 803	0.00	0.00	0.04	0.16	0.00	0.20
		804 – 850	0.00	0.00	0.00	0.01	0.00	0.01
PBT		650 – 699	0.18	0.03	0.00	0.00	0.00	0.21
	Decision	700 – 724	0.03	0.13	0.06	0.01	0.00	0.24
	Accuracy	725 – 749	0.01	0.06	0.15	0.06	0.00	0.29
		750 – 803	0.01	0.00	0.03	0.21	0.00	0.24
		804 – 850	0.00	0.00	0.00	0.01	0.00	0.01
		650 – 699	0.17	0.04	0.00	0.00	0.00	0.21
	Decision	700 – 724	0.06	0.11	0.06	0.01	0.00	0.24
	Consistency	725 – 749	0.02	0.07	0.12	0.08	0.00	0.29
		750 – 803	0.00	0.00	0.05	0.19	0.00	0.24
		804 – 850	0.00	0.00	0.00	0.01	0.00	0.01

Appendix 10.1: IRT Results for Spring 2015 English Language Arts/Literacy (ELA/L)

Table A.10.1 CBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
CBT	E03	All Items	300	111	0.84	1.45	-1.45	8.21	0.50	0.20	0.10	1.03
CBT	E03	Reading	192	93	0.73	1.55	-1.45	8.21	0.46	0.19	0.10	1.03
CBT	E03	Writing	108	18	1.41	0.53	0.67	2.32	0.72	0.10	0.42	0.81
CBT	E03	All PBA	228	75	0.79	1.26	-1.45	6.37	0.55	0.19	0.10	1.03
CBT	E03	All EOY	72	36	0.95	1.81	-0.82	8.21	0.41	0.19	0.10	0.90
CBT	E04	All Items	326	124	0.44	0.96	-1.30	3.36	0.49	0.21	0.12	0.94
CBT	E04	Reading	218	106	0.42	1.03	-1.30	3.35	0.44	0.17	0.12	0.94
CBT	E04	Writing	108	18	0.58	0.22	0.24	0.95	0.83	0.06	0.70	0.90
CBT	E04	All PBA	244	83	0.34	0.89	-1.30	3.30	0.54	0.22	0.16	0.94
CBT	E04	All EOY	82	41	0.65	1.08	-1.23	3.35	0.41	0.15	0.12	0.87
CBT	E05	All Items	414	138	0.78	0.96	-1.35	4.80	0.54	0.25	0.07	1.04
CBT	E05	Reading	222	106	0.70	1.07	-1.35	4.80	0.45	0.21	0.07	1.04
CBT	E05	Writing	192	32	1.05	0.36	0.49	1.64	0.83	0.10	0.61	1.02
CBT	E05	All PBA	342	102	0.76	0.98	-1.35	4.80	0.59	0.26	0.07	1.04
CBT	E05	All EOY	72	36	0.84	0.91	-0.64	3.09	0.38	0.14	0.14	0.69
CBT	E06	All Items	425	157	0.56	1.07	-2.17	6.16	0.45	0.19	0.11	0.90
CBT	E06	Reading	290	139	0.53	1.13	-2.17	6.16	0.41	0.16	0.11	0.86
CBT	E06	Writing	135	18	0.79	0.35	0.25	1.43	0.74	0.09	0.57	0.90
CBT	E06	All PBA	279	84	0.46	0.96	-1.35	3.16	0.49	0.21	0.11	0.90
CBT	E06	All EOY	146	73	0.67	1.18	-2.17	6.16	0.40	0.16	0.15	0.80
CBT	E07	All Items	405	147	0.47	0.97	-1.63	3.03	0.46	0.20	0.12	0.91
CBT	E07	Reading	270	129	0.43	1.01	-1.63	3.03	0.42	0.17	0.12	0.86
CBT	E07	Writing	135	18	0.74	0.45	-0.13	1.48	0.77	0.09	0.59	0.91
CBT	E07	All PBA	275	82	0.52	0.94	-1.36	2.86	0.48	0.23	0.12	0.91
CBT	E07	All EOY	130	65	0.40	1.00	-1.63	3.03	0.44	0.16	0.13	0.83
CBT	E08	All Items	407	148	0.64	1.31	-1.56	6.40	0.45	0.22	0.06	1.21
CBT	E08	Reading	272	130	0.66	1.39	-1.56	6.40	0.41	0.20	0.06	1.21
CBT	E08	Writing	135	18	0.43	0.39	-0.08	1.22	0.77	0.08	0.60	0.90
CBT	E08	All PBA	279	84	0.46	1.12	-1.56	5.08	0.50	0.24	0.06	1.21
CBT	E08	All EOY	128	64	0.86	1.51	-1.55	6.40	0.39	0.16	0.08	0.78

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
CBT	E09	All Items	546	199	0.82	1.34	-1.29	9.59	0.47	0.22	0.06	0.98
CBT	E09	Reading	366	175	0.79	1.42	-1.29	9.59	0.42	0.17	0.06	0.90
CBT	E09	Writing	180	24	1.01	0.34	0.43	1.59	0.87	0.05	0.78	0.98
CBT	E09	All PBA	372	112	0.59	0.88	-1.29	4.29	0.53	0.24	0.15	0.98
CBT	E09	All EOY	174	87	1.12	1.72	-0.82	9.59	0.40	0.16	0.06	0.74
CBT	E10	All Items	534	193	0.66	0.90	-2.02	3.78	0.48	0.18	0.15	0.86
CBT	E10	Reading	354	169	0.68	0.95	-2.02	3.78	0.44	0.16	0.15	0.83
CBT	E10	Writing	180	24	0.55	0.32	-0.01	1.08	0.73	0.09	0.51	0.86
CBT	E10	All PBA	360	106	0.64	0.90	-2.02	3.78	0.50	0.20	0.15	0.86
CBT	E10	All EOY	174	87	0.68	0.90	-0.99	2.83	0.46	0.16	0.16	0.83
CBT	E11	All Items	419	154	1.14	1.31	-1.02	7.76	0.48	0.20	0.09	0.95
CBT	E11	Reading	284	136	1.21	1.37	-1.02	7.76	0.44	0.19	0.09	0.95
CBT	E11	Writing	135	18	0.65	0.37	0.00	1.20	0.74	0.08	0.59	0.87
CBT	E11	All PBA	279	84	0.90	1.26	-1.02	7.76	0.52	0.22	0.10	0.95
CBT	E11	All EOY	140	70	1.43	1.32	-0.82	7.61	0.43	0.17	0.09	0.84

Note: ELA/L Grade 9 CBT contained 15 field test items used to scale Fall block to Spring.

Table A.10.2 PBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E03	All Items	202	75	1.07	1.74	-0.74	10.44	0.47	0.19	0.08	0.89
PBT	E03	Reading	130	63	1.03	1.88	-0.74	10.44	0.43	0.18	0.08	0.89
PBT	E03	Writing	72	12	1.26	0.52	0.57	2.06	0.68	0.05	0.57	0.75
PBT	E03	All PBA	152	50	0.65	0.86	-0.74	3.05	0.54	0.16	0.13	0.89
PBT	E03	All EOY	50	25	1.90	2.59	-0.65	10.44	0.34	0.16	0.08	0.61
PBT	E04	All Items	232	86	0.63	1.09	-1.30	4.51	0.46	0.21	0.12	0.86
PBT	E04	Reading	148	72	0.66	1.18	-1.30	4.51	0.39	0.16	0.11	0.80
PBT	E04	Writing	84	14	0.44	0.32	0.13	1.21	0.79	0.05	0.67	0.86
PBT	E04	All PBA	184	62	0.55	1.09	-1.30	4.51	0.48	0.23	0.11	0.86
PBT	E04	All EOY	48	24	0.82	1.08	-1.30	2.75	0.40	0.12	0.15	0.57
PBT	E05	All Items	251	87	0.83	1.05	-1.35	5.31	0.49	0.22	0.07	0.89
PBT	E05	Reading	143	69	0.82	1.17	-1.35	5.31	0.43	0.19	0.07	0.85
PBT	E05	Writing	108	18	0.89	0.30	0.43	1.49	0.73	0.11	0.52	0.89
PBT	E05	All PBA	203	63	0.72	0.98	-1.35	5.31	0.53	0.22	0.07	0.89
PBT	E05	All EOY	48	24	1.13	1.20	-0.67	3.75	0.38	0.17	0.11	0.76
PBT	E06	All Items	274	100	0.60	1.03	-1.50	5.58	0.45	0.17	0.14	0.80
PBT	E06	Reading	184	88	0.57	1.09	-1.50	5.58	0.42	0.15	0.14	0.77
PBT	E06	Writing	90	12	0.76	0.42	0.09	1.41	0.66	0.09	0.50	0.80
PBT	E06	All PBA	186	56	0.49	0.89	-1.50	3.01	0.50	0.17	0.17	0.80
PBT	E06	All EOY	88	44	0.72	1.19	-1.11	5.58	0.39	0.14	0.14	0.77
PBT	E07	All Items	270	98	0.47	0.98	-1.68	4.36	0.44	0.17	0.11	0.87
PBT	E07	Reading	180	86	0.42	1.03	-1.68	4.36	0.42	0.16	0.11	0.87
PBT	E07	Writing	90	12	0.75	0.45	0.09	1.39	0.64	0.09	0.53	0.78
PBT	E07	All PBA	182	54	0.56	0.95	-1.68	3.83	0.43	0.17	0.11	0.78
PBT	E07	All EOY	88	44	0.35	1.03	-1.42	4.36	0.46	0.16	0.14	0.87
PBT	E08	All Items	272	99	0.63	1.67	-1.26	12.48	0.45	0.20	0.04	1.01
PBT	E08	Reading	182	87	0.69	1.76	-1.26	12.48	0.42	0.19	0.04	1.01
PBT	E08	Writing	90	12	0.25	0.48	-0.21	1.29	0.67	0.07	0.56	0.78
PBT	E08	All PBA	186	56	0.41	1.12	-1.22	5.71	0.48	0.19	0.04	0.78
PBT	E08	All EOY	86	43	0.93	2.17	-1.26	12.48	0.42	0.19	0.10	1.01

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
PBT	E09	All Items	274	100	0.86	1.47	-1.03	8.42	0.46	0.20	0.09	0.93
PBT	E09	Reading	184	88	0.86	1.56	-1.03	8.42	0.41	0.16	0.09	0.88
PBT	E09	Writing	90	12	0.85	0.44	0.25	1.55	0.79	0.07	0.72	0.93
PBT	E09	All PBA	186	56	0.57	0.81	-0.93	2.33	0.50	0.21	0.13	0.93
PBT	E09	All EOY	88	44	1.21	1.97	-1.03	8.42	0.40	0.16	0.09	0.73
PBT	E10	All Items	284	105	0.71	0.87	-1.03	3.57	0.49	0.18	0.15	0.95
PBT	E10	Reading	194	93	0.74	0.91	-1.03	3.57	0.47	0.17	0.15	0.95
PBT	E10	Writing	90	12	0.49	0.35	-0.01	1.02	0.71	0.07	0.59	0.79
PBT	E10	All PBA	186	56	0.84	0.78	-0.65	2.99	0.50	0.19	0.18	0.95
PBT	E10	All EOY	98	49	0.57	0.95	-1.03	3.57	0.49	0.17	0.15	0.87
PBT	E11	All Items	325	119	1.03	1.03	-0.76	5.71	0.45	0.18	0.09	1.00
PBT	E11	Reading	220	105	1.09	1.07	-0.76	5.71	0.42	0.17	0.09	1.00
PBT	E11	Writing	105	14	0.54	0.43	-0.16	1.21	0.62	0.09	0.43	0.75
PBT	E11	All PBA	217	65	0.80	0.88	-0.76	2.93	0.46	0.19	0.09	1.00
PBT	E11	All EOY	108	54	1.30	1.14	-0.63	5.71	0.43	0.16	0.12	0.85

Table A.10.3 CBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	E03	All Items	300	111	0.012	0.026	0.003	0.206	0.003	0.002	0.000	0.008
CBT	E03	Reading	192	93	0.013	0.028	0.003	0.206	0.003	0.001	0.000	0.006
CBT	E03	Writing	108	18	0.008	0.002	0.005	0.012	0.006	0.001	0.004	0.008
CBT	E03	All PBA	228	75	0.010	0.022	0.003	0.180	0.004	0.002	0.000	0.008
CBT	E03	All EOY	72	36	0.015	0.034	0.003	0.206	0.002	0.001	0.001	0.005
CBT	E04	All Items	326	124	0.007	0.006	0.003	0.054	0.003	0.001	0.001	0.007
CBT	E04	Reading	218	106	0.008	0.006	0.003	0.054	0.002	0.001	0.001	0.007
CBT	E04	Writing	108	18	0.005	0.001	0.004	0.006	0.005	0.001	0.005	0.007
CBT	E04	All PBA	244	83	0.006	0.003	0.003	0.021	0.003	0.002	0.001	0.007
CBT	E04	All EOY	82	41	0.010	0.009	0.003	0.054	0.002	0.001	0.001	0.004
CBT	E05	All Items	414	138	0.008	0.006	0.003	0.038	0.004	0.003	0.000	0.012
CBT	E05	Reading	222	106	0.008	0.007	0.003	0.038	0.003	0.002	0.000	0.011
CBT	E05	Writing	192	32	0.007	0.001	0.005	0.010	0.008	0.002	0.006	0.012
CBT	E05	All PBA	342	102	0.008	0.006	0.003	0.038	0.005	0.003	0.000	0.012
CBT	E05	All EOY	72	36	0.009	0.007	0.004	0.036	0.002	0.001	0.001	0.003
CBT	E06	All Items	425	157	0.007	0.005	0.003	0.032	0.003	0.001	0.001	0.009
CBT	E06	Reading	290	139	0.008	0.005	0.003	0.032	0.002	0.001	0.001	0.009
CBT	E06	Writing	135	18	0.005	0.001	0.003	0.007	0.005	0.001	0.004	0.009
CBT	E06	All PBA	279	84	0.007	0.005	0.003	0.024	0.003	0.002	0.001	0.009
CBT	E06	All EOY	146	73	0.008	0.005	0.003	0.032	0.002	0.001	0.001	0.004
CBT	E07	All Items	405	147	0.007	0.004	0.003	0.026	0.002	0.001	0.001	0.006
CBT	E07	Reading	270	129	0.007	0.005	0.003	0.026	0.002	0.001	0.001	0.006
CBT	E07	Writing	135	18	0.004	0.001	0.003	0.005	0.005	0.001	0.003	0.006
CBT	E07	All PBA	275	82	0.007	0.004	0.003	0.025	0.003	0.002	0.001	0.006
CBT	E07	All EOY	130	65	0.007	0.005	0.003	0.026	0.002	0.001	0.001	0.004
CBT	E08	All Items	407	148	0.010	0.017	0.003	0.171	0.003	0.001	0.001	0.007
CBT	E08	Reading	272	130	0.011	0.018	0.003	0.171	0.002	0.001	0.001	0.006
CBT	E08	Writing	135	18	0.004	0.000	0.003	0.005	0.005	0.001	0.004	0.007
CBT	E08	All PBA	279	84	0.008	0.008	0.003	0.046	0.003	0.002	0.001	0.007
CBT	E08	All EOY	128	64	0.013	0.024	0.003	0.171	0.002	0.001	0.001	0.005

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	E09	All Items	546	199	0.013	0.026	0.004	0.283	0.004	0.002	0.001	0.009
CBT	E09	Reading	366	175	0.013	0.027	0.004	0.283	0.003	0.001	0.001	0.009
CBT	E09	Writing	180	24	0.006	0.001	0.004	0.008	0.008	0.001	0.006	0.009
CBT	E09	All PBA	372	112	0.009	0.008	0.004	0.079	0.004	0.003	0.001	0.009
CBT	E09	All EOY	174	87	0.018	0.037	0.004	0.283	0.003	0.001	0.001	0.005
CBT	E10	All Items	534	193	0.011	0.009	0.005	0.078	0.004	0.002	0.001	0.010
CBT	E10	Reading	354	169	0.012	0.009	0.005	0.078	0.004	0.002	0.001	0.009
CBT	E10	Writing	180	24	0.007	0.001	0.005	0.008	0.008	0.001	0.005	0.010
CBT	E10	All PBA	360	106	0.011	0.010	0.005	0.078	0.005	0.003	0.001	0.010
CBT	E10	All EOY	174	87	0.011	0.008	0.005	0.044	0.004	0.001	0.001	0.008
CBT	E11	All Items	419	154	0.018	0.032	0.005	0.285	0.005	0.002	0.001	0.010
CBT	E11	Reading	284	136	0.020	0.034	0.005	0.285	0.004	0.002	0.001	0.010
CBT	E11	Writing	135	18	0.007	0.001	0.006	0.009	0.009	0.001	0.006	0.010
CBT	E11	All PBA	279	84	0.015	0.032	0.005	0.285	0.005	0.003	0.001	0.010
CBT	E11	All EOY	140	70	0.022	0.032	0.005	0.250	0.004	0.001	0.002	0.009

Table A.10.4 PBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E03	All Items	202	75	0.016	0.028	0.004	0.178	0.004	0.002	0.001	0.009
PBT	E03	Reading	130	63	0.017	0.030	0.004	0.178	0.003	0.001	0.001	0.008
PBT	E03	Writing	72	12	0.009	0.002	0.007	0.014	0.007	0.001	0.006	0.009
PBT	E03	All PBA	152	50	0.009	0.006	0.004	0.047	0.005	0.002	0.002	0.009
PBT	E03	All EOY	50	25	0.028	0.045	0.005	0.178	0.003	0.001	0.001	0.005
PBT	E04	All Items	232	86	0.013	0.011	0.005	0.064	0.004	0.002	0.001	0.010
PBT	E04	Reading	148	72	0.013	0.011	0.005	0.064	0.003	0.001	0.001	0.009
PBT	E04	Writing	84	14	0.008	0.001	0.006	0.009	0.008	0.001	0.006	0.010
PBT	E04	All PBA	184	62	0.012	0.011	0.005	0.064	0.004	0.003	0.001	0.010
PBT	E04	All EOY	48	24	0.013	0.010	0.005	0.048	0.003	0.001	0.001	0.005
PBT	E05	All Items	251	87	0.012	0.010	0.005	0.071	0.005	0.003	0.001	0.014
PBT	E05	Reading	143	69	0.013	0.011	0.005	0.071	0.004	0.002	0.001	0.014
PBT	E05	Writing	108	18	0.010	0.001	0.007	0.011	0.010	0.002	0.007	0.014
PBT	E05	All PBA	203	63	0.011	0.008	0.005	0.060	0.006	0.004	0.001	0.014
PBT	E05	All EOY	48	24	0.016	0.015	0.005	0.071	0.003	0.001	0.002	0.005
PBT	E06	All Items	274	100	0.012	0.009	0.005	0.052	0.004	0.002	0.001	0.009
PBT	E06	Reading	184	88	0.012	0.010	0.005	0.052	0.004	0.001	0.001	0.008
PBT	E06	Writing	90	12	0.008	0.001	0.006	0.010	0.008	0.001	0.006	0.009
PBT	E06	All PBA	186	56	0.010	0.008	0.005	0.052	0.005	0.002	0.002	0.009
PBT	E06	All EOY	88	44	0.013	0.010	0.005	0.051	0.003	0.001	0.001	0.005
PBT	E07	All Items	270	98	0.011	0.009	0.005	0.071	0.004	0.002	0.001	0.009
PBT	E07	Reading	180	86	0.012	0.009	0.005	0.071	0.003	0.001	0.001	0.008
PBT	E07	Writing	90	12	0.008	0.001	0.007	0.009	0.008	0.001	0.006	0.009
PBT	E07	All PBA	182	54	0.011	0.007	0.005	0.049	0.004	0.002	0.001	0.009
PBT	E07	All EOY	88	44	0.011	0.010	0.005	0.071	0.004	0.001	0.002	0.007
PBT	E08	All Items	272	99	0.014	0.018	0.005	0.117	0.004	0.002	0.001	0.010
PBT	E08	Reading	182	87	0.015	0.019	0.005	0.117	0.004	0.002	0.001	0.008
PBT	E08	Writing	90	12	0.008	0.001	0.006	0.010	0.008	0.001	0.006	0.010
PBT	E08	All PBA	186	56	0.013	0.017	0.005	0.116	0.005	0.002	0.001	0.010
PBT	E08	All EOY	86	43	0.016	0.019	0.005	0.117	0.004	0.001	0.001	0.007

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E09	All Items	274	100	0.021	0.051	0.007	0.414	0.005	0.003	0.001	0.013
PBT	E09	Reading	184	88	0.023	0.054	0.007	0.414	0.005	0.002	0.001	0.012
PBT	E09	Writing	90	12	0.009	0.001	0.007	0.011	0.011	0.001	0.009	0.013
PBT	E09	All PBA	186	56	0.013	0.009	0.007	0.067	0.006	0.003	0.002	0.013
PBT	E09	All EOY	88	44	0.032	0.075	0.007	0.414	0.005	0.002	0.001	0.008
PBT	E10	All Items	284	105	0.025	0.017	0.011	0.100	0.010	0.004	0.003	0.021
PBT	E10	Reading	194	93	0.026	0.018	0.011	0.100	0.009	0.003	0.003	0.019
PBT	E10	Writing	90	12	0.016	0.002	0.014	0.020	0.019	0.002	0.014	0.021
PBT	E10	All PBA	186	56	0.025	0.017	0.011	0.090	0.011	0.005	0.004	0.021
PBT	E10	All EOY	98	49	0.025	0.017	0.011	0.100	0.009	0.003	0.003	0.017
PBT	E11	All Items	325	119	0.055	0.068	0.017	0.663	0.016	0.007	0.006	0.042
PBT	E11	Reading	220	105	0.058	0.071	0.017	0.663	0.014	0.006	0.006	0.032
PBT	E11	Writing	105	14	0.031	0.005	0.022	0.039	0.028	0.007	0.018	0.042
PBT	E11	All PBA	217	65	0.047	0.039	0.017	0.241	0.017	0.009	0.006	0.042
PBT	E11	All EOY	108	54	0.065	0.090	0.017	0.663	0.015	0.004	0.009	0.026

Table A.10.5 CBT IRT Model Fit for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	E03	All Items	292	107	2394	1951	267	8589	0.13	0.05	0.05	0.25
CBT	E03	Reading	184	89	2745	1955	305	8589	0.14	0.05	0.05	0.25
CBT	E03	Writing	108	18	659	337	267	1600	0.11	0.02	0.08	0.15
CBT	E03	All PBA	226	74	2132	1996	267	8589	0.13	0.05	0.05	0.25
CBT	E03	All EOY	66	33	2983	1735	375	6898	0.15	0.05	0.05	0.24
CBT	E04	All Items	312	117	2600	2395	234	9876	0.13	0.05	0.04	0.24
CBT	E04	Reading	204	99	2889	2493	234	9876	0.13	0.05	0.04	0.24
CBT	E04	Writing	108	18	1007	344	605	1746	0.12	0.01	0.11	0.15
CBT	E04	All PBA	236	79	2377	2270	303	9876	0.12	0.04	0.04	0.24
CBT	E04	All EOY	76	38	3062	2605	234	9848	0.13	0.06	0.05	0.24
CBT	E05	All Items	404	133	2352	2371	114	9245	0.13	0.04	0.03	0.23
CBT	E05	Reading	212	101	2916	2464	114	9245	0.13	0.05	0.03	0.23
CBT	E05	Writing	192	32	570	202	284	1154	0.12	0.01	0.10	0.16
CBT	E05	All PBA	340	101	2001	2165	114	8261	0.12	0.04	0.03	0.22
CBT	E05	All EOY	64	32	3458	2674	668	9245	0.13	0.05	0.06	0.23
CBT	E06	All Items	409	149	2647	2321	200	9782	0.12	0.05	0.04	0.23
CBT	E06	Reading	274	131	2895	2369	200	9782	0.12	0.05	0.04	0.23
CBT	E06	Writing	135	18	848	268	359	1458	0.11	0.01	0.09	0.14
CBT	E06	All PBA	271	80	2459	2239	259	9782	0.12	0.04	0.04	0.23
CBT	E06	All EOY	138	69	2866	2411	200	9324	0.12	0.05	0.04	0.23
CBT	E07	All Items	387	138	2891	2448	305	9939	0.12	0.05	0.04	0.23
CBT	E07	Reading	252	120	3190	2489	305	9939	0.13	0.05	0.04	0.23
CBT	E07	Writing	135	18	896	291	615	1610	0.11	0.01	0.10	0.13
CBT	E07	All PBA	265	77	2605	2457	305	9835	0.12	0.04	0.04	0.23
CBT	E07	All EOY	122	61	3251	2407	374	9939	0.13	0.05	0.04	0.22
CBT	E08	All Items	393	141	2674	2152	242	9795	0.12	0.04	0.04	0.24
CBT	E08	Reading	258	123	2907	2205	242	9795	0.12	0.05	0.04	0.24
CBT	E08	Writing	135	18	1078	377	638	1850	0.12	0.01	0.10	0.15
CBT	E08	All PBA	271	80	2264	1891	340	7934	0.12	0.04	0.04	0.21
CBT	E08	All EOY	122	61	3211	2362	242	9795	0.13	0.05	0.04	0.24

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	E09	All Items	546	199	1933	1787	157	9426	0.14	0.05	0.04	0.31
CBT	E09	Reading	366	175	2100	1841	157	9426	0.14	0.06	0.04	0.31
CBT	E09	Writing	180	24	713	258	389	1529	0.14	0.01	0.11	0.18
CBT	E09	All PBA	372	112	1774	1698	193	9426	0.14	0.05	0.04	0.31
CBT	E09	All EOY	174	87	2137	1885	157	8191	0.14	0.06	0.05	0.30
CBT	E10	All Items	543	193	1609	1565	223	9725	0.15	0.06	0.06	0.36
CBT	E10	Reading	354	169	1735	1635	223	9725	0.15	0.06	0.06	0.36
CBT	E10	Writing	180	24	734	177	491	1309	0.17	0.01	0.14	0.20
CBT	E10	All PBA	360	106	1465	1283	235	6471	0.15	0.05	0.06	0.29
CBT	E10	All EOY	174	87	1781	1840	223	9725	0.15	0.07	0.06	0.36
CBT	E11	All Items	419	154	1176	1087	144	7407	0.14	0.05	0.05	0.33
CBT	E11	Reading	284	136	1238	1141	144	7407	0.14	0.06	0.05	0.33
CBT	E11	Writing	135	18	714	166	520	1117	0.17	0.01	0.15	0.20
CBT	E11	All PBA	279	84	1256	1195	187	7407	0.15	0.05	0.06	0.33
CBT	E11	All EOY	140	70	1081	940	144	4270	0.14	0.06	0.05	0.28

Table A.10.6 PBT IRT Model Fit for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E03	All Items	202	75	1539	1411	169	5173	0.15	0.06	0.06	0.27
PBT	E03	Reading	130	63	1746	1449	169	5173	0.15	0.06	0.05	0.27
PBT	E03	Writing	72	12	452	186	246	888	0.13	0.02	0.10	0.16
PBT	E03	All PBA	152	50	1333	1304	218	5024	0.14	0.05	0.06	0.26
PBT	E03	All EOY	50	25	1951	1551	169	5173	0.16	0.07	0.05	0.27
PBT	E04	All Items	232	86	1494	1338	172	5967	0.14	0.05	0.05	0.26
PBT	E04	Reading	148	72	1698	1372	172	5967	0.15	0.05	0.05	0.26
PBT	E04	Writing	84	14	447	91	329	585	0.13	0.01	0.12	0.15
PBT	E04	All PBA	184	62	1347	1258	172	5328	0.14	0.05	0.05	0.26
PBT	E04	All EOY	48	24	1874	1487	278	5967	0.15	0.05	0.06	0.25
PBT	E05	All Items	251	87	1284	1188	196	6777	0.14	0.04	0.05	0.31
PBT	E05	Reading	143	69	1537	1212	196	6777	0.14	0.05	0.05	0.31
PBT	E05	Writing	108	18	313	57	217	413	0.13	0.01	0.11	0.14
PBT	E05	All PBA	203	63	1131	1061	196	3903	0.14	0.04	0.05	0.23
PBT	E05	All EOY	48	24	1686	1417	307	6777	0.15	0.05	0.06	0.31
PBT	E06	All Items	274	100	1270	1033	154	4859	0.14	0.05	0.05	0.27
PBT	E06	Reading	184	88	1390	1044	154	4859	0.14	0.05	0.05	0.27
PBT	E06	Writing	90	12	388	111	246	678	0.13	0.02	0.10	0.15
PBT	E06	All PBA	186	56	1078	989	234	4859	0.13	0.05	0.06	0.27
PBT	E06	All EOY	88	44	1514	1048	154	3694	0.15	0.05	0.05	0.25
PBT	E07	All Items	270	98	1202	958	186	3661	0.14	0.05	0.06	0.24
PBT	E07	Reading	180	86	1317	967	186	3661	0.14	0.05	0.06	0.24
PBT	E07	Writing	90	12	378	113	237	588	0.13	0.01	0.11	0.16
PBT	E07	All PBA	182	54	1232	1048	237	3661	0.15	0.05	0.07	0.24
PBT	E07	All EOY	88	44	1165	844	186	3128	0.13	0.05	0.06	0.23
PBT	E08	All Items	270	98	1118	937	160	4105	0.14	0.05	0.05	0.27
PBT	E08	Reading	180	86	1224	953	160	4105	0.14	0.05	0.05	0.27
PBT	E08	Writing	90	12	356	69	261	489	0.13	0.02	0.09	0.16
PBT	E08	All PBA	184	55	980	864	182	4105	0.13	0.04	0.06	0.25
PBT	E08	All EOY	86	43	1294	1006	160	4068	0.14	0.05	0.05	0.27

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E09	All Items	274	100	759	627	123	2799	0.14	0.05	0.06	0.27
PBT	E09	Reading	184	88	822	642	123	2799	0.14	0.05	0.06	0.27
PBT	E09	Writing	90	12	297	64	218	398	0.15	0.02	0.11	0.17
PBT	E09	All PBA	186	56	747	641	183	2799	0.15	0.05	0.07	0.27
PBT	E09	All EOY	88	44	775	615	123	2215	0.14	0.05	0.06	0.25
PBT	E10	All Items	284	105	313	254	68	1544	0.17	0.06	0.08	0.37
PBT	E10	Reading	194	93	331	264	68	1544	0.16	0.06	0.08	0.37
PBT	E10	Writing	90	12	171	22	133	204	0.19	0.02	0.16	0.22
PBT	E10	All PBA	186	56	299	201	106	1007	0.17	0.04	0.09	0.27
PBT	E10	All EOY	98	49	329	304	68	1544	0.16	0.07	0.08	0.37
PBT	E11	All Items	325	119	186	273	62	2984	0.20	0.06	0.10	0.59
PBT	E11	Reading	220	105	193	290	62	2984	0.19	0.06	0.10	0.59
PBT	E11	Writing	105	14	132	34	80	191	0.27	0.02	0.24	0.33
PBT	E11	All PBA	217	65	210	363	62	2984	0.21	0.08	0.10	0.59
PBT	E11	All EOY	108	54	158	71	79	396	0.19	0.03	0.13	0.27

Appendix 10.2: IRT Results for Spring 2015 Mathematics

Table A.10.7 CBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M03	All Items	380	268	0.11	1.43	-3.99	4.18	0.80	0.24	0.13	1.41
CBT	M03	SSMC	54	54	-1.11	1.36	-3.99	1.91	0.77	0.19	0.26	1.24
CBT	M03	CR	326	214	0.42	1.27	-2.23	4.18	0.80	0.25	0.13	1.41
CBT	M03	All PBA	168	79	0.40	1.46	-3.99	3.07	0.73	0.23	0.27	1.22
CBT	M03	PBA Type I	57	48	-0.21	1.49	-3.99	3.07	0.82	0.23	0.27	1.22
CBT	M03	PBA Type II	63	18	1.45	0.79	0.02	2.96	0.54	0.14	0.33	0.95
CBT	M03	PBA Type III	48	13	1.22	0.67	-0.43	2.24	0.64	0.13	0.45	0.89
CBT	M03	EOY All	212	189	-0.01	1.40	-3.68	4.18	0.82	0.24	0.13	1.41
CBT	M04	All Items	307	199	0.10	1.27	-2.43	4.10	0.77	0.24	0.13	1.45
CBT	M04	SSMC	50	50	-0.83	0.97	-2.43	2.01	0.77	0.24	0.22	1.36
CBT	M04	CR	257	149	0.41	1.20	-2.27	4.10	0.77	0.24	0.13	1.45
CBT	M04	All PBA	146	64	0.23	1.45	-2.43	2.43	0.72	0.23	0.21	1.43
CBT	M04	PBA Type I	45	37	-0.69	1.17	-2.43	2.01	0.80	0.23	0.21	1.43
CBT	M04	PBA Type II	56	16	1.59	0.46	0.65	2.25	0.58	0.11	0.44	0.79
CBT	M04	PBA Type III	45	11	1.35	0.72	-0.01	2.43	0.62	0.24	0.35	1.20
CBT	M04	EOY All	161	135	0.03	1.17	-2.27	4.10	0.79	0.24	0.13	1.45
CBT	M05	All Items	317	206	0.38	1.29	-2.88	4.72	0.69	0.25	0.21	1.54
CBT	M05	SSMC	45	45	-0.61	1.26	-2.74	4.72	0.64	0.27	0.21	1.36
CBT	M05	CR	272	161	0.66	1.15	-2.88	2.89	0.71	0.24	0.27	1.54
CBT	M05	All PBA	142	62	0.41	1.33	-2.88	2.89	0.63	0.23	0.21	1.23
CBT	M05	PBA Type I	48	37	-0.26	1.24	-2.88	2.40	0.66	0.25	0.21	1.23
CBT	M05	PBA Type II	46	13	1.35	0.76	-0.50	2.89	0.56	0.12	0.31	0.71
CBT	M05	PBA Type III	48	12	1.45	0.66	0.39	2.16	0.63	0.23	0.36	1.09
CBT	M05	EOY All	175	144	0.37	1.27	-2.74	4.72	0.72	0.25	0.23	1.54
CBT	M06	All Items	305	192	0.62	1.28	-2.20	4.73	0.77	0.27	0.10	1.47

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M06	SSMC	38	38	0.05	1.60	-2.20	4.72	0.62	0.22	0.10	1.07
CBT	M06	CR	267	154	0.76	1.15	-1.85	4.05	0.81	0.26	0.26	1.47
CBT	M06	All PBA	146	66	0.60	1.25	-2.20	4.72	0.73	0.23	0.11	1.33
CBT	M06	PBA Type I	46	39	0.04	1.29	-2.20	4.72	0.80	0.24	0.11	1.33
CBT	M06	PBA Type II	52	15	1.41	0.58	-0.03	2.52	0.57	0.15	0.29	0.90
CBT	M06	PBA Type III	48	12	1.40	0.64	-0.06	2.05	0.68	0.20	0.47	1.17
CBT	M06	EOY All	159	126	0.63	1.29	-1.85	4.64	0.79	0.28	0.10	1.47
CBT	M07	All Items	386	238	1.13	1.58	-3.26	9.04	0.78	0.32	0.08	1.72
CBT	M07	SSMC	48	48	-0.35	1.12	-3.26	4.44	0.58	0.26	0.08	1.20
CBT	M07	CR	338	190	1.51	1.45	-1.83	9.04	0.82	0.32	0.11	1.72
CBT	M07	All PBA	172	77	1.04	1.41	-1.81	6.77	0.71	0.26	0.15	1.46
CBT	M07	PBA Type I	56	46	0.82	1.69	-1.81	6.77	0.79	0.30	0.15	1.46
CBT	M07	PBA Type II	56	16	1.21	0.79	-0.84	2.41	0.60	0.13	0.33	0.75
CBT	M07	PBA Type III	60	15	1.56	0.66	0.21	2.77	0.62	0.15	0.41	1.00
CBT	M07	EOY All	214	161	1.17	1.65	-3.26	9.04	0.81	0.34	0.08	1.72
CBT	M08	All Items	303	195	1.33	1.44	-2.22	7.57	0.71	0.33	0.13	1.76
CBT	M08	SSMC	45	45	0.46	1.52	-2.22	5.30	0.49	0.27	0.13	1.73
CBT	M08	CR	258	150	1.58	1.31	-1.41	7.57	0.78	0.31	0.16	1.76
CBT	M08	All PBA	121	58	1.22	1.51	-2.22	3.44	0.67	0.27	0.13	1.39
CBT	M08	PBA Type I	39	36	0.81	1.68	-2.22	3.44	0.73	0.30	0.13	1.39
CBT	M08	PBA Type II	46	13	2.15	0.83	0.66	3.22	0.60	0.22	0.20	0.92
CBT	M08	PBA Type III	36	9	1.55	0.73	0.14	2.36	0.57	0.09	0.45	0.77
CBT	M08	EOY All	182	137	1.37	1.41	-1.28	7.57	0.73	0.35	0.16	1.76
CBT	A1	All Items	482	273	2.04	1.66	-0.92	12.83	0.64	0.36	0.06	2.62
CBT	A1	SSMC	71	71	1.32	1.53	-0.92	6.69	0.47	0.42	0.06	2.62
CBT	A1	CR	411	202	2.30	1.63	-0.54	12.83	0.70	0.32	0.09	1.82
CBT	A1	All PBA	199	93	1.96	1.34	-0.83	7.01	0.63	0.29	0.06	1.47

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	A1	PBA Type I	56	56	1.86	1.60	-0.83	7.01	0.66	0.36	0.06	1.47
CBT	A1	PBA Type II	68	19	2.36	0.61	1.08	3.52	0.57	0.14	0.26	0.90
CBT	A1	PBA Type III	75	18	1.82	0.92	0.18	3.15	0.61	0.17	0.37	1.05
CBT	A1	EOY All	283	180	2.09	1.80	-0.92	12.83	0.64	0.39	0.07	2.62
CBT	GO	All Items	506	286	1.35	1.16	-1.67	5.87	0.75	0.34	0.09	1.97
CBT	GO	SSMC	52	52	0.36	1.24	-1.67	5.64	0.50	0.31	0.09	1.97
CBT	GO	CR	454	234	1.57	1.03	-0.92	5.87	0.80	0.32	0.14	1.75
CBT	GO	All PBA	218	97	1.43	1.14	-1.01	5.64	0.73	0.30	0.10	1.97
CBT	GO	PBA Type I	57	57	1.03	1.28	-1.01	5.64	0.77	0.36	0.10	1.97
CBT	GO	PBA Type II	74	21	2.09	0.53	1.05	2.95	0.71	0.18	0.46	1.08
CBT	GO	PBA Type III	87	19	1.89	0.59	0.65	3.12	0.64	0.21	0.37	1.04
CBT	GO	EOY All	288	189	1.31	1.18	-1.67	5.87	0.75	0.36	0.09	1.75
CBT	A2	All Items	393	200	1.98	1.67	-1.26	10.46	0.69	0.33	0.04	1.79
CBT	A2	SSMC	39	39	1.27	1.69	-1.26	5.86	0.47	0.22	0.08	1.11
CBT	A2	CR	354	161	2.16	1.62	-1.23	10.46	0.74	0.33	0.04	1.79
CBT	A2	All PBA	200	78	2.06	1.46	-0.74	10.46	0.70	0.30	0.08	1.79
CBT	A2	PBA Type I	41	41	1.71	1.83	-0.74	10.46	0.81	0.34	0.08	1.79
CBT	A2	PBA Type II	72	20	2.46	0.78	0.86	4.09	0.63	0.16	0.23	0.86
CBT	A2	PBA Type III	87	17	2.41	0.70	1.15	3.62	0.52	0.21	0.27	0.93
CBT	A2	EOY All	193	122	1.93	1.80	-1.26	10.19	0.68	0.35	0.04	1.76
CBT	M1	All Items	134	71	1.68	1.24	-1.35	4.40	0.67	0.36	0.07	1.62
CBT	M1	SSMC	13	13	1.37	1.51	-1.35	4.02	0.39	0.27	0.07	1.01
CBT	M1	CR	121	58	1.75	1.17	-0.83	4.40	0.74	0.35	0.17	1.62
CBT	M1	All PBA	62	27	1.86	1.13	0.08	4.03	0.65	0.33	0.18	1.56
CBT	M1	PBA Type I	15	15	1.83	1.30	0.08	4.03	0.66	0.43	0.18	1.56
CBT	M1	PBA Type II	20	6	2.25	0.79	1.23	3.22	0.74	0.13	0.61	0.91
CBT	M1	PBA Type III	27	6	1.56	0.96	0.32	2.67	0.53	0.10	0.46	0.71

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
CBT	M1	EOY All	72	44	1.57	1.30	-1.35	4.40	0.68	0.38	0.07	1.62
CBT	M2	All Items	141	77	1.86	1.39	-0.99	6.62	0.72	0.41	0.06	1.99
CBT	M2	SSMC	16	16	0.66	0.96	-0.98	2.45	0.38	0.14	0.06	0.65
CBT	M2	CR	125	61	2.18	1.31	-0.51	6.62	0.81	0.41	0.15	1.99
CBT	M2	All PBA	68	28	1.86	0.91	-0.20	3.82	0.74	0.37	0.25	1.99
CBT	M2	PBA Type I	15	15	1.53	0.79	-0.02	3.14	0.86	0.45	0.25	1.99
CBT	M2	PBA Type II	17	5	2.61	0.86	1.77	3.82	0.63	0.15	0.45	0.84
CBT	M2	PBA Type III	36	8	2.00	0.95	-0.20	2.92	0.60	0.21	0.34	0.95
CBT	M2	EOY All	73	49	1.87	1.60	-0.98	6.62	0.70	0.43	0.06	1.75
CBT	M3	All Items	144	74	1.67	1.33	-1.19	6.31	0.68	0.28	0.19	1.39
CBT	M3	SSMC	18	18	0.51	1.26	-1.19	3.92	0.55	0.17	0.31	0.92
CBT	M3	CR	126	56	2.05	1.12	-0.02	6.31	0.73	0.29	0.19	1.39
CBT	M3	All PBA	75	28	1.54	1.12	-0.81	3.92	0.70	0.25	0.31	1.27
CBT	M3	PBA Type I	14	14	1.12	1.36	-0.81	3.92	0.77	0.28	0.31	1.27
CBT	M3	PBA Type II	25	7	1.87	0.59	0.95	2.49	0.70	0.14	0.55	0.96
CBT	M3	PBA Type III	36	7	2.05	0.65	1.07	2.78	0.58	0.25	0.35	0.97
CBT	M3	EOY All	69	46	1.75	1.45	-1.19	6.31	0.67	0.29	0.19	1.39

Note: M03 through M08 = mathematics grades 3 through 8, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table A.10.8 PBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
PBT	M03	All Items	235	158	0.16	1.44	-4.01	5.33	0.74	0.23	0.17	1.41
PBT	M03	SSMC	55	55	-0.90	1.45	-4.01	2.66	0.72	0.18	0.33	1.14
PBT	M03	CR	180	103	0.73	1.08	-1.25	5.33	0.74	0.26	0.17	1.41
PBT	M03	All PBA	110	48	0.29	1.45	-4.01	3.33	0.67	0.21	0.36	1.13
PBT	M03	PBA Type I	32	27	-0.48	1.36	-4.01	1.58	0.78	0.18	0.50	1.13
PBT	M03	PBA Type II	42	12	1.49	0.86	0.07	3.33	0.50	0.14	0.36	0.88
PBT	M03	PBA Type III	36	9	1.00	0.72	-0.45	1.71	0.55	0.12	0.38	0.78
PBT	M03	EOY All	125	110	0.11	1.45	-3.49	5.33	0.77	0.24	0.17	1.41
PBT	M04	All Items	215	135	0.12	1.27	-2.39	3.34	0.73	0.27	0.15	2.14
PBT	M04	SSMC	54	54	-0.70	1.07	-2.39	1.96	0.78	0.33	0.18	2.14
PBT	M04	CR	161	81	0.67	1.09	-1.76	3.34	0.70	0.22	0.15	1.22
PBT	M04	All PBA	105	46	0.22	1.34	-2.23	2.25	0.68	0.20	0.18	1.09
PBT	M04	PBA Type I	31	26	-0.58	1.22	-2.23	1.74	0.77	0.19	0.18	1.09
PBT	M04	PBA Type II	38	11	1.30	0.52	0.59	2.25	0.54	0.12	0.40	0.75
PBT	M04	PBA Type III	36	9	1.23	0.63	0.14	2.05	0.58	0.14	0.37	0.78
PBT	M04	EOY All	110	89	0.07	1.24	-2.39	3.34	0.76	0.30	0.15	2.14
PBT	M05	All Items	207	135	0.25	1.22	-2.49	5.29	0.65	0.23	0.11	1.34
PBT	M05	SSMC	62	62	-0.38	1.19	-2.48	5.29	0.66	0.27	0.14	1.34
PBT	M05	CR	145	73	0.79	0.95	-1.13	3.54	0.64	0.18	0.11	1.24
PBT	M05	All PBA	96	41	0.54	1.33	-2.48	5.29	0.58	0.21	0.14	1.19
PBT	M05	PBA Type I	31	23	0.08	1.49	-2.48	5.29	0.57	0.24	0.14	1.19
PBT	M05	PBA Type II	35	10	0.88	0.61	-0.36	1.77	0.57	0.13	0.38	0.75
PBT	M05	PBA Type III	30	8	1.44	0.90	0.18	2.40	0.61	0.22	0.34	1.01
PBT	M05	EOY All	111	94	0.13	1.15	-2.36	3.54	0.68	0.23	0.11	1.34
PBT	M06	All Items	211	132	0.63	1.21	-3.43	4.34	0.68	0.26	0.09	1.46
PBT	M06	SSMC	45	45	0.13	1.54	-3.43	4.34	0.58	0.24	0.09	0.98

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
PBT	M06	CR	166	87	0.89	0.90	-0.90	2.99	0.74	0.25	0.30	1.45
PBT	M06	All PBA	98	43	0.64	1.16	-1.92	3.52	0.63	0.22	0.15	1.18
PBT	M06	PBA Type I	26	23	0.15	1.24	-1.92	3.52	0.71	0.24	0.15	1.18
PBT	M06	PBA Type II	42	12	1.12	0.79	-0.69	2.45	0.49	0.15	0.30	0.77
PBT	M06	PBA Type III	30	8	1.32	0.77	-0.12	2.16	0.63	0.14	0.44	0.85
PBT	M06	EOY All	113	89	0.62	1.24	-3.43	4.34	0.71	0.27	0.09	1.45
PBT	M07	All Items	208	130	1.09	1.39	-4.13	5.64	0.64	0.28	0.05	1.43
PBT	M07	SSMC	56	56	0.45	1.56	-4.13	5.64	0.52	0.26	0.05	1.23
PBT	M07	CR	152	74	1.57	1.02	-0.95	5.45	0.72	0.26	0.21	1.43
PBT	M07	All PBA	102	47	0.98	1.19	-1.64	3.16	0.59	0.21	0.22	1.42
PBT	M07	PBA Type I	33	28	0.73	1.32	-1.64	3.16	0.62	0.26	0.22	1.42
PBT	M07	PBA Type II	39	11	1.04	0.78	-0.95	2.00	0.56	0.09	0.46	0.68
PBT	M07	PBA Type III	30	8	1.76	0.86	0.22	2.88	0.53	0.17	0.38	0.94
PBT	M07	EOY All	106	83	1.15	1.49	-4.13	5.64	0.66	0.31	0.05	1.43
PBT	M08	All Items	188	116	1.23	1.42	-2.66	5.16	0.63	0.28	0.14	1.49
PBT	M08	SSMC	47	47	0.50	1.35	-2.66	3.67	0.55	0.27	0.14	1.49
PBT	M08	CR	141	69	1.72	1.25	-1.03	5.16	0.69	0.26	0.21	1.40
PBT	M08	All PBA	96	46	1.27	1.47	-2.66	3.67	0.61	0.27	0.14	1.49
PBT	M08	PBA Type I	31	28	0.87	1.62	-2.66	3.67	0.64	0.33	0.14	1.49
PBT	M08	PBA Type II	38	11	2.20	0.86	0.50	3.18	0.59	0.20	0.24	0.93
PBT	M08	PBA Type III	27	7	1.39	0.87	-0.17	2.45	0.52	0.08	0.45	0.66
PBT	M08	EOY All	92	70	1.20	1.40	-0.93	5.16	0.65	0.28	0.19	1.40
PBT	A1	All Items	239	133	2.05	1.77	-0.83	12.09	0.52	0.24	0.02	1.22
PBT	A1	SSMC	53	53	1.47	1.46	-0.83	5.16	0.43	0.26	0.02	1.22
PBT	A1	CR	186	80	2.44	1.88	-0.33	12.09	0.59	0.21	0.15	1.07
PBT	A1	All PBA	116	51	2.04	1.32	-0.73	5.16	0.51	0.22	0.05	1.04
PBT	A1	PBA Type I	28	28	1.87	1.55	-0.73	5.16	0.48	0.27	0.05	1.04

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
PBT	A1	PBA Type II	46	13	2.51	0.78	1.20	4.00	0.52	0.10	0.32	0.69
PBT	A1	PBA Type III	42	10	1.93	1.10	-0.10	3.23	0.58	0.15	0.39	0.91
PBT	A1	EOY All	123	82	2.06	2.03	-0.83	12.09	0.53	0.26	0.02	1.22
PBT	GO	All Items	253	135	1.34	1.28	-1.41	5.87	0.67	0.29	0.13	1.97
PBT	GO	SSMC	41	41	0.56	1.41	-1.41	5.87	0.59	0.33	0.13	1.97
PBT	GO	CR	212	94	1.68	1.07	-0.22	4.35	0.71	0.27	0.25	1.48
PBT	GO	All PBA	110	46	1.47	1.16	-0.75	4.32	0.67	0.29	0.13	1.97
PBT	GO	PBA Type I	24	24	0.84	1.12	-0.75	4.32	0.72	0.37	0.13	1.97
PBT	GO	PBA Type II	38	11	2.47	0.70	1.19	3.63	0.66	0.13	0.45	0.86
PBT	GO	PBA Type III	48	11	1.85	0.70	0.50	2.77	0.56	0.19	0.32	0.84
PBT	GO	EOY All	143	89	1.27	1.34	-1.41	5.87	0.67	0.29	0.20	1.48
PBT	A2	All Items	266	135	2.19	1.67	-1.05	11.65	0.55	0.27	0.00	1.17
PBT	A2	SSMC	35	35	1.43	1.58	-1.05	6.80	0.46	0.24	0.10	1.10
PBT	A2	CR	231	100	2.45	1.62	-0.73	11.65	0.58	0.28	0.00	1.17
PBT	A2	All PBA	137	51	2.37	1.09	-0.63	6.01	0.60	0.22	0.20	0.99
PBT	A2	PBA Type I	25	25	2.23	1.24	-0.63	6.01	0.68	0.24	0.22	0.99
PBT	A2	PBA Type II	46	13	2.47	0.90	1.04	4.28	0.54	0.14	0.22	0.67
PBT	A2	PBA Type III	66	13	2.56	0.98	0.82	3.94	0.48	0.21	0.20	0.96
PBT	A2	EOY All	129	84	2.07	1.93	-1.05	11.65	0.52	0.30	0.00	1.17
PBT	M1	All Items	118	65	1.83	1.26	-1.05	5.36	0.56	0.29	0.00	1.42
PBT	M1	SSMC	21	21	1.68	1.34	-1.05	3.72	0.42	0.26	0.11	1.02
PBT	M1	CR	97	44	1.91	1.23	-0.14	5.35	0.62	0.28	0.00	1.42
PBT	M1	All PBA	49	23	2.04	1.23	0.19	5.00	0.49	0.23	0.11	0.96
PBT	M1	PBA Type I	14	14	2.11	1.46	0.19	5.00	0.44	0.27	0.11	0.96
PBT	M1	PBA Type II	14	4	1.96	0.54	1.23	2.53	0.68	0.07	0.60	0.76
PBT	M1	PBA Type III	21	5	1.91	1.08	0.54	3.12	0.48	0.10	0.37	0.62
PBT	M1	EOY All	69	42	1.72	1.28	-1.05	5.35	0.59	0.31	0.00	1.42

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
PBT	M2	All Items	131	74	1.59	1.22	-1.39	3.95	0.61	0.24	0.13	1.34
PBT	M2	SSMC	26	26	0.78	0.98	-1.39	3.73	0.48	0.17	0.13	0.84
PBT	M2	CR	105	48	2.03	1.11	-0.51	3.95	0.68	0.25	0.26	1.34
PBT	M2	All PBA	59	26	1.89	1.14	-0.51	3.94	0.58	0.21	0.20	0.98
PBT	M2	PBA Type I	15	15	1.61	1.08	-0.04	3.73	0.62	0.24	0.20	0.98
PBT	M2	PBA Type II	17	5	2.75	0.86	1.80	3.94	0.53	0.12	0.41	0.72
PBT	M2	PBA Type III	27	6	1.87	1.26	-0.51	3.01	0.50	0.14	0.31	0.65
PBT	M2	EOY All	72	48	1.42	1.24	-1.39	3.95	0.62	0.26	0.13	1.34
PBT	M3	All Items	120	66	2.04	1.62	-0.81	6.72	0.56	0.26	0.13	1.31
PBT	M3	SSMC	23	23	1.47	1.74	-0.81	5.56	0.53	0.27	0.18	1.31
PBT	M3	CR	97	43	2.35	1.49	-0.07	6.72	0.58	0.26	0.13	1.14
PBT	M3	All PBA	52	20	2.37	1.54	-0.81	5.56	0.53	0.17	0.20	0.92
PBT	M3	PBA Type I	10	10	2.64	2.01	-0.81	5.56	0.52	0.15	0.20	0.72
PBT	M3	PBA Type II	18	5	1.86	0.92	0.83	2.97	0.55	0.13	0.42	0.75
PBT	M3	PBA Type III	24	5	2.34	0.87	0.93	3.21	0.51	0.27	0.25	0.92
PBT	M3	EOY All	68	46	1.90	1.66	-0.81	6.72	0.58	0.29	0.13	1.31

Note: M03 through M08 = mathematics grades 3 through 8, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table A.10.9 CBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M03	All Items	380	268	0.011	0.010	0.002	0.081	0.007	0.003	0.001	0.018
CBT	M03	SSMC	54	54	0.016	0.013	0.005	0.078	0.009	0.003	0.004	0.018
CBT	M03	CR	326	214	0.010	0.009	0.002	0.081	0.007	0.003	0.001	0.014
CBT	M03	All PBA	168	79	0.010	0.010	0.002	0.078	0.006	0.003	0.002	0.018
CBT	M03	PBA Type I	57	48	0.013	0.012	0.003	0.078	0.008	0.003	0.003	0.018
CBT	M03	PBA Type II	63	18	0.007	0.002	0.003	0.013	0.003	0.001	0.002	0.005
CBT	M03	PBA Type III	48	13	0.006	0.004	0.002	0.017	0.004	0.002	0.002	0.007
CBT	M03	EOY All	212	189	0.012	0.010	0.003	0.081	0.008	0.003	0.001	0.017
CBT	M04	All Items	307	199	0.009	0.010	0.003	0.128	0.006	0.002	0.002	0.013
CBT	M04	SSMC	50	50	0.010	0.007	0.003	0.038	0.006	0.002	0.003	0.013
CBT	M04	CR	257	149	0.008	0.011	0.003	0.128	0.006	0.002	0.002	0.012
CBT	M04	All PBA	146	64	0.008	0.007	0.003	0.051	0.005	0.002	0.002	0.012
CBT	M04	PBA Type I	45	37	0.009	0.008	0.003	0.051	0.006	0.002	0.004	0.012
CBT	M04	PBA Type II	56	16	0.007	0.003	0.003	0.016	0.004	0.001	0.002	0.007
CBT	M04	PBA Type III	45	11	0.005	0.002	0.003	0.012	0.004	0.002	0.002	0.006
CBT	M04	EOY All	161	135	0.009	0.012	0.003	0.128	0.006	0.002	0.002	0.013
CBT	M05	All Items	317	206	0.010	0.012	0.003	0.136	0.006	0.002	0.002	0.015
CBT	M05	SSMC	45	45	0.015	0.021	0.003	0.136	0.006	0.002	0.003	0.015
CBT	M05	CR	272	161	0.009	0.007	0.003	0.038	0.005	0.002	0.002	0.014
CBT	M05	All PBA	142	62	0.010	0.010	0.003	0.051	0.005	0.002	0.002	0.009
CBT	M05	PBA Type I	48	37	0.012	0.012	0.003	0.051	0.005	0.002	0.003	0.009
CBT	M05	PBA Type II	46	13	0.006	0.004	0.003	0.018	0.003	0.000	0.002	0.004
CBT	M05	PBA Type III	48	12	0.006	0.003	0.003	0.015	0.004	0.002	0.002	0.007
CBT	M05	EOY All	175	144	0.011	0.012	0.003	0.136	0.006	0.002	0.002	0.015
CBT	M06	All Items	305	192	0.010	0.019	0.002	0.190	0.006	0.003	0.001	0.021
CBT	M06	SSMC	38	38	0.020	0.040	0.003	0.190	0.005	0.002	0.003	0.009

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M06	CR	267	154	0.008	0.007	0.002	0.063	0.006	0.003	0.001	0.021
CBT	M06	All PBA	146	66	0.009	0.019	0.002	0.159	0.005	0.002	0.001	0.010
CBT	M06	PBA Type I	46	39	0.011	0.025	0.002	0.159	0.006	0.002	0.003	0.010
CBT	M06	PBA Type II	52	15	0.006	0.005	0.003	0.024	0.003	0.001	0.001	0.005
CBT	M06	PBA Type III	48	12	0.005	0.002	0.003	0.010	0.004	0.001	0.002	0.005
CBT	M06	EOY All	159	126	0.011	0.019	0.002	0.190	0.006	0.003	0.001	0.021
CBT	M07	All Items	386	238	0.018	0.043	0.002	0.383	0.007	0.004	0.001	0.028
CBT	M07	SSMC	48	48	0.016	0.023	0.004	0.130	0.005	0.001	0.003	0.008
CBT	M07	CR	338	190	0.018	0.047	0.002	0.383	0.008	0.005	0.001	0.028
CBT	M07	All PBA	172	77	0.011	0.015	0.002	0.093	0.006	0.003	0.001	0.017
CBT	M07	PBA Type I	56	46	0.015	0.018	0.002	0.093	0.007	0.003	0.003	0.017
CBT	M07	PBA Type II	56	16	0.006	0.002	0.003	0.010	0.003	0.001	0.001	0.005
CBT	M07	PBA Type III	60	15	0.007	0.004	0.003	0.014	0.004	0.001	0.002	0.006
CBT	M07	EOY All	214	161	0.021	0.052	0.003	0.383	0.008	0.004	0.001	0.028
CBT	M08	All Items	303	195	0.018	0.029	0.000	0.274	0.007	0.004	0.000	0.033
CBT	M08	SSMC	45	45	0.025	0.039	0.000	0.205	0.005	0.002	0.000	0.009
CBT	M08	CR	258	150	0.015	0.024	0.003	0.274	0.008	0.005	0.002	0.033
CBT	M08	All PBA	121	58	0.014	0.019	0.003	0.135	0.006	0.003	0.002	0.018
CBT	M08	PBA Type I	39	36	0.018	0.023	0.003	0.135	0.007	0.003	0.003	0.018
CBT	M08	PBA Type II	46	13	0.010	0.005	0.004	0.019	0.004	0.002	0.002	0.010
CBT	M08	PBA Type III	36	9	0.006	0.002	0.003	0.008	0.003	0.001	0.002	0.005
CBT	M08	EOY All	182	137	0.019	0.032	0.000	0.274	0.008	0.005	0.000	0.033
CBT	A1	All Items	482	273	0.039	0.092	0.000	1.097	0.008	0.005	0.000	0.033
CBT	A1	SSMC	71	71	0.047	0.092	0.000	0.572	0.005	0.002	0.000	0.009
CBT	A1	CR	411	202	0.036	0.092	0.000	1.097	0.009	0.006	0.000	0.033
CBT	A1	All PBA	199	93	0.032	0.071	0.000	0.572	0.007	0.004	0.000	0.030
CBT	A1	PBA Type I	56	56	0.045	0.090	0.000	0.572	0.009	0.005	0.000	0.030

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	A1	PBA Type II	68	19	0.013	0.006	0.003	0.030	0.005	0.002	0.002	0.008
CBT	A1	PBA Type III	75	18	0.010	0.008	0.003	0.034	0.005	0.003	0.001	0.013
CBT	A1	EOY All	283	180	0.043	0.101	0.000	1.097	0.008	0.006	0.000	0.033
CBT	GO	All Items	506	286	0.024	0.033	0.000	0.306	0.012	0.007	0.000	0.063
CBT	GO	SSMC	52	52	0.034	0.054	0.000	0.306	0.009	0.003	0.000	0.017
CBT	GO	CR	454	234	0.022	0.026	0.005	0.222	0.013	0.008	0.002	0.063
CBT	GO	All PBA	218	97	0.022	0.038	0.000	0.306	0.010	0.006	0.000	0.035
CBT	GO	PBA Type I	57	57	0.028	0.048	0.000	0.306	0.012	0.006	0.000	0.035
CBT	GO	PBA Type II	74	21	0.015	0.005	0.006	0.026	0.009	0.004	0.003	0.019
CBT	GO	PBA Type III	87	19	0.012	0.007	0.005	0.035	0.007	0.005	0.002	0.021
CBT	GO	EOY All	288	189	0.025	0.031	0.005	0.286	0.013	0.008	0.003	0.063
CBT	A2	All Items	393	200	0.046	0.122	0.005	1.086	0.011	0.008	0.001	0.057
CBT	A2	SSMC	39	39	0.054	0.097	0.005	0.438	0.007	0.002	0.005	0.011
CBT	A2	CR	354	161	0.044	0.128	0.005	1.086	0.011	0.009	0.001	0.057
CBT	A2	All PBA	200	78	0.038	0.127	0.005	1.086	0.010	0.007	0.002	0.042
CBT	A2	PBA Type I	41	41	0.055	0.174	0.006	1.086	0.013	0.009	0.005	0.042
CBT	A2	PBA Type II	72	20	0.019	0.018	0.005	0.082	0.007	0.003	0.003	0.012
CBT	A2	PBA Type III	87	17	0.017	0.018	0.006	0.074	0.005	0.003	0.002	0.012
CBT	A2	EOY All	193	122	0.051	0.119	0.005	0.799	0.011	0.009	0.001	0.057
CBT	M1	All Items	134	71	0.050	0.080	0.006	0.527	0.017	0.013	0.003	0.062
CBT	M1	SSMC	13	13	0.106	0.159	0.012	0.527	0.010	0.002	0.007	0.016
CBT	M1	CR	121	58	0.038	0.040	0.006	0.212	0.019	0.014	0.003	0.062
CBT	M1	All PBA	62	27	0.048	0.055	0.007	0.212	0.017	0.014	0.004	0.062
CBT	M1	PBA Type I	15	15	0.068	0.065	0.012	0.212	0.021	0.016	0.007	0.062
CBT	M1	PBA Type II	20	6	0.031	0.027	0.010	0.067	0.016	0.012	0.007	0.032
CBT	M1	PBA Type III	27	6	0.017	0.011	0.007	0.034	0.007	0.003	0.004	0.011
CBT	M1	EOY All	72	44	0.051	0.092	0.006	0.527	0.017	0.013	0.003	0.054

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M2	All Items	141	77	0.068	0.101	0.013	0.674	0.028	0.025	0.003	0.182
CBT	M2	SSMC	16	16	0.077	0.092	0.023	0.407	0.017	0.003	0.009	0.022
CBT	M2	CR	125	61	0.066	0.104	0.013	0.674	0.031	0.028	0.003	0.182
CBT	M2	All PBA	68	28	0.041	0.026	0.015	0.132	0.026	0.032	0.006	0.182
CBT	M2	PBA Type I	15	15	0.049	0.031	0.020	0.132	0.036	0.042	0.012	0.182
CBT	M2	PBA Type II	17	5	0.029	0.009	0.020	0.039	0.015	0.005	0.010	0.022
CBT	M2	PBA Type III	36	8	0.032	0.015	0.015	0.065	0.015	0.009	0.006	0.030
CBT	M2	EOY All	73	49	0.084	0.123	0.013	0.674	0.029	0.021	0.003	0.093
CBT	M3	All Items	144	74	0.070	0.101	0.013	0.821	0.030	0.020	0.007	0.099
CBT	M3	SSMC	18	18	0.066	0.065	0.021	0.313	0.024	0.005	0.016	0.038
CBT	M3	CR	126	56	0.071	0.111	0.013	0.821	0.031	0.023	0.007	0.099
CBT	M3	All PBA	75	28	0.045	0.056	0.013	0.313	0.024	0.015	0.007	0.077
CBT	M3	PBA Type I	14	14	0.062	0.075	0.018	0.313	0.031	0.017	0.016	0.077
CBT	M3	PBA Type II	25	7	0.029	0.014	0.013	0.052	0.019	0.011	0.009	0.041
CBT	M3	PBA Type III	36	7	0.027	0.008	0.017	0.041	0.014	0.008	0.007	0.027
CBT	M3	EOY All	69	46	0.085	0.119	0.021	0.821	0.033	0.022	0.009	0.099

Note: M03 through M08 = mathematics grades 3 through 8, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table A.10.10 PBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	M03	All Items	235	158	0.013	0.013	0.004	0.091	0.008	0.003	0.002	0.019
PBT	M03	SSMC	55	55	0.018	0.015	0.004	0.083	0.009	0.003	0.004	0.019
PBT	M03	CR	180	103	0.011	0.011	0.004	0.091	0.007	0.003	0.002	0.014
PBT	M03	All PBA	110	48	0.011	0.012	0.004	0.083	0.006	0.003	0.002	0.018
PBT	M03	PBA Type I	32	27	0.014	0.015	0.004	0.083	0.008	0.003	0.004	0.018
PBT	M03	PBA Type II	42	12	0.009	0.003	0.004	0.015	0.004	0.001	0.002	0.006
PBT	M03	PBA Type III	36	9	0.006	0.002	0.004	0.009	0.004	0.001	0.003	0.006
PBT	M03	EOY All	125	110	0.014	0.013	0.004	0.091	0.008	0.003	0.002	0.019
PBT	M04	All Items	215	135	0.012	0.009	0.004	0.071	0.008	0.006	0.002	0.061
PBT	M04	SSMC	54	54	0.013	0.007	0.005	0.050	0.010	0.009	0.004	0.061
PBT	M04	CR	161	81	0.011	0.009	0.004	0.071	0.007	0.003	0.002	0.017
PBT	M04	All PBA	105	46	0.011	0.007	0.004	0.050	0.007	0.003	0.002	0.014
PBT	M04	PBA Type I	31	26	0.013	0.009	0.005	0.050	0.009	0.003	0.004	0.014
PBT	M04	PBA Type II	38	11	0.009	0.004	0.004	0.014	0.005	0.001	0.003	0.007
PBT	M04	PBA Type III	36	9	0.008	0.003	0.004	0.013	0.005	0.002	0.002	0.009
PBT	M04	EOY All	110	89	0.013	0.009	0.004	0.071	0.009	0.008	0.003	0.061
PBT	M05	All Items	207	135	0.015	0.023	0.003	0.256	0.007	0.003	0.002	0.017
PBT	M05	SSMC	62	62	0.019	0.032	0.004	0.256	0.008	0.002	0.003	0.014
PBT	M05	CR	145	73	0.011	0.008	0.003	0.052	0.007	0.003	0.002	0.017
PBT	M05	All PBA	96	41	0.017	0.039	0.004	0.256	0.006	0.003	0.002	0.015
PBT	M05	PBA Type I	31	23	0.025	0.052	0.005	0.256	0.007	0.002	0.004	0.015
PBT	M05	PBA Type II	35	10	0.007	0.002	0.005	0.011	0.004	0.001	0.002	0.007
PBT	M05	PBA Type III	30	8	0.009	0.004	0.004	0.014	0.005	0.003	0.002	0.011
PBT	M05	EOY All	111	94	0.014	0.009	0.003	0.052	0.008	0.003	0.002	0.017
PBT	M06	All Items	211	132	0.019	0.027	0.004	0.173	0.008	0.004	0.003	0.026
PBT	M06	SSMC	45	45	0.031	0.043	0.006	0.173	0.008	0.002	0.005	0.016

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	M06	CR	166	87	0.012	0.008	0.004	0.038	0.009	0.005	0.003	0.026
PBT	M06	All PBA	98	43	0.015	0.024	0.004	0.161	0.007	0.004	0.003	0.016
PBT	M06	PBA Type I	26	23	0.019	0.032	0.004	0.161	0.009	0.004	0.003	0.016
PBT	M06	PBA Type II	42	12	0.010	0.006	0.005	0.029	0.004	0.002	0.003	0.009
PBT	M06	PBA Type III	30	8	0.009	0.005	0.005	0.021	0.006	0.002	0.003	0.009
PBT	M06	EOY All	113	89	0.021	0.028	0.005	0.173	0.009	0.004	0.003	0.026
PBT	M07	All Items	208	130	0.030	0.071	0.004	0.769	0.010	0.005	0.003	0.033
PBT	M07	SSMC	56	56	0.044	0.104	0.006	0.769	0.009	0.002	0.005	0.015
PBT	M07	CR	152	74	0.020	0.023	0.004	0.155	0.010	0.007	0.003	0.033
PBT	M07	All PBA	102	47	0.019	0.023	0.004	0.124	0.008	0.005	0.003	0.033
PBT	M07	PBA Type I	33	28	0.025	0.028	0.005	0.124	0.010	0.006	0.004	0.033
PBT	M07	PBA Type II	39	11	0.009	0.002	0.006	0.014	0.005	0.001	0.004	0.008
PBT	M07	PBA Type III	30	8	0.014	0.008	0.004	0.027	0.005	0.002	0.003	0.009
PBT	M07	EOY All	106	83	0.037	0.087	0.005	0.769	0.010	0.005	0.003	0.033
PBT	M08	All Items	188	116	0.022	0.022	0.000	0.156	0.009	0.005	0.000	0.030
PBT	M08	SSMC	47	47	0.025	0.029	0.000	0.156	0.008	0.003	0.000	0.015
PBT	M08	CR	141	69	0.020	0.016	0.004	0.089	0.010	0.006	0.003	0.030
PBT	M08	All PBA	96	46	0.024	0.030	0.000	0.156	0.008	0.004	0.000	0.023
PBT	M08	PBA Type I	31	28	0.030	0.036	0.000	0.156	0.009	0.004	0.000	0.018
PBT	M08	PBA Type II	38	11	0.018	0.012	0.007	0.042	0.008	0.006	0.003	0.023
PBT	M08	PBA Type III	27	7	0.010	0.003	0.006	0.013	0.005	0.002	0.003	0.009
PBT	M08	EOY All	92	70	0.021	0.016	0.004	0.089	0.010	0.005	0.003	0.030
PBT	A1	All Items	239	133	0.062	0.124	0.000	1.117	0.010	0.005	0.000	0.031
PBT	A1	SSMC	53	53	0.070	0.104	0.000	0.513	0.008	0.003	0.000	0.016
PBT	A1	CR	186	80	0.055	0.138	0.005	1.117	0.011	0.007	0.002	0.031
PBT	A1	All PBA	116	51	0.051	0.086	0.000	0.513	0.009	0.005	0.000	0.031
PBT	A1	PBA Type I	28	28	0.077	0.110	0.000	0.513	0.011	0.006	0.000	0.031

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	A1	PBA Type II	46	13	0.024	0.012	0.007	0.054	0.008	0.002	0.004	0.010
PBT	A1	PBA Type III	42	10	0.015	0.008	0.006	0.031	0.008	0.005	0.002	0.020
PBT	A1	EOY All	123	82	0.068	0.145	0.000	1.117	0.010	0.006	0.000	0.029
PBT	GO	All Items	253	135	0.040	0.053	0.000	0.512	0.020	0.011	0.000	0.055
PBT	GO	SSMC	41	41	0.055	0.086	0.000	0.512	0.018	0.008	0.000	0.042
PBT	GO	CR	212	94	0.034	0.026	0.009	0.145	0.021	0.012	0.004	0.055
PBT	GO	All PBA	110	46	0.034	0.045	0.000	0.316	0.018	0.009	0.000	0.044
PBT	GO	PBA Type I	24	24	0.041	0.060	0.000	0.316	0.020	0.008	0.000	0.040
PBT	GO	PBA Type II	38	11	0.031	0.022	0.016	0.089	0.018	0.011	0.007	0.044
PBT	GO	PBA Type III	48	11	0.020	0.010	0.010	0.040	0.013	0.008	0.004	0.031
PBT	GO	EOY All	143	89	0.043	0.057	0.009	0.512	0.022	0.012	0.007	0.055
PBT	A2	All Items	266	135	0.144	0.295	0.013	1.866	0.024	0.015	0.000	0.077
PBT	A2	SSMC	35	35	0.140	0.209	0.013	1.024	0.022	0.006	0.014	0.036
PBT	A2	CR	231	100	0.145	0.320	0.013	1.866	0.025	0.017	0.000	0.077
PBT	A2	All PBA	137	51	0.074	0.090	0.013	0.546	0.023	0.013	0.004	0.061
PBT	A2	PBA Type I	25	25	0.104	0.119	0.029	0.546	0.032	0.012	0.014	0.061
PBT	A2	PBA Type II	46	13	0.045	0.028	0.013	0.115	0.016	0.008	0.007	0.032
PBT	A2	PBA Type III	66	13	0.046	0.032	0.020	0.137	0.014	0.007	0.004	0.029
PBT	A2	EOY All	129	84	0.186	0.362	0.013	1.866	0.024	0.016	0.000	0.077
PBT	M1	All Items	118	65	0.154	0.257	0.016	1.775	0.033	0.020	0.000	0.116
PBT	M1	SSMC	21	21	0.208	0.216	0.029	0.629	0.029	0.007	0.019	0.049
PBT	M1	CR	97	44	0.129	0.272	0.016	1.775	0.035	0.024	0.000	0.116
PBT	M1	All PBA	49	23	0.161	0.196	0.016	0.629	0.028	0.012	0.012	0.059
PBT	M1	PBA Type I	14	14	0.236	0.222	0.029	0.629	0.032	0.013	0.019	0.059
PBT	M1	PBA Type II	14	4	0.035	0.010	0.025	0.047	0.023	0.005	0.017	0.030
PBT	M1	PBA Type III	21	5	0.050	0.031	0.016	0.094	0.019	0.006	0.012	0.026
PBT	M1	EOY All	69	42	0.151	0.286	0.018	1.775	0.036	0.024	0.000	0.116

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	SE of <i>b</i> Estimates				SE of <i>a</i> Estimates			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	M2	All Items	131	74	0.129	0.096	0.042	0.598	0.058	0.034	0.014	0.190
PBT	M2	SSMC	26	26	0.153	0.120	0.049	0.598	0.048	0.012	0.025	0.077
PBT	M2	CR	105	48	0.116	0.078	0.042	0.306	0.064	0.041	0.014	0.190
PBT	M2	All PBA	59	26	0.125	0.115	0.042	0.598	0.049	0.024	0.014	0.122
PBT	M2	PBA Type I	15	15	0.155	0.139	0.049	0.598	0.059	0.025	0.037	0.122
PBT	M2	PBA Type II	17	5	0.101	0.077	0.053	0.237	0.035	0.011	0.025	0.053
PBT	M2	PBA Type III	27	6	0.072	0.030	0.042	0.123	0.033	0.017	0.014	0.055
PBT	M2	EOY All	72	48	0.131	0.084	0.045	0.380	0.063	0.038	0.021	0.190
PBT	M3	All Items	120	66	0.217	0.211	0.026	0.940	0.062	0.037	0.010	0.208
PBT	M3	SSMC	23	23	0.273	0.266	0.058	0.940	0.066	0.023	0.036	0.130
PBT	M3	CR	97	43	0.187	0.171	0.026	0.821	0.059	0.043	0.010	0.208
PBT	M3	All PBA	52	20	0.216	0.243	0.026	0.940	0.046	0.023	0.010	0.091
PBT	M3	PBA Type I	10	10	0.353	0.285	0.052	0.940	0.062	0.018	0.036	0.091
PBT	M3	PBA Type II	18	5	0.077	0.062	0.026	0.183	0.029	0.011	0.021	0.048
PBT	M3	PBA Type III	24	5	0.081	0.035	0.035	0.117	0.032	0.020	0.010	0.057
PBT	M3	EOY All	68	46	0.218	0.198	0.039	0.884	0.069	0.040	0.012	0.208

Note: M03 through M08 = mathematics grades 3 through 8, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table A.10.11 CBT IRT Model Fit for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M03	All Items	378	267	693	1125	41	9550	0.07	0.04	0.03	0.27
CBT	M03	SSMC	54	54	416	612	41	3513	0.06	0.03	0.03	0.14
CBT	M03	CR	324	213	764	1213	57	9550	0.08	0.04	0.03	0.27
CBT	M03	All PBA	168	79	777	1270	49	9550	0.08	0.04	0.03	0.22
CBT	M03	PBA Type I	57	48	392	367	49	1780	0.06	0.02	0.03	0.11
CBT	M03	PBA Type II	63	18	925	591	364	2095	0.10	0.03	0.05	0.18
CBT	M03	PBA Type III	48	13	1991	2691	416	9550	0.12	0.04	0.08	0.22
CBT	M03	EOY All	210	188	658	1060	41	7111	0.07	0.04	0.03	0.27
CBT	M04	All Items	294	193	1242	1440	77	9073	0.09	0.04	0.03	0.26
CBT	M04	SSMC	50	50	892	1365	92	7772	0.07	0.04	0.03	0.26
CBT	M04	CR	244	143	1364	1450	77	9073	0.10	0.04	0.03	0.26
CBT	M04	All PBA	143	63	1345	1023	102	4565	0.10	0.04	0.04	0.21
CBT	M04	PBA Type I	45	37	973	817	102	3743	0.08	0.03	0.04	0.17
CBT	M04	PBA Type II	53	15	1881	1213	506	4565	0.12	0.03	0.08	0.17
CBT	M04	PBA Type III	45	11	1868	882	609	3304	0.12	0.04	0.08	0.21
CBT	M04	EOY All	151	130	1192	1605	77	9073	0.09	0.04	0.03	0.26
CBT	M05	All Items	312	204	1302	1611	68	9390	0.09	0.04	0.03	0.24
CBT	M05	SSMC	45	45	611	537	68	2317	0.07	0.03	0.03	0.15
CBT	M05	CR	267	159	1498	1755	85	9390	0.10	0.04	0.03	0.24
CBT	M05	All PBA	139	61	1577	1638	106	6292	0.10	0.04	0.04	0.19
CBT	M05	PBA Type I	48	37	1143	1417	106	6040	0.08	0.04	0.04	0.17
CBT	M05	PBA Type II	43	12	2746	1970	399	6292	0.13	0.04	0.07	0.18
CBT	M05	PBA Type III	48	12	1746	1421	374	5579	0.12	0.04	0.06	0.19
CBT	M05	EOY All	173	143	1185	1591	68	9390	0.09	0.04	0.03	0.24
CBT	M06	All Items	303	191	1230	1469	85	8343	0.08	0.04	0.03	0.25
CBT	M06	SSMC	38	38	1111	1624	85	7979	0.08	0.05	0.04	0.21

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M06	CR	265	153	1259	1432	100	8343	0.09	0.04	0.03	0.25
CBT	M06	All PBA	146	66	1446	1228	100	5757	0.10	0.04	0.04	0.17
CBT	M06	PBA Type I	46	39	1188	1164	100	5367	0.09	0.03	0.04	0.16
CBT	M06	PBA Type II	52	15	1721	1384	354	5757	0.10	0.03	0.07	0.17
CBT	M06	PBA Type III	48	12	1943	1089	446	3802	0.12	0.03	0.08	0.17
CBT	M06	EOY All	157	125	1116	1575	85	8343	0.08	0.04	0.03	0.25
CBT	M07	All Items	382	236	1105	1356	35	7471	0.09	0.05	0.02	0.26
CBT	M07	SSMC	48	48	885	1165	62	5635	0.08	0.04	0.03	0.19
CBT	M07	CR	334	188	1161	1397	35	7471	0.09	0.05	0.02	0.26
CBT	M07	All PBA	172	77	1413	1453	80	7411	0.10	0.04	0.03	0.20
CBT	M07	PBA Type I	56	46	1078	1222	80	5635	0.09	0.04	0.03	0.20
CBT	M07	PBA Type II	56	16	1996	1753	613	7411	0.12	0.03	0.07	0.17
CBT	M07	PBA Type III	60	15	1817	1560	374	5282	0.13	0.05	0.05	0.20
CBT	M07	EOY All	210	159	956	1284	35	7471	0.08	0.05	0.02	0.26
CBT	M08	All Items	296	193	907	1199	49	7813	0.08	0.04	0.03	0.24
CBT	M08	SSMC	44	44	806	719	49	3308	0.08	0.04	0.03	0.16
CBT	M08	CR	252	149	936	1308	73	7813	0.08	0.04	0.03	0.24
CBT	M08	All PBA	115	57	1316	1506	49	7813	0.09	0.04	0.03	0.22
CBT	M08	PBA Type I	39	36	816	911	49	4494	0.07	0.03	0.03	0.14
CBT	M08	PBA Type II	46	13	1484	1216	160	3724	0.10	0.04	0.04	0.17
CBT	M08	PBA Type III	30	8	3294	2377	893	7813	0.14	0.05	0.08	0.22
CBT	M08	EOY All	181	136	735	1001	73	7374	0.08	0.04	0.03	0.24
CBT	A1	All Items	470	263	805	1055	30	6887	0.09	0.04	0.02	0.23
CBT	A1	SSMC	66	66	501	550	40	3642	0.08	0.03	0.02	0.15
CBT	A1	CR	404	197	907	1160	30	6887	0.09	0.05	0.02	0.23
CBT	A1	All PBA	197	91	968	1217	30	5550	0.10	0.05	0.03	0.23
CBT	A1	PBA Type I	54	54	405	325	30	1547	0.07	0.03	0.03	0.13

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	A1	PBA Type II	68	19	1209	1172	259	3700	0.11	0.04	0.06	0.20
CBT	A1	PBA Type III	75	18	2400	1680	522	5550	0.15	0.05	0.08	0.23
CBT	A1	EOY All	273	172	719	951	40	6887	0.09	0.04	0.02	0.21
CBT	GO	All Items	505	285	421	487	46	3434	0.10	0.05	0.03	0.25
CBT	GO	SSMC	51	51	316	326	46	1348	0.09	0.04	0.03	0.18
CBT	GO	CR	454	234	443	513	46	3434	0.10	0.05	0.03	0.25
CBT	GO	All PBA	217	96	488	556	51	3095	0.11	0.05	0.03	0.25
CBT	GO	PBA Type I	56	56	260	212	51	982	0.08	0.03	0.03	0.17
CBT	GO	PBA Type II	74	21	604	502	191	2364	0.13	0.04	0.05	0.21
CBT	GO	PBA Type III	87	19	1031	853	125	3095	0.15	0.06	0.06	0.25
CBT	GO	EOY All	288	189	387	445	46	3434	0.10	0.04	0.03	0.23
CBT	A2	All Items	373	194	646	834	50	4541	0.10	0.05	0.03	0.31
CBT	A2	SSMC	39	39	394	320	53	1630	0.09	0.03	0.03	0.17
CBT	A2	CR	334	155	709	909	50	4541	0.10	0.06	0.03	0.31
CBT	A2	All PBA	190	76	700	805	57	3995	0.10	0.05	0.03	0.24
CBT	A2	PBA Type I	41	41	301	275	57	1156	0.07	0.03	0.03	0.17
CBT	A2	PBA Type II	68	19	955	1029	132	3995	0.13	0.04	0.07	0.20
CBT	A2	PBA Type III	81	16	1417	838	131	2905	0.16	0.06	0.04	0.24
CBT	A2	EOY All	183	118	611	853	50	4541	0.10	0.05	0.03	0.31
CBT	M1	All Items	134	71	255	254	38	1243	0.11	0.05	0.05	0.31
CBT	M1	SSMC	13	13	147	78	67	335	0.08	0.02	0.05	0.11
CBT	M1	CR	121	58	280	273	38	1243	0.11	0.05	0.05	0.31
CBT	M1	All PBA	62	27	260	273	47	1052	0.11	0.05	0.05	0.25
CBT	M1	PBA Type I	15	15	100	39	47	175	0.07	0.02	0.05	0.10
CBT	M1	PBA Type II	20	6	288	210	73	536	0.11	0.03	0.07	0.13
CBT	M1	PBA Type III	27	6	634	307	271	1052	0.18	0.06	0.12	0.25
CBT	M1	EOY All	72	44	252	244	38	1243	0.11	0.05	0.05	0.31

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
CBT	M2	All Items	141	77	143	128	20	541	0.12	0.05	0.04	0.27
CBT	M2	SSMC	16	16	101	70	40	243	0.11	0.04	0.07	0.18
CBT	M2	CR	125	61	153	138	20	541	0.13	0.05	0.04	0.26
CBT	M2	All PBA	68	28	178	151	20	541	0.13	0.06	0.04	0.26
CBT	M2	PBA Type I	15	15	93	61	20	243	0.10	0.04	0.04	0.18
CBT	M2	PBA Type II	17	5	240	169	70	458	0.14	0.04	0.10	0.18
CBT	M2	PBA Type III	36	8	299	170	111	541	0.20	0.05	0.12	0.26
CBT	M2	EOY All	73	49	122	110	35	475	0.12	0.03	0.06	0.20
CBT	M3	All Items	144	74	130	111	24	535	0.14	0.05	0.07	0.32
CBT	M3	SSMC	18	18	84	43	29	166	0.13	0.03	0.07	0.19
CBT	M3	CR	126	56	145	122	24	535	0.15	0.06	0.07	0.32
CBT	M3	All PBA	75	28	176	137	33	535	0.16	0.06	0.07	0.29
CBT	M3	PBA Type I	14	14	91	46	33	166	0.12	0.04	0.07	0.19
CBT	M3	PBA Type II	25	7	238	159	100	535	0.19	0.03	0.14	0.23
CBT	M3	PBA Type III	36	7	284	138	122	440	0.20	0.07	0.13	0.29
CBT	M3	EOY All	69	46	103	83	24	434	0.14	0.05	0.07	0.32

Note: M03 through M08 = mathematics grades 3 through 8, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Table A.10.12 PBT IRT Model Fit for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	M03	All Items	233	157	560	722	67	5886	0.08	0.04	0.03	0.28
PBT	M03	SSMC	55	55	390	476	67	2572	0.07	0.03	0.03	0.16
PBT	M03	CR	178	102	651	813	82	5886	0.09	0.04	0.04	0.27
PBT	M03	All PBA	110	48	655	904	67	5886	0.09	0.04	0.04	0.27
PBT	M03	PBA Type I	32	27	350	234	67	968	0.07	0.02	0.04	0.13
PBT	M03	PBA Type II	42	12	693	478	222	1755	0.10	0.03	0.06	0.15
PBT	M03	PBA Type III	36	9	1521	1775	373	5886	0.14	0.06	0.08	0.27
PBT	M03	EOY All	123	109	517	626	82	3914	0.08	0.04	0.03	0.20
PBT	M04	All Items	213	134	764	1029	66	8137	0.10	0.05	0.03	0.32
PBT	M04	SSMC	54	54	512	625	74	3523	0.08	0.04	0.03	0.21
PBT	M04	CR	159	80	934	1204	66	8137	0.11	0.05	0.04	0.32
PBT	M04	All PBA	105	46	789	691	91	3613	0.11	0.04	0.04	0.24
PBT	M04	PBA Type I	31	26	560	503	91	1919	0.09	0.03	0.04	0.18
PBT	M04	PBA Type II	38	11	1060	631	237	2223	0.13	0.03	0.09	0.19
PBT	M04	PBA Type III	36	9	1116	1004	421	3613	0.14	0.05	0.08	0.24
PBT	M04	EOY All	108	88	751	1171	66	8137	0.10	0.05	0.03	0.32
PBT	M05	All Items	207	135	721	862	39	6293	0.10	0.04	0.04	0.22
PBT	M05	SSMC	62	62	477	580	55	2845	0.08	0.04	0.04	0.22
PBT	M05	CR	145	73	928	1002	39	6293	0.12	0.04	0.04	0.22
PBT	M05	All PBA	96	41	810	624	86	2315	0.11	0.05	0.05	0.22
PBT	M05	PBA Type I	31	23	684	641	86	2111	0.10	0.04	0.05	0.22
PBT	M05	PBA Type II	35	10	959	440	427	1662	0.13	0.04	0.07	0.21
PBT	M05	PBA Type III	30	8	985	753	196	2315	0.12	0.05	0.07	0.19
PBT	M05	EOY All	111	94	682	947	39	6293	0.09	0.04	0.04	0.22
PBT	M06	All Items	210	131	590	812	65	5419	0.10	0.05	0.04	0.29
PBT	M06	SSMC	44	44	432	616	65	3287	0.08	0.05	0.04	0.23

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	M06	CR	166	87	670	888	69	5419	0.11	0.05	0.04	0.28
PBT	M06	All PBA	98	43	796	1079	65	5419	0.11	0.05	0.04	0.28
PBT	M06	PBA Type I	26	23	610	1087	65	5419	0.09	0.04	0.04	0.21
PBT	M06	PBA Type II	42	12	724	449	262	1540	0.13	0.03	0.08	0.18
PBT	M06	PBA Type III	30	8	1440	1535	248	4902	0.16	0.06	0.11	0.28
PBT	M06	EOY All	112	88	489	627	69	3287	0.09	0.05	0.04	0.23
PBT	M07	All Items	208	130	553	581	39	2512	0.10	0.05	0.04	0.24
PBT	M07	SSMC	56	56	396	425	39	2474	0.09	0.04	0.04	0.20
PBT	M07	CR	152	74	672	654	52	2512	0.11	0.05	0.04	0.24
PBT	M07	All PBA	102	47	611	579	52	2512	0.11	0.05	0.04	0.22
PBT	M07	PBA Type I	33	28	437	529	52	1952	0.09	0.04	0.04	0.17
PBT	M07	PBA Type II	39	11	864	428	204	1592	0.15	0.03	0.11	0.19
PBT	M07	PBA Type III	30	8	875	747	199	2512	0.14	0.05	0.06	0.22
PBT	M07	EOY All	106	83	520	584	39	2508	0.10	0.05	0.04	0.24
PBT	M08	All Items	186	114	480	629	47	4449	0.10	0.04	0.03	0.27
PBT	M08	SSMC	45	45	360	321	57	1493	0.08	0.03	0.04	0.14
PBT	M08	CR	141	69	558	759	47	4449	0.10	0.05	0.03	0.27
PBT	M08	All PBA	94	44	459	491	47	2859	0.10	0.04	0.03	0.24
PBT	M08	PBA Type I	29	26	275	216	47	765	0.08	0.03	0.03	0.14
PBT	M08	PBA Type II	38	11	447	242	80	838	0.11	0.02	0.08	0.13
PBT	M08	PBA Type III	27	7	1159	847	443	2859	0.17	0.05	0.11	0.24
PBT	M08	EOY All	92	70	493	705	69	4449	0.09	0.04	0.04	0.27
PBT	A1	All Items	234	128	364	579	53	5843	0.10	0.05	0.04	0.33
PBT	A1	SSMC	48	48	309	827	57	5843	0.09	0.05	0.04	0.33
PBT	A1	CR	186	80	397	359	53	1535	0.11	0.05	0.04	0.23
PBT	A1	All PBA	114	49	319	285	72	1535	0.10	0.04	0.04	0.23
PBT	A1	PBA Type I	26	26	175	110	72	560	0.08	0.02	0.04	0.13

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	A1	PBA Type II	46	13	319	186	124	685	0.11	0.03	0.06	0.18
PBT	A1	PBA Type III	42	10	695	372	182	1535	0.15	0.04	0.09	0.23
PBT	A1	EOY All	120	79	392	703	53	5843	0.10	0.05	0.04	0.33
PBT	GO	All Items	252	134	151	131	37	822	0.13	0.05	0.05	0.30
PBT	GO	SSMC	40	40	113	133	37	822	0.11	0.04	0.05	0.20
PBT	GO	CR	212	94	168	128	40	679	0.14	0.05	0.06	0.30
PBT	GO	All PBA	109	45	177	151	38	679	0.14	0.05	0.07	0.30
PBT	GO	PBA Type I	23	23	101	64	38	314	0.11	0.03	0.07	0.19
PBT	GO	PBA Type II	38	11	184	103	63	431	0.15	0.03	0.11	0.19
PBT	GO	PBA Type III	48	11	328	204	121	679	0.20	0.05	0.11	0.30
PBT	GO	EOY All	143	89	139	119	37	822	0.13	0.05	0.05	0.23
PBT	A2	All Items	265	134	159	296	17	2358	0.15	0.08	0.06	0.70
PBT	A2	SSMC	35	35	76	46	27	244	0.12	0.03	0.07	0.19
PBT	A2	CR	230	99	189	338	17	2358	0.16	0.09	0.06	0.70
PBT	A2	All PBA	137	51	126	89	17	372	0.15	0.06	0.06	0.31
PBT	A2	PBA Type I	25	25	62	33	17	173	0.10	0.03	0.06	0.15
PBT	A2	PBA Type II	46	13	192	91	61	372	0.19	0.02	0.14	0.23
PBT	A2	PBA Type III	66	13	186	72	50	319	0.19	0.07	0.08	0.31
PBT	A2	EOY All	128	83	179	369	27	2358	0.15	0.09	0.07	0.70
PBT	M1	All Items	118	65	90	82	20	483	0.18	0.06	0.08	0.35
PBT	M1	SSMC	21	21	48	11	28	66	0.14	0.03	0.09	0.22
PBT	M1	CR	97	44	110	93	20	483	0.20	0.07	0.08	0.35
PBT	M1	All PBA	49	23	86	71	28	337	0.18	0.06	0.10	0.30
PBT	M1	PBA Type I	14	14	48	10	28	64	0.15	0.03	0.10	0.22
PBT	M1	PBA Type II	14	4	108	43	72	165	0.21	0.03	0.18	0.25
PBT	M1	PBA Type III	21	5	173	105	79	337	0.25	0.06	0.16	0.30
PBT	M1	EOY All	69	42	93	88	20	483	0.18	0.07	0.08	0.35

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	M2	All Items	131	74	43	25	4	111	0.23	0.05	0.10	0.37
PBT	M2	SSMC	26	26	30	13	4	63	0.19	0.04	0.10	0.29
PBT	M2	CR	105	48	50	27	11	111	0.24	0.05	0.13	0.37
PBT	M2	All PBA	59	26	45	24	19	111	0.23	0.05	0.15	0.32
PBT	M2	PBA Type I	15	15	31	11	19	63	0.20	0.04	0.15	0.29
PBT	M2	PBA Type II	17	5	72	31	42	111	0.26	0.03	0.22	0.29
PBT	M2	PBA Type III	27	6	59	20	34	94	0.28	0.03	0.25	0.32
PBT	M2	EOY All	72	48	42	26	4	103	0.22	0.05	0.10	0.37
PBT	M3	All Items	119	65	53	36	14	228	0.22	0.05	0.14	0.38
PBT	M3	SSMC	23	23	36	20	14	104	0.20	0.04	0.14	0.27
PBT	M3	CR	96	42	62	39	15	228	0.23	0.05	0.15	0.38
PBT	M3	All PBA	52	20	75	48	30	228	0.22	0.06	0.15	0.38
PBT	M3	PBA Type I	10	10	51	21	30	104	0.18	0.03	0.15	0.26
PBT	M3	PBA Type II	18	5	87	33	45	123	0.25	0.05	0.18	0.30
PBT	M3	PBA Type III	24	5	110	77	41	228	0.27	0.07	0.21	0.38
PBT	M3	EOY All	67	45	43	24	14	101	0.22	0.04	0.14	0.32

Note: M03 through M08 = mathematics grades 3 through 8, A1 = Algebra I, GO = Geometry, A2 = Algebra II, M1 = Integrated Mathematics I, M2 = Integrated Mathematics II, M3 = Integrated Mathematics III.

Appendix 12.1: Form Composition

Table A.12.1 Form Composition for ELA/L Grade 3

Claims	Subclaims	Number of Items	Number of Points
Reading	Reading Literary Text	13 - 16	27 - 33
	Reading Informational Text	9 - 11	19 - 23
	Vocabulary	5 - 9	10 - 18
	Claim Total	31	64
Writing	Written Expression	3	27
	Knowledge of Conventions	3	9
	Claim Total	6	36
SUMMATIVE TOTAL		37	100

Note: This table is identical to Table 12.1 in Section 12.

Table A.12.2 Form Composition for ELA/L Grade 4

Claims	Subclaims	Number of Items	Number of Points
Reading	Reading Literary Text	12 - 14	25 - 29
	Reading Informational Text	12 - 14	25 - 29
	Vocabulary	6 - 9	12 - 18
	Claim Total	33 - 34	68 - 70
Writing	Written Expression	3	27
	Knowledge of Conventions	3	9
	Claim Total	6	36
SUMMATIVE TOTAL		39 - 40	104 - 106

Table A.12.3 Form Composition for ELA/L Grade 5

Claims	Subclaims	Number of Items	Number of Points
Reading	Reading Literary Text	11 - 14	23 - 29
	Reading Informational Text	12 - 13	25 - 27
	Vocabulary	8 - 10	16 - 20
	Claim Total	33 - 34	68 - 70
Writing	Written Expression	3	27
	Knowledge of Conventions	3	9
	Claim Total	6	36
SUMMATIVE TOTAL		39 - 40	104 - 106

Table A.12.4 Form Composition for ELA/L Grade 6

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	13 - 17	28 - 36
	Reading Informational Text	16 - 23	34 - 48
	Vocabulary	7 - 12	14 - 24
	Claim Total	44	92
Writing			
	Written Expression	3	36
	Knowledge of Conventions	3	9
	Claim Total	6	45
SUMMATIVE TOTAL		50	137

Table A.12.5 Form Composition for ELA/L Grade 7

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	13 - 19	28 - 40
	Reading Informational Text	15 - 21	32 - 44
	Vocabulary	8 - 13	16 - 26
	Claim Total	42 - 44	88 - 92
Writing			
	Written Expression	3	36
	Knowledge of Conventions	3	9
	Claim Total	6	45
SUMMATIVE TOTAL		48 - 50	133 - 137

Table A.12.6 Form Composition for ELA/L Grade 8

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	12 - 19	26 - 40
	Reading Informational Text	15 - 22	32 - 46
	Vocabulary	8 - 10	16 - 20
	Claim Total	43 - 44	90 - 92
Writing			
	Written Expression	3	36
	Knowledge of Conventions	3	9
	Claim Total	6	45
SUMMATIVE TOTAL		49 - 50	135 - 137

Table A.12.7 Form Composition for ELA/L Grade 9

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	12 - 20	26 - 42
	Reading Informational Text	14 - 22	30 - 46
	Vocabulary	8 - 11	16 - 22
	Claim Total	43 - 44	90 - 92
Writing			
	Written Expression	3	36
	Knowledge of Conventions	3	9
	Claim Total	6	45
SUMMATIVE TOTAL		49 - 50	135 - 137

Table A.12.8 Form Composition for ELA/L Grade 10

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	13 - 20	28 - 42
	Reading Informational Text	15 - 21	32 - 44
	Vocabulary	8 - 11	16 - 22
	Claim Total	43 - 44	90 - 92
Writing			
	Written Expression	3	36
	Knowledge of Conventions	3	9
	Claim Total	6	45
SUMMATIVE TOTAL		49 - 50	135 - 137

Table A.12.9 Form Composition for ELA/L Grade 11

Claims	Subclaims	Number of Items	Number of Points
Reading			
	Reading Literary Text	12 - 20	26 - 42
	Reading Informational Text	14 - 21	30 - 44
	Vocabulary	8 - 13	16 - 26
	Claim Total	43 - 44	90 - 92
Writing			
	Written Expression	3	36
	Knowledge of Conventions	3	9
	Claim Total	6	45
SUMMATIVE TOTAL		49 - 50	135 - 137

Table A.12.10 Form Composition for Mathematics Grade 3

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	36	43
	Additional & Supporting Content	12 - 13	12 - 13
	Expressing Mathematical Reasoning	3 - 4	11 - 14
	Modeling and Applications	3	12
TOTAL		54 - 56	78 - 82

Note: This table is identical to Table 12.3 in Section 12.

Table A.12.11 Form Composition for Mathematics Grade 4

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	38	47
	Additional & Supporting Content	8	9
	Expressing Mathematical Reasoning	4	14
	Modeling and Applications	3	12
TOTAL		53	82

Table A.12.12 Form Composition for Mathematics Grade 5

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	35	44
	Additional & Supporting Content	9 - 10	11 - 12
	Expressing Mathematical Reasoning	4	14
	Modeling and Applications	3	12
TOTAL		51 - 52	81 - 82

Table A.12.13 Form Composition for Mathematics Grade 6

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	27 - 28	35 - 36
	Additional & Supporting Content	16	20
	Expressing Mathematical Reasoning	4	14
	Modeling and Applications	3	12
TOTAL		50 - 51	81 - 82

Table A.12.14 Form Composition for Mathematics Grade 7

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	33 - 34	41 - 43
	Additional & Supporting Content	8 - 9	11 - 13
	Expressing Mathematical Reasoning	4	14
	Modeling and Applications	3	12
TOTAL		48 - 50	78 - 82

Table A.12.15 Form Composition for Mathematics Grade 8

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	35 - 36	41 - 43
	Additional & Supporting Content	8	13
	Expressing Mathematical Reasoning	4	14
	Modeling and Applications	3	12
TOTAL		50 - 51	80 - 82

Table A.12.16 Form Composition for Algebra I

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	25 - 28	32 - 37
	Additional & Supporting Content	17 - 18	28 - 32
	Expressing Mathematical Reasoning	3 - 4	11 - 14
	Modeling and Applications	4	18
TOTAL		50 - 53	93 - 97

Table A.12.17 Form Composition for Geometry

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	25 - 26	38 - 39
	Additional & Supporting Content	17 - 19	24 - 26
	Expressing Mathematical Reasoning	4	13 - 14
	Modeling and Applications	4	18
TOTAL		51 - 53	95 - 97

Table A.12.18 Form Composition for Algebra II

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	25 - 27	34 - 40
	Additional & Supporting Content	16 - 19	21 - 31
	Expressing Mathematical Reasoning	4 - 5	14 - 18
	Modeling and Applications	5	22 - 24
TOTAL		51 - 54	96 - 107

Table A.12.19 Form Composition for Integrated Mathematics I

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	27 - 28	38 - 39
	Additional & Supporting Content	13 - 14	22 - 23
	Expressing Mathematical Reasoning	3 - 4	11 - 14
	Modeling and Applications	4	18
TOTAL		47 - 50	89 - 94

Table A.12.20 Form Composition for Integrated Mathematics II

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	28 - 30	37 - 42
	Additional & Supporting Content	14	23 - 25
	Expressing Mathematical Reasoning	4	14
	Modeling and Applications	3 - 4	15 - 18
TOTAL		49 - 52	91 - 97

Table A.12.21 Form Composition for Integrated Mathematics III

	Subclaims	Number of Items	Number of Points
Mathematics			
	Major Content	24 - 28	33 - 37
	Additional & Supporting Content	14 - 16	21 - 26
	Expressing Mathematical Reasoning	4 - 5	15 - 18
	Modeling and Applications	5	23 - 24
TOTAL		49 - 54	96 - 104

Appendix 12.2: Scaling Constants and Associated Information

Table A.12.22 Threshold Scores and Scaling Constants for ELA/L Grades 3 to 8

PARCC Assessment	Threshold Cut	Raw Score	Theta	Scale Score	A	B
Grade 3 ELA	Level 2 Cut	17	-0.9769	700	36.224	735.3872
	Level 3 Cut	28	-0.2867	726		
	Level 4 Cut	40	0.4034	750		
	Level 5 Cut	67	2.0652	810		
Grade 4 ELA	Level 2 Cut	19	-1.3276	700	30.7863	740.8719
	Level 3 Cut	33	-0.5122	725		
	Level 4 Cut	50	0.2965	750		
	Level 5 Cut	76	1.6011	790		
Grade 5 ELA	Level 2 Cut	15	-1.3768	700	29.469	740.5729
	Level 3 Cut	28	-0.5285	726		
	Level 4 Cut	44	0.3199	750		
	Level 5 Cut	77	1.9854	799		
Grade 6 ELA	Level 2 Cut	25	-1.3649	700	28.7257	739.2078
	Level 3 Cut	43	-0.4787	725		
	Level 4 Cut	64	0.3757	750		
	Level 5 Cut	98	1.7686	790		
Grade 7 ELA	Level 2 Cut	24	-1.1752	700	33.8845	739.8211
	Level 3 Cut	40	-0.4375	725		
	Level 4 Cut	60	0.3004	750		
	Level 5 Cut	88	1.3373	785		
Grade 8 ELA	Level 2 Cut	27	-1.1431	700	34.3289	739.2413
	Level 3 Cut	43	-0.4203	725		
	Level 4 Cut	63	0.3134	750		
	Level 5 Cut	96	1.5827	794		

Note: Raw score values are the cuts for a single form from the performance level setting. The Level 3 cut scores were set to the midpoint between the Level 2 and Level 4 cut scores on the scale score metric.

Table A.12.23 Threshold Scores and Scaling Constants for Mathematics Grades 3 to 8

PARCC Assessment	Threshold Cut	Raw Score	Theta	Scale Score	A	B
Grade 3 Mathematics	Level 2 Cut	15	-1.2403	700	31.849	739.503
	Level 3 Cut	24	-0.4553	727		
	Level 4 Cut	33	0.3296	750		
	Level 5 Cut	56	1.5902	790		
Grade 4 Mathematics	Level 2 Cut	15	-1.2763	700	29.558	737.725
	Level 3 Cut	26	-0.4305	727		
	Level 4 Cut	37	0.4153	750		
	Level 5 Cut	63	1.9791	796		
Grade 5 Mathematics	Level 2 Cut	12	-1.273	700	28.981	736.892
	Level 3 Cut	21	-0.417	725		
	Level 4 Cut	34	0.4523	750		
	Level 5 Cut	58	1.8156	790		
Grade 6 Mathematics	Level 2 Cut	11	-1.2700	700	28.658	736.396
	Level 3 Cut	22	-0.3896	725		
	Level 4 Cut	37	0.4747	750		
	Level 5 Cut	63	1.8136	788		
Grade 7 Mathematics	Level 2 Cut	8	-1.3483	700	25.678	734.622
	Level 3 Cut	15	-0.3733	725		
	Level 4 Cut	26	0.5989	750		
	Level 5 Cut	54	2.0131	786		
Grade 8 Mathematics	Level 2 Cut	11	-0.8417	700	33.562	728.249
	Level 3 Cut	18	-0.0968	728		
	Level 4 Cut	26	0.6481	750		
	Level 5 Cut	55	2.177	801		

Note: Raw score values are the cuts for a single form from the performance level setting. The Level 3 cut scores were set to the midpoint between the Level 2 and Level 4 cut scores on the scale score metric.

Table A.12.24 Threshold Scores and Scaling Constants for High School ELA

PARCC Assessment	Threshold Cut	Raw Score	Theta	Scale Score	A	B
Grade 9 ELA	Level 2 Cut	20	-1.0779	700	34.256	736.9245
	Level 3 Cut	35	-0.3481	726		
	Level 4 Cut	53	0.3817	750		
	Level 5 Cut	86	1.5891	791		
Grade 10 ELA	Level 2 Cut	29	-0.8354	700	42.4556	735.4674
	Level 3 Cut	44	-0.2465	725		
	Level 4 Cut	62	0.3423	750		
	Level 5 Cut	92	1.3758	794		
Grade 11 ELA	Level 2 Cut	19	-1.0889	700	36.6354	739.8923
	Level 3 Cut	33	-0.4065	726		
	Level 4 Cut	51	0.2759	750		
	Level 5 Cut	85	1.4112	792		

Note: Raw score values are the cuts for a single form from the performance level setting. The Level 3 cut scores were set to the midpoint between the Level 2 and Level 4 cut scores on the scale score metric.

Table A.12.25 Threshold Scores and Scaling Constants for High School Mathematics

PARCC Assessment	Threshold Cut	Raw Score	Theta	Scale Score	A	B
Algebra I	Level 2 Cut	9	-1.1045	700	30.3269	733.4961
	Level 3 Cut	15	-0.2802	728		
	Level 4 Cut	22	0.5442	750		
	Level 5 Cut	53	2.3716	805		
Algebra II	Level 2 Cut	12	-0.5384	700	35.6176	719.1765
	Level 3 Cut	19	0.1635	726		
	Level 4 Cut	28	0.8654	750		
	Level 5 Cut	62	2.4823	808		
Geometry	Level 2 Cut	9	-1.2478	700	25.2016	731.4466
	Level 3 Cut	16	-0.2558	726		
	Level 4 Cut	27	0.7362	750		
	Level 5 Cut	56	2.0459	783		
Integrated Mathematics I	Level 2 Cut	11	-0.8966	700	33.2624	729.823
	Level 3 Cut	18	-0.145	726		
	Level 4 Cut	28	0.6066	750		
	Level 5 Cut	57	2.072	799		
Integrated Mathematics II	Level 2 Cut	11	-0.9298	700	29.3118	727.254
	Level 3 Cut	17	-0.0769	725		
	Level 4 Cut	26	0.776	750		
	Level 5 Cut	47	1.9669	785		
Integrated Mathematics III	Level 2 Cut	12	-0.4075	700	39.5288	716.108
	Level 3 Cut	19	0.2249	726		
	Level 4 Cut	28	0.8574	750		
	Level 5 Cut	61	2.2163	804		

Note: Raw score values are the cuts for a single form from the performance level setting. The Level 3 cut scores were set to the midpoint between the Level 2 and Level 4 cut scores on the scale score metric.

Table A.12.26 Scaling Constants for Reading and Writing Grades 3 to 8

	Reading		Writing	
	A_R	B_R	A_W	B_W
Grade 3 ELA	14.4896	44.1549	7.2448	32.0774
Grade 4 ELA	12.3145	46.3488	6.1573	33.1744
Grade 5 ELA	11.7876	46.2291	5.8938	33.1146
Grade 6 ELA	11.4903	45.6831	5.7451	32.8416
Grade 7 ELA	13.5538	45.9284	6.7769	32.9642
Grade 8 ELA	13.7315	45.6965	6.8658	32.8483
Grade 9 ELA	13.7024	44.7698	6.8512	32.3849
Grade 10 ELA	16.9823	44.1870	8.4911	32.0935
Grade 11 ELA	14.6542	45.9569	7.3271	32.9785

Appendix 12.3: Raw-to-Scale Conversion Tables for Performance Level Setting (PLS) Forms

Table A.12.27 Conversion Table for Performance Level Setting Form: ELA/L Grade 3

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	2.5	518.0432	650	15
1	-5.7004	2.3	528.8959	650	15
2	-4.0919	1.2	587.1622	650	15
3	-3.3588	0.9	613.7180	650	15
4	-2.9022	0.7	630.2579	650	15
5	-2.5750	0.6	642.1104	650	15
6	-2.3214	0.5	651.2968	651	15
7	-2.1144	0.5	658.7952	659	15
8	-1.9395	0.4	665.1308	665	15
9	-1.7878	0.4	670.6259	671	14.8
10	-1.6535	0.4	675.4908	675	14.0
11	-1.5327	0.4	679.8667	680	13.2
12	-1.4227	0.3	683.8513	684	12.6
13	-1.3214	0.3	687.5208	688	12.1
14	-1.2273	0.3	690.9295	691	11.7
15	-1.1391	0.3	694.1244	694	11.3
16	-1.0559	0.3	697.1383	697	11.0
17	-0.9769	0.3	700.0000	700	10.7
18	-0.9016	0.3	702.7276	703	10.5
19	-0.8294	0.3	705.3430	705	10.2
20	-0.7599	0.3	707.8606	708	10.1
21	-0.6926	0.3	710.2985	710	9.9
22	-0.6274	0.3	712.6603	713	9.8
23	-0.5639	0.3	714.9605	715	9.6
24	-0.5018	0.3	717.2100	717	9.5
25	-0.4411	0.3	719.4088	719	9.4
26	-0.3814	0.3	721.5714	722	9.4
27	-0.3227	0.3	723.6977	724	9.3
28	-0.2648	0.3	725.7951	726	9.3
29	-0.2075	0.3	727.8707	728	9.2
30	-0.1509	0.3	729.9210	730	9.2
31	-0.0946	0.3	731.9604	732	9.2
32	-0.0388	0.3	733.9817	734	9.2
33	0.0167	0.3	735.9921	736	9.2
34	0.0721	0.3	737.9990	738	9.2
35	0.1273	0.3	739.9985	740	9.2
36	0.1825	0.3	741.9981	742	9.2
37	0.2376	0.3	743.9940	744	9.3
38	0.2928	0.3	745.9936	746	9.3
39	0.3480	0.3	747.9932	748	9.3
40	0.4034	0.3	750.0000	750	9.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
41	0.4590	0.3	752.0140	752	9.4
42	0.5149	0.3	754.0389	754	9.5
43	0.5709	0.3	756.0675	756	9.5
44	0.6273	0.3	758.1105	758	9.6
45	0.6840	0.3	760.1644	760	9.7
46	0.7410	0.3	762.2292	762	9.7
47	0.7983	0.3	764.3048	764	9.8
48	0.8561	0.3	766.3986	766	9.9
49	0.9142	0.3	768.5032	769	9.9
50	0.9728	0.3	770.6259	771	10.0
51	1.0318	0.3	772.7631	773	10.1
52	1.0913	0.3	774.9185	775	10.2
53	1.1513	0.3	777.0919	777	10.2
54	1.2118	0.3	779.2834	779	10.3
55	1.2728	0.3	781.4931	781	10.4
56	1.3344	0.3	783.7245	784	10.5
57	1.3966	0.3	785.9776	786	10.6
58	1.4595	0.3	788.2561	788	10.7
59	1.5230	0.3	790.5564	791	10.8
60	1.5873	0.3	792.8856	793	10.9
61	1.6525	0.3	795.2474	795	11.0
62	1.7185	0.3	797.6381	798	11.1
63	1.7855	0.3	800.0652	800	11.2
64	1.8536	0.3	802.5320	803	11.3
65	1.9228	0.3	805.0387	805	11.4
66	1.9933	0.3	807.5925	808	11.6
67	2.0652	0.3	810.1970	810	11.7
68	2.1386	0.3	812.8558	813	11.9
69	2.2137	0.3	815.5763	816	12.0
70	2.2906	0.3	818.3619	818	12.2
71	2.3695	0.3	821.2200	821	12.4
72	2.4505	0.3	824.1541	824	12.6
73	2.5339	0.4	827.1752	827	12.8
74	2.6199	0.4	830.2905	830	13.0
75	2.7088	0.4	833.5108	834	13.3
76	2.8007	0.4	836.8398	837	13.6
77	2.8960	0.4	840.2919	840	13.9
78	2.9952	0.4	843.8853	844	14.2
79	3.0985	0.4	847.6273	848	14.5
80	3.2064	0.4	851.5358	850	14.5
81	3.3197	0.4	855.6400	850	14.5
82	3.4389	0.4	859.9579	850	14.5
83	3.5651	0.5	864.5294	850	14.5
84	3.6992	0.5	869.3870	850	14.5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
85	3.8427	0.5	874.5852	850	14.5
86	3.9975	0.5	880.1926	850	14.5
87	4.1660	0.5	886.2964	850	14.5
88	4.3514	0.6	893.0123	850	14.5
89	4.5582	0.6	900.5034	850	14.5
90	4.7929	0.7	909.0052	850	14.5
91	5.0649	0.7	918.8581	850	14.5
92	5.3888	0.8	930.5911	850	14.5
93	5.7881	1	945.0553	850	14.5
94	6.3022	1.1	963.6781	850	14.5
95	6.9979	1.4	988.8791	850	14.5
96	7.9842	1.8	1024.6069	850	14.5
97	9.4219	2.4	1076.6861	850	14.5
98	10.000	2.7	1097.6272	850	14.5
99	10.000	2.7	1097.6272	850	14.5
100	10.000	2.7	1097.6272	850	14.5

Table A.12.28 Conversion Table for Performance Level Setting Form: ELA/L Grade 4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	1.9	556.1541	650	15
1	-5.5863	1.6	568.8904	650	15
2	-4.4430	1.0	604.0884	650	15
3	-3.8331	0.8	622.8649	650	15
4	-3.4200	0.7	635.5828	650	15
5	-3.1085	0.6	645.1727	650	15
6	-2.8589	0.5	652.8569	653	15
7	-2.6509	0.5	659.2605	659	15
8	-2.4726	0.5	664.7497	665	14.1
9	-2.3165	0.4	669.5554	670	13.2
10	-2.1776	0.4	673.8317	674	12.5
11	-2.0523	0.4	677.6892	678	11.9
12	-1.9381	0.4	681.2050	681	11.4
13	-1.8330	0.4	684.4406	684	10.9
14	-1.7355	0.3	687.4423	687	10.5
15	-1.6443	0.3	690.2500	690	10.2
16	-1.5587	0.3	692.8853	693	9.9
17	-1.4778	0.3	695.3759	695	9.6
18	-1.4009	0.3	697.7434	698	9.4
19	-1.3276	0.3	700.0000	700	9.2
20	-1.2575	0.3	702.1581	702	9.0
21	-1.1901	0.3	704.2331	704	8.8
22	-1.1251	0.3	706.2342	706	8.7
23	-1.0622	0.3	708.1707	708	8.5
24	-1.0013	0.3	710.0456	710	8.4
25	-0.9422	0.3	711.8650	712	8.3
26	-0.8845	0.3	713.6414	714	8.2
27	-0.8282	0.3	715.3747	715	8.1
28	-0.7732	0.3	717.0679	717	8.0
29	-0.7192	0.3	718.7304	719	8.0
30	-0.6662	0.3	720.3621	720	7.9
31	-0.6142	0.3	721.9630	722	7.8
32	-0.5628	0.3	723.5454	724	7.8
33	-0.5122	0.3	725.1032	725	7.7
34	-0.4623	0.2	726.6394	727	7.7
35	-0.4128	0.2	728.1633	728	7.7
36	-0.3639	0.2	729.6688	730	7.6
37	-0.3154	0.2	731.1619	731	7.6
38	-0.2673	0.2	732.6427	733	7.6
39	-0.2194	0.2	734.1174	734	7.6
40	-0.1719	0.2	735.5797	736	7.5
41	-0.1246	0.2	737.0359	737	7.5
42	-0.0775	0.2	738.4860	738	7.5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	-0.0305	0.2	739.9329	740	7.5
44	0.0163	0.2	741.3737	741	7.5
45	0.0631	0.2	742.8145	743	7.5
46	0.1098	0.2	744.2522	744	7.5
47	0.1564	0.2	745.6869	746	7.5
48	0.2031	0.2	747.1246	747	7.5
49	0.2498	0.2	748.5623	749	7.6
50	0.2965	0.2	750.0000	750	7.6
51	0.3433	0.2	751.4408	751	7.6
52	0.3902	0.2	752.8847	753	7.6
53	0.4372	0.2	754.3317	754	7.6
54	0.4844	0.2	755.7848	756	7.6
55	0.5317	0.2	757.2410	757	7.7
56	0.5791	0.2	758.7002	759	7.7
57	0.6268	0.3	760.1688	760	7.7
58	0.6746	0.3	761.6403	762	7.7
59	0.7227	0.3	763.1212	763	7.7
60	0.7710	0.3	764.6081	765	7.8
61	0.8196	0.3	766.1044	766	7.8
62	0.8685	0.3	767.6098	768	7.8
63	0.9177	0.3	769.1245	769	7.9
64	0.9673	0.3	770.6515	771	7.9
65	1.0172	0.3	772.1877	772	7.9
66	1.0675	0.3	773.7363	774	8.0
67	1.1182	0.3	775.2971	775	8.0
68	1.1694	0.3	776.8734	777	8.1
69	1.2211	0.3	778.4651	778	8.1
70	1.2733	0.3	780.0721	780	8.1
71	1.3262	0.3	781.7007	782	8.2
72	1.3796	0.3	783.3447	783	8.3
73	1.4338	0.3	785.0133	785	8.3
74	1.4887	0.3	786.7035	787	8.4
75	1.5445	0.3	788.4213	788	8.4
76	1.6011	0.3	790.1638	790	8.5
77	1.6588	0.3	791.9402	792	8.6
78	1.7175	0.3	793.7474	794	8.7
79	1.7775	0.3	795.5945	796	8.8
80	1.8388	0.3	797.4817	797	8.9
81	1.9016	0.3	799.4151	799	9.0
82	1.9661	0.3	801.4008	801	9.1
83	2.0324	0.3	803.4420	803	9.3
84	2.1008	0.3	805.5478	806	9.4
85	2.1715	0.3	807.7244	808	9.6
86	2.2448	0.3	809.9810	810	9.8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	2.3212	0.3	812.3331	812	10.0
88	2.4009	0.3	814.7867	815	10.3
89	2.4846	0.3	817.3635	817	10.6
90	2.5728	0.4	820.0789	820	10.9
91	2.6663	0.4	822.9574	823	11.2
92	2.7659	0.4	826.0237	826	11.6
93	2.8730	0.4	829.3209	829	12.1
94	2.9889	0.4	832.8891	833	12.7
95	3.1157	0.4	836.7928	837	13.3
96	3.2558	0.5	841.1059	841	14.1
97	3.4127	0.5	845.9363	846	15
98	3.5915	0.5	851.4409	850	15
99	3.7993	0.6	857.8383	850	15
100	4.0472	0.6	865.4702	850	15
101	4.3532	0.7	874.8908	850	15
102	4.7494	0.9	887.0884	850	15
103	5.3000	1.1	904.0393	850	15
104	6.1618	1.5	930.5709	850	15
105	7.8929	2.5	983.8651	850	15
106	10.000	4.2	1048.7349	850	15

Table A.12.29 Conversion Table for Performance Level Setting Form: ELA/L Grade 5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	2.2	563.7589	650	15
1	-5.7665	2.0	570.6399	650	15
2	-4.4023	1.3	610.8415	650	15
3	-3.6956	0.9	631.6673	650	15
4	-3.2311	0.8	645.3556	650	15
5	-2.8900	0.7	655.4075	655	15
6	-2.6225	0.6	663.2904	663	15
7	-2.4033	0.5	669.7501	670	15
8	-2.2179	0.5	675.2136	675	14.3
9	-2.0573	0.5	679.9463	680	13.3
10	-1.9155	0.4	684.1250	684	12.5
11	-1.7883	0.4	687.8735	688	11.8
12	-1.6728	0.4	691.2772	691	11.2
13	-1.5667	0.4	694.4038	694	10.7
14	-1.4685	0.4	697.2977	697	10.3
15	-1.3768	0.3	700.0000	700	10.0
16	-1.2906	0.3	702.5402	703	9.7
17	-1.2091	0.3	704.9419	705	9.4
18	-1.1316	0.3	707.2258	707	9.2
19	-1.0576	0.3	709.4065	709	9.0
20	-0.9867	0.3	711.4958	711	8.8
21	-0.9185	0.3	713.5056	714	8.6
22	-0.8526	0.3	715.4476	715	8.5
23	-0.7888	0.3	717.3278	717	8.3
24	-0.7268	0.3	719.1548	719	8.2
25	-0.6665	0.3	720.9318	721	8.1
26	-0.6077	0.3	722.6646	723	8.0
27	-0.5502	0.3	724.3591	724	7.9
28	-0.4939	0.3	726.0182	726	7.9
29	-0.4386	0.3	727.6478	728	7.8
30	-0.3842	0.3	729.2509	729	7.7
31	-0.3307	0.3	730.8275	731	7.7
32	-0.2780	0.3	732.3805	732	7.6
33	-0.2259	0.3	733.9159	734	7.6
34	-0.1744	0.3	735.4335	735	7.6
35	-0.1235	0.3	736.9335	737	7.5
36	-0.0730	0.3	738.4217	738	7.5
37	-0.0229	0.3	739.8981	740	7.5
38	0.0268	0.3	741.3627	741	7.5
39	0.0762	0.3	742.8184	743	7.4
40	0.1253	0.3	744.2654	744	7.4
41	0.1742	0.3	745.7064	746	7.4
42	0.2229	0.3	747.1415	747	7.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.2715	0.3	748.5737	749	7.4
44	0.3199	0.3	750.0000	750	7.4
45	0.3683	0.3	751.4263	751	7.4
46	0.4165	0.3	752.8467	753	7.4
47	0.4648	0.3	754.2701	754	7.4
48	0.5129	0.3	755.6876	756	7.4
49	0.5611	0.3	757.1080	757	7.4
50	0.6093	0.3	758.5284	759	7.4
51	0.6575	0.3	759.9488	760	7.4
52	0.7058	0.3	761.3721	761	7.4
53	0.7541	0.3	762.7955	763	7.4
54	0.8024	0.3	764.2188	764	7.4
55	0.8509	0.3	765.6481	766	7.4
56	0.8995	0.3	767.0803	767	7.4
57	0.9482	0.3	768.5154	769	7.4
58	0.9970	0.3	769.9535	770	7.4
59	1.0459	0.3	771.3945	771	7.5
60	1.0950	0.3	772.8415	773	7.5
61	1.1443	0.3	774.2943	774	7.5
62	1.1939	0.3	775.7559	776	7.5
63	1.2436	0.3	777.2205	777	7.5
64	1.2936	0.3	778.6940	779	7.5
65	1.3439	0.3	780.1763	780	7.6
66	1.3945	0.3	781.6674	782	7.6
67	1.4455	0.3	783.1703	783	7.6
68	1.4968	0.3	784.6821	785	7.7
69	1.5486	0.3	786.2086	786	7.7
70	1.6009	0.3	787.7498	788	7.7
71	1.6537	0.3	789.3058	789	7.8
72	1.7071	0.3	790.8794	791	7.8
73	1.7611	0.3	792.4708	792	7.9
74	1.8159	0.3	794.0857	794	7.9
75	1.8715	0.3	795.7241	796	8.0
76	1.9279	0.3	797.3862	797	8.0
77	1.9854	0.3	799.0807	799	8.1
78	2.0439	0.3	800.8046	801	8.2
79	2.1036	0.3	802.5639	803	8.3
80	2.1647	0.3	804.3644	804	8.4
81	2.2273	0.3	806.2092	806	8.5
82	2.2915	0.3	808.1011	808	8.6
83	2.3576	0.3	810.0490	810	8.8
84	2.4259	0.3	812.0617	812	8.9
85	2.4966	0.3	814.1452	814	9.1
86	2.5701	0.3	816.3112	816	9.3

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	2.6468	0.3	818.5714	819	9.5
88	2.7272	0.3	820.9408	821	9.8
89	2.8118	0.3	823.4338	823	10.1
90	2.9016	0.4	826.0802	826	10.4
91	2.9974	0.4	828.9033	829	10.8
92	3.1004	0.4	831.9386	832	11.3
93	3.2123	0.4	835.2362	835	11.9
94	3.3350	0.4	838.8520	839	12.6
95	3.4714	0.5	842.8716	843	13.4
96	3.6252	0.5	847.4039	847	14.4
97	3.8017	0.5	852.6052	850	14.4
98	4.0085	0.6	858.6994	850	14.4
99	4.2567	0.7	866.0136	850	14.4
100	4.5635	0.8	875.0547	850	14.4
101	4.9556	0.9	886.6095	850	14.4
102	5.4767	1.1	901.9658	850	14.4
103	6.2052	1.4	923.4339	850	14.4
104	7.3100	1.9	955.9913	850	14.4
105	9.3372	3.0	1015.7308	850	14.4
106	10.000	3.5	1035.2629	850	14.4

Table A.12.30 Conversion Table for Performance Level Setting Form: ELA/L Grade 6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	1.3	566.8536	650	14.5
1	-6.0000	1.3	566.8536	650	14.5
2	-5.2704	1.1	587.8119	650	14.5
3	-4.5982	0.9	607.1213	650	14.5
4	-4.1239	0.8	620.7459	650	14.5
5	-3.7628	0.7	631.1187	650	14.5
6	-3.4739	0.6	639.4176	650	14.5
7	-3.2345	0.5	646.2945	650	14.5
8	-3.0306	0.5	652.1517	652	14.5
9	-2.8535	0.5	657.2390	657	13.5
10	-2.6970	0.4	661.7346	662	12.8
11	-2.5568	0.4	665.7619	666	12.1
12	-2.4298	0.4	669.4101	669	11.5
13	-2.3137	0.4	672.7451	673	11.0
14	-2.2067	0.4	675.8188	676	10.6
15	-2.1074	0.4	678.6713	679	10.2
16	-2.0146	0.3	681.3370	681	9.8
17	-1.9274	0.3	683.8419	684	9.5
18	-1.8452	0.3	686.2031	686	9.3
19	-1.7672	0.3	688.4437	688	9.0
20	-1.6929	0.3	690.5781	691	8.8
21	-1.6220	0.3	692.6147	693	8.6
22	-1.5541	0.3	694.5652	695	8.4
23	-1.4887	0.3	696.4439	696	8.3
24	-1.4258	0.3	698.2507	698	8.1
25	-1.3649	0.3	700.0001	700	8.0
26	-1.3060	0.3	701.6920	702	7.9
27	-1.2489	0.3	703.3323	703	7.8
28	-1.1933	0.3	704.9294	705	7.7
29	-1.1391	0.3	706.4864	706	7.6
30	-1.0862	0.3	708.0059	708	7.5
31	-1.0346	0.3	709.4882	709	7.4
32	-0.9840	0.3	710.9417	711	7.3
33	-0.9345	0.3	712.3636	712	7.2
34	-0.8858	0.2	713.7626	714	7.2
35	-0.8380	0.2	715.1357	715	7.1
36	-0.7910	0.2	716.4858	716	7.1
37	-0.7446	0.2	717.8186	718	7.0
38	-0.6990	0.2	719.1285	719	7.0
39	-0.6539	0.2	720.4241	720	6.9
40	-0.6094	0.2	721.7024	722	6.9
41	-0.5654	0.2	722.9663	723	6.9
42	-0.5218	0.2	724.2187	724	6.8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	-0.4787	0.2	725.4568	725	6.8
44	-0.4360	0.2	726.6834	727	6.8
45	-0.3936	0.2	727.9014	728	6.8
46	-0.3515	0.2	729.1107	729	6.7
47	-0.3098	0.2	730.3086	730	6.7
48	-0.2683	0.2	731.5007	732	6.7
49	-0.2270	0.2	732.6871	733	6.7
50	-0.1860	0.2	733.8648	734	6.7
51	-0.1452	0.2	735.0368	735	6.7
52	-0.1045	0.2	736.2060	736	6.7
53	-0.0640	0.2	737.3694	737	6.7
54	-0.0237	0.2	738.5270	739	6.6
55	0.0166	0.2	739.6846	740	6.6
56	0.0567	0.2	740.8365	741	6.6
57	0.0967	0.2	741.9856	742	6.6
58	0.1367	0.2	743.1346	743	6.6
59	0.1766	0.2	744.2808	744	6.6
60	0.2165	0.2	745.4269	745	6.6
61	0.2563	0.2	746.5702	747	6.6
62	0.2961	0.2	747.7135	748	6.6
63	0.3359	0.2	748.8568	749	6.7
64	0.3757	0.2	750.0000	750	6.7
65	0.4155	0.2	751.1433	751	6.7
66	0.4553	0.2	752.2866	752	6.7
67	0.4951	0.2	753.4299	753	6.7
68	0.5349	0.2	754.5732	755	6.7
69	0.5748	0.2	755.7193	756	6.7
70	0.6147	0.2	756.8655	757	6.7
71	0.6546	0.2	758.0116	758	6.7
72	0.6946	0.2	759.1607	759	6.7
73	0.7347	0.2	760.3126	760	6.7
74	0.7748	0.2	761.4645	761	6.8
75	0.8149	0.2	762.6164	763	6.8
76	0.8552	0.2	763.7740	764	6.8
77	0.8955	0.2	764.9317	765	6.8
78	0.9358	0.2	766.0893	766	6.8
79	0.9763	0.2	767.2527	767	6.8
80	1.0168	0.2	768.4161	768	6.8
81	1.0574	0.2	769.5824	770	6.8
82	1.0981	0.2	770.7515	771	6.9
83	1.1389	0.2	771.9235	772	6.9
84	1.1798	0.2	773.0984	773	6.9
85	1.2207	0.2	774.2733	774	6.9
86	1.2618	0.2	775.4539	775	6.9

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	1.3031	0.2	776.6403	777	6.9
88	1.3444	0.2	777.8266	778	7.0
89	1.3859	0.2	779.0187	779	7.0
90	1.4275	0.2	780.2137	780	7.0
91	1.4694	0.2	781.4173	781	7.0
92	1.5114	0.2	782.6238	783	7.1
93	1.5536	0.2	783.8360	784	7.1
94	1.5960	0.2	785.0540	785	7.1
95	1.6387	0.2	786.2806	786	7.1
96	1.6817	0.2	787.5158	788	7.2
97	1.7249	0.3	788.7568	789	7.2
98	1.7686	0.3	790.0121	790	7.3
99	1.8126	0.3	791.2760	791	7.3
100	1.8570	0.3	792.5514	793	7.3
101	1.9019	0.3	793.8412	794	7.4
102	1.9473	0.3	795.1454	795	7.5
103	1.9932	0.3	796.4639	796	7.5
104	2.0398	0.3	797.8025	798	7.6
105	2.0872	0.3	799.1641	799	7.6
106	2.1353	0.3	800.5458	801	7.7
107	2.1843	0.3	801.9533	802	7.8
108	2.2342	0.3	803.3868	803	7.9
109	2.2853	0.3	804.8546	805	8.0
110	2.3376	0.3	806.3570	806	8.1
111	2.3912	0.3	807.8967	808	8.2
112	2.4464	0.3	809.4824	809	8.4
113	2.5032	0.3	811.1140	811	8.5
114	2.5620	0.3	812.8030	813	8.7
115	2.6229	0.3	814.5524	815	8.9
116	2.6863	0.3	816.3736	816	9.0
117	2.7525	0.3	818.2753	818	9.3
118	2.8218	0.3	820.2660	820	9.5
119	2.8947	0.3	822.3601	822	9.8
120	2.9718	0.4	824.5748	825	10.1
121	3.0537	0.4	826.9275	827	10.4
122	3.1412	0.4	829.4410	829	10.8
123	3.2353	0.4	832.1441	832	11.3
124	3.3373	0.4	835.0741	835	11.8
125	3.4485	0.4	838.2684	838	12.3
126	3.5711	0.5	841.7901	842	13.0
127	3.7077	0.5	845.7141	846	13.8
128	3.8617	0.5	850.1378	850	13.8
129	4.0381	0.6	855.2050	850	13.8
130	4.2439	0.6	861.1168	850	13.8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
131	4.4898	0.7	868.1804	850	13.8
132	4.7928	0.8	876.8843	850	13.8
133	5.1826	0.9	888.0816	850	13.8
134	5.7162	1.1	903.4096	850	13.8
135	6.5236	1.4	926.6028	850	13.8
136	8.0053	2.1	969.1656	850	13.8
137	10.000	3.4	1026.4648	850	13.8

Table A.12.31 Conversion Table for Performance Level Setting Form: ELA/L Grade 7

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	1.9	536.5141	650	15
1	-6.0000	1.9	536.5141	650	15
2	-4.7838	1.2	577.7244	650	15
3	-4.0880	0.9	601.3013	650	15
4	-3.6310	0.8	616.7865	650	15
5	-3.2948	0.7	628.1784	650	15
6	-3.0302	0.6	637.1443	650	15
7	-2.8128	0.5	644.5108	650	15
8	-2.6284	0.5	650.7591	651	15
9	-2.4685	0.5	656.1772	656	15
10	-2.3272	0.4	660.9651	661	14.4
11	-2.2007	0.4	665.2515	665	13.7
12	-2.0860	0.4	669.1380	669	13.0
13	-1.9810	0.4	672.6959	673	12.4
14	-1.8842	0.4	675.9759	676	11.9
15	-1.7942	0.3	679.0255	679	11.5
16	-1.7100	0.3	681.8786	682	11.1
17	-1.6310	0.3	684.5555	685	10.8
18	-1.5563	0.3	687.0867	687	10.5
19	-1.4854	0.3	689.4891	689	10.2
20	-1.4179	0.3	691.7763	692	9.9
21	-1.3535	0.3	693.9584	694	9.7
22	-1.2917	0.3	696.0525	696	9.5
23	-1.2324	0.3	698.0618	698	9.3
24	-1.1752	0.3	700.0000	700	9.1
25	-1.1199	0.3	701.8738	702	9.0
26	-1.0665	0.3	703.6833	704	8.8
27	-1.0146	0.3	705.4419	705	8.7
28	-0.9642	0.3	707.1497	707	8.6
29	-0.9151	0.2	708.8134	709	8.5
30	-0.8673	0.2	710.4331	710	8.4
31	-0.8205	0.2	712.0189	712	8.3
32	-0.7748	0.2	713.5674	714	8.2
33	-0.7300	0.2	715.0854	715	8.1
34	-0.6861	0.2	716.5729	717	8.0
35	-0.6430	0.2	718.0334	718	7.9
36	-0.6007	0.2	719.4667	719	7.9
37	-0.5590	0.2	720.8797	721	7.8
38	-0.5179	0.2	722.2723	722	7.8
39	-0.4774	0.2	723.6446	724	7.7
40	-0.4375	0.2	724.9966	725	7.7
41	-0.3980	0.2	726.3351	726	7.6
42	-0.3590	0.2	727.6566	728	7.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	-0.3203	0.2	728.9679	729	7.5
44	-0.2821	0.2	730.2623	730	7.5
45	-0.2442	0.2	731.5465	732	7.5
46	-0.2066	0.2	732.8206	733	7.5
47	-0.1693	0.2	734.0845	734	7.4
48	-0.1322	0.2	735.3416	735	7.4
49	-0.0954	0.2	736.5885	737	7.4
50	-0.0588	0.2	737.8287	738	7.4
51	-0.0223	0.2	739.0655	739	7.4
52	0.0139	0.2	740.2921	740	7.4
53	0.0501	0.2	741.5187	742	7.4
54	0.0861	0.2	742.7386	743	7.4
55	0.1220	0.2	743.9550	744	7.4
56	0.1578	0.2	745.1681	745	7.4
57	0.1935	0.2	746.3778	746	7.4
58	0.2292	0.2	747.5874	748	7.4
59	0.2648	0.2	748.7937	749	7.4
60	0.3004	0.2	750.0000	750	7.4
61	0.3360	0.2	751.2063	751	7.4
62	0.3716	0.2	752.4126	752	7.4
63	0.4072	0.2	753.6189	754	7.4
64	0.4428	0.2	754.8252	755	7.4
65	0.4785	0.2	756.0348	756	7.4
66	0.5142	0.2	757.2445	757	7.4
67	0.5500	0.2	758.4576	758	7.5
68	0.5858	0.2	759.6706	760	7.5
69	0.6218	0.2	760.8905	761	7.5
70	0.6578	0.2	762.1103	762	7.5
71	0.6940	0.2	763.3369	763	7.5
72	0.7302	0.2	764.5636	765	7.6
73	0.7666	0.2	765.7970	766	7.6
74	0.8032	0.2	767.0371	767	7.6
75	0.8399	0.2	768.2807	768	7.6
76	0.8768	0.2	769.5310	770	7.7
77	0.9139	0.2	770.7881	771	7.7
78	0.9511	0.2	772.0486	772	7.7
79	0.9886	0.2	773.3193	773	7.7
80	1.0263	0.2	774.5968	775	7.8
81	1.0642	0.2	775.8810	776	7.8
82	1.1023	0.2	777.1720	777	7.8
83	1.1407	0.2	778.4731	778	7.9
84	1.1794	0.2	779.7845	780	7.9
85	1.2184	0.2	781.1060	781	8.0
86	1.2577	0.2	782.4376	782	8.0

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	1.2973	0.2	783.7795	784	8.0
88	1.3373	0.2	785.1348	785	8.1
89	1.3776	0.2	786.5004	787	8.1
90	1.4183	0.2	787.8795	788	8.2
91	1.4594	0.2	789.2721	789	8.2
92	1.5009	0.2	790.6783	791	8.3
93	1.5429	0.2	792.1015	792	8.3
94	1.5854	0.2	793.5416	794	8.4
95	1.6283	0.2	794.9952	795	8.5
96	1.6718	0.3	796.4692	796	8.5
97	1.7159	0.3	797.9635	798	8.6
98	1.7606	0.3	799.4782	799	8.7
99	1.8059	0.3	801.0131	801	8.7
100	1.8520	0.3	802.5752	803	8.8
101	1.8988	0.3	804.1610	804	8.9
102	1.9464	0.3	805.7739	806	9.0
103	1.9949	0.3	807.4173	807	9.1
104	2.0444	0.3	809.0946	809	9.2
105	2.0950	0.3	810.8091	811	9.3
106	2.1467	0.3	812.5610	813	9.4
107	2.1997	0.3	814.3568	814	9.5
108	2.2540	0.3	816.1968	816	9.7
109	2.3100	0.3	818.0943	818	9.8
110	2.3676	0.3	820.0460	820	10.0
111	2.4272	0.3	822.0656	822	10.2
112	2.4889	0.3	824.1562	824	10.4
113	2.5530	0.3	826.3282	826	10.6
114	2.6199	0.3	828.5951	829	10.9
115	2.6898	0.3	830.9636	831	11.2
116	2.7634	0.3	833.4575	833	11.5
117	2.8410	0.3	836.0870	836	11.8
118	2.9233	0.4	838.8757	839	12.2
119	3.0112	0.4	841.8541	842	12.7
120	3.1056	0.4	845.0528	845	13.2
121	3.2076	0.4	848.5090	849	13.8
122	3.3189	0.4	852.2804	850	13.8
123	3.4413	0.5	856.4278	850	13.8
124	3.5774	0.5	861.0395	850	13.8
125	3.7307	0.5	866.2340	850	13.8
126	3.9058	0.6	872.1672	850	13.8
127	4.1093	0.6	879.0627	850	13.8
128	4.3505	0.7	887.2356	850	13.8
129	4.6440	0.8	897.1807	850	13.8
130	5.0126	0.9	909.6705	850	13.8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
131	5.4954	1.0	926.0300	850	13.8
132	6.1659	1.3	948.7495	850	13.8
133	7.1859	1.7	983.3117	850	13.8
134	9.0569	2.7	1046.7096	850	13.8
135	10.000	3.3	1078.6661	850	13.8

Table A.12.32 Conversion Table for Performance Level Setting Form: ELA/L Grade 8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	1.8	533.2679	650	15
1	-6.0000	1.8	533.2679	650	15
2	-5.6481	1.6	545.3482	650	15
3	-4.7448	1.2	576.3575	650	15
4	-4.1705	0.9	596.0726	650	15
5	-3.7567	0.8	610.2779	650	15
6	-3.4362	0.7	621.2803	650	15
7	-3.1760	0.6	630.2127	650	15
8	-2.9575	0.6	637.7136	650	15
9	-2.7696	0.5	644.1640	650	15
10	-2.6050	0.5	649.8145	650	15
11	-2.4585	0.5	654.8437	655	15
12	-2.3267	0.4	659.3682	659	14.9
13	-2.2068	0.4	663.4843	663	14.2
14	-2.0969	0.4	667.2570	667	13.6
15	-1.9954	0.4	670.7414	671	13.0
16	-1.9011	0.4	673.9786	674	12.5
17	-1.8129	0.4	677.0064	677	12.1
18	-1.7301	0.3	679.8489	680	11.7
19	-1.6520	0.3	682.5300	683	11.3
20	-1.5780	0.3	685.0703	685	11.0
21	-1.5076	0.3	687.4871	687	10.7
22	-1.4405	0.3	689.7905	690	10.4
23	-1.3763	0.3	691.9944	692	10.2
24	-1.3147	0.3	694.1091	694	10.0
25	-1.2555	0.3	696.1414	696	9.8
26	-1.1983	0.3	698.1050	698	9.6
27	-1.1431	0.3	699.9999	700	9.5
28	-1.0897	0.3	701.8331	702	9.3
29	-1.0378	0.3	703.6148	704	9.2
30	-0.9873	0.3	705.3484	705	9.0
31	-0.9382	0.3	707.0339	707	8.9
32	-0.8903	0.3	708.6783	709	8.8
33	-0.8435	0.3	710.2849	710	8.7
34	-0.7977	0.3	711.8571	712	8.6
35	-0.7529	0.2	713.3951	713	8.5
36	-0.7089	0.2	714.9055	715	8.4
37	-0.6658	0.2	716.3851	716	8.4
38	-0.6233	0.2	717.8441	718	8.3
39	-0.5816	0.2	719.2756	719	8.2
40	-0.5404	0.2	720.6900	721	8.2
41	-0.4999	0.2	722.0803	722	8.1
42	-0.4598	0.2	723.4569	723	8.1

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	-0.4203	0.2	724.8129	725	8.0
44	-0.3812	0.2	726.1551	726	8.0
45	-0.3425	0.2	727.4837	727	8.0
46	-0.3042	0.2	728.7984	729	7.9
47	-0.2663	0.2	730.0995	730	7.9
48	-0.2287	0.2	731.3903	731	7.9
49	-0.1914	0.2	732.6707	733	7.9
50	-0.1543	0.2	733.9444	734	7.8
51	-0.1175	0.2	735.2077	735	7.8
52	-0.0809	0.2	736.4641	736	7.8
53	-0.0445	0.2	737.7137	738	7.8
54	-0.0082	0.2	738.9598	739	7.8
55	0.0279	0.2	740.1991	740	7.8
56	0.0638	0.2	741.4315	741	7.8
57	0.0997	0.2	742.6639	743	7.8
58	0.1354	0.2	743.8894	744	7.8
59	0.1711	0.2	745.1150	745	7.8
60	0.2067	0.2	746.3371	746	7.8
61	0.2423	0.2	747.5592	748	7.8
62	0.2778	0.2	748.7779	749	7.8
63	0.3134	0.2	750.0000	750	7.8
64	0.3490	0.2	751.2221	751	7.8
65	0.3845	0.2	752.4408	752	7.8
66	0.4202	0.2	753.6663	754	7.8
67	0.4559	0.2	754.8918	755	7.8
68	0.4916	0.2	756.1174	756	7.8
69	0.5274	0.2	757.3464	757	7.9
70	0.5634	0.2	758.5822	759	7.9
71	0.5994	0.2	759.8180	760	7.9
72	0.6356	0.2	761.0607	761	7.9
73	0.6719	0.2	762.3069	762	8.0
74	0.7084	0.2	763.5599	764	8.0
75	0.7450	0.2	764.8163	765	8.0
76	0.7818	0.2	766.0796	766	8.0
77	0.8188	0.2	767.3498	767	8.1
78	0.8560	0.2	768.6268	769	8.1
79	0.8935	0.2	769.9142	770	8.1
80	0.9312	0.2	771.2084	771	8.2
81	0.9691	0.2	772.5094	773	8.2
82	1.0074	0.2	773.8242	774	8.2
83	1.0459	0.2	775.1459	775	8.3
84	1.0847	0.2	776.4779	776	8.3
85	1.1238	0.2	777.8201	778	8.4
86	1.1633	0.2	779.1761	779	8.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	1.2032	0.2	780.5458	781	8.5
88	1.2434	0.2	781.9259	782	8.5
89	1.2841	0.2	783.3230	783	8.6
90	1.3252	0.3	784.7340	785	8.6
91	1.3668	0.3	786.1620	786	8.7
92	1.4088	0.3	787.6039	788	8.7
93	1.4514	0.3	789.0663	789	8.8
94	1.4946	0.3	790.5493	791	8.9
95	1.5383	0.3	792.0494	792	9.0
96	1.5827	0.3	793.5737	794	9.0
97	1.6278	0.3	795.1219	795	9.1
98	1.6736	0.3	796.6941	797	9.2
99	1.7202	0.3	798.2939	798	9.3
100	1.7677	0.3	799.9245	800	9.4
101	1.8162	0.3	801.5894	802	9.5
102	1.8656	0.3	803.2853	803	9.6
103	1.9161	0.3	805.0189	805	9.7
104	1.9678	0.3	806.7937	807	9.8
105	2.0208	0.3	808.6131	809	10.0
106	2.0753	0.3	810.4841	810	10.1
107	2.1312	0.3	812.4031	812	10.3
108	2.1889	0.3	814.3838	814	10.5
109	2.2485	0.3	816.4298	816	10.6
110	2.3100	0.3	818.5411	819	10.8
111	2.3739	0.3	820.7347	821	11.1
112	2.4403	0.3	823.0141	823	11.3
113	2.5094	0.3	825.3862	825	11.6
114	2.5817	0.3	827.8682	828	11.8
115	2.6575	0.4	830.4704	830	12.1
116	2.7371	0.4	833.2029	833	12.5
117	2.8212	0.4	836.0900	836	12.8
118	2.9103	0.4	839.1487	839	13.2
119	3.0050	0.4	842.3996	842	13.7
120	3.1062	0.4	845.8737	846	14.2
121	3.2149	0.4	849.6053	850	14.2
122	3.3322	0.4	853.6321	850	14.2
123	3.4598	0.5	858.0124	850	14.2
124	3.5994	0.5	862.8047	850	14.2
125	3.7534	0.5	868.0914	850	14.2
126	3.9250	0.6	873.9822	850	14.2
127	4.1184	0.6	880.6214	850	14.2
128	4.3395	0.6	888.2116	850	14.2
129	4.5964	0.7	897.0307	850	14.2
130	4.9015	0.8	907.5044	850	14.2

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
131	5.2737	0.9	920.2816	850	14.2
132	5.7436	1.1	936.4128	850	14.2
133	6.3643	1.3	957.7207	850	14.2
134	7.2370	1.7	987.6795	850	14.2
135	8.5901	2.3	1034.1300	850	14.2
136	10.000	3.1	1082.5303	850	14.2
137	10.000	3.1	1082.5303	850	14.2

Table A.12.33 Conversion Table for Performance Level Setting Form: ELA/L Grade 9

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	1.9	531.3885	650	15
1	-5.4896	1.6	548.8728	650	15
2	-4.3236	1.1	588.8153	650	15
3	-3.6708	0.9	611.1776	650	15
4	-3.2296	0.7	626.2913	650	15
5	-2.9022	0.6	637.5067	650	15
6	-2.6441	0.6	646.3482	650	15
7	-2.4323	0.5	653.6036	654	15
8	-2.2529	0.5	659.7492	660	15
9	-2.0975	0.4	665.0725	665	15
10	-1.9605	0.4	669.7656	670	14.2
11	-1.8378	0.4	673.9688	674	13.4
12	-1.7266	0.4	677.7781	678	12.8
13	-1.6249	0.4	681.2619	681	12.2
14	-1.5310	0.3	684.4786	684	11.7
15	-1.4437	0.3	687.4691	687	11.3
16	-1.3620	0.3	690.2678	690	10.9
17	-1.2852	0.3	692.8987	693	10.6
18	-1.2125	0.3	695.3891	695	10.3
19	-1.1436	0.3	697.7493	698	10.0
20	-1.0779	0.3	700.0000	700	9.8
21	-1.0150	0.3	702.1547	702	9.5
22	-0.9547	0.3	704.2203	704	9.3
23	-0.8966	0.3	706.2106	706	9.2
24	-0.8406	0.3	708.1289	708	9.0
25	-0.7865	0.3	709.9822	710	8.8
26	-0.7340	0.3	711.7806	712	8.7
27	-0.6830	0.3	713.5277	714	8.6
28	-0.6334	0.2	715.2267	715	8.5
29	-0.5850	0.2	716.8847	717	8.4
30	-0.5378	0.2	718.5016	719	8.3
31	-0.4916	0.2	720.0843	720	8.2
32	-0.4464	0.2	721.6326	722	8.1
33	-0.4020	0.2	723.1536	723	8.0
34	-0.3585	0.2	724.6437	725	7.9
35	-0.3156	0.2	726.1133	726	7.9
36	-0.2734	0.2	727.5589	728	7.8
37	-0.2319	0.2	728.9805	729	7.8
38	-0.1909	0.2	730.3850	730	7.7
39	-0.1504	0.2	731.7724	732	7.7
40	-0.1104	0.2	733.1426	733	7.6
41	-0.0708	0.2	734.4992	734	7.6
42	-0.0316	0.2	735.8420	736	7.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.0072	0.2	737.1711	737	7.5
44	0.0457	0.2	738.4900	738	7.5
45	0.0839	0.2	739.7986	740	7.5
46	0.1218	0.2	741.0969	741	7.4
47	0.1595	0.2	742.3883	742	7.4
48	0.1970	0.2	743.6729	744	7.4
49	0.2342	0.2	744.9473	745	7.4
50	0.2713	0.2	746.2182	746	7.4
51	0.3083	0.2	747.4856	747	7.4
52	0.3451	0.2	748.7462	749	7.4
53	0.3817	0.2	750.0000	750	7.4
54	0.4183	0.2	751.2538	751	7.3
55	0.4548	0.2	752.5041	753	7.3
56	0.4912	0.2	753.7510	754	7.3
57	0.5276	0.2	754.9980	755	7.3
58	0.5638	0.2	756.2380	756	7.3
59	0.6001	0.2	757.4815	757	7.3
60	0.6363	0.2	758.7216	759	7.3
61	0.6725	0.2	759.9617	760	7.3
62	0.7087	0.2	761.2017	761	7.3
63	0.7449	0.2	762.4418	762	7.3
64	0.7810	0.2	763.6784	764	7.3
65	0.8172	0.2	764.9185	765	7.3
66	0.8534	0.2	766.1586	766	7.4
67	0.8896	0.2	767.3986	767	7.4
68	0.9259	0.2	768.6421	769	7.4
69	0.9622	0.2	769.8856	770	7.4
70	0.9985	0.2	771.1291	771	7.4
71	1.0348	0.2	772.3726	772	7.4
72	1.0712	0.2	773.6195	774	7.4
73	1.1077	0.2	774.8699	775	7.4
74	1.1442	0.2	776.1202	776	7.4
75	1.1808	0.2	777.3740	777	7.4
76	1.2174	0.2	778.6278	779	7.4
77	1.2541	0.2	779.8849	780	7.5
78	1.2909	0.2	781.1456	781	7.5
79	1.3278	0.2	782.4096	782	7.5
80	1.3648	0.2	783.6771	784	7.5
81	1.4019	0.2	784.9480	785	7.5
82	1.4390	0.2	786.2189	786	7.5
83	1.4764	0.2	787.5001	788	7.5
84	1.5138	0.2	788.7812	789	7.6
85	1.5514	0.2	790.0693	790	7.6
86	1.5891	0.2	791.3607	791	7.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	1.6271	0.2	792.6624	793	7.6
88	1.6652	0.2	793.9676	794	7.7
89	1.7035	0.2	795.2796	795	7.7
90	1.7420	0.2	796.5985	797	7.7
91	1.7808	0.2	797.9276	798	7.8
92	1.8199	0.2	799.2670	799	7.8
93	1.8593	0.2	800.6167	801	7.9
94	1.8990	0.2	801.9766	802	7.9
95	1.9390	0.2	803.3469	803	7.9
96	1.9795	0.2	804.7343	805	8.0
97	2.0204	0.2	806.1353	806	8.1
98	2.0617	0.2	807.5501	808	8.1
99	2.1036	0.2	808.9854	809	8.2
100	2.1461	0.2	810.4413	810	8.3
101	2.1892	0.2	811.9177	812	8.3
102	2.2330	0.2	813.4181	813	8.4
103	2.2775	0.2	814.9425	815	8.5
104	2.3229	0.3	816.4978	816	8.6
105	2.3693	0.3	818.0872	818	8.7
106	2.4166	0.3	819.7075	820	8.8
107	2.4651	0.3	821.3690	821	9.0
108	2.5149	0.3	823.0749	823	9.1
109	2.5661	0.3	824.8288	825	9.2
110	2.6188	0.3	826.6341	827	9.4
111	2.6733	0.3	828.5011	829	9.6
112	2.7297	0.3	830.4331	830	9.8
113	2.7883	0.3	832.4405	832	10.0
114	2.8494	0.3	834.5335	835	10.2
115	2.9132	0.3	836.7191	837	10.5
116	2.9803	0.3	839.0177	839	10.8
117	3.0509	0.3	841.4361	841	11.1
118	3.1258	0.3	844.0019	844	11.5
119	3.2054	0.3	846.7287	847	11.9
120	3.2907	0.4	849.6507	850	11.9
121	3.3826	0.4	852.7988	850	11.9
122	3.4824	0.4	856.2176	850	11.9
123	3.5917	0.4	859.9618	850	11.9
124	3.7125	0.4	864.0999	850	11.9
125	3.8476	0.5	868.7279	850	11.9
126	4.0008	0.5	873.9759	850	11.9
127	4.1772	0.5	880.0187	850	11.9
128	4.3844	0.6	887.1165	850	11.9
129	4.6337	0.7	895.6565	850	11.9
130	4.9429	0.8	906.2485	850	11.9

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
131	5.3421	0.9	919.9235	850	11.9
132	5.8878	1.1	938.6170	850	11.9
133	6.7060	1.4	966.6452	850	11.9
134	8.2067	2.2	1018.0532	850	11.9
135	10.000	3.5	1079.4845	850	11.9

Table A.12.34 Conversion Table for Performance Level Setting Form: ELA/L Grade 10

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	1.9	480.7338	650	15
1	-5.4329	1.6	504.8104	650	15
2	-4.3822	1.0	549.4185	650	15
3	-3.8069	0.8	573.8432	650	15
4	-3.4127	0.7	590.5792	650	15
5	-3.1134	0.6	603.2861	650	15
6	-2.8725	0.5	613.5137	650	15
7	-2.6709	0.5	622.0727	650	15
8	-2.4976	0.5	629.4303	650	15
9	-2.3455	0.4	635.8878	650	15
10	-2.2101	0.4	641.6363	650	15
11	-2.0879	0.4	646.8244	650	15
12	-1.9765	0.4	651.5539	652	15
13	-1.8742	0.4	655.8971	656	15
14	-1.7795	0.3	659.9177	660	14.4
15	-1.6913	0.3	663.6622	664	13.9
16	-1.6088	0.3	667.1648	667	13.5
17	-1.5312	0.3	670.4594	670	13.0
18	-1.4579	0.3	673.5714	674	12.7
19	-1.3884	0.3	676.5220	677	12.3
20	-1.3224	0.3	679.3241	679	12.0
21	-1.2594	0.3	681.9988	682	11.7
22	-1.1991	0.3	684.5589	685	11.5
23	-1.1413	0.3	687.0128	687	11.2
24	-1.0858	0.3	689.3691	689	11.0
25	-1.0323	0.3	691.6405	692	10.8
26	-0.9807	0.3	693.8312	694	10.6
27	-0.9307	0.2	695.9540	696	10.5
28	-0.8823	0.2	698.0088	698	10.3
29	-0.8354	0.2	700.0000	700	10.1
30	-0.7898	0.2	701.9360	702	10.0
31	-0.7454	0.2	703.8210	704	9.9
32	-0.7021	0.2	705.6593	706	9.7
33	-0.6598	0.2	707.4552	707	9.6
34	-0.6185	0.2	709.2086	709	9.5
35	-0.5780	0.2	710.9281	711	9.4
36	-0.5384	0.2	712.6093	713	9.3
37	-0.4996	0.2	714.2566	714	9.3
38	-0.4614	0.2	715.8784	716	9.2
39	-0.4239	0.2	717.4705	717	9.1
40	-0.3870	0.2	719.0371	719	9.0
41	-0.3506	0.2	720.5825	721	9.0
42	-0.3148	0.2	722.1024	722	8.9

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	-0.2794	0.2	723.6053	724	8.9
44	-0.2445	0.2	725.0870	725	8.8
45	-0.2099	0.2	726.5560	727	8.8
46	-0.1758	0.2	728.0037	728	8.7
47	-0.1420	0.2	729.4387	729	8.7
48	-0.1084	0.2	730.8652	731	8.7
49	-0.0752	0.2	732.2747	732	8.6
50	-0.0422	0.2	733.6758	734	8.6
51	-0.0095	0.2	735.0641	735	8.6
52	0.0231	0.2	736.4481	736	8.6
53	0.0554	0.2	737.8194	738	8.6
54	0.0877	0.2	739.1908	739	8.6
55	0.1197	0.2	740.5493	741	8.5
56	0.1517	0.2	741.9079	742	8.5
57	0.1836	0.2	743.2622	743	8.5
58	0.2154	0.2	744.6123	745	8.5
59	0.2471	0.2	745.9582	746	8.5
60	0.2789	0.2	747.3083	747	8.5
61	0.3106	0.2	748.6541	749	8.6
62	0.3423	0.2	750.0000	750	8.6
63	0.3740	0.2	751.3458	751	8.6
64	0.4058	0.2	752.6959	753	8.6
65	0.4376	0.2	754.0460	754	8.6
66	0.4695	0.2	755.4003	755	8.6
67	0.5015	0.2	756.7589	757	8.6
68	0.5336	0.2	758.1217	758	8.7
69	0.5658	0.2	759.4888	759	8.7
70	0.5982	0.2	760.8643	761	8.7
71	0.6307	0.2	762.2441	762	8.8
72	0.6634	0.2	763.6324	764	8.8
73	0.6962	0.2	765.0250	765	8.8
74	0.7293	0.2	766.4303	766	8.9
75	0.7626	0.2	767.8440	768	8.9
76	0.7961	0.2	769.2663	769	8.9
77	0.8298	0.2	770.6971	771	9.0
78	0.8638	0.2	772.1405	772	9.0
79	0.8981	0.2	773.5968	774	9.1
80	0.9327	0.2	775.0657	775	9.1
81	0.9676	0.2	776.5474	777	9.2
82	1.0028	0.2	778.0419	778	9.2
83	1.0383	0.2	779.5490	780	9.3
84	1.0742	0.2	781.0732	781	9.3
85	1.1104	0.2	782.6101	783	9.4
86	1.1470	0.2	784.1640	784	9.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	1.1840	0.2	785.7348	786	9.5
88	1.2215	0.2	787.3269	787	9.6
89	1.2593	0.2	788.9317	789	9.6
90	1.2977	0.2	790.5620	791	9.7
91	1.3365	0.2	792.2093	792	9.8
92	1.3758	0.2	793.8778	794	9.8
93	1.4156	0.2	795.5675	796	9.9
94	1.4560	0.2	797.2828	797	10.0
95	1.4969	0.2	799.0192	799	10.1
96	1.5385	0.2	800.7853	801	10.2
97	1.5807	0.2	802.5770	803	10.2
98	1.6235	0.2	804.3941	804	10.3
99	1.6671	0.2	806.2451	806	10.4
100	1.7115	0.2	808.1302	808	10.5
101	1.7566	0.3	810.0449	810	10.7
102	1.8027	0.3	812.0021	812	10.8
103	1.8496	0.3	813.9933	814	10.9
104	1.8976	0.3	816.0311	816	11.0
105	1.9466	0.3	818.1115	818	11.2
106	1.9968	0.3	820.2427	820	11.3
107	2.0482	0.3	822.4250	822	11.5
108	2.1010	0.3	824.6666	825	11.6
109	2.1553	0.3	826.9720	827	11.8
110	2.2112	0.3	829.3452	829	12.0
111	2.2688	0.3	831.7907	832	12.2
112	2.3285	0.3	834.3253	834	12.5
113	2.3902	0.3	836.9448	837	12.7
114	2.4544	0.3	839.6704	840	13.0
115	2.5212	0.3	842.5065	843	13.3
116	2.5909	0.3	845.4656	845	13.6
117	2.6640	0.3	848.5691	849	13.9
118	2.7408	0.3	851.8297	850	13.9
119	2.8218	0.3	855.2686	850	13.9
120	2.9077	0.4	858.9155	850	13.9
121	2.9991	0.4	862.7960	850	13.9
122	3.0969	0.4	866.9481	850	13.9
123	3.2021	0.4	871.4145	850	13.9
124	3.3161	0.4	876.2544	850	13.9
125	3.4403	0.4	881.5274	850	13.9
126	3.5769	0.5	887.3268	850	13.9
127	3.7286	0.5	893.7674	850	13.9
128	3.8988	0.5	900.9933	850	13.9
129	4.0924	0.6	909.2127	850	13.9
130	4.3161	0.6	918.7100	850	13.9

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
131	4.5800	0.7	929.9140	850	13.9
132	4.9000	0.8	943.4998	850	13.9
133	5.3031	0.9	960.6137	850	13.9
134	5.8413	1.0	983.4633	850	13.9
135	6.6365	1.3	1017.2240	850	13.9
136	8.1029	2.1	1079.4809	850	13.9
137	10.000	3.6	1160.0234	850	13.9

Table A.12.35 Conversion Table for Performance Level Setting Form: ELA/L Grade 11

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-6.0000	2.4	520.0799	650	15
1	-4.8745	1.5	561.3130	650	15
2	-3.8597	1.0	598.4906	650	15
3	-3.3200	0.7	618.2628	650	15
4	-2.9561	0.6	631.5944	650	15
5	-2.6823	0.5	641.6252	650	15
6	-2.4628	0.5	649.6666	650	15
7	-2.2792	0.4	656.3929	656	15
8	-2.1212	0.4	662.1813	662	15
9	-1.9823	0.4	667.2699	667	14.2
10	-1.8580	0.4	671.8237	672	13.5
11	-1.7454	0.4	675.9489	676	12.9
12	-1.6424	0.3	679.7223	680	12.4
13	-1.5472	0.3	683.2100	683	11.9
14	-1.4588	0.3	686.4486	686	11.5
15	-1.3761	0.3	689.4783	689	11.2
16	-1.2983	0.3	692.3286	692	10.9
17	-1.2248	0.3	695.0213	695	10.6
18	-1.1552	0.3	697.5711	698	10.3
19	-1.0889	0.3	700.0000	700	10.1
20	-1.0257	0.3	702.3154	702	9.9
21	-0.9652	0.3	704.5318	705	9.7
22	-0.9071	0.3	706.6603	707	9.5
23	-0.8512	0.3	708.7082	709	9.3
24	-0.7974	0.3	710.6792	711	9.2
25	-0.7454	0.2	712.5843	713	9.1
26	-0.6951	0.2	714.4270	714	8.9
27	-0.6463	0.2	716.2148	716	8.8
28	-0.5989	0.2	717.9514	718	8.7
29	-0.5528	0.2	719.6403	720	8.6
30	-0.5080	0.2	721.2815	721	8.5
31	-0.4642	0.2	722.8861	723	8.4
32	-0.4214	0.2	724.4541	724	8.3
33	-0.3795	0.2	725.9892	726	8.2
34	-0.3385	0.2	727.4912	727	8.1
35	-0.2983	0.2	728.9640	729	8.1
36	-0.2588	0.2	730.4111	730	8.0
37	-0.2201	0.2	731.8288	732	8.0
38	-0.1819	0.2	733.2283	733	7.9
39	-0.1443	0.2	734.6058	735	7.9
40	-0.1072	0.2	735.9650	736	7.8
41	-0.0706	0.2	737.3058	737	7.8
42	-0.0345	0.2	738.6284	739	7.7

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.0012	0.2	739.9363	740	7.7
44	0.0365	0.2	741.2295	741	7.7
45	0.0715	0.2	742.5117	743	7.6
46	0.1062	0.2	743.7830	744	7.6
47	0.1406	0.2	745.0432	745	7.6
48	0.1747	0.2	746.2925	746	7.6
49	0.2087	0.2	747.5381	748	7.5
50	0.2424	0.2	748.7727	749	7.5
51	0.2759	0.2	750.0000	750	7.5
52	0.3092	0.2	751.2200	751	7.5
53	0.3424	0.2	752.4363	752	7.5
54	0.3755	0.2	753.6489	754	7.5
55	0.4085	0.2	754.8579	755	7.5
56	0.4414	0.2	756.0632	756	7.5
57	0.4742	0.2	757.2648	757	7.5
58	0.5070	0.2	758.4664	758	7.5
59	0.5397	0.2	759.6644	760	7.5
60	0.5724	0.2	760.8624	761	7.5
61	0.6051	0.2	762.0604	762	7.5
62	0.6378	0.2	763.2584	763	7.5
63	0.6705	0.2	764.4563	764	7.5
64	0.7032	0.2	765.6543	766	7.5
65	0.7360	0.2	766.8560	767	7.5
66	0.7688	0.2	768.0576	768	7.5
67	0.8017	0.2	769.2629	769	7.5
68	0.8346	0.2	770.4682	770	7.5
69	0.8676	0.2	771.6772	772	7.5
70	0.9007	0.2	772.8898	773	7.5
71	0.9338	0.2	774.1024	774	7.6
72	0.9671	0.2	775.3224	775	7.6
73	1.0004	0.2	776.5424	777	7.6
74	1.0339	0.2	777.7696	778	7.6
75	1.0675	0.2	779.0006	779	7.6
76	1.1012	0.2	780.2352	780	7.6
77	1.1350	0.2	781.4735	781	7.6
78	1.1690	0.2	782.7191	783	7.7
79	1.2031	0.2	783.9683	784	7.7
80	1.2374	0.2	785.2249	785	7.7
81	1.2718	0.2	786.4852	786	7.7
82	1.3064	0.2	787.7528	788	7.7
83	1.3411	0.2	789.0240	789	7.8
84	1.3761	0.2	790.3063	790	7.8
85	1.4112	0.2	791.5922	792	7.8
86	1.4465	0.2	792.8854	793	7.8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	1.4821	0.2	794.1896	794	7.9
88	1.5179	0.2	795.5012	796	7.9
89	1.5539	0.2	796.8200	797	7.9
90	1.5902	0.2	798.1499	798	7.9
91	1.6268	0.2	799.4908	799	8.0
92	1.6637	0.2	800.8426	801	8.0
93	1.7009	0.2	802.2055	802	8.0
94	1.7385	0.2	803.5829	804	8.1
95	1.7765	0.2	804.9751	805	8.1
96	1.8149	0.2	806.3819	806	8.2
97	1.8537	0.2	807.8033	808	8.2
98	1.8931	0.2	809.2468	809	8.3
99	1.9329	0.2	810.7049	811	8.3
100	1.9733	0.2	812.1849	812	8.4
101	2.0144	0.2	813.6906	814	8.5
102	2.0561	0.2	815.2183	815	8.5
103	2.0986	0.2	816.7754	817	8.6
104	2.1419	0.2	818.3617	818	8.7
105	2.1860	0.2	819.9773	820	8.8
106	2.2311	0.2	821.6295	822	8.9
107	2.2773	0.2	823.3221	823	9.0
108	2.3246	0.2	825.0550	825	9.1
109	2.3732	0.3	826.8354	827	9.3
110	2.4232	0.3	828.6672	829	9.4
111	2.4747	0.3	830.5539	831	9.5
112	2.5279	0.3	832.5029	833	9.7
113	2.5830	0.3	834.5215	835	9.9
114	2.6401	0.3	836.6134	837	10.1
115	2.6995	0.3	838.7896	839	10.3
116	2.7615	0.3	841.0610	841	10.5
117	2.8263	0.3	843.4349	843	10.8
118	2.8944	0.3	845.9298	846	11.1
119	2.9660	0.3	848.5529	849	11.4
120	3.0417	0.3	851.3262	850	11.4
121	3.1221	0.3	854.2717	850	11.4
122	3.2078	0.3	857.4113	850	11.4
123	3.2996	0.4	860.7745	850	11.4
124	3.3985	0.4	864.3977	850	11.4
125	3.5058	0.4	868.3287	850	11.4
126	3.6231	0.4	872.6260	850	11.4
127	3.7523	0.4	877.3593	850	11.4
128	3.8961	0.5	882.6275	850	11.4
129	4.0583	0.5	888.5697	850	11.4
130	4.2439	0.5	895.3693	850	11.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
131	4.4606	0.6	903.3082	850	11.4
132	4.7207	0.6	912.8370	850	11.4
133	5.0448	0.7	924.7106	850	11.4
134	5.4736	0.9	940.4198	850	11.4
135	6.1022	1.1	963.4488	850	11.4
136	7.2558	1.7	1005.7114	850	11.4
137	10.000	4.3	1106.2463	850	11.4

Table A.12.36 Conversion Table for Performance Level Setting Form: Mathematics Grade 3

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	14.8	421.0105	650	13.4
1	-4.3334	0.9	601.4872	650	13.4
2	-3.6635	0.6	622.8230	650	13.4
3	-3.2563	0.5	635.7920	650	13.4
4	-2.9514	0.5	645.5028	650	13.4
5	-2.7023	0.4	653.4364	653	13.4
6	-2.4888	0.4	660.2362	660	12.4
7	-2.3007	0.4	666.2270	666	11.7
8	-2.1316	0.3	671.6127	672	11.0
9	-1.9774	0.3	676.5239	677	10.5
10	-1.8353	0.3	681.0497	681	10.0
11	-1.7028	0.3	685.2697	685	9.7
12	-1.5784	0.3	689.2317	689	9.3
13	-1.4606	0.3	692.9836	693	9.1
14	-1.3482	0.3	696.5634	697	8.8
15	-1.2403	0.3	699.9999	700	8.6
16	-1.1362	0.3	703.3154	703	8.5
17	-1.0351	0.3	706.5354	707	8.4
18	-0.9365	0.3	709.6757	710	8.3
19	-0.8401	0.3	712.7460	713	8.2
20	-0.7456	0.3	715.7557	716	8.1
21	-0.6527	0.3	718.7145	719	8.1
22	-0.5613	0.3	721.6255	722	8.0
23	-0.4714	0.2	724.4888	724	8.0
24	-0.3830	0.2	727.3043	727	7.9
25	-0.2962	0.2	730.0688	730	7.9
26	-0.2109	0.2	732.7855	733	7.8
27	-0.1275	0.2	735.4417	735	7.7
28	-0.0460	0.2	738.0374	738	7.6
29	0.0335	0.2	740.5694	741	7.6
30	0.1108	0.2	743.0314	743	7.5
31	0.1860	0.2	745.4265	745	7.4
32	0.2589	0.2	747.7483	748	7.3
33	0.3296	0.2	750.0000	750	7.2
34	0.3982	0.2	752.1849	752	7.1
35	0.4646	0.2	754.2996	754	7.0
36	0.5291	0.2	756.3539	756	6.9
37	0.5917	0.2	758.3477	758	6.8
38	0.6525	0.2	760.2841	760	6.7
39	0.7117	0.2	762.1696	762	6.6
40	0.7694	0.2	764.0073	764	6.5
41	0.8256	0.2	765.7972	766	6.4
42	0.8806	0.2	767.5489	768	6.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.9345	0.2	769.2656	769	6.3
44	0.9873	0.2	770.9472	771	6.3
45	1.0393	0.2	772.6034	773	6.2
46	1.0904	0.2	774.2309	774	6.2
47	1.1410	0.2	775.8424	776	6.2
48	1.1910	0.2	777.4349	777	6.1
49	1.2406	0.2	779.0146	779	6.1
50	1.2899	0.2	780.5848	781	6.1
51	1.3392	0.2	782.1549	782	6.1
52	1.3886	0.2	783.7283	784	6.1
53	1.4382	0.2	785.3080	785	6.2
54	1.4882	0.2	786.9005	787	6.2
55	1.5388	0.2	788.5120	789	6.3
56	1.5902	0.2	790.1491	790	6.3
57	1.6427	0.2	791.8212	792	6.4
58	1.6965	0.2	793.5347	794	6.5
59	1.7519	0.2	795.2991	795	6.6
60	1.8092	0.2	797.1241	797	6.7
61	1.8687	0.2	799.0191	799	6.9
62	1.9308	0.2	800.9969	801	7.1
63	1.9960	0.2	803.0735	803	7.3
64	2.0648	0.2	805.2647	805	7.5
65	2.1377	0.2	807.5865	808	7.7
66	2.2154	0.3	810.0612	810	8.0
67	2.2987	0.3	812.7143	813	8.3
68	2.3885	0.3	815.5743	816	8.7
69	2.4860	0.3	818.6796	819	9.1
70	2.5925	0.3	822.0716	822	9.6
71	2.7098	0.3	825.8075	826	10.1
72	2.8403	0.3	829.9638	830	10.8
73	2.9874	0.4	834.6488	835	11.6
74	3.1559	0.4	840.0154	840	12.6
75	3.3529	0.4	846.2897	846	13.9
76	3.5902	0.5	853.8475	850	13.9
77	3.8879	0.6	863.3290	850	13.9
78	4.2830	0.7	875.9126	850	13.9
79	4.8491	0.9	893.9425	850	13.9
80	5.7384	1.3	922.2659	850	13.9
81	7.3154	1.9	972.4921	850	13.9
82	15.000	12.8	1217.2405	850	13.9

Table A.12.37 Conversion Table for Performance Level Setting Form: Mathematics Grade 4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	17.1	442.1466	650	12.8
1	-4.0819	0.9	617.0726	650	12.8
2	-3.3997	0.6	637.2369	650	12.8
3	-3.0042	0.5	648.9271	650	12.8
4	-2.7224	0.4	657.2564	657	12.8
5	-2.5016	0.4	663.7828	664	11.4
6	-2.3182	0.4	669.2037	669	10.4
7	-2.1599	0.3	673.8827	674	9.7
8	-2.0194	0.3	678.0356	678	9.1
9	-1.8920	0.3	681.8012	682	8.7
10	-1.7746	0.3	685.2713	685	8.4
11	-1.6649	0.3	688.5138	689	8.1
12	-1.5613	0.3	691.5760	692	7.9
13	-1.4626	0.3	694.4934	694	7.7
14	-1.3678	0.3	697.2954	697	7.6
15	-1.2763	0.3	700.0000	700	7.5
16	-1.1874	0.3	702.6277	703	7.4
17	-1.1007	0.2	705.1903	705	7.3
18	-1.0158	0.2	707.6998	708	7.3
19	-0.9323	0.2	710.1679	710	7.2
20	-0.8501	0.2	712.5975	713	7.2
21	-0.7688	0.2	715.0006	715	7.2
22	-0.6885	0.2	717.3741	717	7.2
23	-0.6089	0.2	719.7269	720	7.2
24	-0.5301	0.2	722.0560	722	7.2
25	-0.4519	0.2	724.3674	724	7.2
26	-0.3745	0.2	726.6552	727	7.2
27	-0.2979	0.2	728.9193	729	7.1
28	-0.2221	0.2	731.1598	731	7.1
29	-0.1471	0.2	733.3766	733	7.1
30	-0.0731	0.2	735.5639	736	7.1
31	0.0000	0.2	737.7246	738	7.1
32	0.0720	0.2	739.8528	740	7.0
33	0.1429	0.2	741.9484	742	7.0
34	0.2127	0.2	744.0115	744	6.9
35	0.2813	0.2	746.0392	746	6.9
36	0.3489	0.2	748.0373	748	6.8
37	0.4153	0.2	750.0000	750	6.8
38	0.4806	0.2	751.9301	752	6.7
39	0.5448	0.2	753.8277	754	6.7
40	0.6080	0.2	755.6957	756	6.7
41	0.6703	0.2	757.5372	758	6.6
42	0.7316	0.2	759.3491	759	6.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.7920	0.2	761.1344	761	6.5
44	0.8517	0.2	762.8990	763	6.5
45	0.9107	0.2	764.6429	765	6.5
46	0.9690	0.2	766.3661	766	6.4
47	1.0269	0.2	768.0775	768	6.4
48	1.0843	0.2	769.7741	770	6.4
49	1.1414	0.2	771.4619	771	6.4
50	1.1982	0.2	773.1408	773	6.4
51	1.2550	0.2	774.8196	775	6.4
52	1.3117	0.2	776.4956	776	6.4
53	1.3686	0.2	778.1774	778	6.4
54	1.4257	0.2	779.8652	780	6.4
55	1.4833	0.2	781.5677	782	6.4
56	1.5413	0.2	783.2820	783	6.5
57	1.6002	0.2	785.0230	785	6.5
58	1.6598	0.2	786.7846	787	6.6
59	1.7206	0.2	788.5818	789	6.6
60	1.7827	0.2	790.4173	790	6.7
61	1.8463	0.2	792.2972	792	6.8
62	1.9116	0.2	794.2273	794	6.9
63	1.9791	0.2	796.2224	796	7.0
64	2.0490	0.2	798.2885	798	7.1
65	2.1217	0.2	800.4374	800	7.3
66	2.1976	0.3	802.6808	803	7.5
67	2.2773	0.3	805.0366	805	7.7
68	2.3613	0.3	807.5194	808	7.9
69	2.4504	0.3	810.1530	810	8.2
70	2.5452	0.3	812.9551	813	8.5
71	2.6466	0.3	815.9523	816	8.9
72	2.7556	0.3	819.1741	819	9.3
73	2.8733	0.3	822.6530	823	9.7
74	3.0011	0.3	826.4305	826	10.2
75	3.1411	0.4	830.5686	831	10.8
76	3.2964	0.4	835.1589	835	11.6
77	3.4724	0.4	840.3611	840	12.5
78	3.6792	0.5	846.4737	846	13.9
79	3.9381	0.5	854.1262	850	13.9
80	4.3004	0.7	864.8350	850	13.9
81	4.9439	1.0	883.8554	850	13.9
82	15.000	93.2	1181.0916	850	13.9

Table A.12.38 Conversion Table for Performance Level Setting Form: Mathematics Grade 5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	9.6	447.0871	650	16.5
1	-4.6046	1.3	603.4485	650	16.5
2	-3.6649	0.8	630.6815	650	16.5
3	-3.1278	0.7	646.2469	650	16.5
4	-2.7508	0.6	657.1725	657	16.5
5	-2.4589	0.5	665.6319	666	14.6
6	-2.2194	0.5	672.5728	673	13.2
7	-2.0152	0.4	678.4906	678	12.2
8	-1.8361	0.4	683.6810	684	11.4
9	-1.6760	0.4	688.3208	688	10.8
10	-1.5306	0.4	692.5345	693	10.3
11	-1.3970	0.3	696.4063	696	9.8
12	-1.2730	0.3	699.9999	700	9.4
13	-1.1571	0.3	703.3588	703	9.1
14	-1.0481	0.3	706.5176	707	8.8
15	-0.9450	0.3	709.5055	710	8.5
16	-0.8471	0.3	712.3427	712	8.3
17	-0.7538	0.3	715.0466	715	8.1
18	-0.6646	0.3	717.6317	718	7.9
19	-0.5789	0.3	720.1153	720	7.7
20	-0.4965	0.3	722.5033	723	7.6
21	-0.4170	0.3	724.8072	725	7.4
22	-0.3401	0.3	727.0358	727	7.3
23	-0.2655	0.2	729.1978	729	7.2
24	-0.1931	0.2	731.2960	731	7.1
25	-0.1225	0.2	733.3420	733	7.0
26	-0.0537	0.2	735.3358	735	6.9
27	0.0136	0.2	737.2862	737	6.8
28	0.0795	0.2	739.1960	739	6.7
29	0.1442	0.2	741.0711	741	6.7
30	0.2077	0.2	742.9113	743	6.6
31	0.2702	0.2	744.7226	745	6.6
32	0.3317	0.2	746.5049	747	6.5
33	0.3924	0.2	748.2640	748	6.5
34	0.4523	0.2	750.0000	750	6.4
35	0.5116	0.2	751.7185	752	6.4
36	0.5703	0.2	753.4197	753	6.4
37	0.6284	0.2	755.1034	755	6.4
38	0.6861	0.2	756.7756	757	6.3
39	0.7433	0.2	758.4333	758	6.3
40	0.8002	0.2	760.0823	760	6.3
41	0.8568	0.2	761.7226	762	6.3
42	0.9131	0.2	763.3542	763	6.2

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.9693	0.2	764.9829	765	6.2
44	1.0253	0.2	766.6058	767	6.2
45	1.0812	0.2	768.2258	768	6.2
46	1.1370	0.2	769.8429	770	6.2
47	1.1928	0.2	771.4600	771	6.2
48	1.2487	0.2	773.0801	773	6.2
49	1.3046	0.2	774.7001	775	6.2
50	1.3606	0.2	776.3230	776	6.2
51	1.4168	0.2	777.9517	778	6.2
52	1.4731	0.2	779.5833	780	6.2
53	1.5295	0.2	781.2178	781	6.2
54	1.5862	0.2	782.8610	783	6.2
55	1.6431	0.2	784.5100	785	6.2
56	1.7003	0.2	786.1676	786	6.2
57	1.7577	0.2	787.8311	788	6.2
58	1.8156	0.2	789.5091	790	6.2
59	1.8739	0.2	791.1987	791	6.2
60	1.9328	0.2	792.9056	793	6.2
61	1.9925	0.2	794.6357	795	6.3
62	2.0533	0.2	796.3978	796	6.4
63	2.1154	0.2	798.1974	798	6.4
64	2.1794	0.2	800.0522	800	6.6
65	2.2458	0.2	801.9765	802	6.7
66	2.3151	0.2	803.9849	804	6.9
67	2.3881	0.2	806.1004	806	7.1
68	2.4657	0.3	808.3493	808	7.4
69	2.5490	0.3	810.7634	811	7.7
70	2.6391	0.3	813.3745	813	8.1
71	2.7377	0.3	816.2320	816	8.6
72	2.8464	0.3	819.3822	819	9.1
73	2.9677	0.3	822.8975	823	9.7
74	3.1045	0.4	826.8621	827	10.5
75	3.2610	0.4	831.3975	831	11.4
76	3.4432	0.4	836.6778	837	12.5
77	3.6601	0.5	842.9636	843	13.9
78	3.9269	0.5	850.6956	850	13.9
79	4.2722	0.6	860.7026	850	13.9
80	4.7611	0.8	874.8712	850	13.9
81	5.6079	1.2	899.4118	850	13.9
82	15.000	38.9	1171.5996	850	13.9

Table A.12.39 Conversion Table for Performance Level Setting Form: Mathematics Grade 6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	15.2	449.8140	650	16.2
1	-3.9636	1.1	622.8064	650	16.2
2	-3.1760	0.7	645.3776	650	16.2
3	-2.7271	0.6	658.2422	658	16.2
4	-2.4090	0.5	667.3584	667	13.9
5	-2.1615	0.4	674.4513	674	12.3
6	-1.9583	0.4	680.2746	680	11.2
7	-1.7856	0.4	685.2239	685	10.3
8	-1.6350	0.3	689.5398	690	9.6
9	-1.5012	0.3	693.3743	693	9.1
10	-1.3804	0.3	696.8362	697	8.6
11	-1.2700	0.3	700.0001	700	8.2
12	-1.1678	0.3	702.9290	703	7.9
13	-1.0725	0.3	705.6601	706	7.6
14	-0.9829	0.3	708.2279	708	7.4
15	-0.8979	0.3	710.6638	711	7.2
16	-0.8170	0.2	712.9823	713	7.0
17	-0.7395	0.2	715.2033	715	6.8
18	-0.6650	0.2	717.3383	717	6.7
19	-0.5931	0.2	719.3988	719	6.6
20	-0.5233	0.2	721.3992	721	6.5
21	-0.4556	0.2	723.3393	723	6.4
22	-0.3896	0.2	725.2308	725	6.3
23	-0.3251	0.2	727.0792	727	6.2
24	-0.2620	0.2	728.8876	729	6.2
25	-0.2001	0.2	730.6615	731	6.1
26	-0.1394	0.2	732.4010	732	6.0
27	-0.0797	0.2	734.1119	734	6.0
28	-0.0209	0.2	735.7970	736	5.9
29	0.0371	0.2	737.4592	737	5.9
30	0.0942	0.2	739.0956	739	5.9
31	0.1506	0.2	740.7119	741	5.8
32	0.2062	0.2	742.3053	742	5.8
33	0.2612	0.2	743.8815	744	5.8
34	0.3155	0.2	745.4377	745	5.7
35	0.3692	0.2	746.9766	747	5.7
36	0.4222	0.2	748.4955	748	5.7
37	0.4747	0.2	750.0000	750	5.6
38	0.5267	0.2	751.4903	751	5.6
39	0.5781	0.2	752.9633	753	5.6
40	0.6290	0.2	754.4220	754	5.6
41	0.6794	0.2	755.8664	756	5.6
42	0.7295	0.2	757.3022	757	5.5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	0.7791	0.2	758.7236	759	5.5
44	0.8285	0.2	760.1393	760	5.5
45	0.8777	0.2	761.5493	762	5.5
46	0.9267	0.2	762.9536	763	5.5
47	0.9755	0.2	764.3521	764	5.5
48	1.0244	0.2	765.7535	766	5.5
49	1.0733	0.2	767.1548	767	5.5
50	1.1224	0.2	768.5620	769	5.6
51	1.1716	0.2	769.9719	770	5.6
52	1.2212	0.2	771.3934	771	5.6
53	1.2712	0.2	772.8263	773	5.6
54	1.3216	0.2	774.2707	774	5.7
55	1.3726	0.2	775.7322	776	5.7
56	1.4243	0.2	777.2139	777	5.7
57	1.4767	0.2	778.7156	779	5.8
58	1.5300	0.2	780.2430	780	5.8
59	1.5843	0.2	781.7992	782	5.9
60	1.6397	0.2	783.3869	783	5.9
61	1.6963	0.2	785.0089	785	6.0
62	1.7542	0.2	786.6682	787	6.1
63	1.8136	0.2	788.3705	788	6.1
64	1.8746	0.2	790.1187	790	6.2
65	1.9374	0.2	791.9184	792	6.3
66	2.0023	0.2	793.7783	794	6.4
67	2.0695	0.2	795.7041	796	6.5
68	2.1394	0.2	797.7074	798	6.6
69	2.2125	0.2	799.8023	800	6.7
70	2.2896	0.2	802.0118	802	6.9
71	2.3717	0.3	804.3647	804	7.2
72	2.4604	0.3	806.9066	807	7.5
73	2.5579	0.3	809.7008	810	7.9
74	2.6674	0.3	812.8389	813	8.5
75	2.7934	0.3	816.4498	816	9.2
76	2.9430	0.4	820.7371	821	10.2
77	3.1271	0.4	826.0131	826	11.6
78	3.3645	0.5	832.8165	833	13.6
79	3.6916	0.6	842.1906	842	16.7
80	4.1992	0.8	856.7375	850	16.7
81	5.2729	1.4	887.5078	850	16.7
82	15.000	15.8	1166.2690	850	16.7

Table A.12.40 Conversion Table for Performance Level Setting Form: Mathematics Grade 7

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	6.2	477.8425	650	20.0
1	-6.0057	2.9	580.4077	650	20.0
2	-3.8634	1.3	635.4175	650	20.0
3	-2.9716	0.8	658.3171	658	20.0
4	-2.4450	0.6	671.8390	672	16.2
5	-2.0742	0.5	681.3604	681	13.6
6	-1.7867	0.5	688.7428	689	11.9
7	-1.5504	0.4	694.8105	695	10.8
8	-1.3483	0.4	700.0000	700	9.9
9	-1.1706	0.4	704.5630	705	9.3
10	-1.0111	0.3	708.6586	709	8.7
11	-0.8654	0.3	712.3998	712	8.3
12	-0.7307	0.3	715.8587	716	8.0
13	-0.6047	0.3	719.0941	719	7.7
14	-0.4860	0.3	722.1420	722	7.5
15	-0.3733	0.3	725.0359	725	7.3
16	-0.2657	0.3	727.7989	728	7.1
17	-0.1628	0.3	730.4411	730	7.0
18	-0.0640	0.3	732.9781	733	6.8
19	0.0309	0.3	735.4149	735	6.7
20	0.1222	0.3	737.7593	738	6.5
21	0.2099	0.2	740.0113	740	6.4
22	0.2942	0.2	742.1759	742	6.3
23	0.3751	0.2	744.2533	744	6.1
24	0.4528	0.2	746.2485	746	6.0
25	0.5273	0.2	748.1615	748	5.9
26	0.5989	0.2	750.0000	750	5.8
27	0.6675	0.2	751.7615	752	5.6
28	0.7335	0.2	753.4562	753	5.5
29	0.7969	0.2	755.0842	755	5.4
30	0.8581	0.2	756.6557	757	5.3
31	0.9172	0.2	758.1733	758	5.2
32	0.9743	0.2	759.6395	760	5.1
33	1.0297	0.2	761.0620	761	5.0
34	1.0835	0.2	762.4435	762	5.0
35	1.1360	0.2	763.7916	764	4.9
36	1.1871	0.2	765.1037	765	4.8
37	1.2372	0.2	766.3902	766	4.8
38	1.2862	0.2	767.6484	768	4.7
39	1.3343	0.2	768.8835	769	4.7
40	1.3815	0.2	770.0955	770	4.6
41	1.4281	0.2	771.2921	771	4.6
42	1.4740	0.2	772.4707	772	4.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.5193	0.2	773.6339	774	4.5
44	1.5643	0.2	774.7894	775	4.5
45	1.6088	0.2	775.9321	776	4.5
46	1.6532	0.2	777.0722	777	4.5
47	1.6973	0.2	778.2046	778	4.5
48	1.7415	0.2	779.3396	779	4.5
49	1.7857	0.2	780.4745	780	4.5
50	1.8301	0.2	781.6146	782	4.5
51	1.8750	0.2	782.7676	783	4.5
52	1.9203	0.2	783.9308	784	4.6
53	1.9663	0.2	785.1120	785	4.6
54	2.0131	0.2	786.3137	786	4.7
55	2.0610	0.2	787.5437	788	4.7
56	2.1100	0.2	788.8019	789	4.8
57	2.1605	0.2	790.0986	790	4.9
58	2.2126	0.2	791.4364	791	5.0
59	2.2665	0.2	792.8205	793	5.1
60	2.3226	0.2	794.2610	794	5.2
61	2.3810	0.2	795.7606	796	5.4
62	2.4422	0.2	797.3321	797	5.5
63	2.5065	0.2	798.9832	799	5.7
64	2.5743	0.2	800.7241	801	5.9
65	2.6461	0.2	802.5678	803	6.1
66	2.7225	0.2	804.5296	805	6.3
67	2.8043	0.3	806.6300	807	6.6
68	2.8924	0.3	808.8923	809	6.9
69	2.9878	0.3	811.3419	811	7.3
70	3.0922	0.3	814.0227	814	7.7
71	3.2073	0.3	816.9782	817	8.2
72	3.3360	0.3	820.2830	820	8.7
73	3.4818	0.4	824.0268	824	9.4
74	3.6502	0.4	828.3510	828	10.3
75	3.8494	0.4	833.4660	833	11.4
76	4.0928	0.5	839.7160	840	13.0
77	4.4038	0.6	847.7018	848	15.2
78	4.8289	0.7	858.6175	850	15.2
79	5.4749	1.0	875.2054	850	15.2
80	6.6447	1.6	905.2434	850	15.2
81	9.4242	3.7	976.6152	850	15.2
82	15.000	8.1	1119.7900	850	15.2

Table A.12.41 Conversion Table for Performance Level Setting Form: Mathematics Grade 8

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	9.1	392.6327	650	18.3
1	-4.3997	1.5	580.5877	650	18.3
2	-3.3280	0.9	616.5557	650	18.3
3	-2.7351	0.7	636.4544	650	18.3
4	-2.3239	0.6	650.2549	650	18.3
5	-2.0068	0.5	660.8973	661	18.3
6	-1.7466	0.5	669.6300	670	16.6
7	-1.5240	0.5	677.1008	677	15.4
8	-1.3280	0.4	683.6789	684	14.5
9	-1.1517	0.4	689.5958	690	13.7
10	-0.9906	0.4	695.0026	695	13.1
11	-0.8417	0.4	699.9999	700	12.6
12	-0.7026	0.4	704.6683	705	12.2
13	-0.5720	0.4	709.0515	709	11.8
14	-0.4487	0.3	713.1896	713	11.5
15	-0.3319	0.3	717.1096	717	11.2
16	-0.2209	0.3	720.8349	721	10.9
17	-0.1152	0.3	724.3824	724	10.6
18	-0.0144	0.3	727.7654	728	10.3
19	0.0818	0.3	730.9940	731	10.1
20	0.1737	0.3	734.0783	734	9.8
21	0.2616	0.3	737.0284	737	9.5
22	0.3456	0.3	739.8476	740	9.3
23	0.4261	0.3	742.5493	743	9.1
24	0.5032	0.3	745.1369	745	8.8
25	0.5771	0.3	747.6171	748	8.6
26	0.6481	0.3	750.0000	750	8.4
27	0.7164	0.2	752.2922	752	8.2
28	0.7822	0.2	754.5006	755	8.0
29	0.8456	0.2	756.6284	757	7.9
30	0.9070	0.2	758.6891	759	7.7
31	0.9664	0.2	760.6826	761	7.6
32	1.0241	0.2	762.6191	763	7.4
33	1.0802	0.2	764.5019	765	7.3
34	1.1350	0.2	766.3411	766	7.2
35	1.1885	0.2	768.1367	768	7.1
36	1.2408	0.2	769.8919	770	7.0
37	1.2923	0.2	771.6204	772	6.9
38	1.3429	0.2	773.3186	773	6.8
39	1.3927	0.2	774.9899	775	6.7
40	1.4420	0.2	776.6445	777	6.7
41	1.4908	0.2	778.2823	778	6.6
42	1.5393	0.2	779.9101	780	6.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.5874	0.2	781.5244	782	6.5
44	1.6354	0.2	783.1353	783	6.5
45	1.6833	0.2	784.7429	785	6.4
46	1.7312	0.2	786.3505	786	6.4
47	1.7792	0.2	787.9615	788	6.4
48	1.8274	0.2	789.5792	790	6.4
49	1.8759	0.2	791.2069	791	6.4
50	1.9247	0.2	792.8447	793	6.4
51	1.9739	0.2	794.4959	794	6.4
52	2.0236	0.2	796.1640	796	6.4
53	2.0740	0.2	797.8555	798	6.4
54	2.1251	0.2	799.5705	800	6.4
55	2.1770	0.2	801.3123	801	6.5
56	2.2299	0.2	803.0877	803	6.5
57	2.2838	0.2	804.8967	805	6.6
58	2.3389	0.2	806.7459	807	6.6
59	2.3955	0.2	808.6455	809	6.7
60	2.4537	0.2	810.5988	811	6.8
61	2.5137	0.2	812.6125	813	6.9
62	2.5758	0.2	814.6967	815	7.1
63	2.6404	0.2	816.8647	817	7.2
64	2.7079	0.2	819.1302	819	7.4
65	2.7788	0.2	821.5097	822	7.6
66	2.8536	0.2	824.0201	824	7.9
67	2.9330	0.2	826.6849	827	8.2
68	3.0180	0.3	829.5376	830	8.6
69	3.1096	0.3	832.6119	833	9.0
70	3.2092	0.3	835.9546	836	9.5
71	3.3185	0.3	839.6229	840	10.1
72	3.4397	0.3	843.6905	844	10.9
73	3.5757	0.4	848.2549	848	11.8
74	3.7303	0.4	853.4435	850	11.8
75	3.9088	0.4	859.4343	850	11.8
76	4.1191	0.5	866.4923	850	11.8
77	4.3729	0.5	875.0102	850	11.8
78	4.6906	0.6	885.6727	850	11.8
79	5.1106	0.8	899.7686	850	11.8
80	5.7222	1.0	920.2949	850	11.8
81	6.8276	1.6	957.3939	850	11.8
82	15.000	17.2	1231.6727	850	11.8

Table A.12.42 Conversion Table for Performance Level Setting Form: Algebra I

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	7.4	430.2271	650	20.0
1	-6.1183	2.7	547.9470	650	20.0
2	-4.0393	1.4	610.9967	650	20.0
3	-3.1329	1.0	638.4850	650	20.0
4	-2.5610	0.8	655.8289	656	20.0
5	-2.1445	0.7	668.4601	668	20.0
6	-1.8162	0.6	678.4164	678	17.7
7	-1.5441	0.5	686.6683	687	16.1
8	-1.3103	0.5	693.7588	694	14.8
9	-1.1045	0.5	700.0000	700	13.8
10	-0.9197	0.4	705.6045	706	13.0
11	-0.7513	0.4	710.7115	711	12.3
12	-0.5962	0.4	715.4152	715	11.8
13	-0.4521	0.4	719.7853	720	11.3
14	-0.3170	0.4	723.8825	724	10.9
15	-0.1899	0.3	727.7370	728	10.5
16	-0.0696	0.3	731.3853	731	10.2
17	0.0447	0.3	734.8517	735	9.9
18	0.1536	0.3	738.1543	738	9.6
19	0.2576	0.3	741.3083	741	9.4
20	0.3572	0.3	744.3289	744	9.2
21	0.4526	0.3	747.2221	747	9.0
22	0.5442	0.3	750.0000	750	8.8
23	0.6322	0.3	752.6688	753	8.6
24	0.7167	0.3	755.2314	755	8.4
25	0.7981	0.3	757.7000	758	8.3
26	0.8763	0.3	760.0716	760	8.1
27	0.9517	0.3	762.3582	762	8.0
28	1.0244	0.3	764.5630	765	7.8
29	1.0946	0.3	766.6919	767	7.7
30	1.1624	0.2	768.7481	769	7.5
31	1.2280	0.2	770.7375	771	7.4
32	1.2916	0.2	772.6663	773	7.3
33	1.3533	0.2	774.5375	775	7.2
34	1.4132	0.2	776.3541	776	7.1
35	1.4717	0.2	778.1282	778	7.0
36	1.5286	0.2	779.8538	780	6.9
37	1.5843	0.2	781.5430	782	6.8
38	1.6388	0.2	783.1958	783	6.7
39	1.6922	0.2	784.8153	785	6.6
40	1.7446	0.2	786.4044	786	6.6
41	1.7962	0.2	787.9693	788	6.5
42	1.8469	0.2	789.5069	790	6.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.8970	0.2	791.0262	791	6.4
44	1.9464	0.2	792.5244	793	6.3
45	1.9952	0.2	794.0043	794	6.3
46	2.0435	0.2	795.4691	795	6.2
47	2.0913	0.2	796.9187	797	6.2
48	2.1388	0.2	798.3593	798	6.1
49	2.1859	0.2	799.7877	800	6.1
50	2.2326	0.2	801.2039	801	6.0
51	2.2792	0.2	802.6172	803	6.0
52	2.3255	0.2	804.0213	804	6.0
53	2.3716	0.2	805.4194	805	6.0
54	2.4176	0.2	806.8144	807	5.9
55	2.4635	0.2	808.2064	808	5.9
56	2.5093	0.2	809.5954	810	5.9
57	2.5552	0.2	810.9874	811	5.9
58	2.6011	0.2	812.3794	812	5.9
59	2.6470	0.2	813.7714	814	5.9
60	2.6931	0.2	815.1695	815	5.9
61	2.7393	0.2	816.5706	817	5.9
62	2.7857	0.2	817.9777	818	5.9
63	2.8323	0.2	819.3910	819	5.9
64	2.8791	0.2	820.8103	821	5.9
65	2.9263	0.2	822.2417	822	6.0
66	2.9737	0.2	823.6792	824	6.0
67	3.0216	0.2	825.1319	825	6.0
68	3.0698	0.2	826.5936	827	6.0
69	3.1184	0.2	828.0675	828	6.1
70	3.1676	0.2	829.5596	830	6.1
71	3.2174	0.2	831.0699	831	6.1
72	3.2679	0.2	832.6014	833	6.2
73	3.3193	0.2	834.1602	834	6.2
74	3.3717	0.2	835.7493	836	6.3
75	3.4254	0.2	837.3779	837	6.3
76	3.4807	0.2	839.0549	839	6.4
77	3.5380	0.2	840.7927	841	6.5
78	3.5978	0.2	842.6062	843	6.7
79	3.6608	0.2	844.5168	845	6.9
80	3.7277	0.2	846.5457	847	7.1
81	3.7993	0.2	848.7171	849	7.4
82	3.8769	0.3	851.0705	850	7.4
83	3.9617	0.3	853.6422	850	7.4
84	4.0553	0.3	856.4808	850	7.4
85	4.1595	0.3	859.6408	850	7.4
86	4.2767	0.3	863.1952	850	7.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	4.4094	0.4	867.2195	850	7.4
88	4.5613	0.4	871.8262	850	7.4
89	4.7370	0.4	877.1546	850	7.4
90	4.9436	0.5	883.4202	850	7.4
91	5.1915	0.5	890.9382	850	7.4
92	5.4985	0.6	900.2486	850	7.4
93	5.8955	0.7	912.2883	850	7.4
94	6.4424	0.9	928.8741	850	7.4
95	7.2748	1.2	954.1182	850	7.4
96	8.8244	2.0	1001.1128	850	7.4
97	15.000	9.1	1188.3996	850	7.4

Table A.12.43 Conversion Table for Performance Level Setting Form: Algebra II

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	9.4	363.0005	650	20.0
1	-4.3378	1.5	564.6745	650	20.0
2	-3.2325	1.0	604.0426	650	20.0
3	-2.6144	0.8	626.0578	650	20.0
4	-2.1834	0.7	641.4090	650	20.0
5	-1.8513	0.6	653.2376	653	20.0
6	-1.5804	0.5	662.8864	663	19.5
7	-1.3511	0.5	671.0536	671	18.0
8	-1.1519	0.5	678.1486	678	16.8
9	-0.9753	0.4	684.4387	684	15.8
10	-0.8164	0.4	690.0983	690	15.0
11	-0.6716	0.4	695.2557	695	14.3
12	-0.5384	0.4	700.0000	700	13.7
13	-0.4146	0.4	704.4094	704	13.2
14	-0.2988	0.4	708.5340	709	12.8
15	-0.1897	0.3	712.4198	712	12.4
16	-0.0865	0.3	716.0956	716	12.0
17	0.0116	0.3	719.5897	720	11.7
18	0.1053	0.3	722.9270	723	11.4
19	0.1949	0.3	726.1184	726	11.1
20	0.2809	0.3	729.1815	729	10.8
21	0.3635	0.3	732.1235	732	10.5
22	0.4431	0.3	734.9587	735	10.3
23	0.5198	0.3	737.6905	738	10.0
24	0.5937	0.3	740.3227	740	9.8
25	0.6651	0.3	742.8658	743	9.6
26	0.7341	0.3	745.3234	745	9.3
27	0.8008	0.3	747.6991	748	9.1
28	0.8654	0.3	750.0000	750	8.9
29	0.9279	0.2	752.2261	752	8.7
30	0.9886	0.2	754.3881	754	8.5
31	1.0476	0.2	756.4895	756	8.4
32	1.1049	0.2	758.5304	759	8.2
33	1.1607	0.2	760.5178	761	8.1
34	1.2152	0.2	762.4590	762	7.9
35	1.2685	0.2	764.3574	764	7.8
36	1.3206	0.2	766.2131	766	7.7
37	1.3718	0.2	768.0367	768	7.6
38	1.4220	0.2	769.8247	770	7.5
39	1.4714	0.2	771.5842	772	7.4
40	1.5200	0.2	773.3153	773	7.3
41	1.5679	0.2	775.0213	775	7.2
42	1.6151	0.2	776.7025	777	7.1

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.6618	0.2	778.3658	778	7.1
44	1.7079	0.2	780.0078	780	7.0
45	1.7535	0.2	781.6320	782	6.9
46	1.7986	0.2	783.2383	783	6.9
47	1.8432	0.2	784.8269	785	6.8
48	1.8874	0.2	786.4012	786	6.8
49	1.9312	0.2	787.9612	788	6.7
50	1.9746	0.2	789.5070	790	6.7
51	2.0177	0.2	791.0421	791	6.6
52	2.0605	0.2	792.5666	793	6.6
53	2.1030	0.2	794.0803	794	6.6
54	2.1453	0.2	795.5869	796	6.5
55	2.1874	0.2	797.0864	797	6.5
56	2.2294	0.2	798.5824	799	6.5
57	2.2714	0.2	800.0783	800	6.5
58	2.3133	0.2	801.5707	802	6.5
59	2.3553	0.2	803.0666	803	6.5
60	2.3974	0.2	804.5661	805	6.5
61	2.4397	0.2	806.0728	806	6.5
62	2.4823	0.2	807.5901	808	6.5
63	2.5253	0.2	809.1216	809	6.5
64	2.5687	0.2	810.6674	811	6.6
65	2.6127	0.2	812.2346	812	6.6
66	2.6574	0.2	813.8267	814	6.6
67	2.7029	0.2	815.4473	815	6.7
68	2.7493	0.2	817.1000	817	6.8
69	2.7967	0.2	818.7882	819	6.8
70	2.8453	0.2	820.5193	821	6.9
71	2.8952	0.2	822.2966	822	7.0
72	2.9465	0.2	824.1238	824	7.2
73	2.9995	0.2	826.0115	826	7.3
74	3.0543	0.2	827.9633	828	7.4
75	3.1111	0.2	829.9864	830	7.6
76	3.1700	0.2	832.0843	832	7.8
77	3.2312	0.2	834.2641	834	8.0
78	3.2949	0.2	836.5329	837	8.2
79	3.3613	0.2	838.8979	839	8.4
80	3.4307	0.2	841.3698	841	8.7
81	3.5033	0.3	843.9556	844	8.9
82	3.5795	0.3	846.6697	847	9.2
83	3.6596	0.3	849.5227	850	9.2
84	3.7442	0.3	852.5359	850	9.2
85	3.8340	0.3	855.7344	850	9.2
86	3.9299	0.3	859.1501	850	9.2

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	4.0328	0.3	862.8152	850	9.2
88	4.1442	0.3	866.7830	850	9.2
89	4.2657	0.4	871.1105	850	9.2
90	4.3992	0.4	875.8654	850	9.2
91	4.5473	0.4	881.1404	850	9.2
92	4.7132	0.4	887.0494	850	9.2
93	4.9013	0.5	893.7490	850	9.2
94	5.1184	0.5	901.4816	850	9.2
95	5.3748	0.6	910.6140	850	9.2
96	5.6881	0.7	921.7730	850	9.2
97	6.0892	0.8	936.0592	850	9.2
98	6.6384	1.0	955.6204	850	9.2
99	7.4770	1.3	985.4893	850	9.2
100	9.0828	2.1	1042.6840	850	9.2
101	15.000	8.1	1253.4405	850	9.2

Table A.12.44 Conversion Table for Performance Level Setting Form: Geometry

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	8.1	479.4306	650	20.0
1	-5.5752	2.1	590.9426	650	20.0
2	-3.9226	1.1	632.5908	650	20.0
3	-3.1489	0.9	652.0893	652	20.0
4	-2.6400	0.7	664.9144	665	17.7
5	-2.2570	0.6	674.5666	675	15.4
6	-1.9465	0.6	682.3917	682	13.9
7	-1.6830	0.5	689.0323	689	12.7
8	-1.4528	0.5	694.8337	695	11.9
9	-1.2478	0.4	700.0000	700	11.1
10	-1.0629	0.4	704.6598	705	10.5
11	-0.8946	0.4	708.9012	709	9.9
12	-0.7403	0.4	712.7899	713	9.4
13	-0.5978	0.4	716.3811	716	8.9
14	-0.4656	0.3	719.7127	720	8.5
15	-0.3422	0.3	722.8226	723	8.2
16	-0.2263	0.3	725.7435	726	7.9
17	-0.1171	0.3	728.4955	728	7.6
18	-0.0136	0.3	731.1039	731	7.4
19	0.0848	0.3	733.5837	734	7.2
20	0.1787	0.3	735.9501	736	7.0
21	0.2685	0.3	738.2132	738	6.8
22	0.3546	0.3	740.3831	740	6.7
23	0.4372	0.3	742.4647	742	6.5
24	0.5165	0.3	744.4632	744	6.4
25	0.5927	0.2	746.3836	746	6.3
26	0.6659	0.2	748.2283	748	6.1
27	0.7362	0.2	750.0000	750	6.0
28	0.8037	0.2	751.7011	752	5.9
29	0.8686	0.2	753.3367	753	5.8
30	0.9310	0.2	754.9093	755	5.6
31	0.9909	0.2	756.4189	756	5.5
32	1.0486	0.2	757.8730	758	5.4
33	1.1042	0.2	759.2742	759	5.3
34	1.1577	0.2	760.6225	761	5.2
35	1.2094	0.2	761.9254	762	5.1
36	1.2594	0.2	763.1855	763	5.0
37	1.3077	0.2	764.4027	764	4.9
38	1.3546	0.2	765.5847	766	4.8
39	1.4001	0.2	766.7314	767	4.8
40	1.4443	0.2	767.8453	768	4.7
41	1.4874	0.2	768.9315	769	4.6
42	1.5294	0.2	769.9899	770	4.5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.5705	0.2	771.0257	771	4.5
44	1.6107	0.2	772.0388	772	4.4
45	1.6500	0.2	773.0292	773	4.4
46	1.6886	0.2	774.0020	774	4.3
47	1.7265	0.2	774.9572	775	4.3
48	1.7638	0.2	775.8972	776	4.2
49	1.8005	0.2	776.8221	777	4.2
50	1.8367	0.2	777.7344	778	4.1
51	1.8724	0.2	778.6341	779	4.1
52	1.9078	0.2	779.5262	780	4.1
53	1.9427	0.2	780.4057	780	4.0
54	1.9774	0.2	781.2802	781	4.0
55	2.0117	0.2	782.1447	782	4.0
56	2.0459	0.2	783.0066	783	4.0
57	2.0798	0.2	783.8609	784	3.9
58	2.1136	0.2	784.7127	785	3.9
59	2.1474	0.2	785.5645	786	3.9
60	2.1811	0.2	786.4138	786	3.9
61	2.2148	0.2	787.2631	787	3.9
62	2.2486	0.2	788.1149	788	3.9
63	2.2825	0.2	788.9693	789	3.9
64	2.3166	0.2	789.8286	790	3.9
65	2.3510	0.2	790.6956	791	3.9
66	2.3857	0.2	791.5701	792	3.9
67	2.4208	0.2	792.4546	792	4.0
68	2.4564	0.2	793.3518	793	4.0
69	2.4925	0.2	794.2616	794	4.0
70	2.5294	0.2	795.1915	795	4.1
71	2.5671	0.2	796.1416	796	4.1
72	2.6057	0.2	797.1144	797	4.2
73	2.6455	0.2	798.1174	798	4.2
74	2.6865	0.2	799.1507	799	4.3
75	2.7291	0.2	800.2243	800	4.4
76	2.7734	0.2	801.3407	801	4.5
77	2.8197	0.2	802.5076	803	4.6
78	2.8683	0.2	803.7323	804	4.7
79	2.9197	0.2	805.0277	805	4.8
80	2.9742	0.2	806.4012	806	5.0
81	3.0325	0.2	807.8705	808	5.2
82	3.0950	0.2	809.4456	809	5.4
83	3.1627	0.2	811.1517	811	5.6
84	3.2364	0.2	813.0091	813	5.9
85	3.3173	0.2	815.0479	815	6.2
86	3.4069	0.3	817.3059	817	6.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	3.5070	0.3	819.8286	820	7.0
88	3.6201	0.3	822.6789	823	7.5
89	3.7498	0.3	825.9476	826	8.1
90	3.9014	0.4	829.7681	830	8.9
91	4.0832	0.4	834.3498	834	9.9
92	4.3105	0.5	840.0781	840	11.4
93	4.6138	0.5	847.7217	848	13.8
94	5.0666	0.7	859.1330	850	13.8
95	5.9232	1.2	880.7207	850	13.8
96	15.000	25.0	1109.4706	850	13.8

Table A.12.45 Conversion Table for Performance Level Setting Form: Integrated Mathematics I

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	6.2	397.1990	650	20.0
1	-6.4170	2.8	516.3782	650	20.0
2	-4.3152	1.5	586.2891	650	20.0
3	-3.3198	1.0	619.3985	650	20.0
4	-2.7035	0.8	639.8981	650	20.0
5	-2.2690	0.7	654.3506	654	20.0
6	-1.9356	0.6	665.4403	665	19.0
7	-1.6644	0.5	674.4611	674	17.1
8	-1.4345	0.5	682.1081	682	15.7
9	-1.2341	0.4	688.7739	689	14.7
10	-1.0563	0.4	694.6879	695	13.8
11	-0.8966	0.4	699.9999	700	13.1
12	-0.7522	0.4	704.8030	705	12.5
13	-0.6209	0.4	709.1704	709	11.9
14	-0.5010	0.3	713.1585	713	11.4
15	-0.3911	0.3	716.8141	717	10.9
16	-0.2898	0.3	720.1836	720	10.5
17	-0.1957	0.3	723.3135	723	10.1
18	-0.1078	0.3	726.2373	726	9.8
19	-0.0248	0.3	728.9981	729	9.5
20	0.0540	0.3	731.6192	732	9.3
21	0.1295	0.3	734.1305	734	9.1
22	0.2025	0.3	736.5586	737	8.9
23	0.2733	0.3	738.9136	739	8.7
24	0.3425	0.3	741.2154	741	8.5
25	0.4103	0.3	743.4706	743	8.4
26	0.4768	0.2	745.6825	746	8.3
27	0.5423	0.2	747.8612	748	8.1
28	0.6066	0.2	750.0000	750	8.0
29	0.6698	0.2	752.1022	752	7.9
30	0.7319	0.2	754.1678	754	7.7
31	0.7926	0.2	756.1868	756	7.6
32	0.8521	0.2	758.1659	758	7.5
33	0.9103	0.2	760.1018	760	7.3
34	0.9671	0.2	761.9911	762	7.2
35	1.0227	0.2	763.8405	764	7.1
36	1.0771	0.2	765.6499	766	7.0
37	1.1303	0.2	767.4195	767	6.9
38	1.1824	0.2	769.1525	769	6.8
39	1.2335	0.2	770.8522	771	6.7
40	1.2837	0.2	772.5219	773	6.6
41	1.3330	0.2	774.1618	774	6.6
42	1.3816	0.2	775.7783	776	6.5

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.4294	0.2	777.3683	777	6.4
44	1.4767	0.2	778.9416	779	6.4
45	1.5234	0.2	780.4949	780	6.3
46	1.5697	0.2	782.0350	782	6.3
47	1.6155	0.2	783.5584	784	6.3
48	1.6611	0.2	785.0752	785	6.2
49	1.7064	0.2	786.5820	787	6.2
50	1.7516	0.2	788.0854	788	6.2
51	1.7967	0.2	789.5856	790	6.2
52	1.8418	0.2	791.0857	791	6.2
53	1.8872	0.2	792.5958	793	6.2
54	1.9327	0.2	794.1092	794	6.2
55	1.9786	0.2	795.6360	796	6.3
56	2.0250	0.2	797.1794	797	6.3
57	2.0720	0.2	798.7427	799	6.4
58	2.1196	0.2	800.3260	800	6.4
59	2.1681	0.2	801.9392	802	6.5
60	2.2176	0.2	803.5857	804	6.5
61	2.2681	0.2	805.2654	805	6.6
62	2.3199	0.2	806.9884	807	6.7
63	2.3731	0.2	808.7580	809	6.8
64	2.4277	0.2	810.5741	811	6.9
65	2.4841	0.2	812.4501	812	7.0
66	2.5423	0.2	814.3860	814	7.1
67	2.6024	0.2	816.3851	816	7.3
68	2.6647	0.2	818.4573	818	7.4
69	2.7294	0.2	820.6094	821	7.6
70	2.7965	0.2	822.8413	823	7.7
71	2.8663	0.2	825.1630	825	7.9
72	2.9391	0.2	827.5845	828	8.1
73	3.0151	0.3	830.1125	830	8.3
74	3.0947	0.3	832.7601	833	8.6
75	3.1786	0.3	835.5509	836	8.8
76	3.2674	0.3	838.5046	839	9.2
77	3.3623	0.3	841.6612	842	9.6
78	3.4648	0.3	845.0706	845	10.1
79	3.5771	0.3	848.8059	849	10.7
80	3.7018	0.3	852.9538	850	10.7
81	3.8427	0.4	857.6404	850	10.7
82	4.0046	0.4	863.0256	850	10.7
83	4.1939	0.5	869.3222	850	10.7
84	4.4193	0.5	876.8195	850	10.7
85	4.6930	0.6	885.9234	850	10.7
86	5.0335	0.7	897.2493	850	10.7

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	5.4703	0.8	911.7783	850	10.7
88	6.0520	1.0	931.1270	850	10.7
89	6.8669	1.3	958.2326	850	10.7
90	8.1295	1.8	1000.2297	850	10.7
91	10.7743	3.5	1088.2021	850	10.7
92	15.000	7.4	1228.7590	850	10.7

Table A.12.46 Conversion Table for Performance Level Setting Form: Integrated Mathematics II

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	6.6	434.1360	650	20.0
1	-6.0191	2.3	550.8233	650	20.0
2	-4.2559	1.3	602.5059	650	20.0
3	-3.3887	1.0	627.9251	650	20.0
4	-2.8196	0.8	644.6064	650	20.0
5	-2.3961	0.7	657.0200	657	20.0
6	-2.0573	0.6	666.9508	667	18.1
7	-1.7731	0.6	675.2812	675	16.5
8	-1.5267	0.5	682.5037	683	15.3
9	-1.3080	0.5	688.9142	689	14.4
10	-1.1105	0.5	694.7032	695	13.6
11	-0.9298	0.4	699.9999	700	13.0
12	-0.7628	0.4	704.8950	705	12.4
13	-0.6075	0.4	709.4471	709	11.9
14	-0.4622	0.4	713.7061	714	11.5
15	-0.3258	0.4	717.7042	718	11.0
16	-0.1972	0.4	721.4737	721	10.6
17	-0.0759	0.4	725.0292	725	10.3
18	0.0389	0.3	728.3942	728	9.9
19	0.1477	0.3	731.5834	732	9.6
20	0.2510	0.3	734.6113	735	9.3
21	0.3491	0.3	737.4867	737	9.0
22	0.4426	0.3	740.2274	740	8.7
23	0.5316	0.3	742.8362	743	8.5
24	0.6167	0.3	745.3306	745	8.2
25	0.6981	0.3	747.7166	748	8.0
26	0.7760	0.3	750.0000	750	7.8
27	0.8508	0.3	752.1925	752	7.6
28	0.9226	0.3	754.2971	754	7.4
29	0.9918	0.2	756.3254	756	7.3
30	1.0585	0.2	758.2805	758	7.1
31	1.1229	0.2	760.1682	760	7.0
32	1.1852	0.2	761.9943	762	6.9
33	1.2456	0.2	763.7648	764	6.7
34	1.3042	0.2	765.4824	765	6.6
35	1.3612	0.2	767.1532	767	6.5
36	1.4168	0.2	768.7830	769	6.4
37	1.4710	0.2	770.3717	770	6.4
38	1.5240	0.2	771.9252	772	6.3
39	1.5760	0.2	773.4494	773	6.2
40	1.6270	0.2	774.9443	775	6.2
41	1.6772	0.2	776.4158	776	6.1
42	1.7266	0.2	777.8638	778	6.1

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.7755	0.2	779.2971	779	6.0
44	1.8239	0.2	780.7158	781	6.0
45	1.8718	0.2	782.1198	782	5.9
46	1.9194	0.2	783.5151	784	5.9
47	1.9669	0.2	784.9074	785	5.9
48	2.0142	0.2	786.2938	786	5.9
49	2.0614	0.2	787.6773	788	5.8
50	2.1087	0.2	789.0638	789	5.8
51	2.1561	0.2	790.4532	790	5.8
52	2.2037	0.2	791.8484	792	5.8
53	2.2515	0.2	793.2495	793	5.8
54	2.2997	0.2	794.6623	795	5.8
55	2.3482	0.2	796.0840	796	5.8
56	2.3973	0.2	797.5232	798	5.8
57	2.4468	0.2	798.9741	799	5.9
58	2.4970	0.2	800.4456	800	5.9
59	2.5479	0.2	801.9375	802	5.9
60	2.5995	0.2	803.4500	803	5.9
61	2.6520	0.2	804.9889	805	6.0
62	2.7054	0.2	806.5541	807	6.0
63	2.7600	0.2	808.1546	808	6.0
64	2.8157	0.2	809.7872	810	6.1
65	2.8729	0.2	811.4639	811	6.2
66	2.9315	0.2	813.1815	813	6.2
67	2.9920	0.2	814.9549	815	6.3
68	3.0546	0.2	816.7898	817	6.4
69	3.1195	0.2	818.6922	819	6.6
70	3.1871	0.2	820.6736	821	6.7
71	3.2580	0.2	822.7518	823	6.9
72	3.3327	0.2	824.9414	825	7.2
73	3.4119	0.3	827.2629	827	7.5
74	3.4963	0.3	829.7368	830	7.8
75	3.5869	0.3	832.3925	832	8.2
76	3.6846	0.3	835.2563	835	8.7
77	3.7906	0.3	838.3633	838	9.3
78	3.9061	0.3	841.7488	842	9.9
79	4.0324	0.4	845.4509	845	10.6
80	4.1707	0.4	849.5047	850	10.6
81	4.3224	0.4	853.9513	850	10.6
82	4.4890	0.4	858.8347	850	10.6
83	4.6717	0.5	864.1899	850	10.6
84	4.8722	0.5	870.0670	850	10.6
85	5.0924	0.5	876.5214	850	10.6
86	5.3349	0.6	883.6295	850	10.6

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	5.6038	0.6	891.5115	850	10.6
88	5.9060	0.7	900.3695	850	10.6
89	6.2543	0.7	910.5788	850	10.6
90	6.6715	0.8	922.8077	850	10.6
91	7.1999	1.0	938.2960	850	10.6
92	7.9228	1.2	959.4855	850	10.6
93	9.0366	1.7	992.1330	850	10.6
94	11.2797	3.0	1057.8823	850	10.6
95	15.000	6.6	1166.9310	850	10.6

Table A.12.47 Conversion Table for Performance Level Setting Form: Integrated Mathematics III

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
0	-10.000	11.0	320.8200	650	20.0
1	-3.9857	1.4	558.5581	650	20.0
2	-2.9620	0.9	599.0237	650	20.0
3	-2.3774	0.7	622.1322	650	20.0
4	-1.9684	0.6	638.2995	650	20.0
5	-1.6530	0.6	650.7669	651	20.0
6	-1.3955	0.5	660.9456	661	19.8
7	-1.1771	0.5	669.5786	670	18.2
8	-0.9874	0.4	677.0773	677	17.0
9	-0.8195	0.4	683.7141	684	16.1
10	-0.6689	0.4	689.6672	690	15.2
11	-0.5324	0.4	695.0629	695	14.5
12	-0.4075	0.4	700.0000	700	13.9
13	-0.2924	0.3	704.5498	705	13.4
14	-0.1855	0.3	708.7754	709	12.9
15	-0.0856	0.3	712.7243	713	12.5
16	0.0083	0.3	716.4361	716	12.1
17	0.0969	0.3	719.9383	720	11.8
18	0.1810	0.3	723.2627	723	11.5
19	0.2611	0.3	726.4290	726	11.2
20	0.3377	0.3	729.4569	729	10.9
21	0.4110	0.3	732.3543	732	10.7
22	0.4816	0.3	735.1451	735	10.5
23	0.5495	0.3	737.8291	738	10.2
24	0.6151	0.3	740.4222	740	10.0
25	0.6786	0.2	742.9322	743	9.9
26	0.7400	0.2	745.3593	745	9.7
27	0.7996	0.2	747.7152	748	9.5
28	0.8574	0.2	750.0000	750	9.3
29	0.9136	0.2	752.2215	752	9.2
30	0.9683	0.2	754.3837	754	9.0
31	1.0215	0.2	756.4867	756	8.8
32	1.0733	0.2	758.5343	759	8.7
33	1.1238	0.2	760.5305	761	8.5
34	1.1730	0.2	762.4753	762	8.4
35	1.2210	0.2	764.3727	764	8.3
36	1.2678	0.2	766.2226	766	8.1
37	1.3135	0.2	768.0291	768	8.0
38	1.3581	0.2	769.7921	770	7.9
39	1.4017	0.2	771.5155	772	7.7
40	1.4444	0.2	773.2034	773	7.6
41	1.4862	0.2	774.8557	775	7.5
42	1.5271	0.2	776.4724	776	7.4

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
43	1.5672	0.2	778.0575	778	7.3
44	1.6066	0.2	779.6150	780	7.2
45	1.6453	0.2	781.1447	781	7.2
46	1.6834	0.2	782.6508	783	7.1
47	1.7210	0.2	784.1371	784	7.0
48	1.7580	0.2	785.5996	786	7.0
49	1.7945	0.2	787.0424	787	6.9
50	1.8307	0.2	788.4734	788	6.8
51	1.8665	0.2	789.8885	790	6.8
52	1.9020	0.2	791.2918	791	6.8
53	1.9373	0.2	792.6871	793	6.7
54	1.9723	0.2	794.0707	794	6.7
55	2.0072	0.2	795.4502	795	6.7
56	2.0420	0.2	796.8258	797	6.7
57	2.0768	0.2	798.2014	798	6.7
58	2.1115	0.2	799.5731	800	6.6
59	2.1463	0.2	800.9487	801	6.6
60	2.1812	0.2	802.3282	802	6.7
61	2.2163	0.2	803.7157	804	6.7
62	2.2516	0.2	805.1110	805	6.7
63	2.2872	0.2	806.5183	807	6.7
64	2.3232	0.2	807.9413	808	6.7
65	2.3597	0.2	809.3841	809	6.8
66	2.3967	0.2	810.8467	811	6.8
67	2.4343	0.2	812.3330	812	6.9
68	2.4727	0.2	813.8509	814	7.0
69	2.5120	0.2	815.4043	815	7.0
70	2.5522	0.2	816.9934	817	7.1
71	2.5935	0.2	818.6259	819	7.2
72	2.6360	0.2	820.3059	820	7.3
73	2.6798	0.2	822.0373	822	7.5
74	2.7252	0.2	823.8319	824	7.6
75	2.7723	0.2	825.6937	826	7.7
76	2.8213	0.2	827.6306	828	7.9
77	2.8723	0.2	829.6466	830	8.1
78	2.9256	0.2	831.7535	832	8.3
79	2.9815	0.2	833.9631	834	8.5
80	3.0402	0.2	836.2835	836	8.7
81	3.1020	0.2	838.7263	839	9.0
82	3.1674	0.2	841.3115	841	9.2
83	3.2367	0.2	844.0509	844	9.6
84	3.3106	0.3	846.9720	847	9.9
85	3.3896	0.3	850.0948	850	9.9
86	3.4747	0.3	853.4587	850	9.9

Raw Score	IRT Theta	Theta CSEM	Unrounded Scale Score	Rounded Scale Score	Scale Score CSEM
87	3.5668	0.3	857.0993	850	9.9
88	3.6674	0.3	861.0759	850	9.9
89	3.7783	0.3	865.4597	850	9.9
90	3.9019	0.3	870.3454	850	9.9
91	4.0417	0.4	875.8716	850	9.9
92	4.2023	0.4	882.2199	850	9.9
93	4.3905	0.4	889.6592	850	9.9
94	4.6167	0.5	898.6006	850	9.9
95	4.8970	0.6	909.6805	850	9.9
96	5.2601	0.7	924.0334	850	9.9
97	5.7624	0.9	943.8888	850	9.9
98	6.5405	1.2	974.6461	850	9.9
99	8.0537	1.9	1034.4611	850	9.9
100	15.000	9.7	1309.0400	850	9.9

Appendix 12.4: IRT Test Characteristic Curves, Information Curves and CSEM Curves

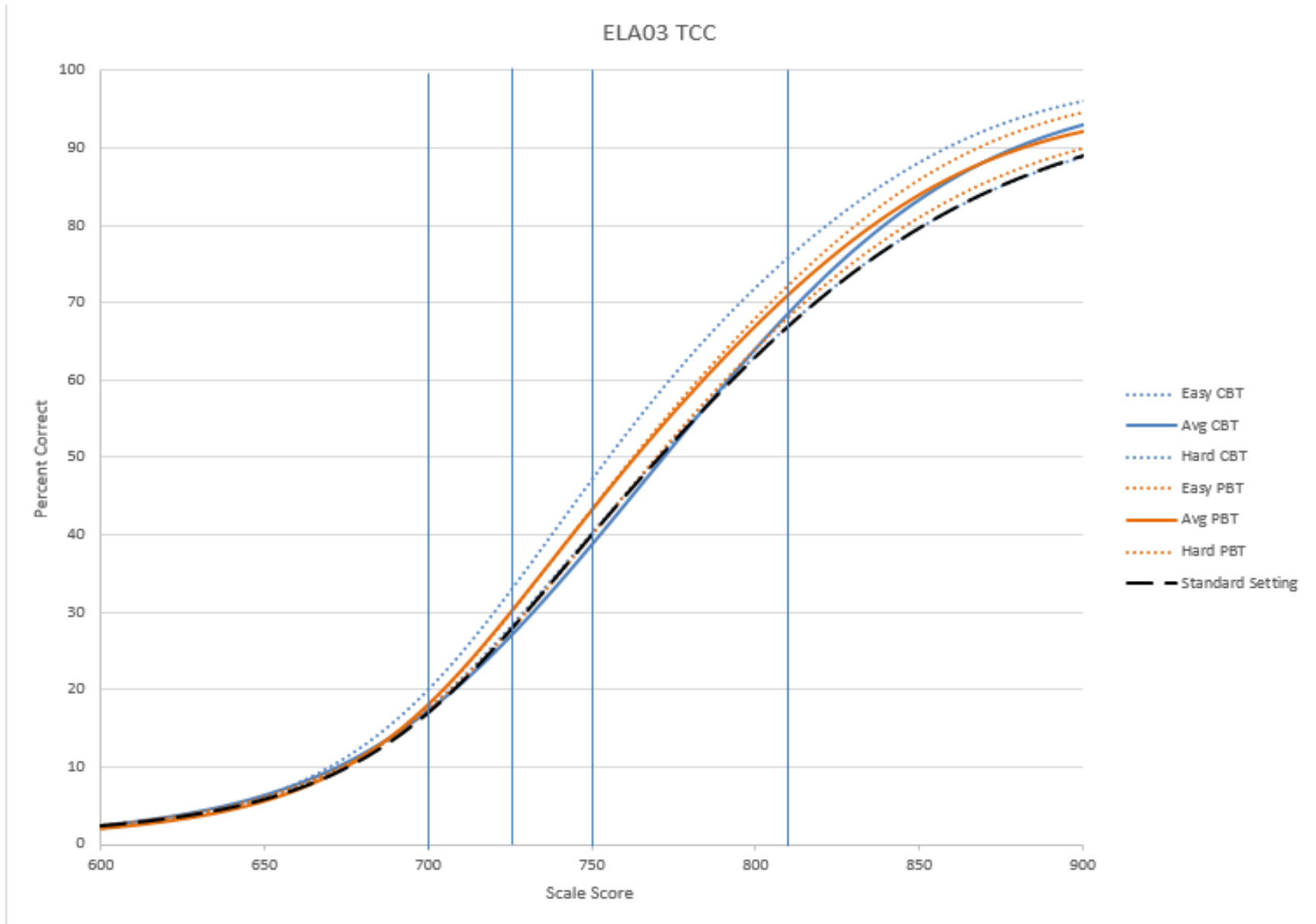


Figure A.12.1 Test Characteristic Curves ELA/L Grade 3

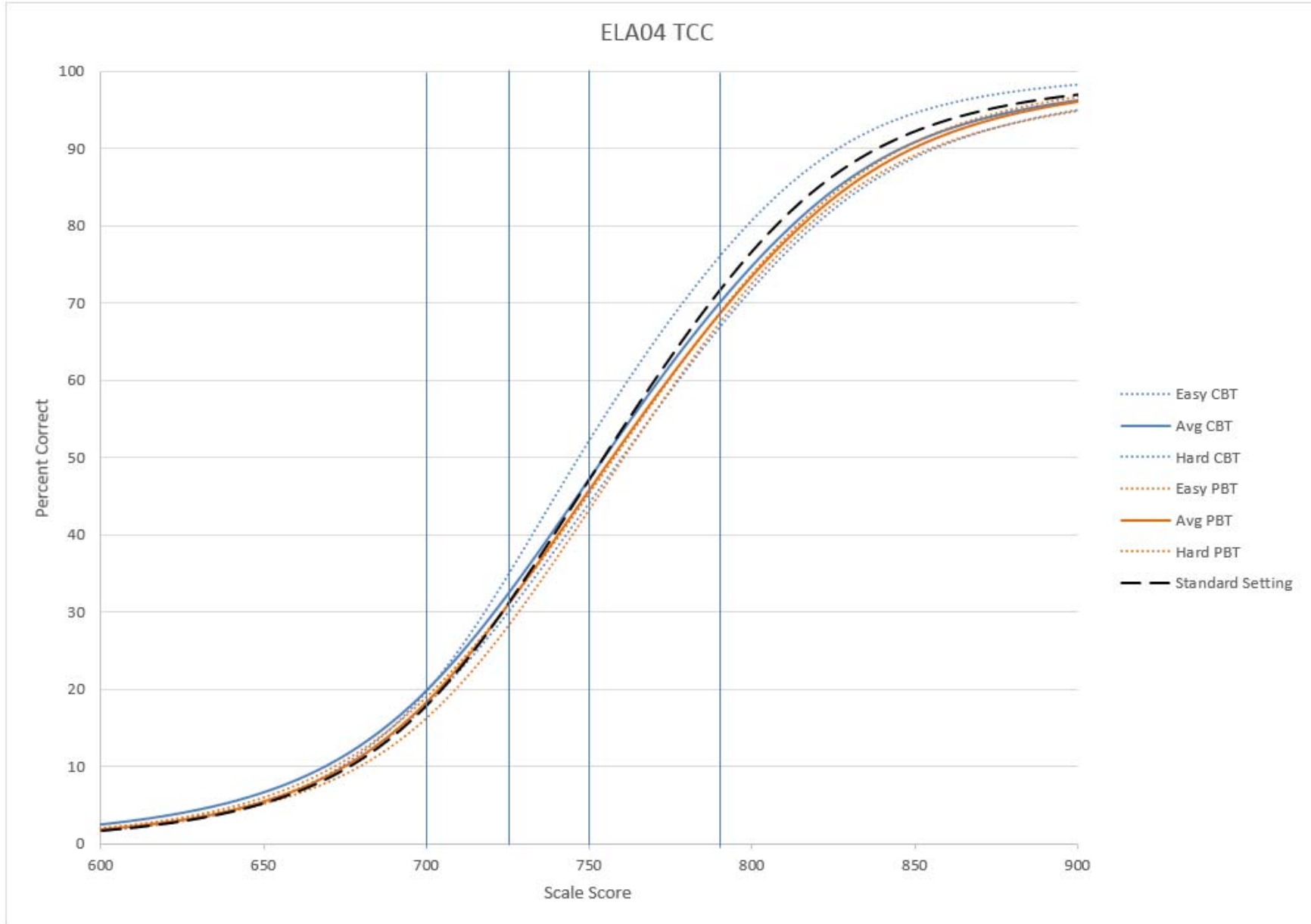


Figure A.12.2 Test Characteristic Curves ELA/L Grade 4

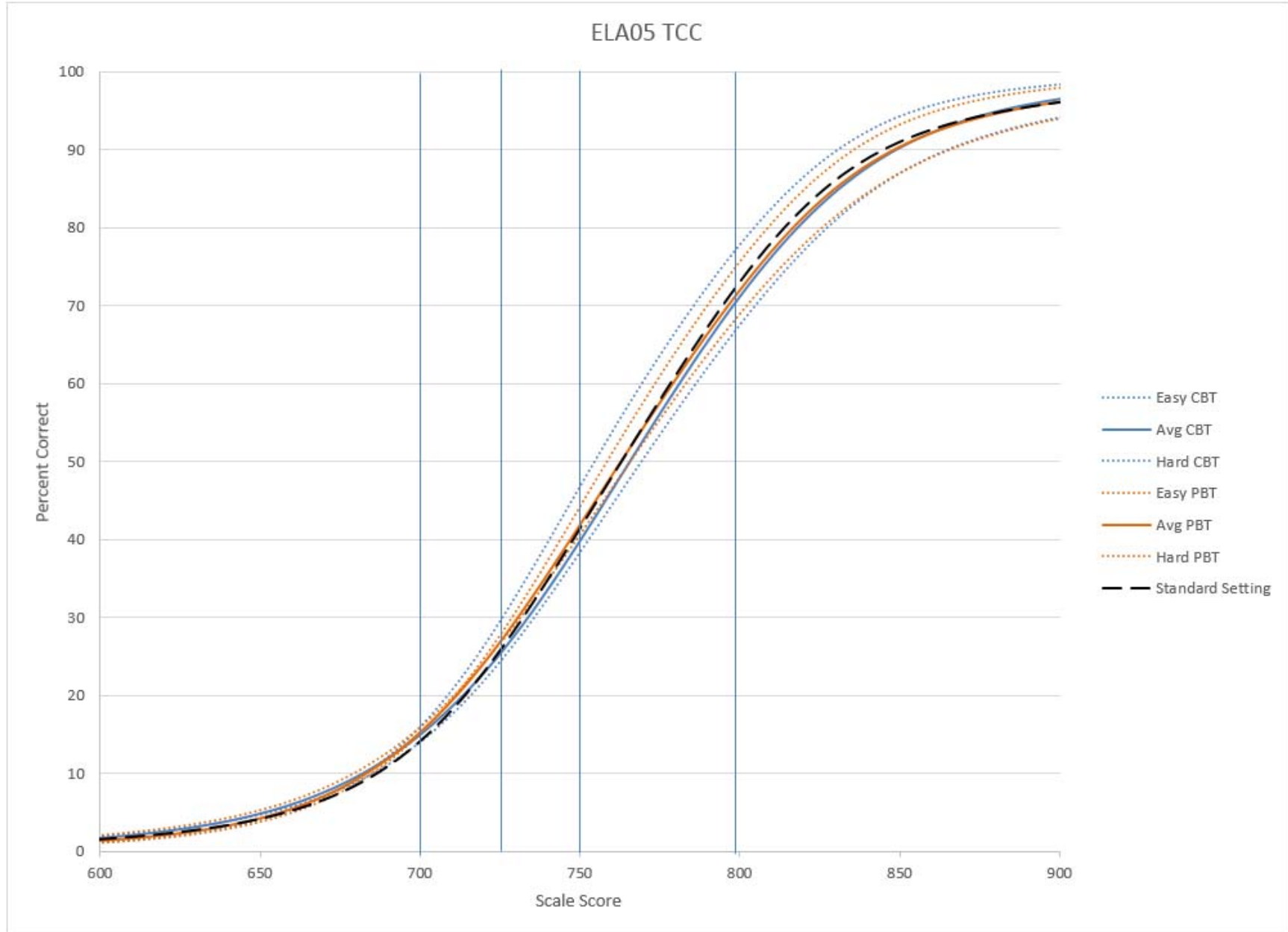


Figure A.12.3 Test Characteristic Curves ELA/L Grade 5

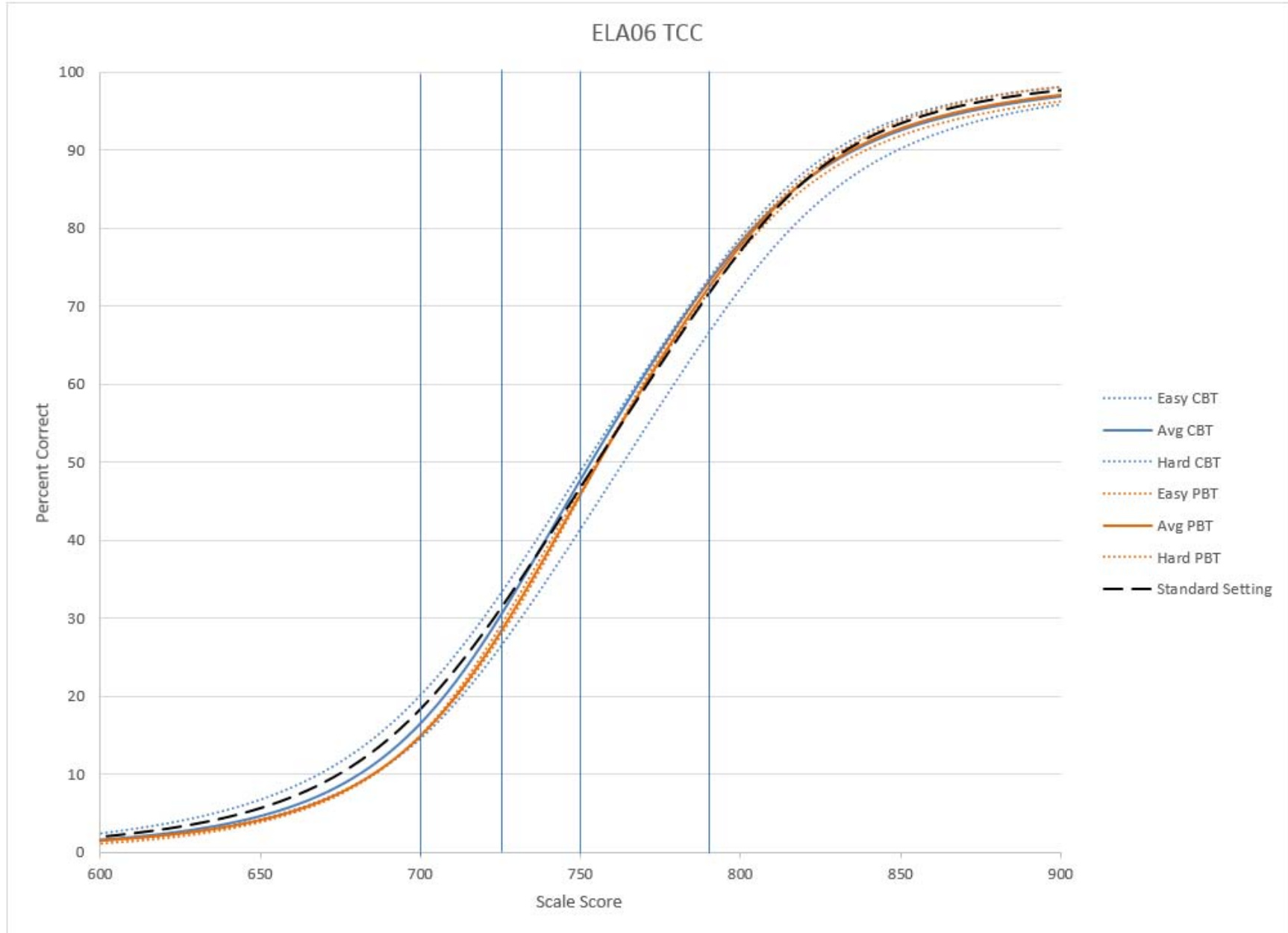


Figure A.12.4 Test Characteristic Curves ELA/L Grade 6

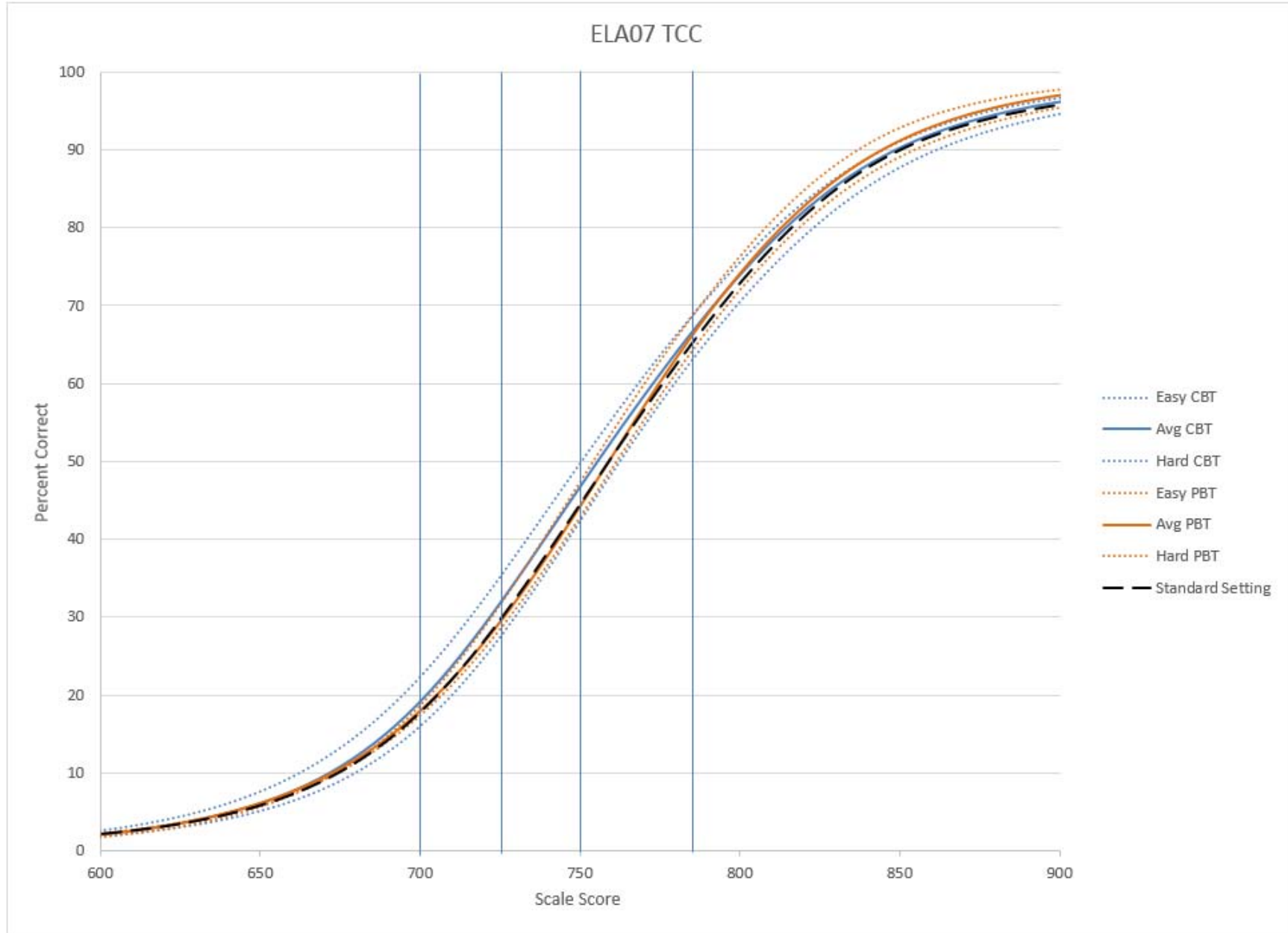


Figure A.12.5 Test Characteristic Curves ELA/L Grade 7

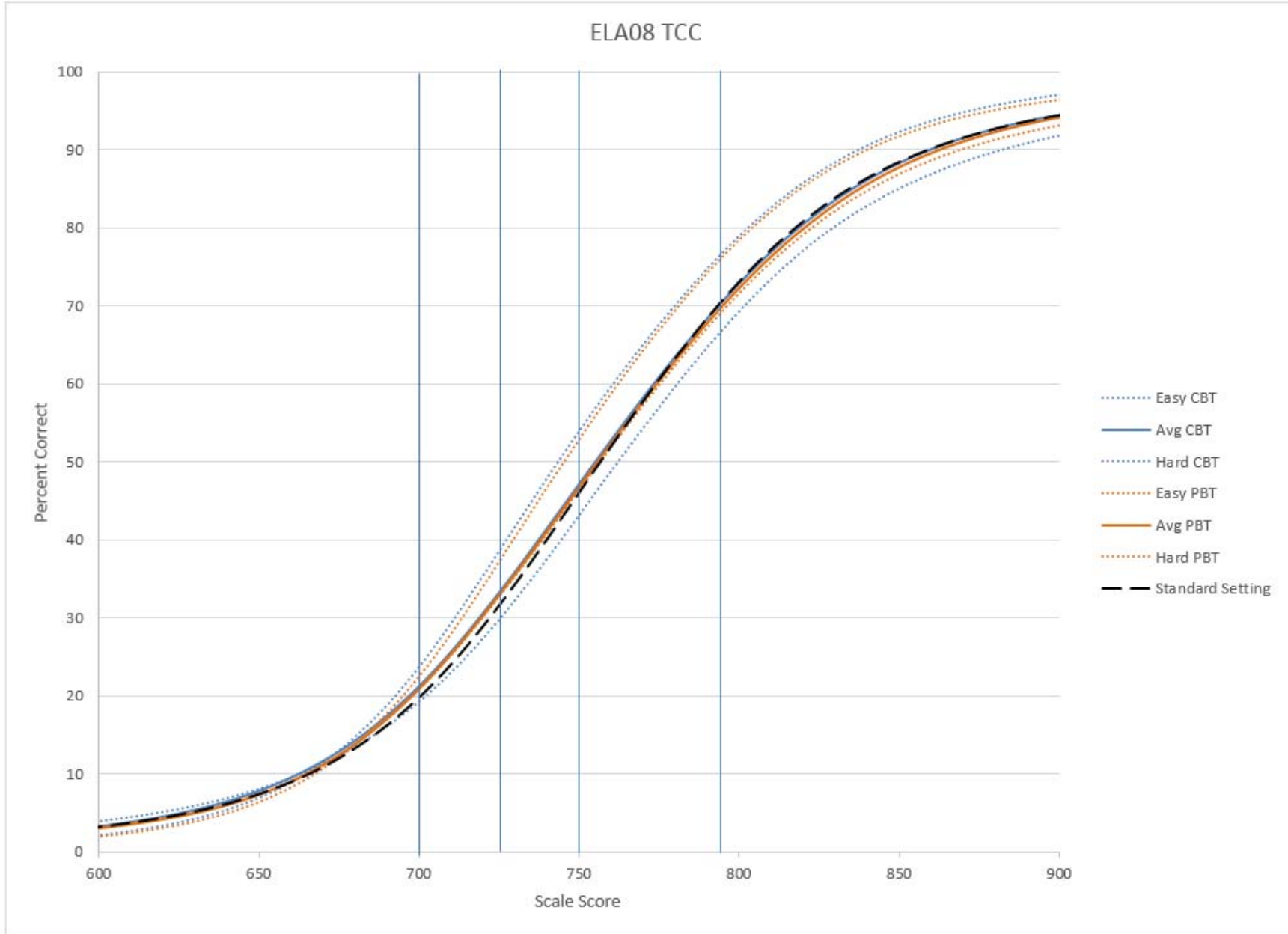


Figure A.12.6 Test Characteristic Curves ELA/L Grade 8

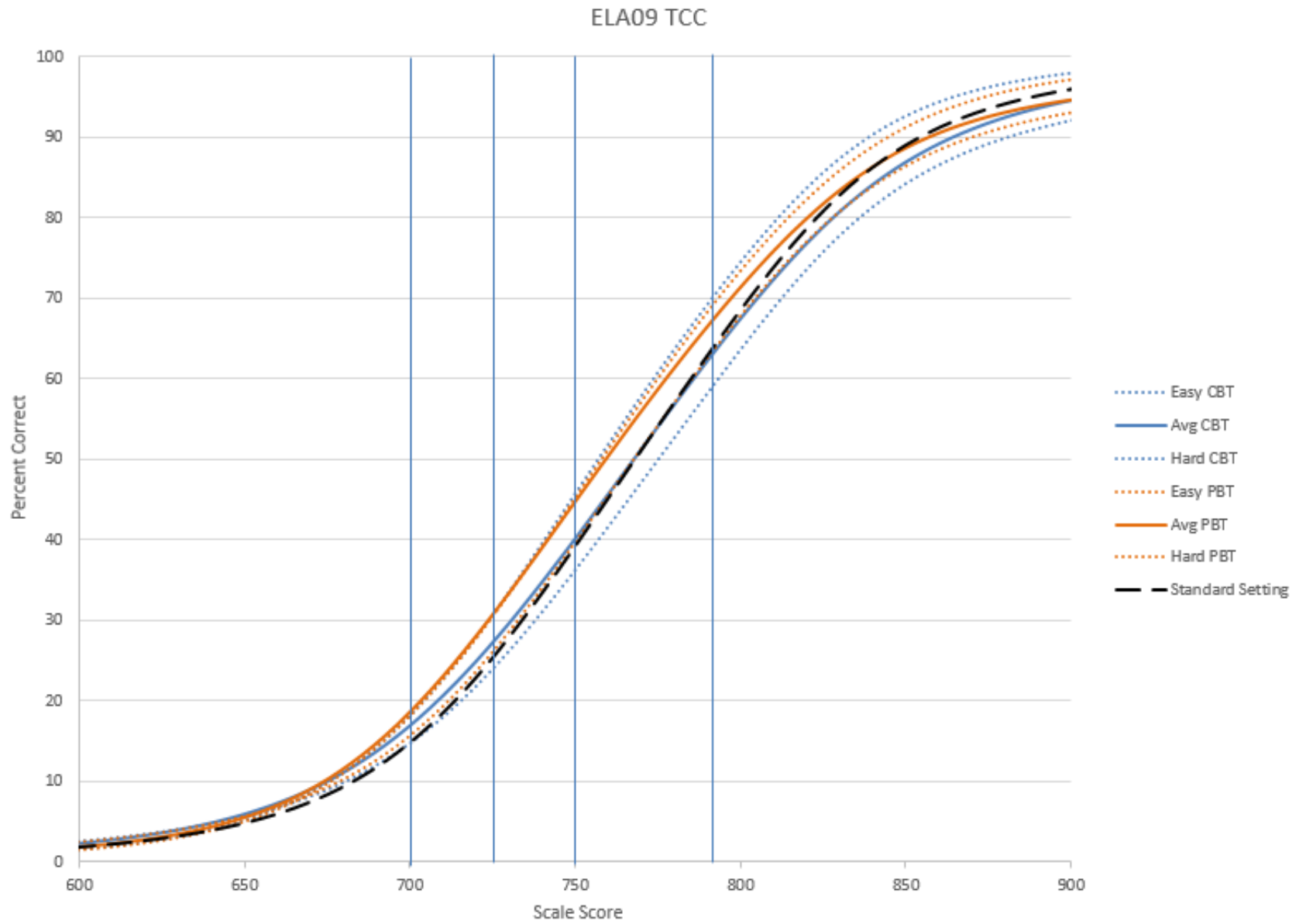


Figure A.12.7 Test Characteristic Curves ELA/L Grade 9

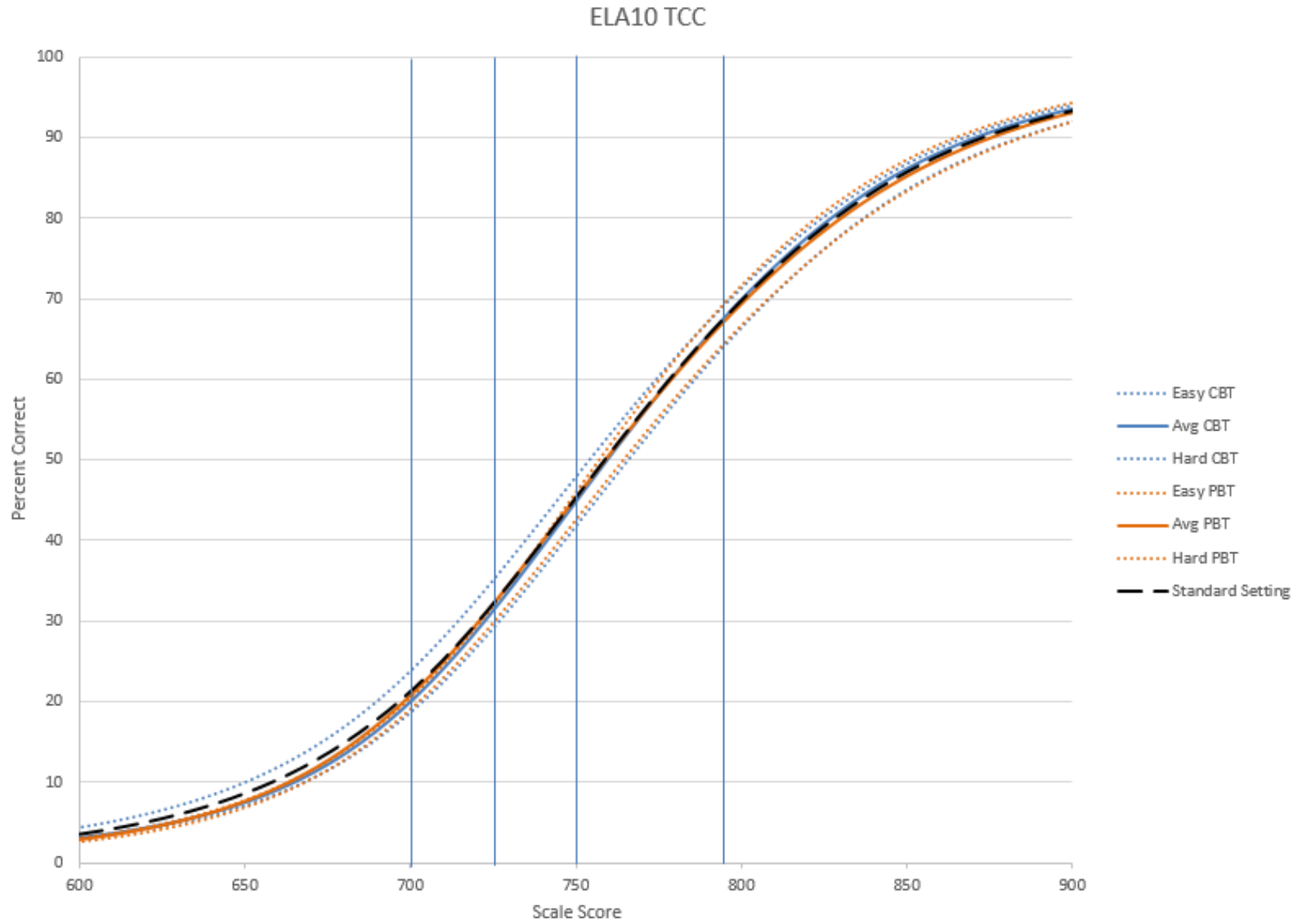


Figure A.12.8 Test Characteristic Curves ELA/L Grade 10

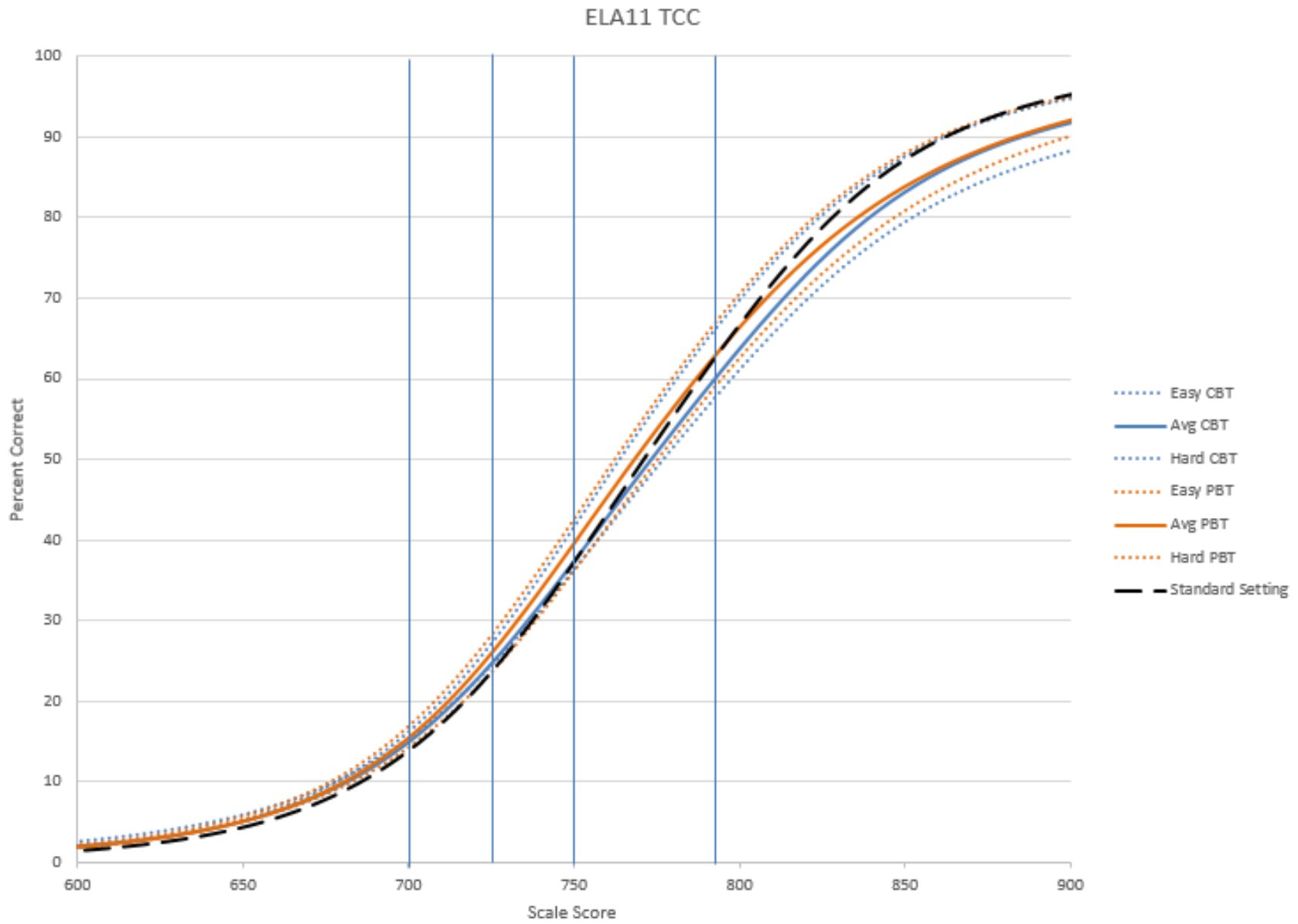


Figure A.12.9 Test Characteristic Curves ELA/L Grade 11

April 18, 2016

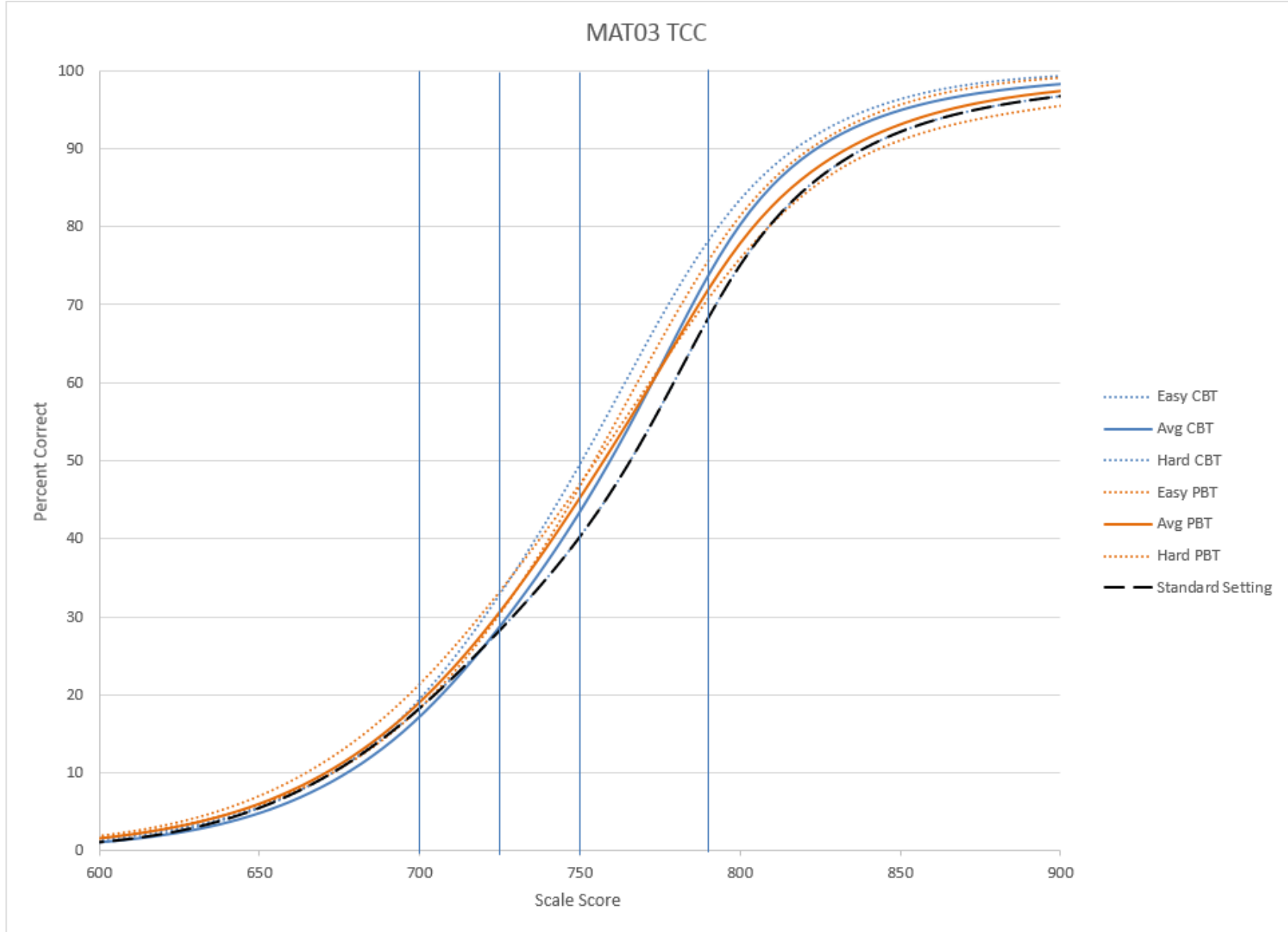


Figure A.12.10 Test Characteristic Curves Mathematics Grade 3

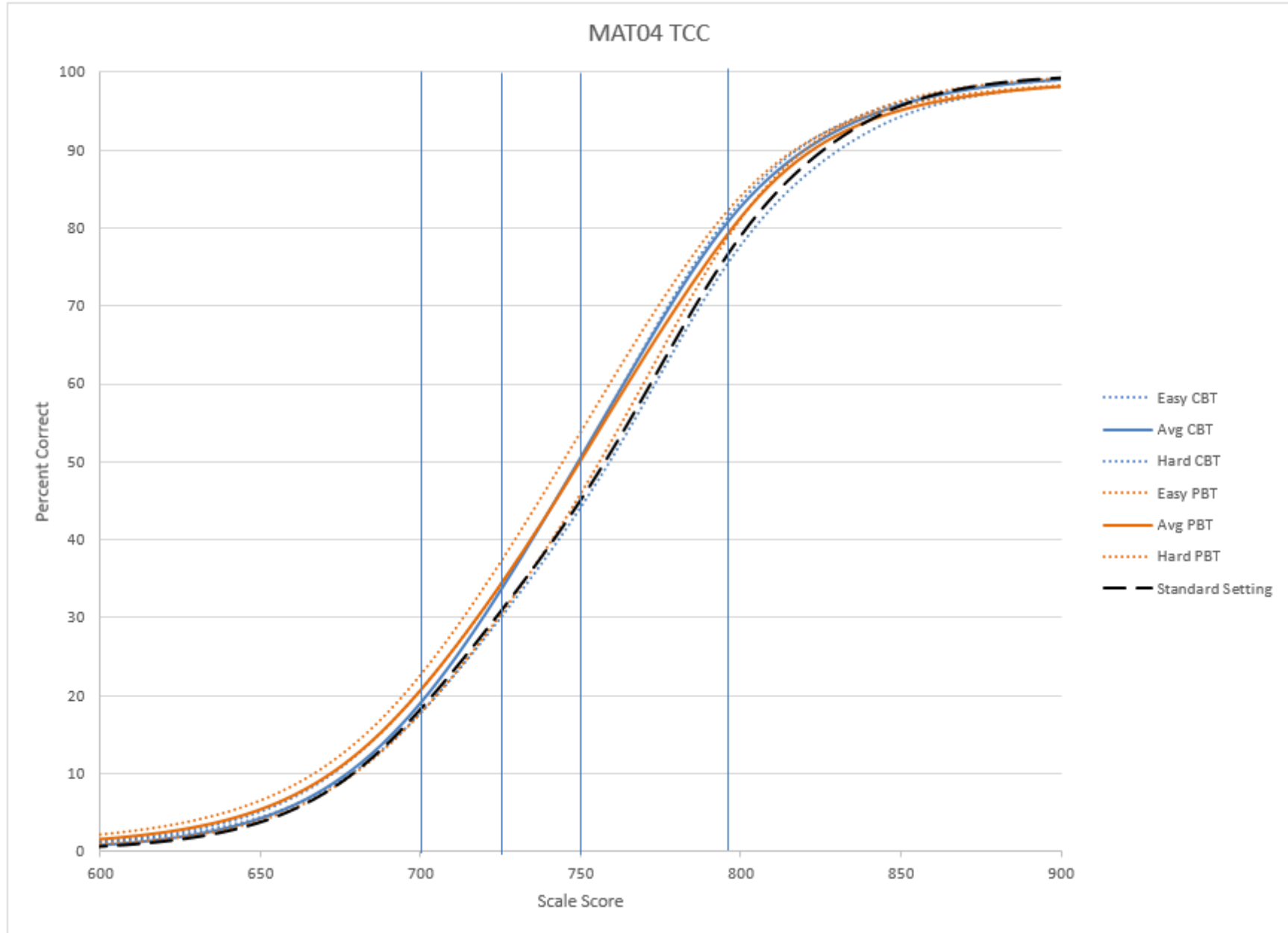


Figure A.12.11 Test Characteristic Curves Mathematics Grade 4

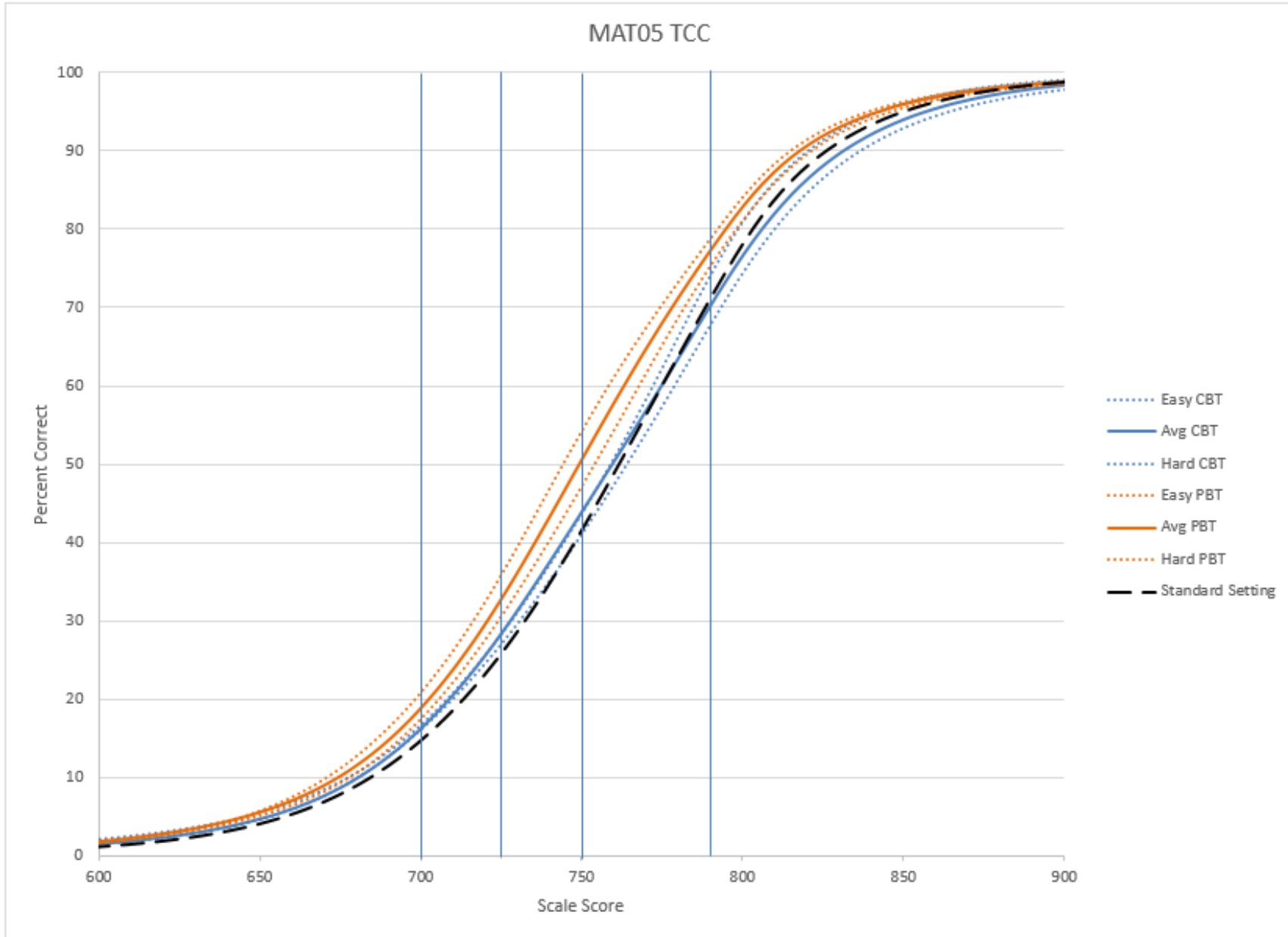


Figure A.12.12 Test Characteristic Curves Mathematics Grade 5

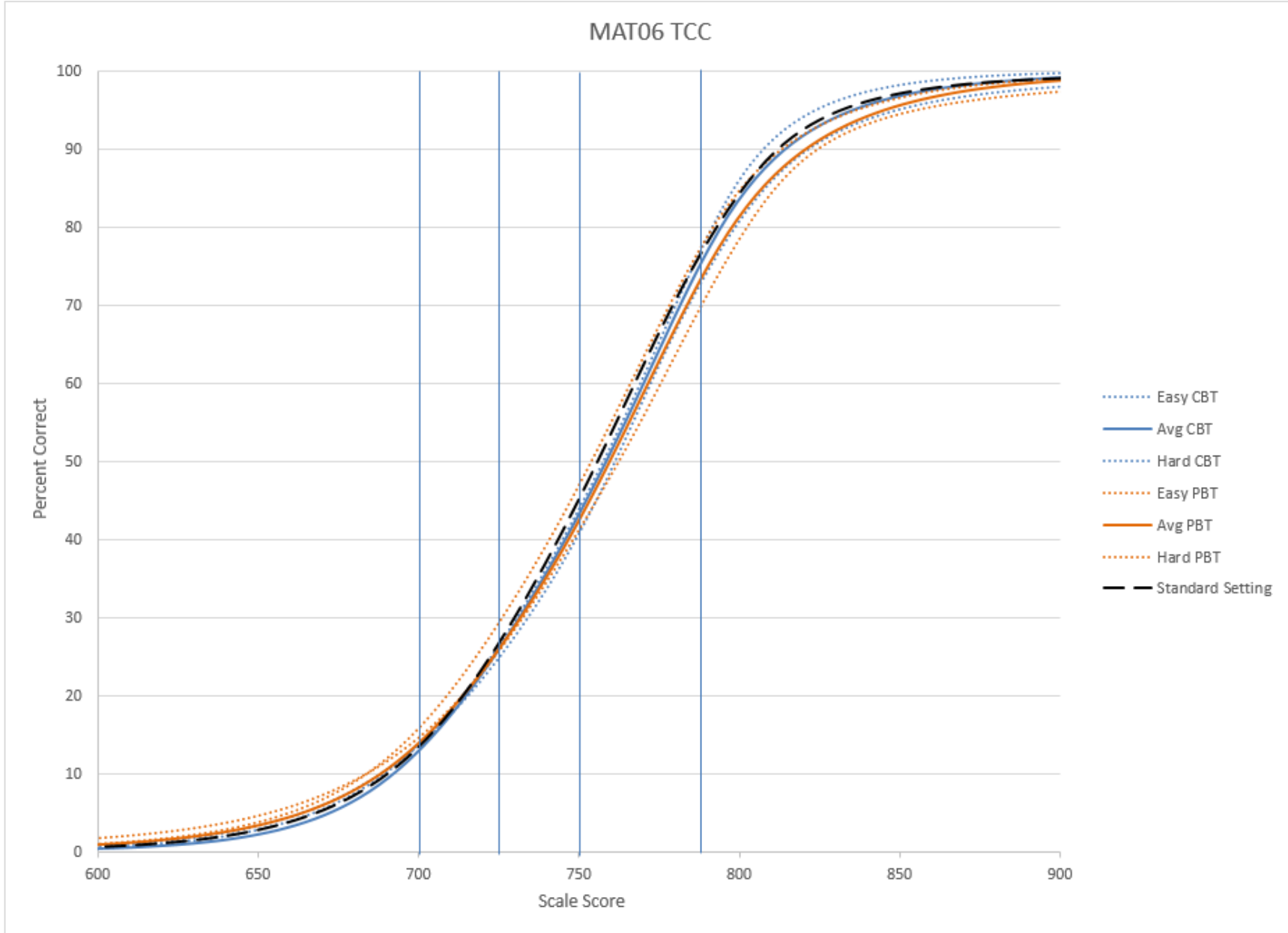


Figure A.12.13 Test Characteristic Curves Mathematics Grade 6

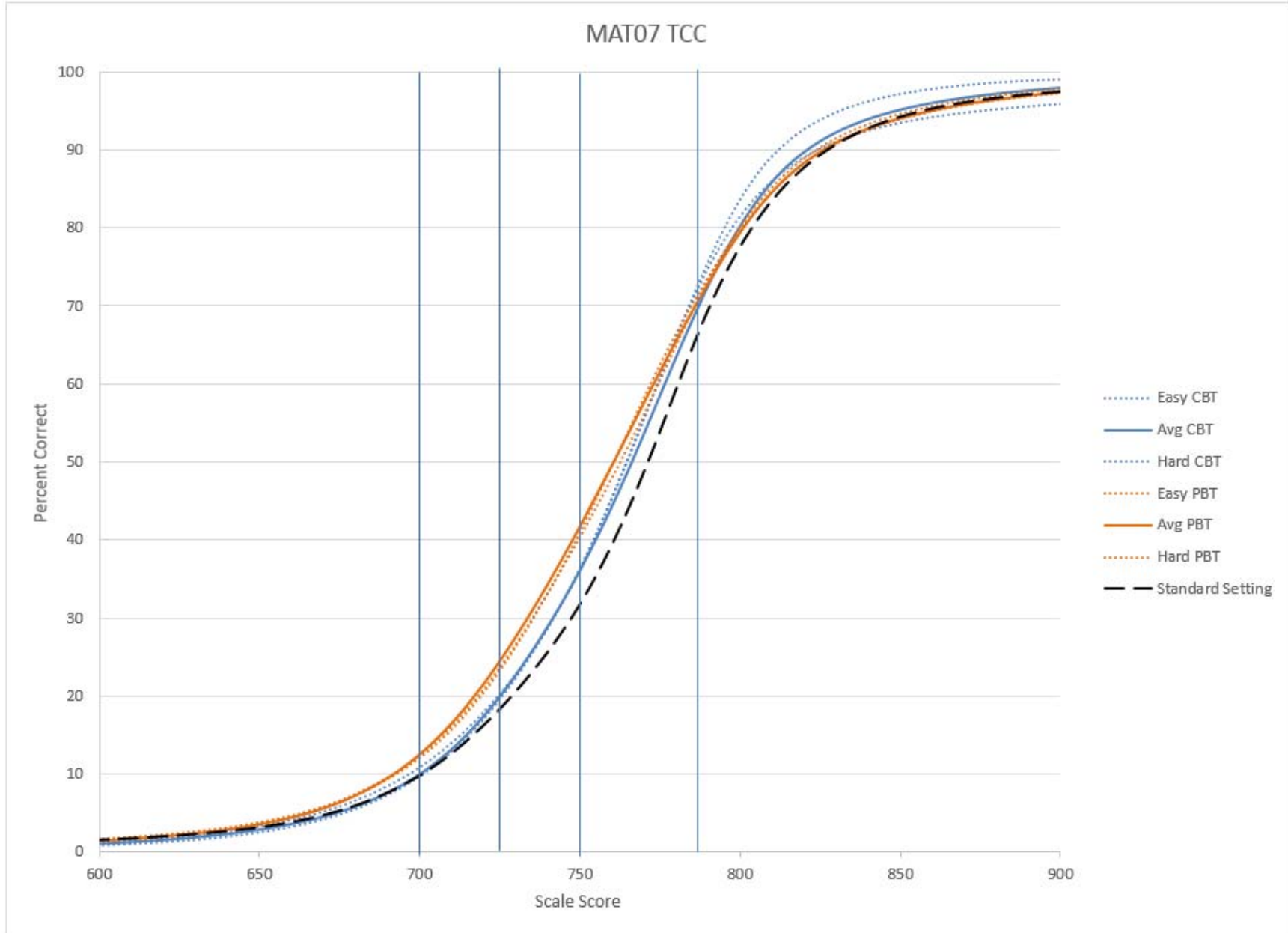


Figure A.12.14 Test Characteristic Curves Mathematics Grade 7

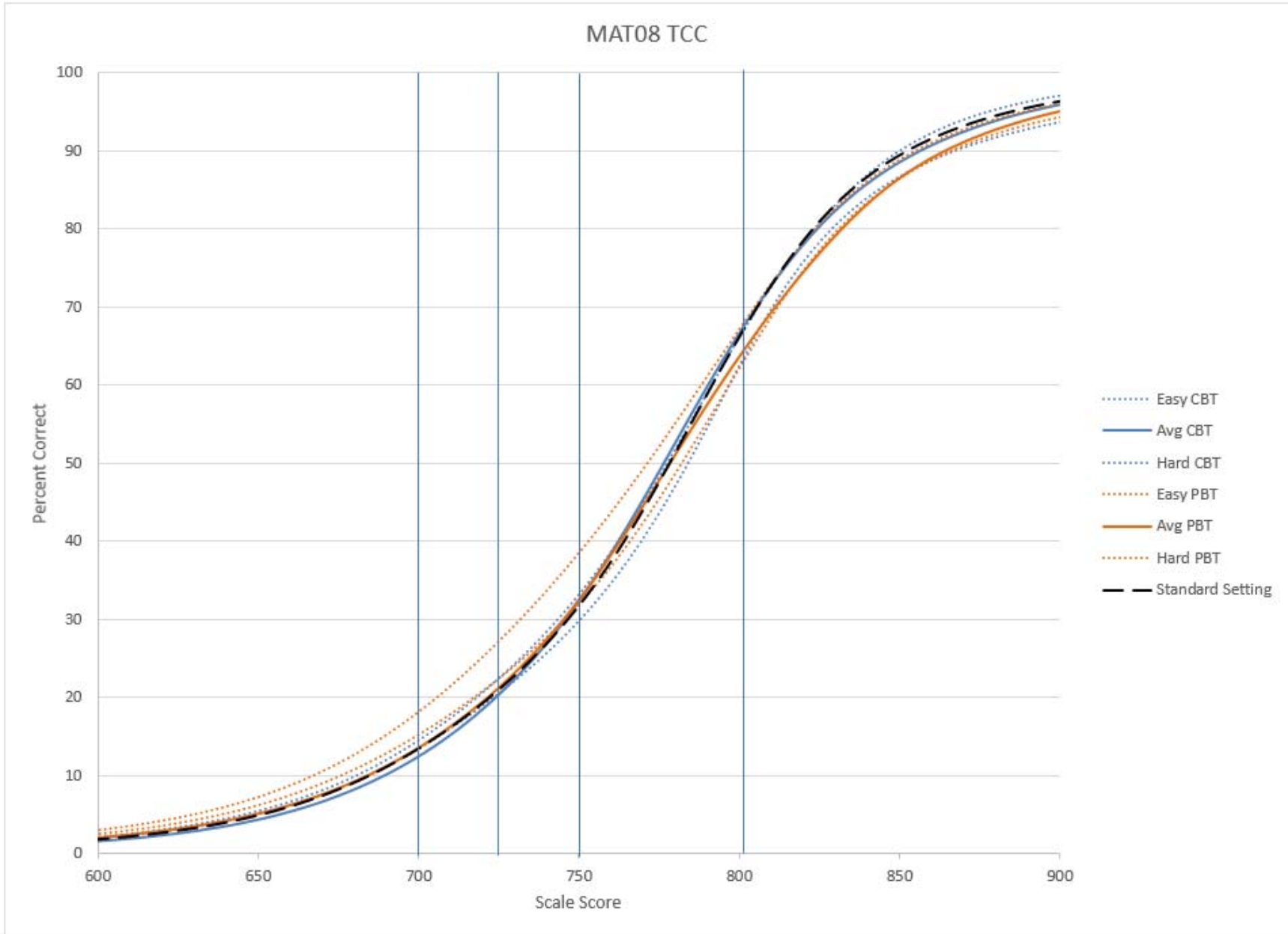


Figure A.12.15 Test Characteristic Curves Mathematics Grade 8

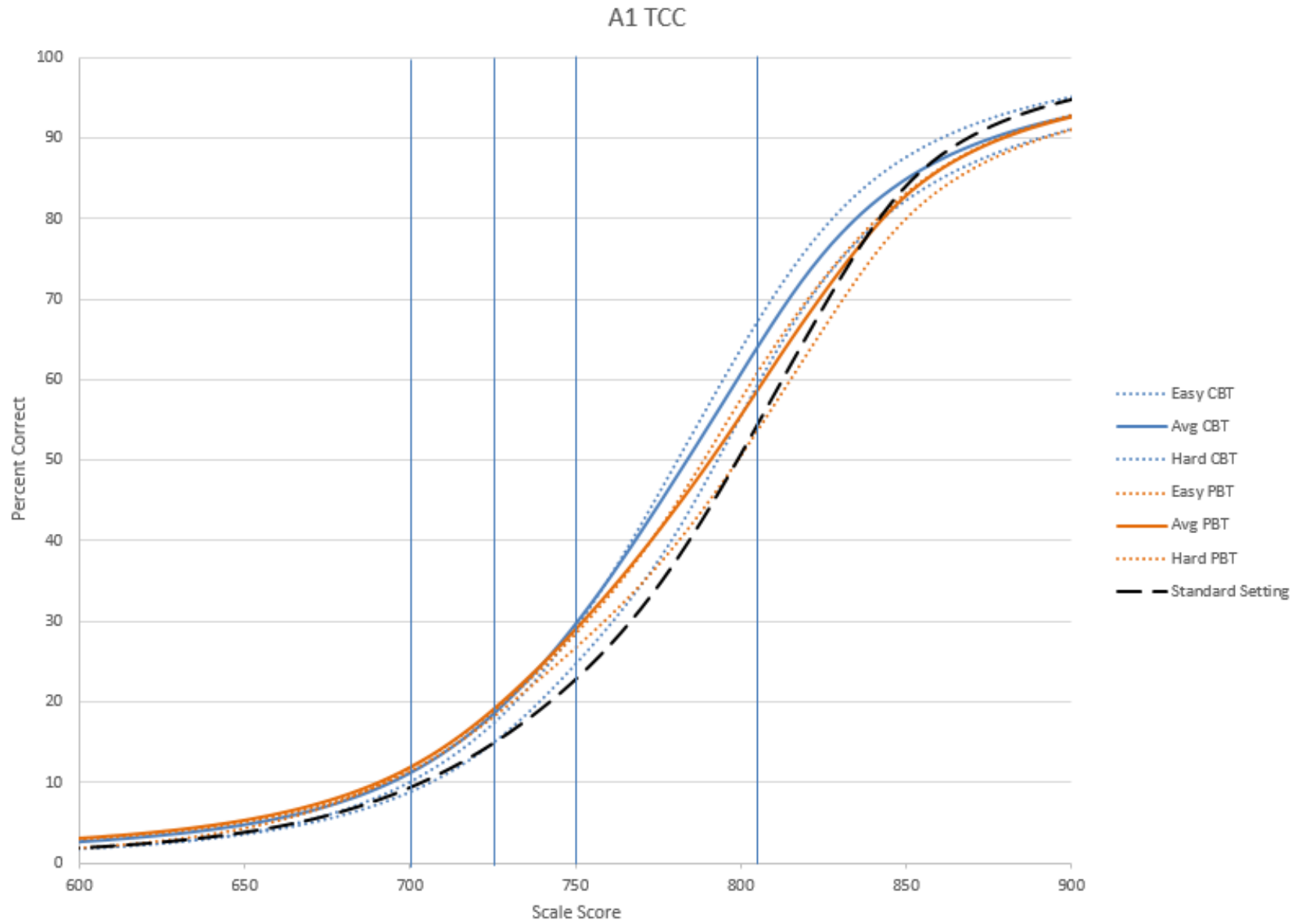


Figure A.12.16 Test Characteristic Curves Algebra I

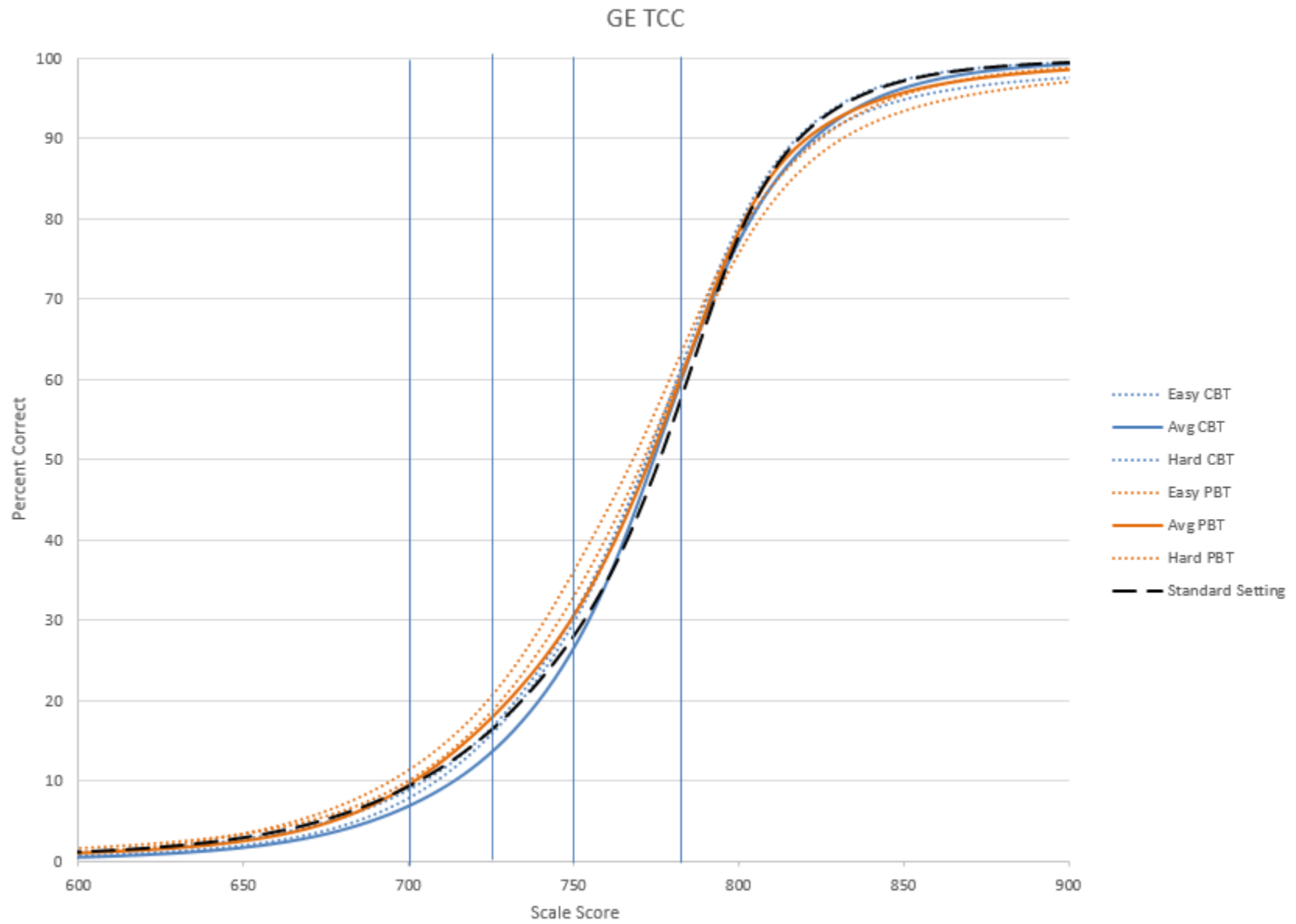


Figure A.12.17 Test Characteristic Curves Geometry

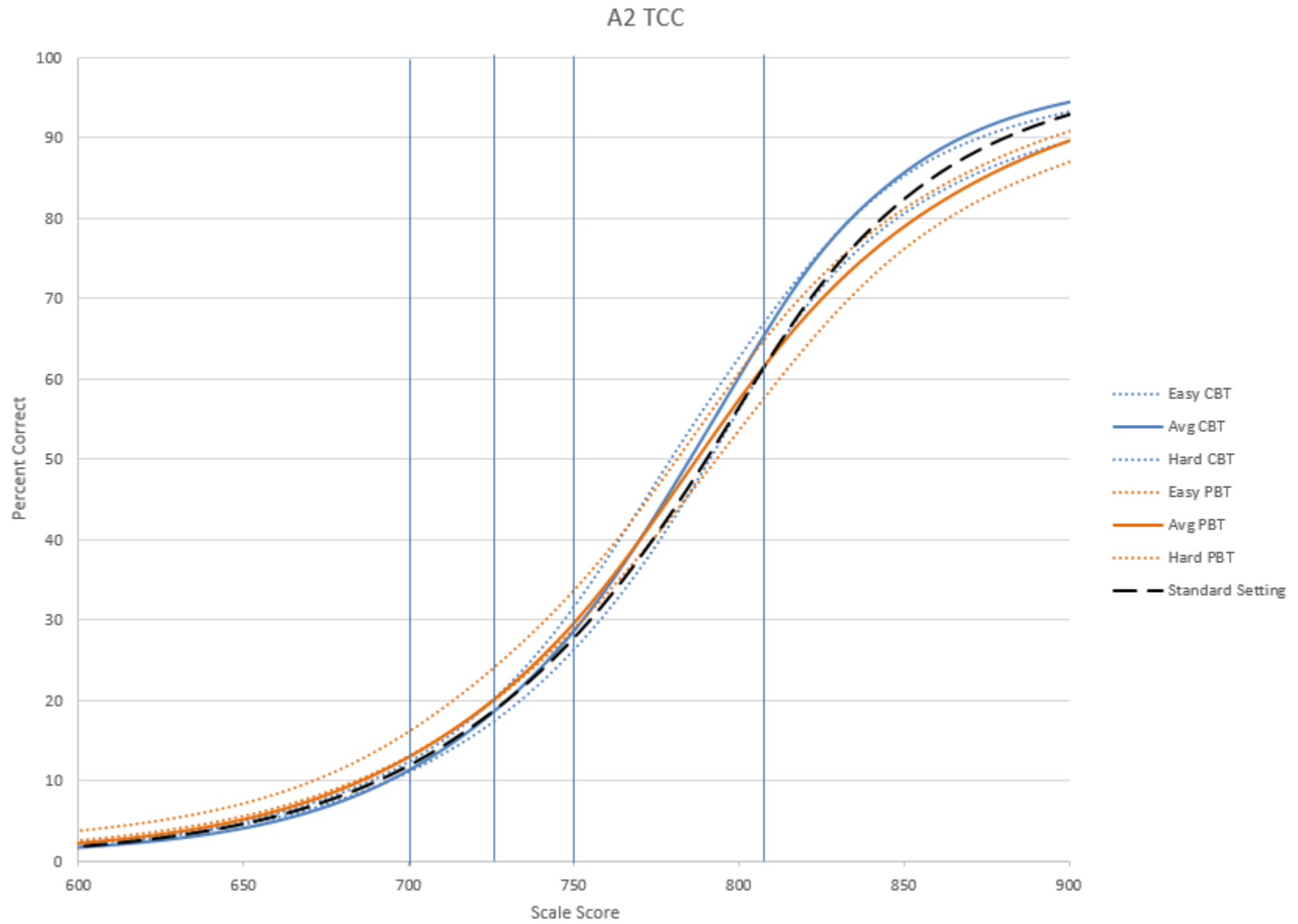


Figure A.12.18 Test Characteristic Curves Algebra II

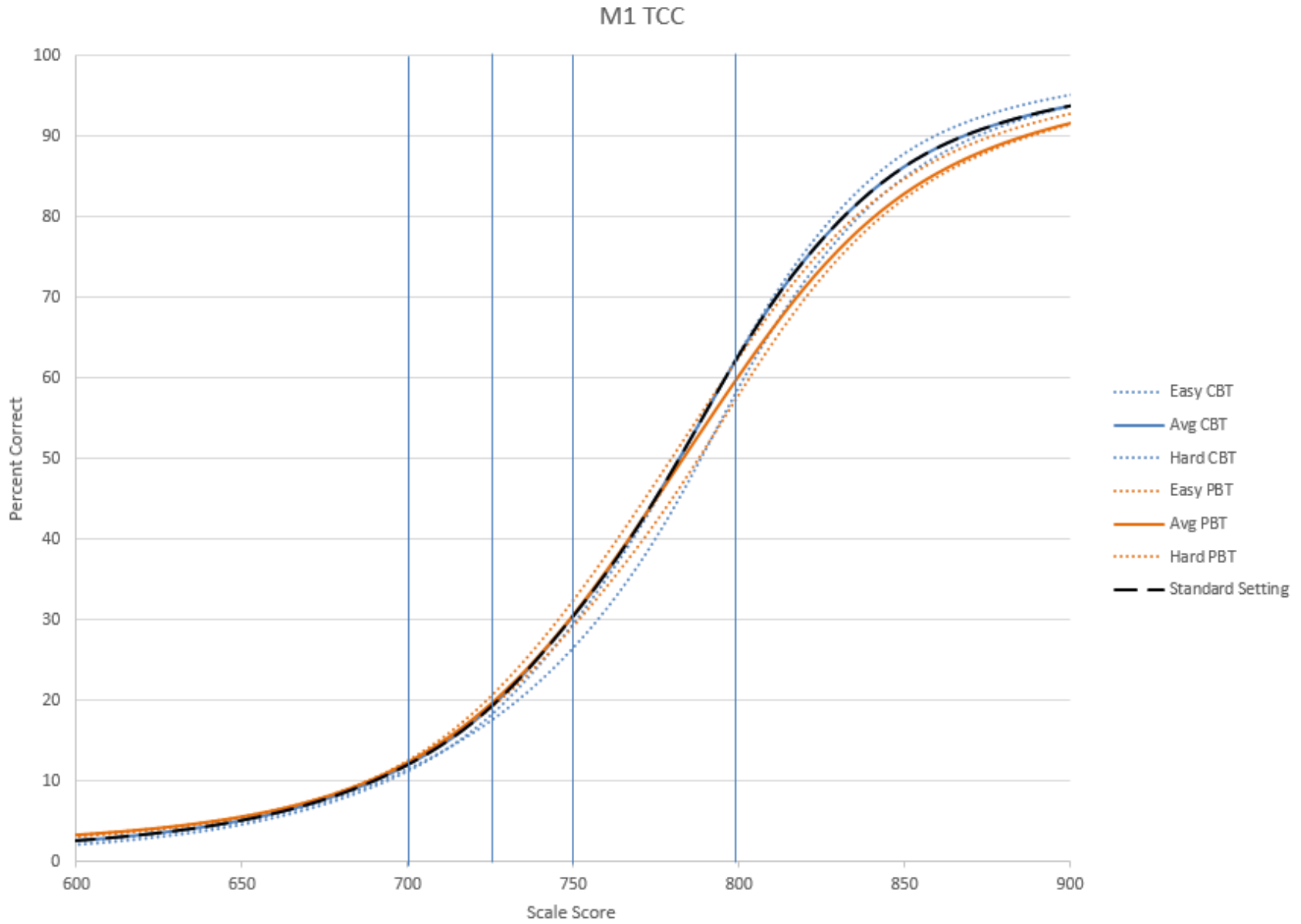


Figure A.12.19 Test Characteristic Curves Integrated Mathematics I

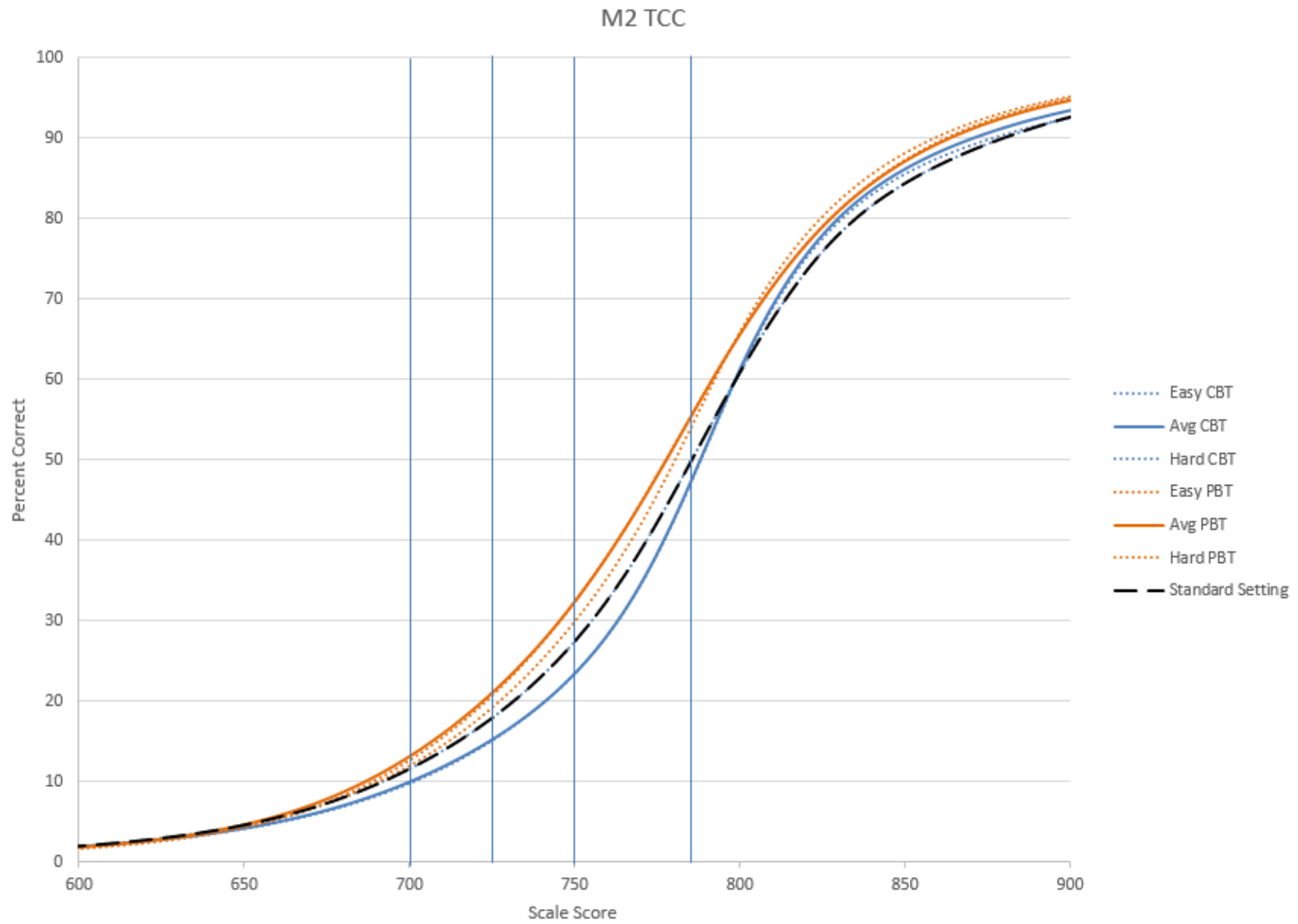


Figure A.12.20 Test Characteristic Curves Integrated Mathematics II

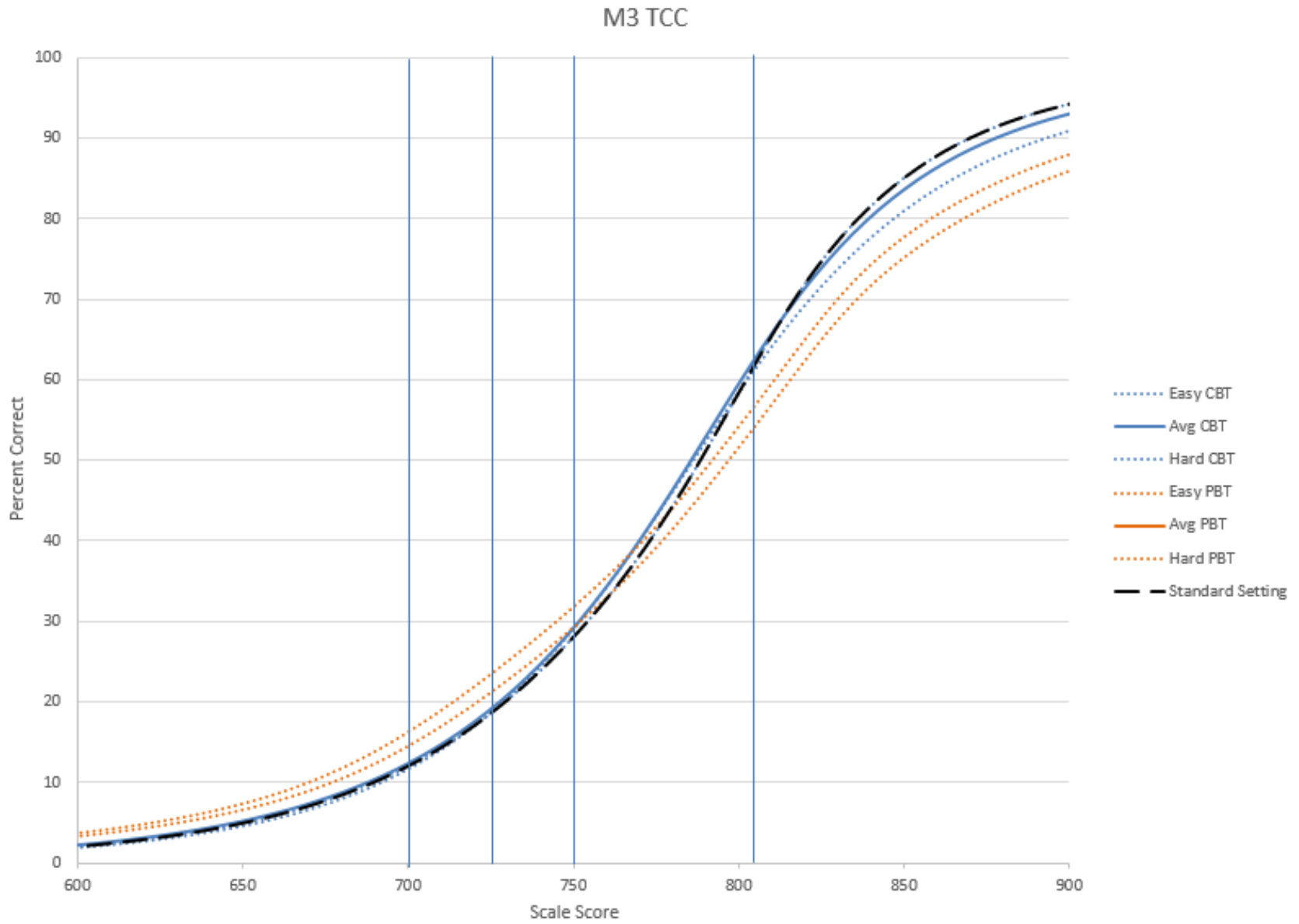


Figure A.12.21 Test Characteristic Curves Integrated Mathematics III

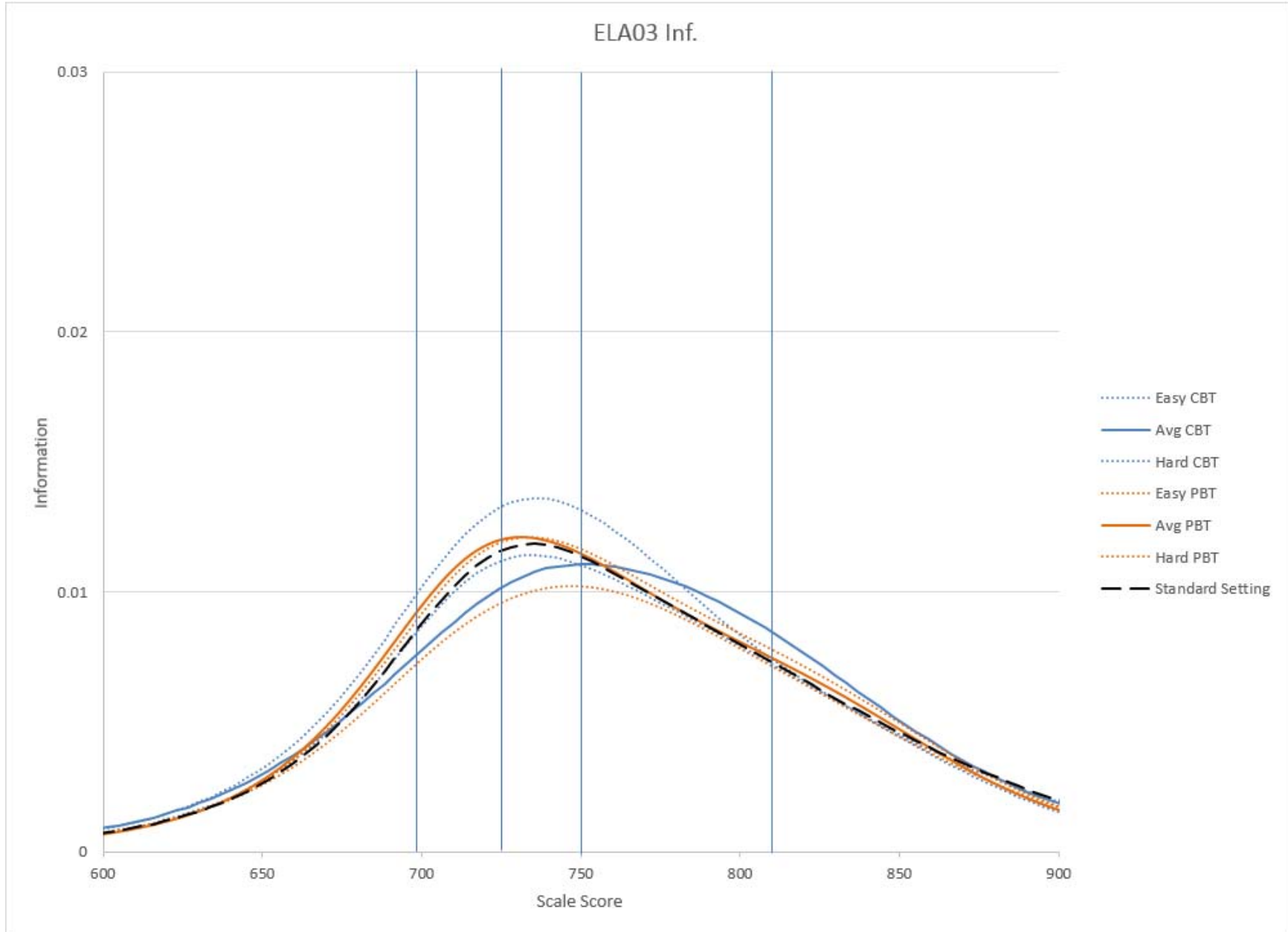


Figure A.12.22 Test Information Curves ELA/L Grade 3

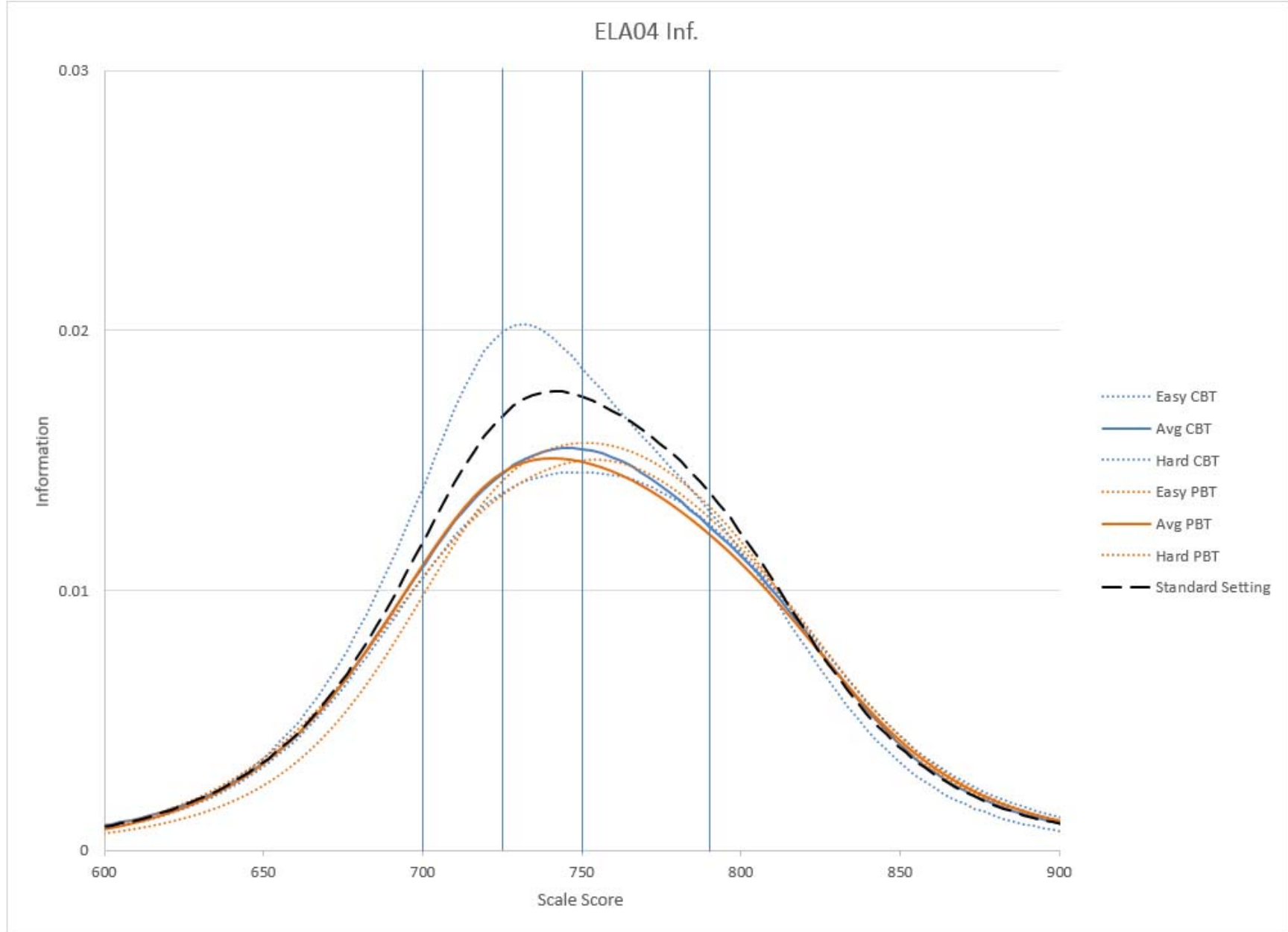


Figure A.12.23 Test Information Curves ELA/L Grade 4

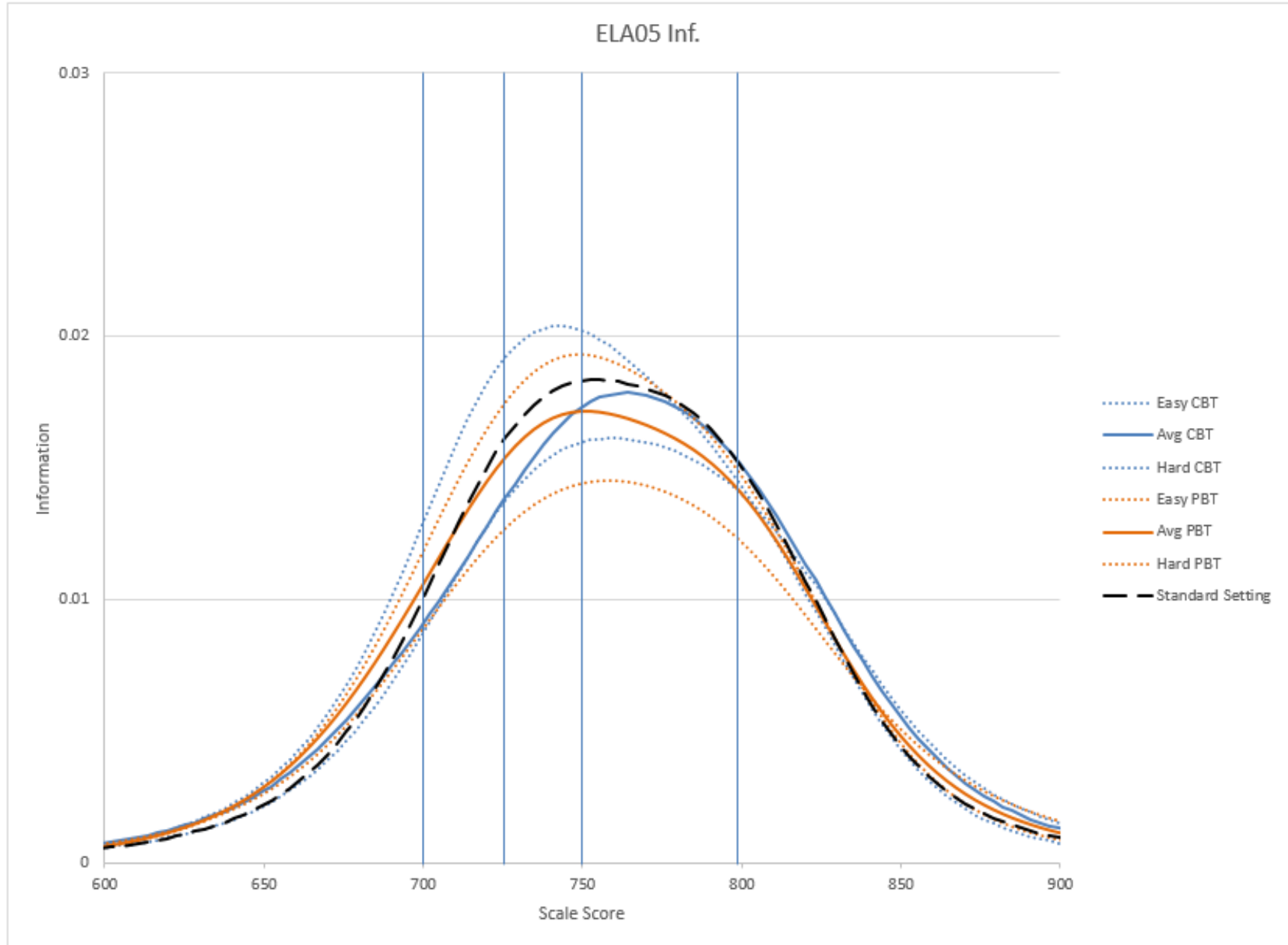


Figure A.12.24 Test Information Curves ELA/L Grade 5

April 18, 2016

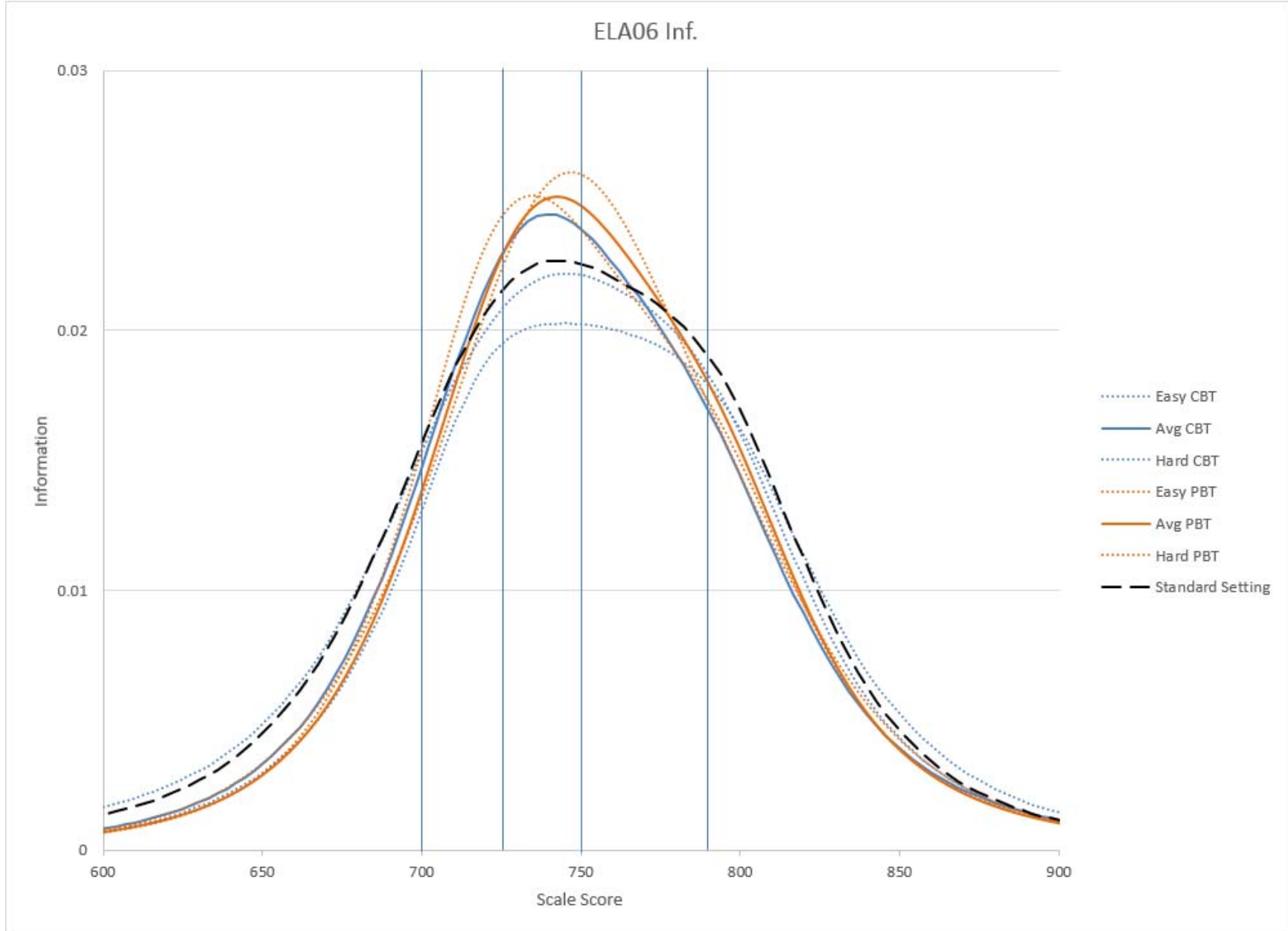


Figure A.12.25 Test Information Curves ELA/L Grade 6

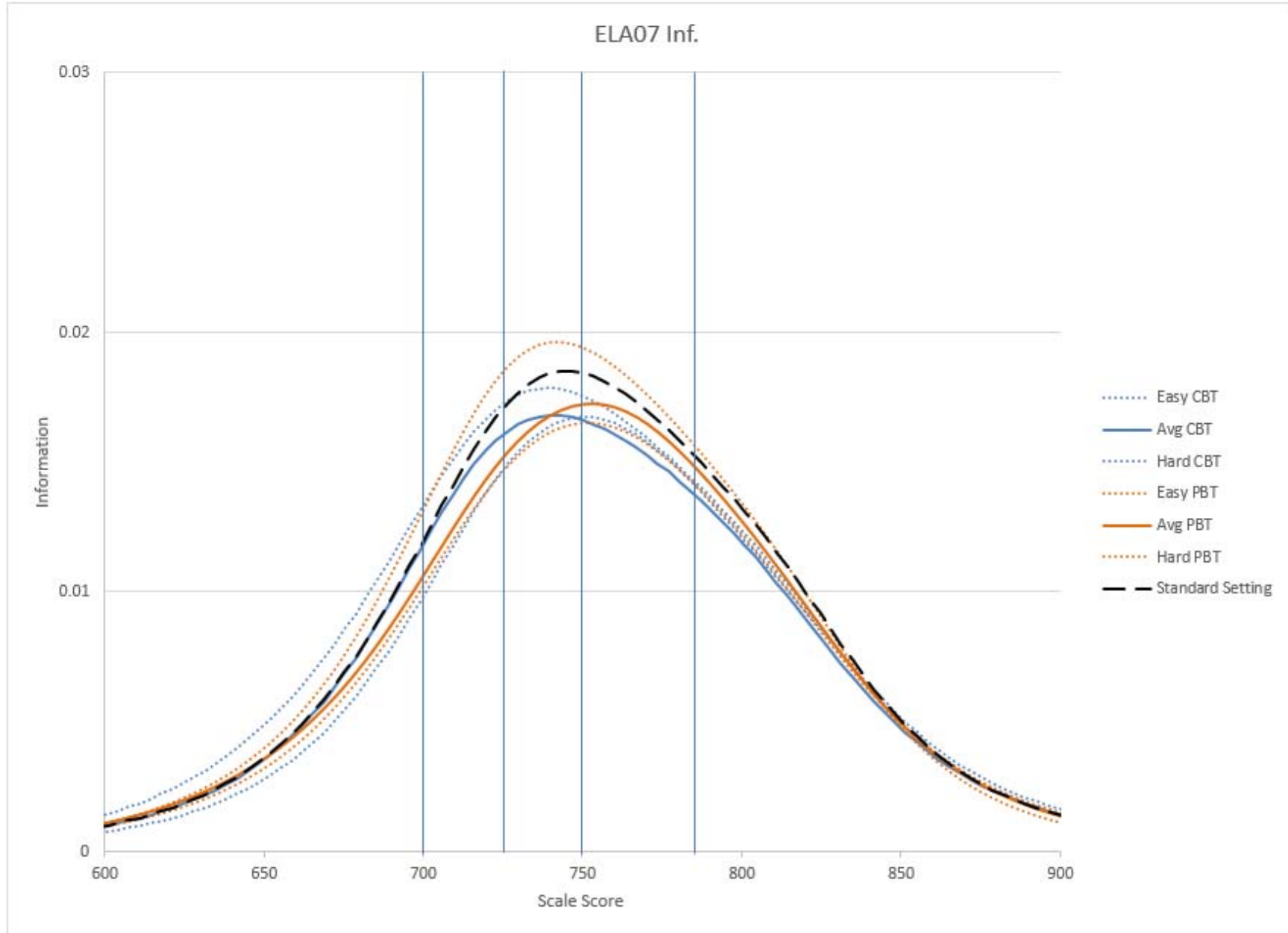


Figure A.12.26 Test Information Curves ELA/L Grade 7

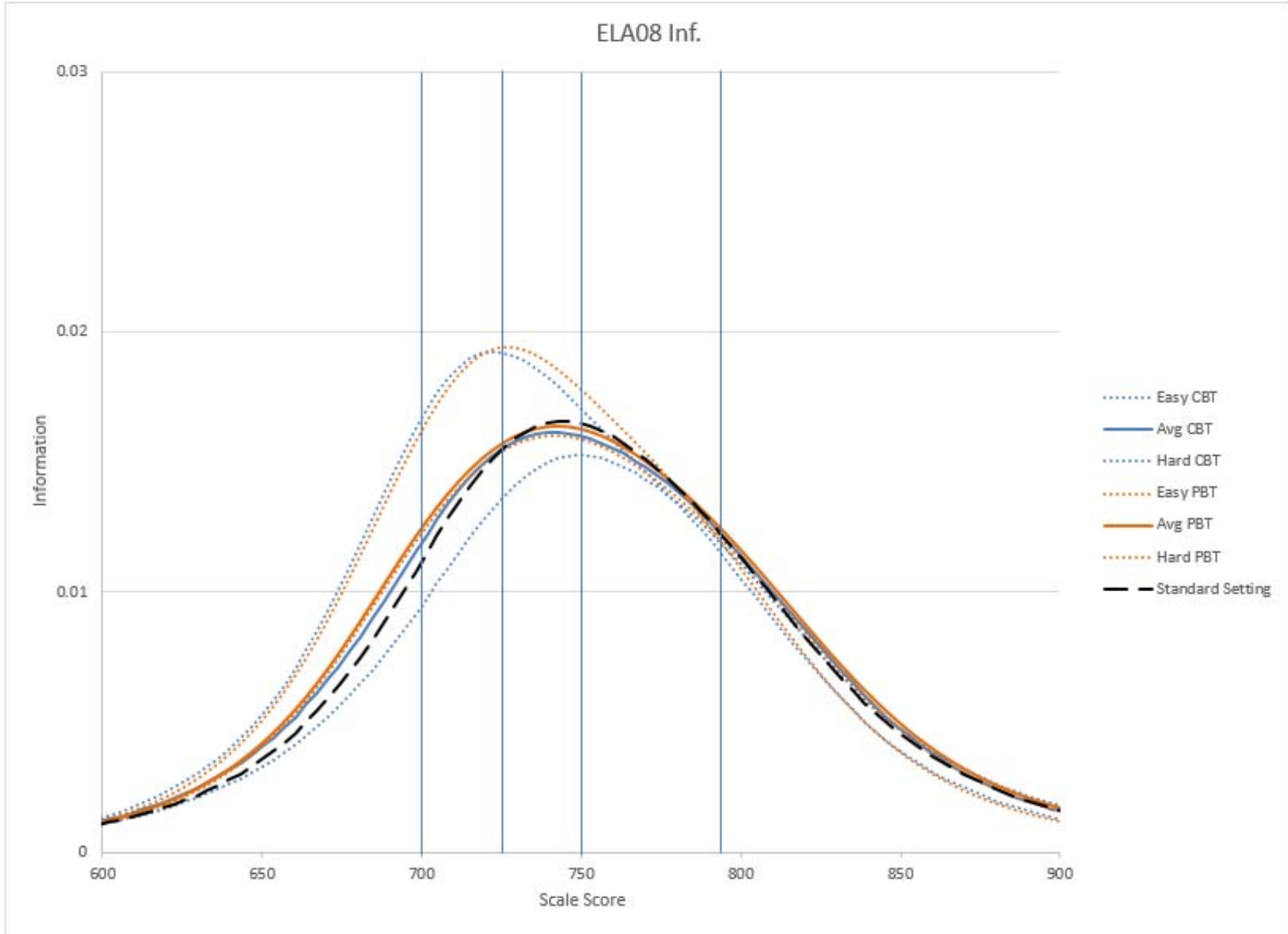


Figure A.12.27 Test Information Curves ELA/L Grade 8

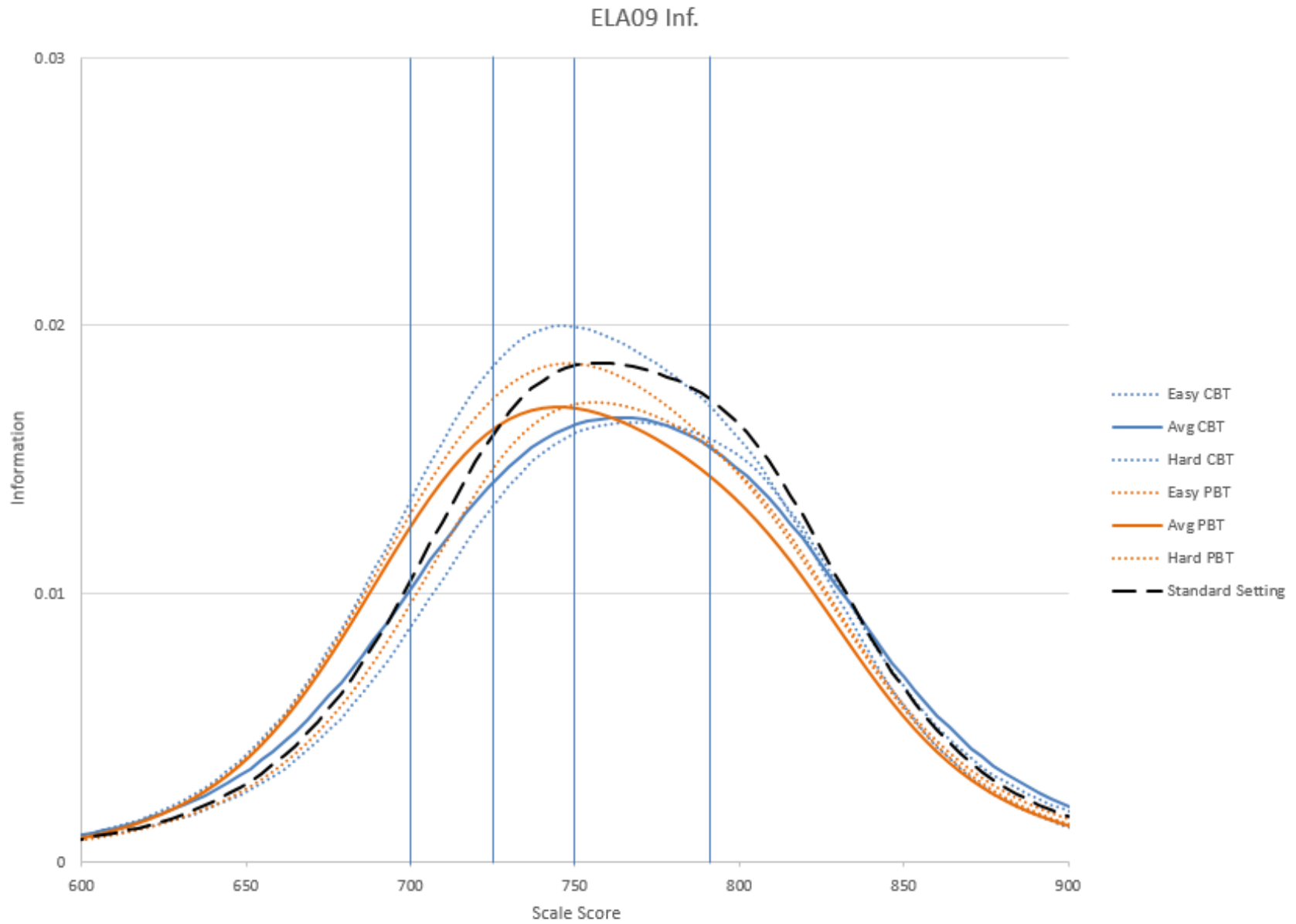


Figure A.12.28 Test Information Curves ELA/L Grade 9

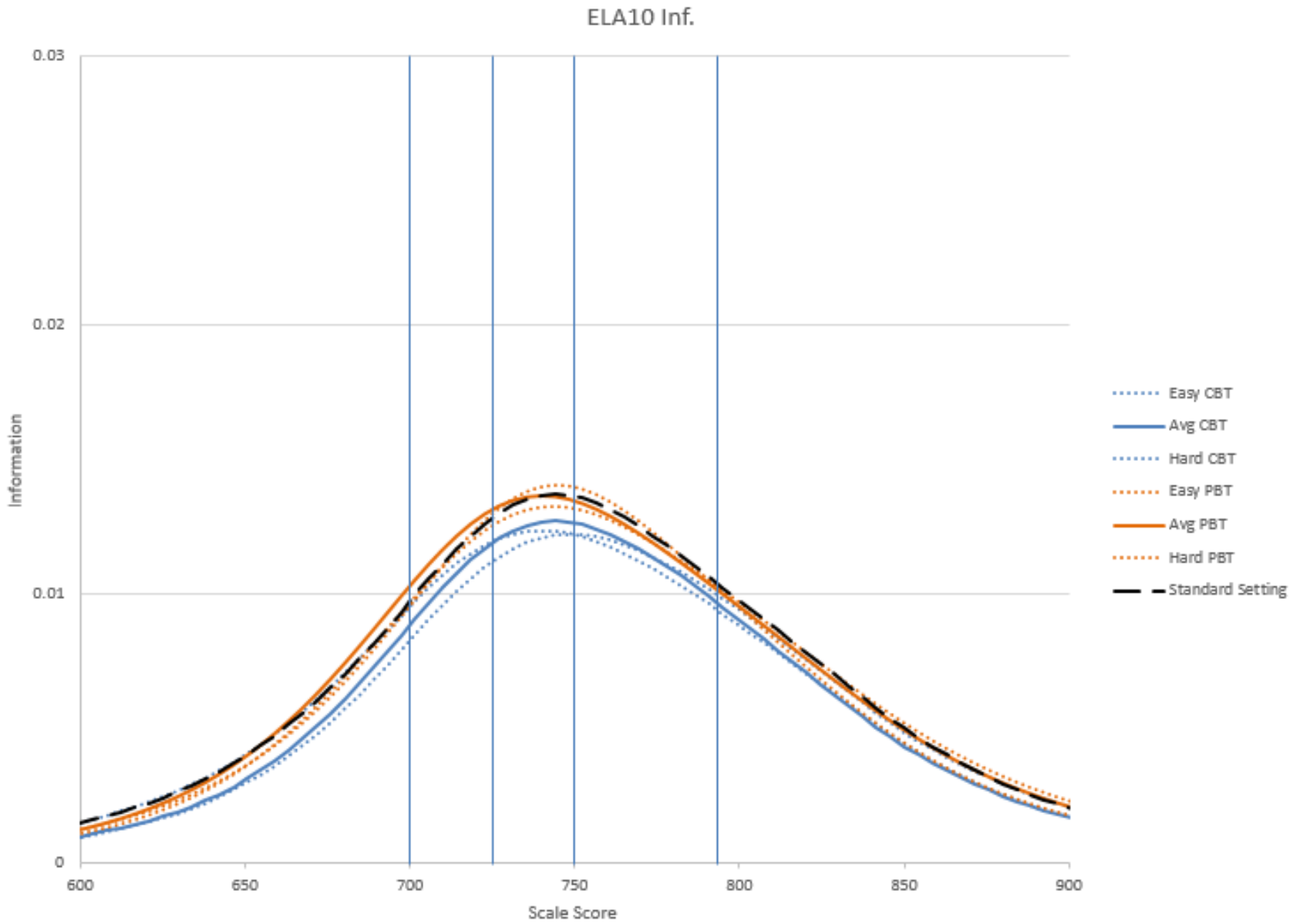


Figure A.12.29 Test Information Curves ELA/L Grade 10

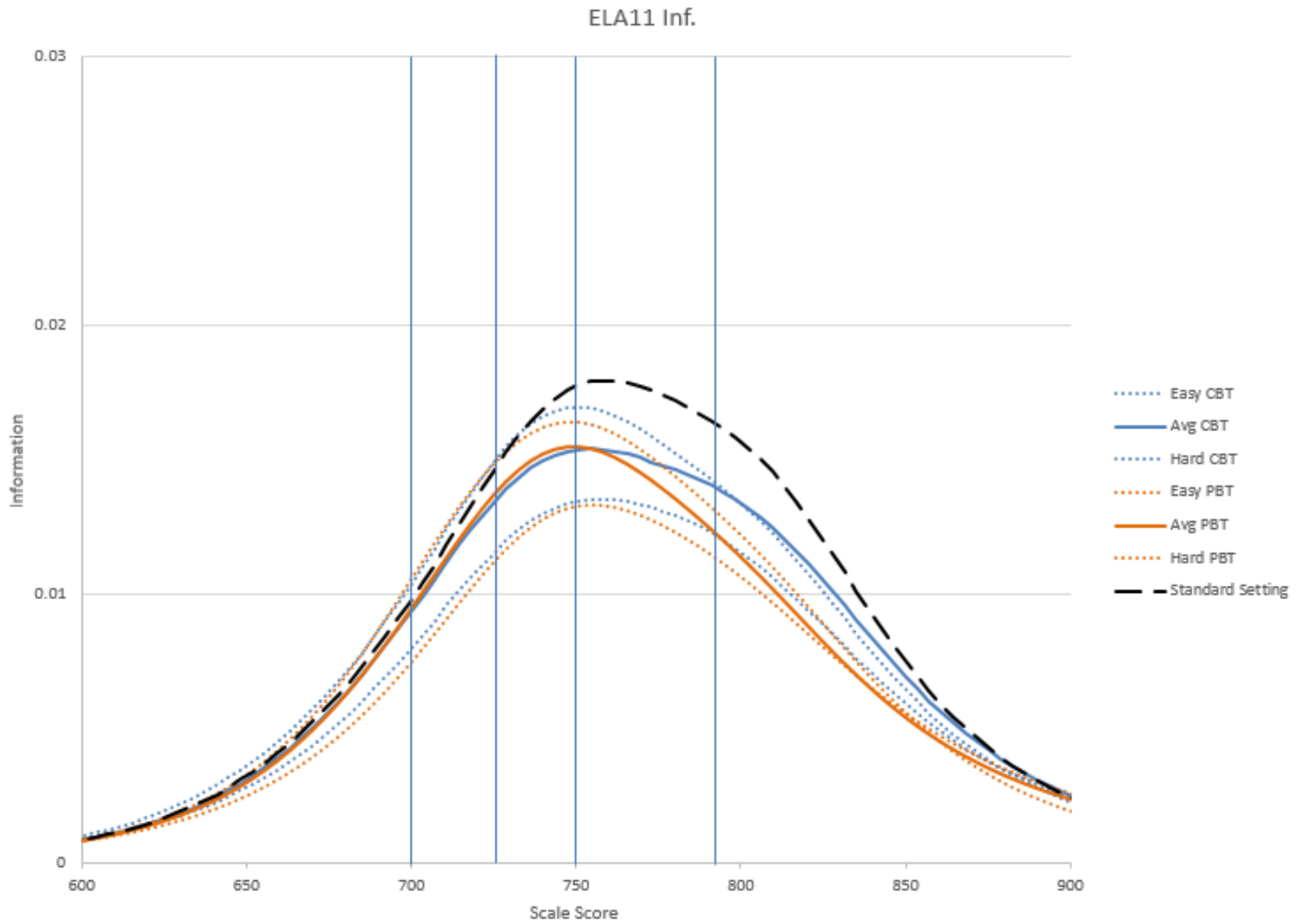


Figure A.12.30 Test Information Curves ELA/L Grade 11

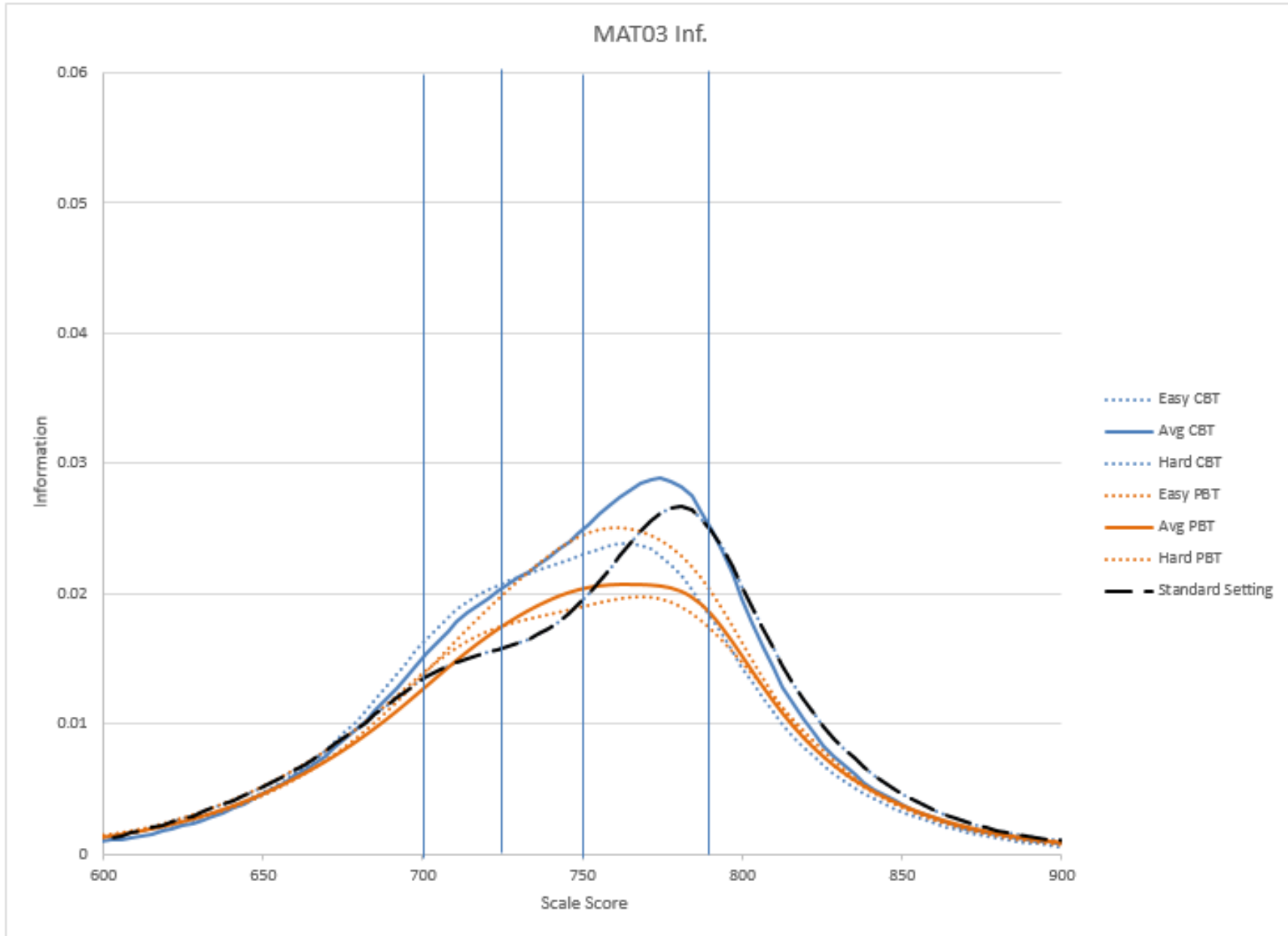


Figure A.12.31 Test Information Curves Mathematics Grade 3

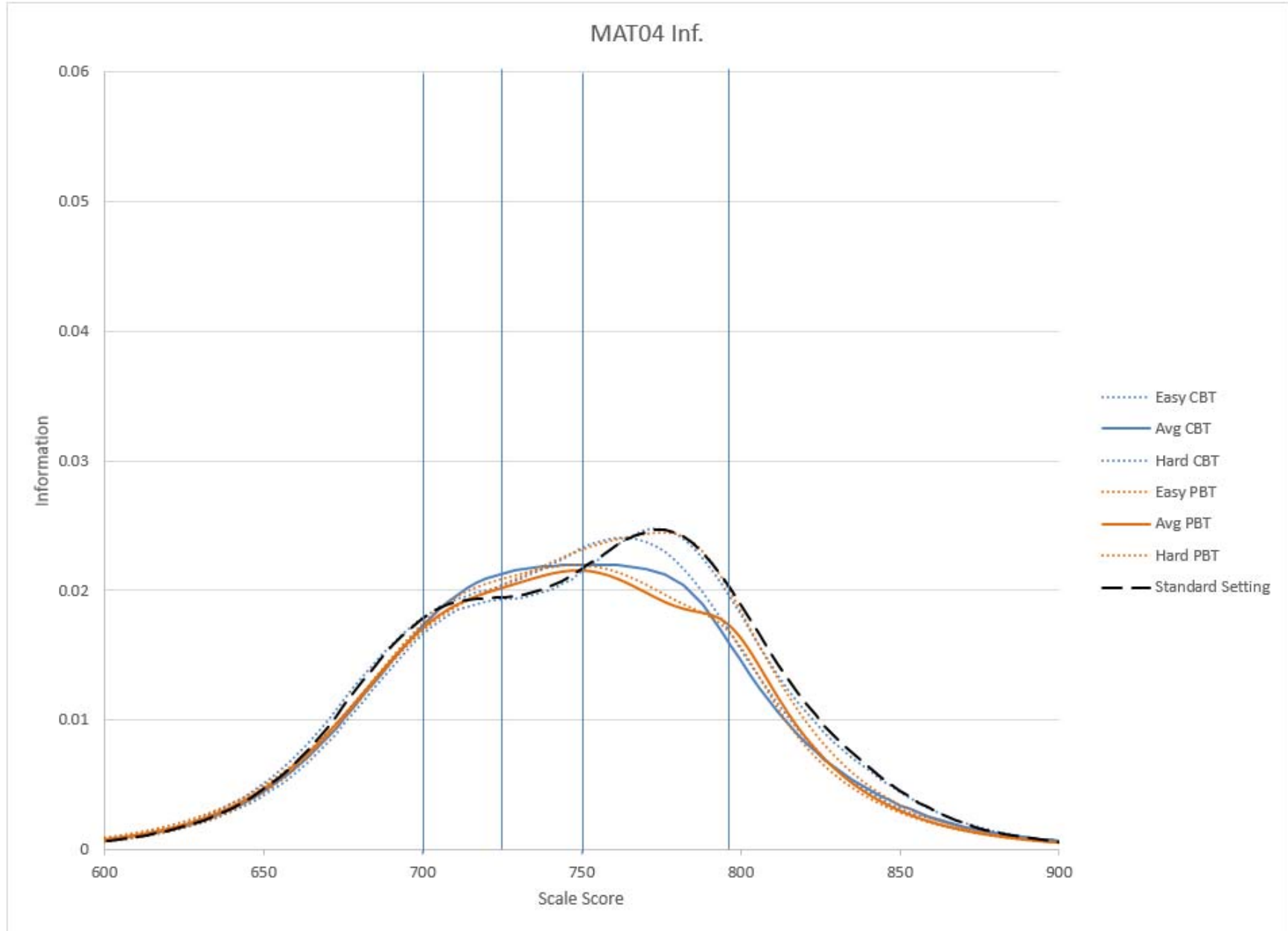


Figure A.12.32 Test Information Curves Mathematics Grade 4

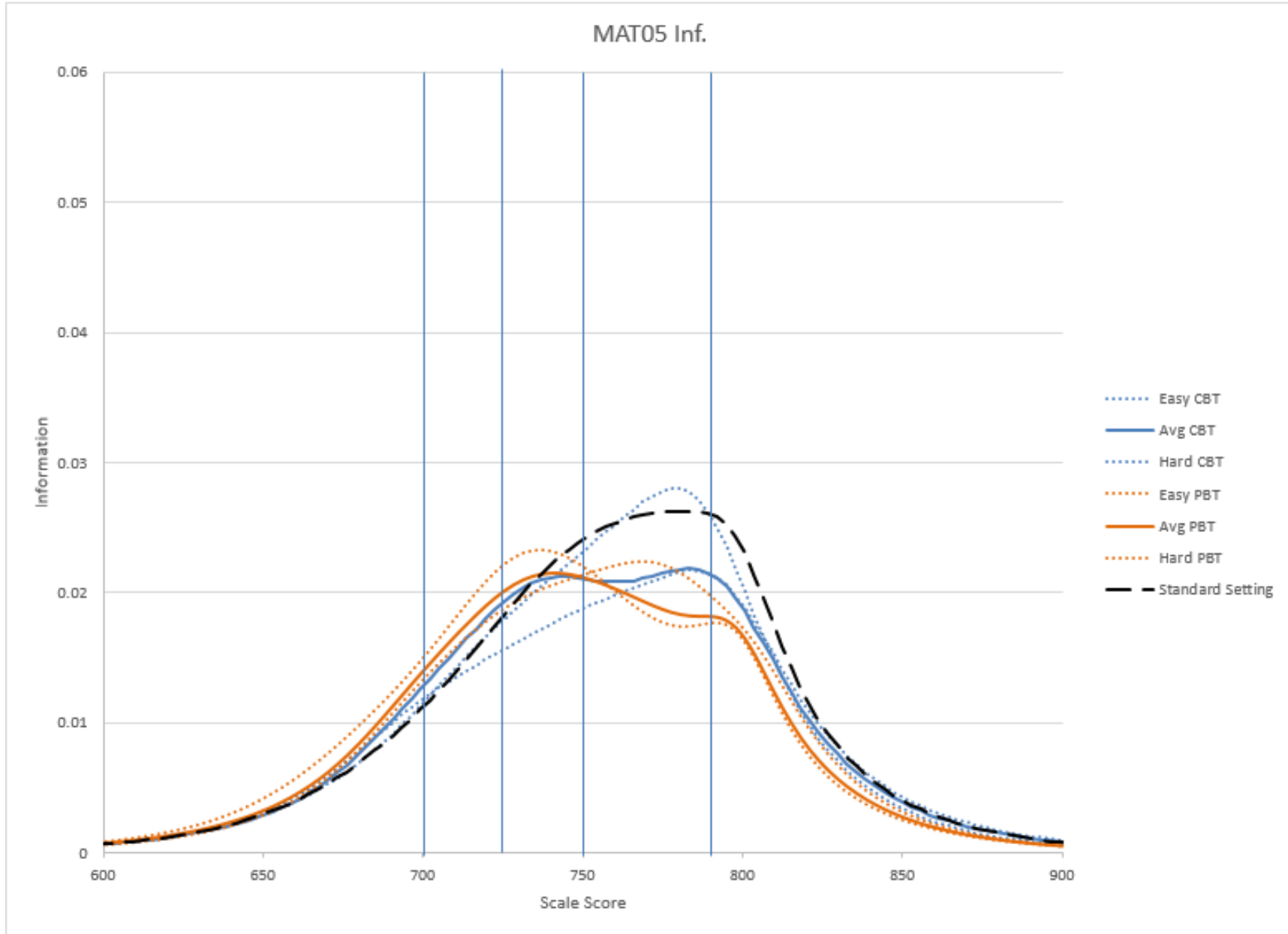


Figure A.12.33 Test Information Curves Mathematics Grade 5

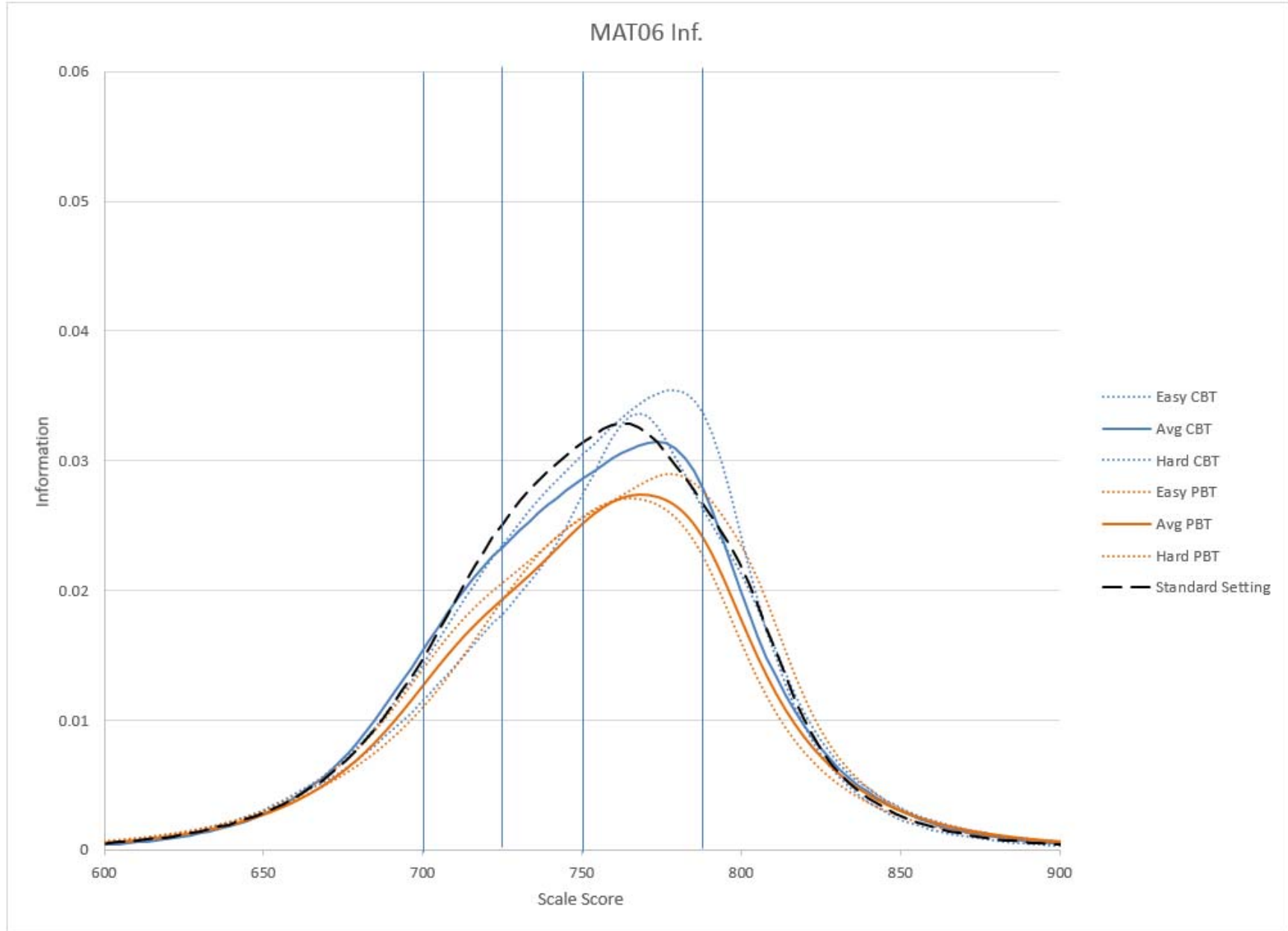


Figure A.12.34 Test Information Curves Mathematics Grade 6

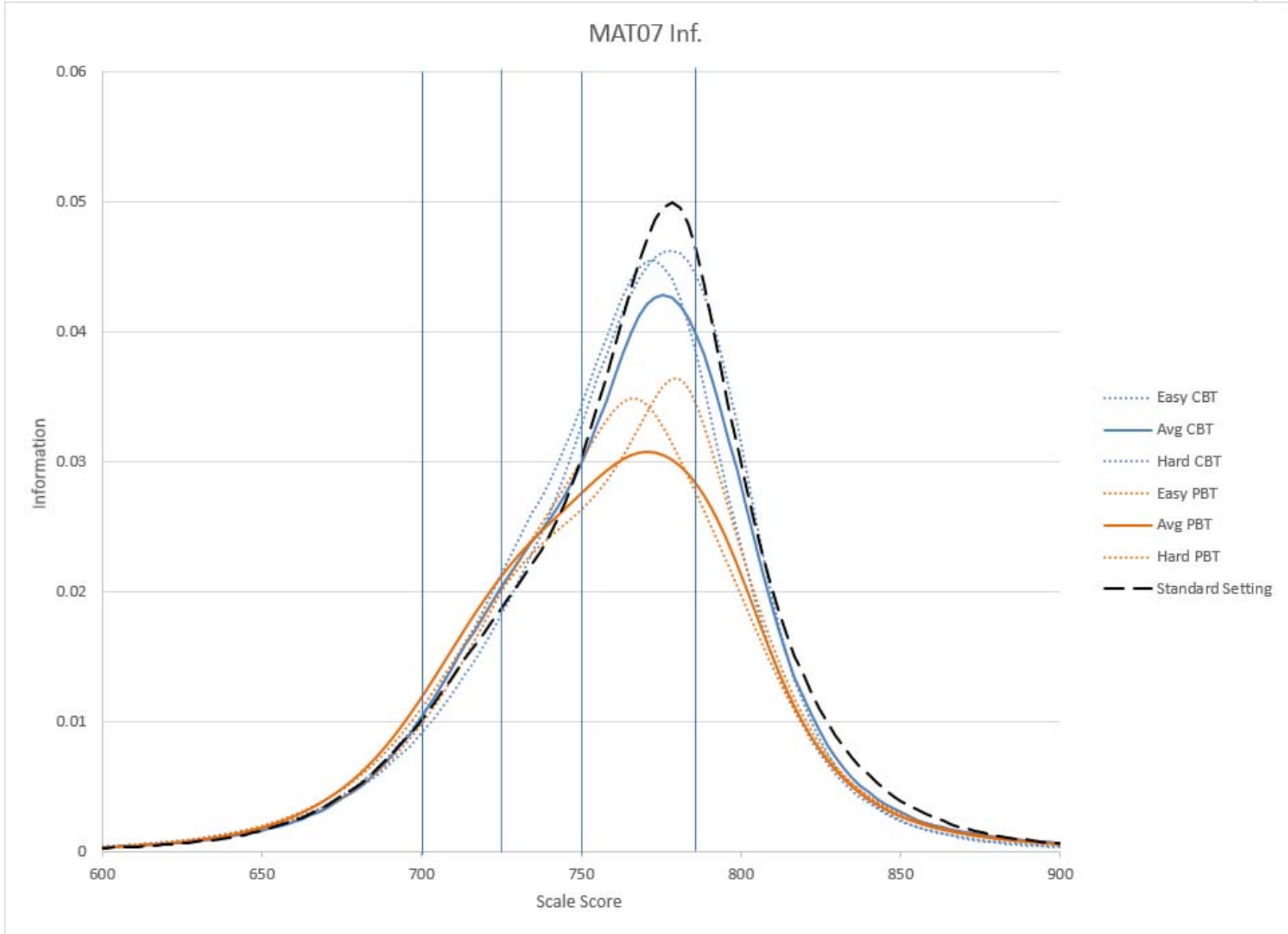


Figure A.12.35 Test Information Curves Mathematics Grade 7

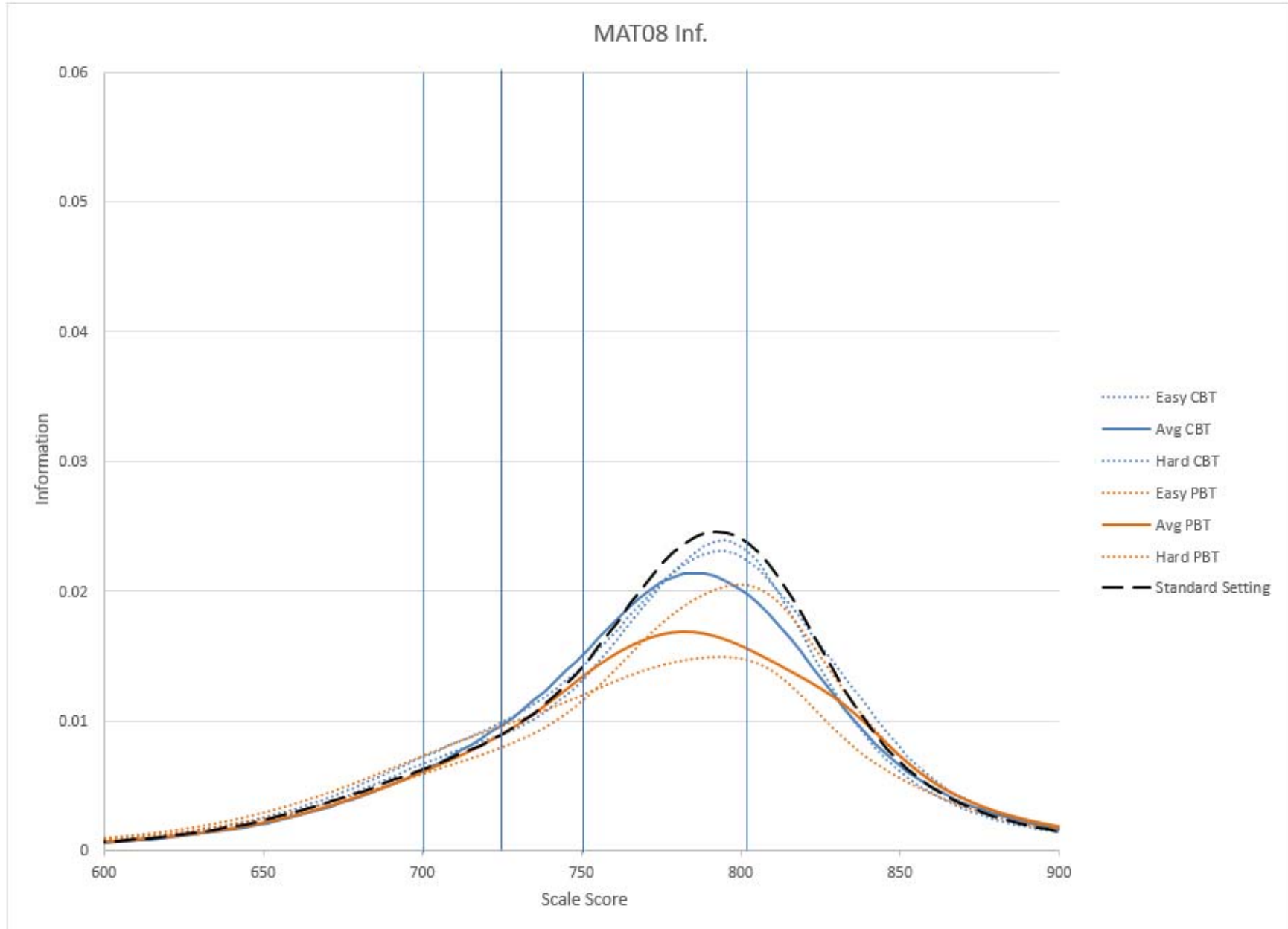


Figure A.12.36 Test Information Curves Mathematics Grade 8

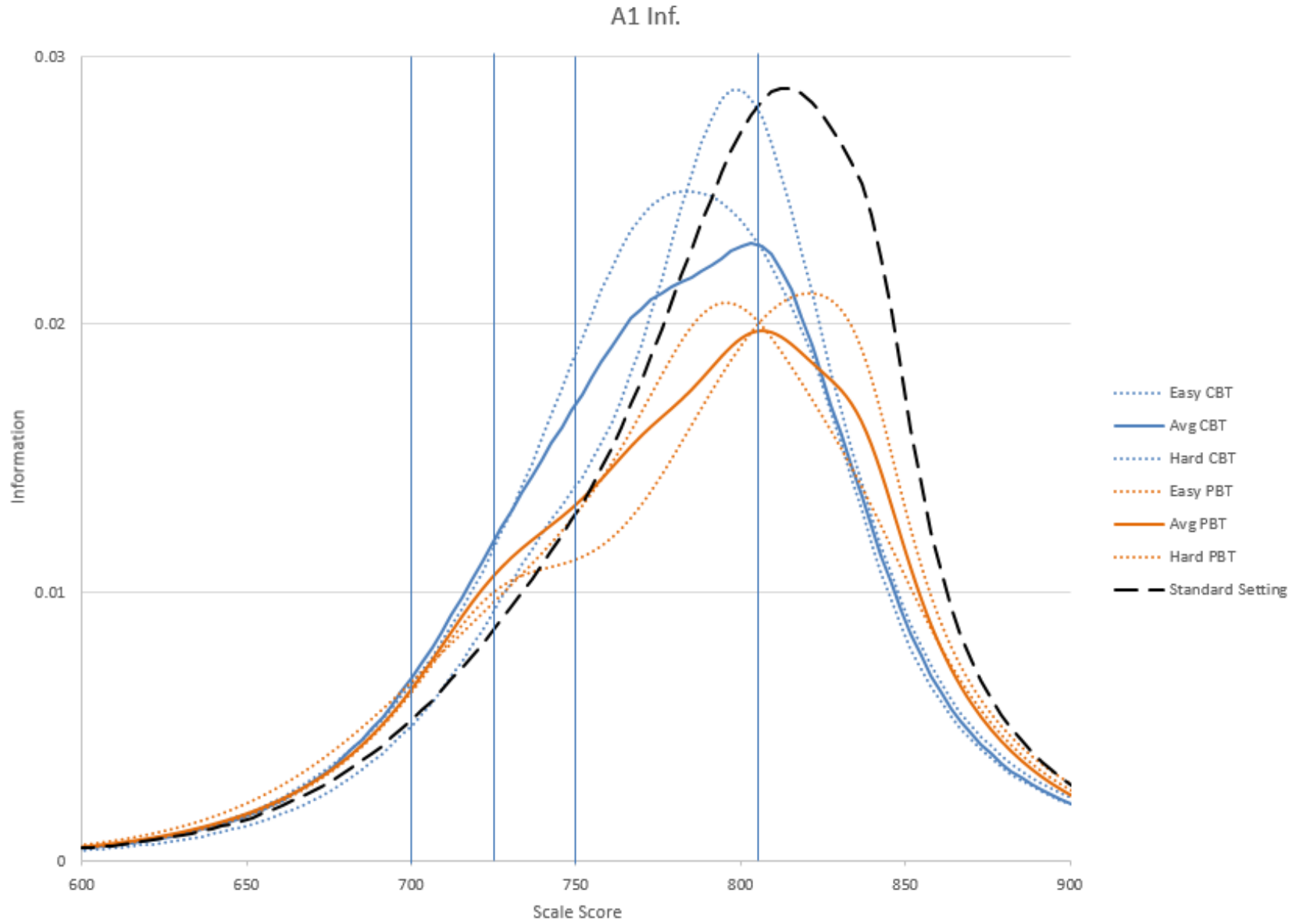


Figure A.12.37 Test Information Curves Algebra I

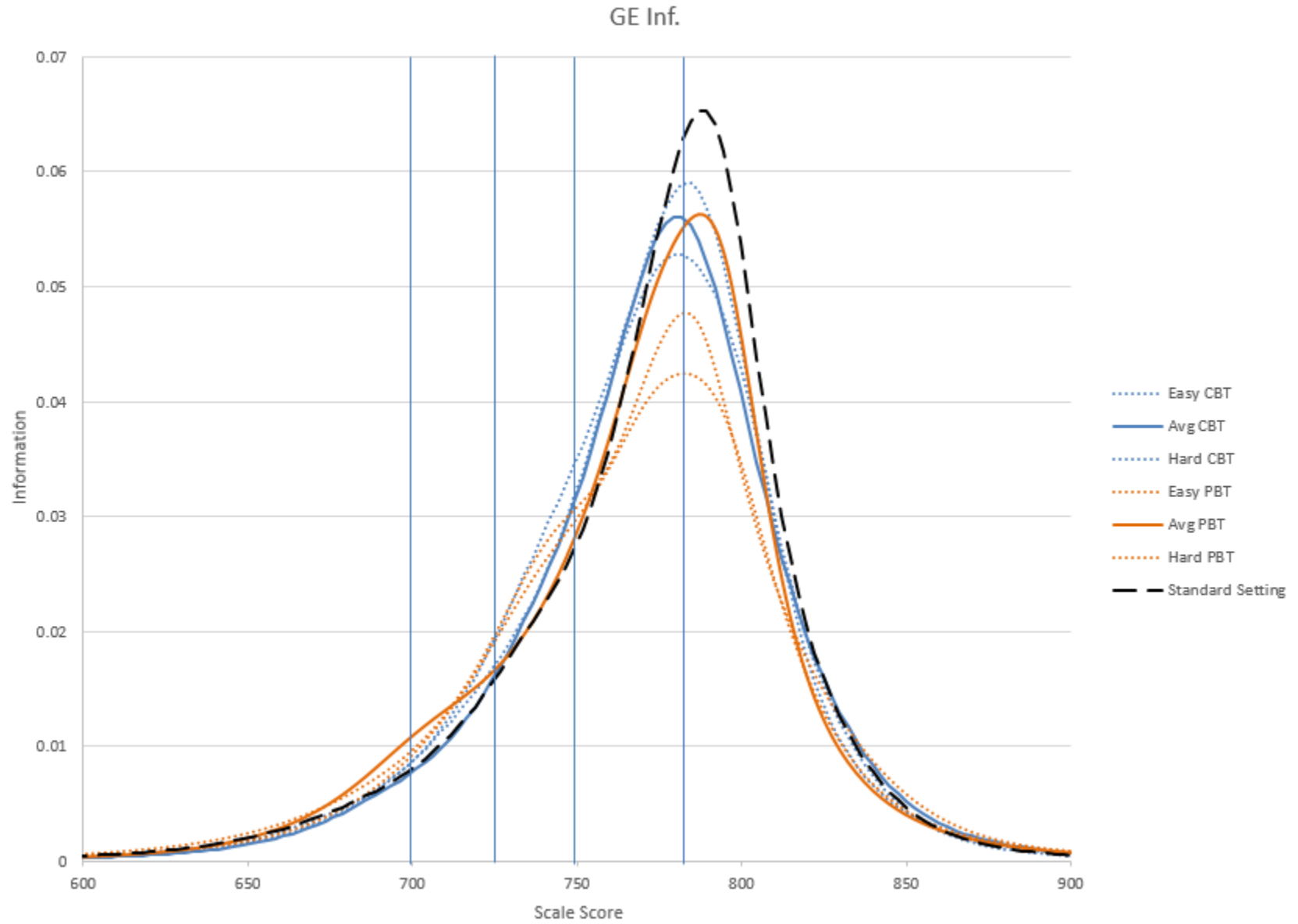


Figure A.12.38 Test Information Curves Geometry

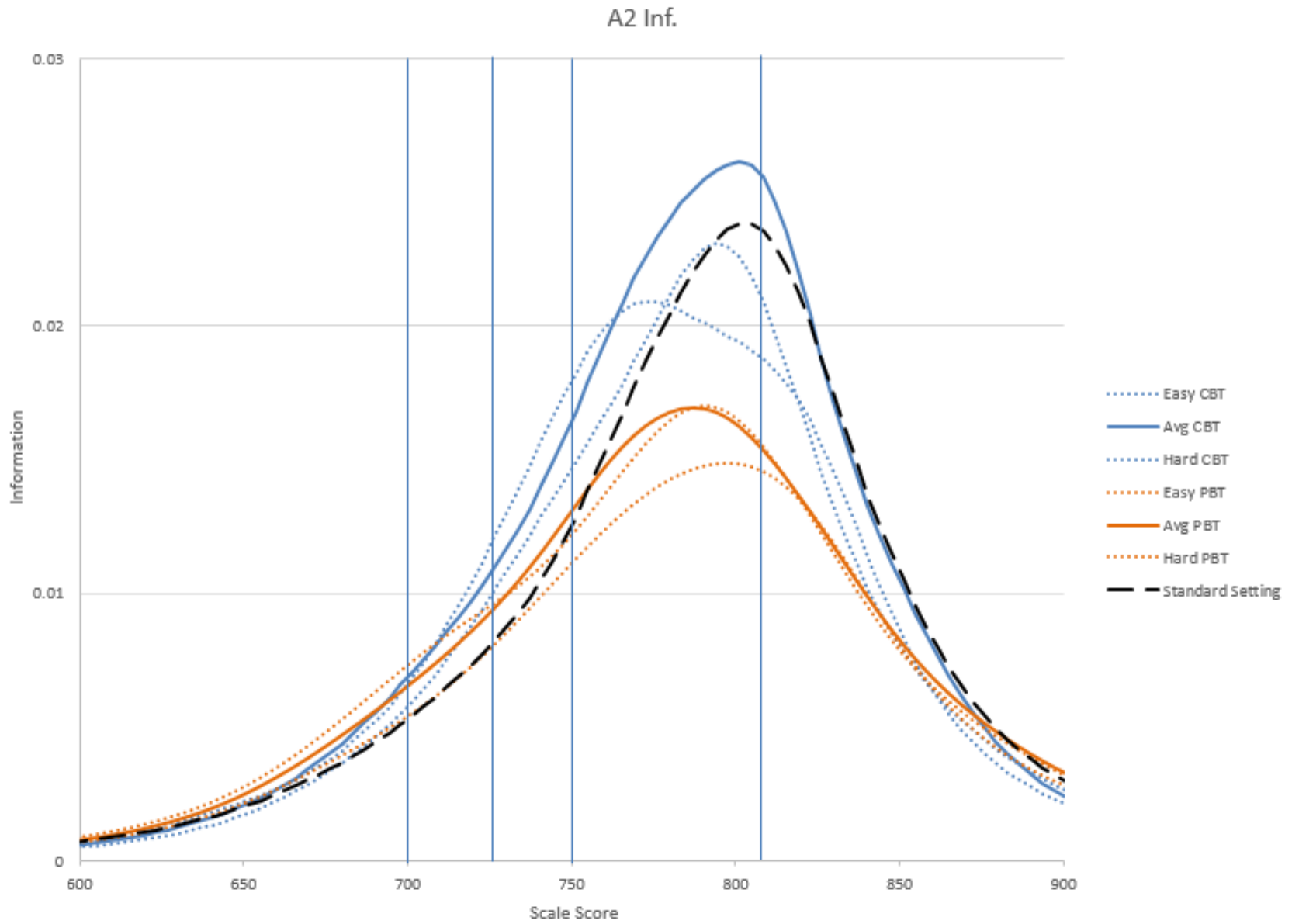


Figure A.12.39 Test Information Curves Algebra II

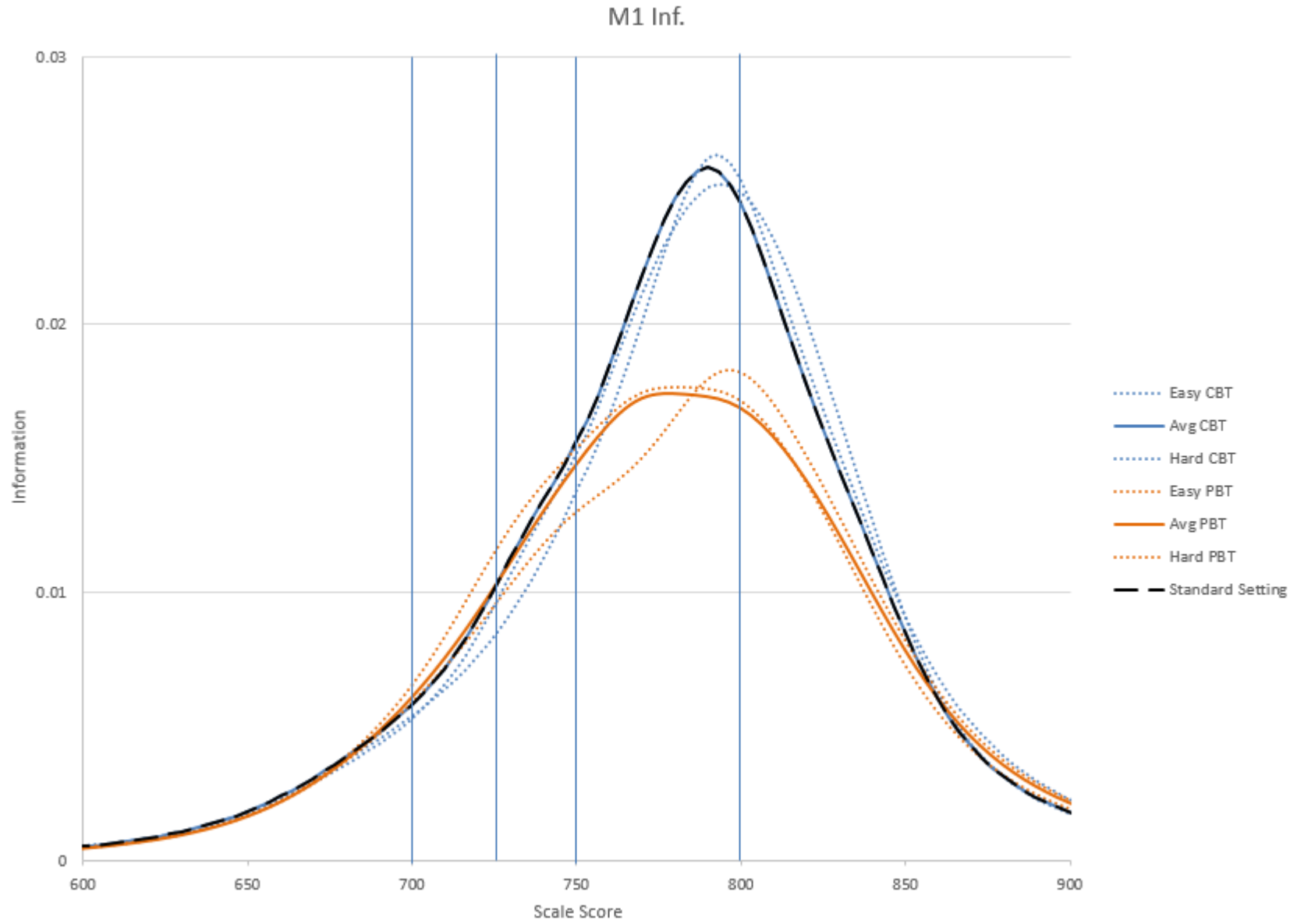


Figure A.12.40 Test Information Curves Integrated Mathematics I

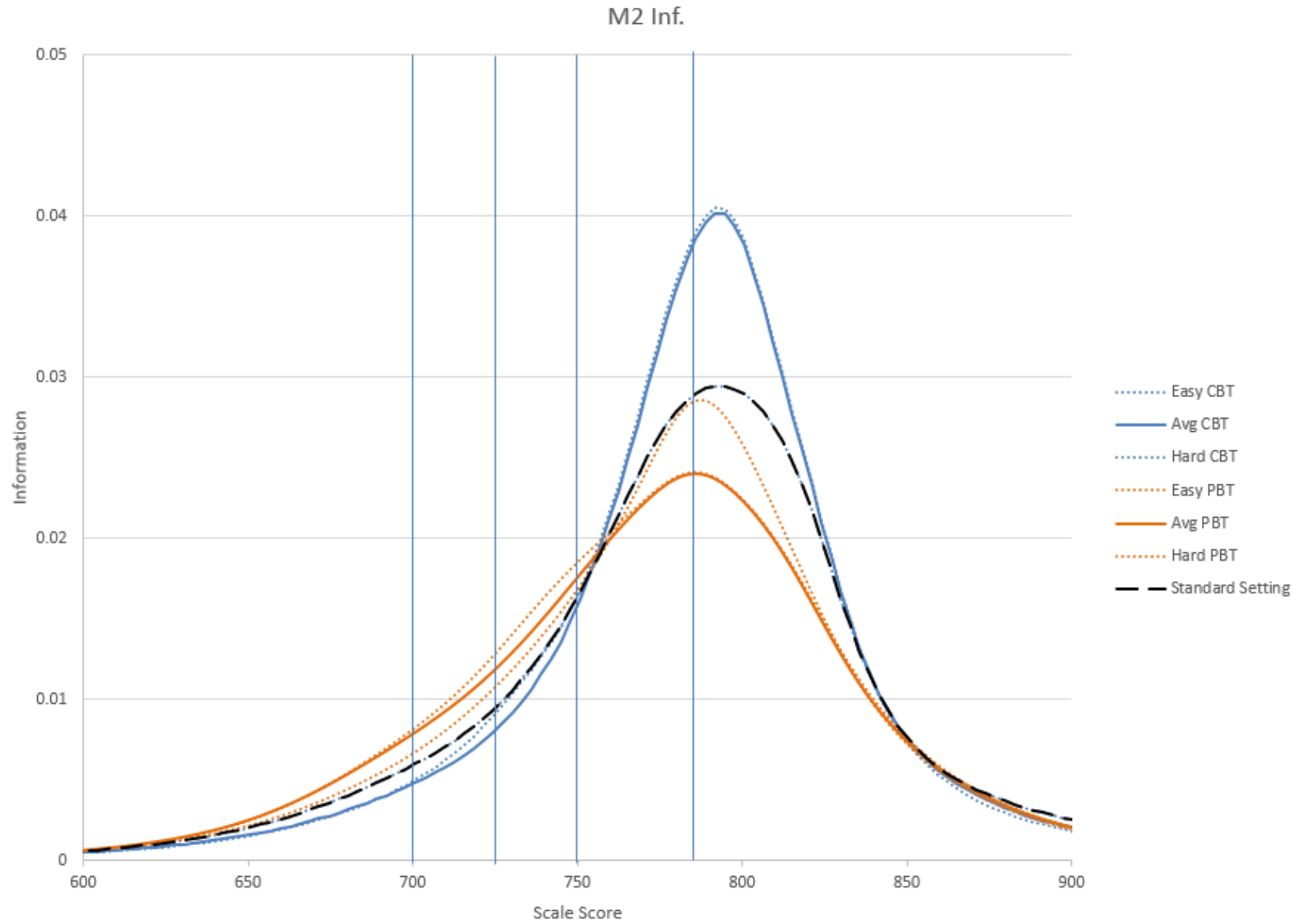


Figure A.12.41 Test Information Curves Integrated Mathematics II

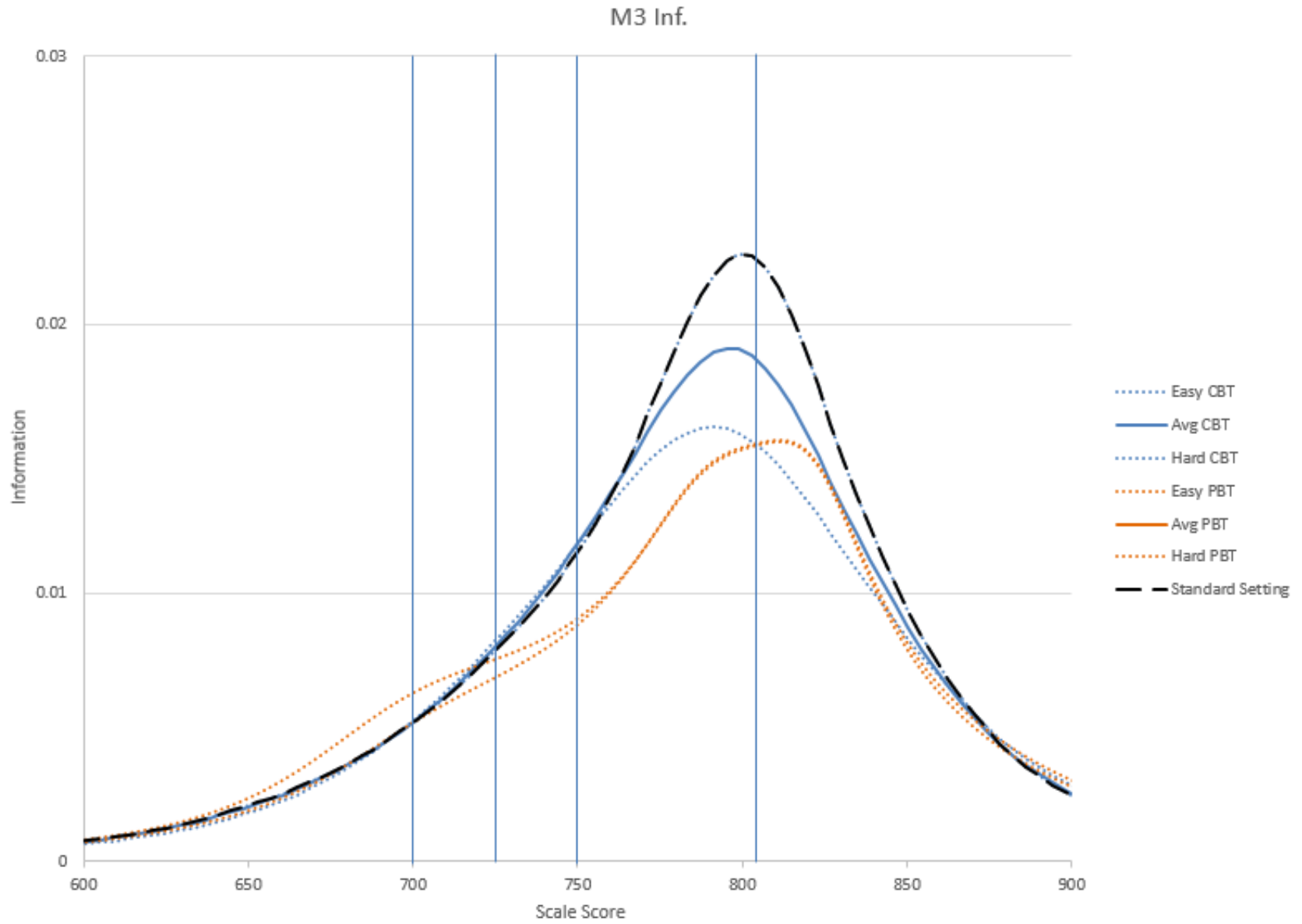


Figure A.12.42 Test Information Curves Integrated Mathematics III

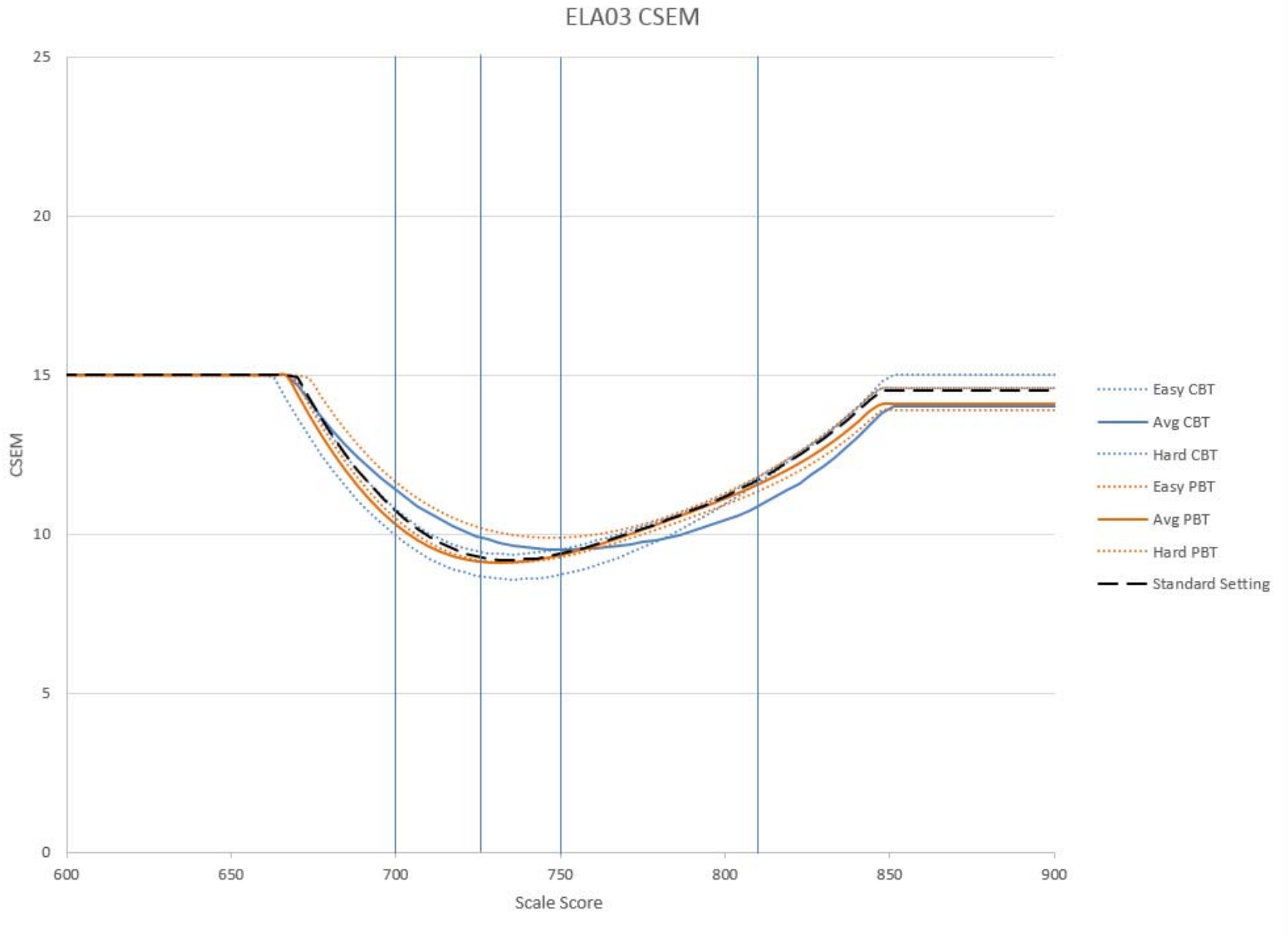


Figure A.12.43 CSEM Curves ELA/L Grade 3

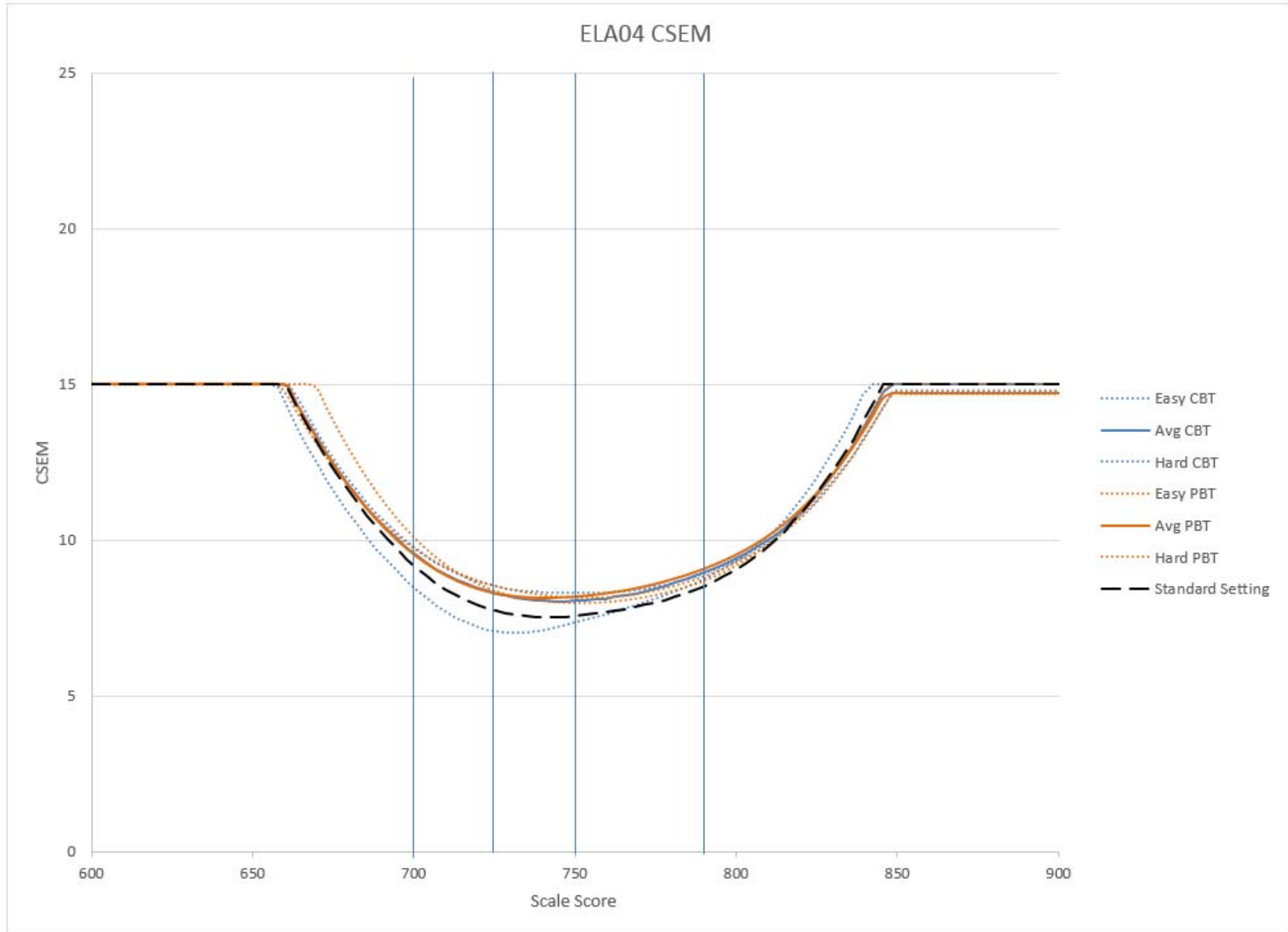


Figure A.12.44 CSEM Curves ELA/L Grade 4

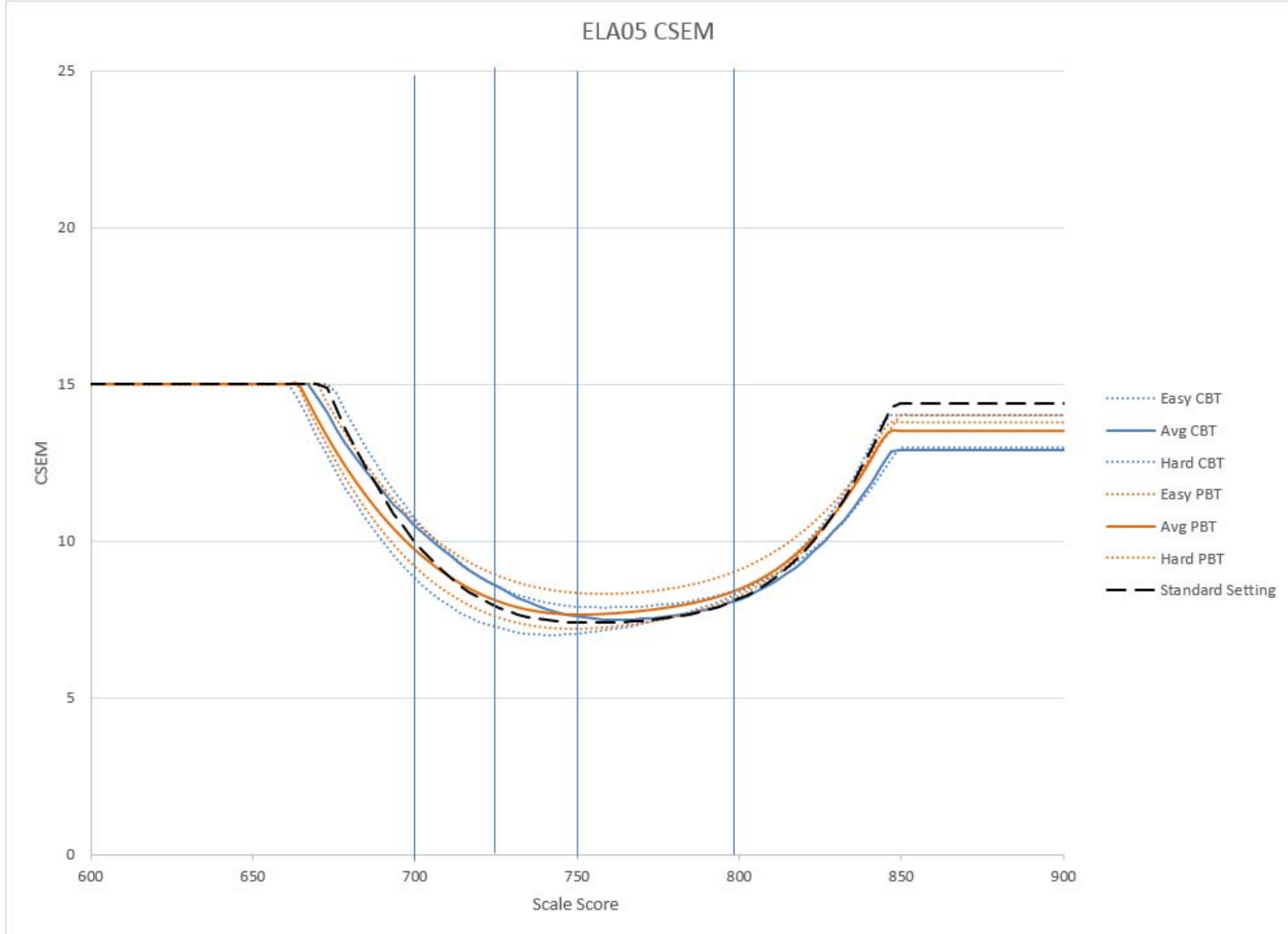


Figure A.12.45 CSEM Curves ELA/L Grade 5

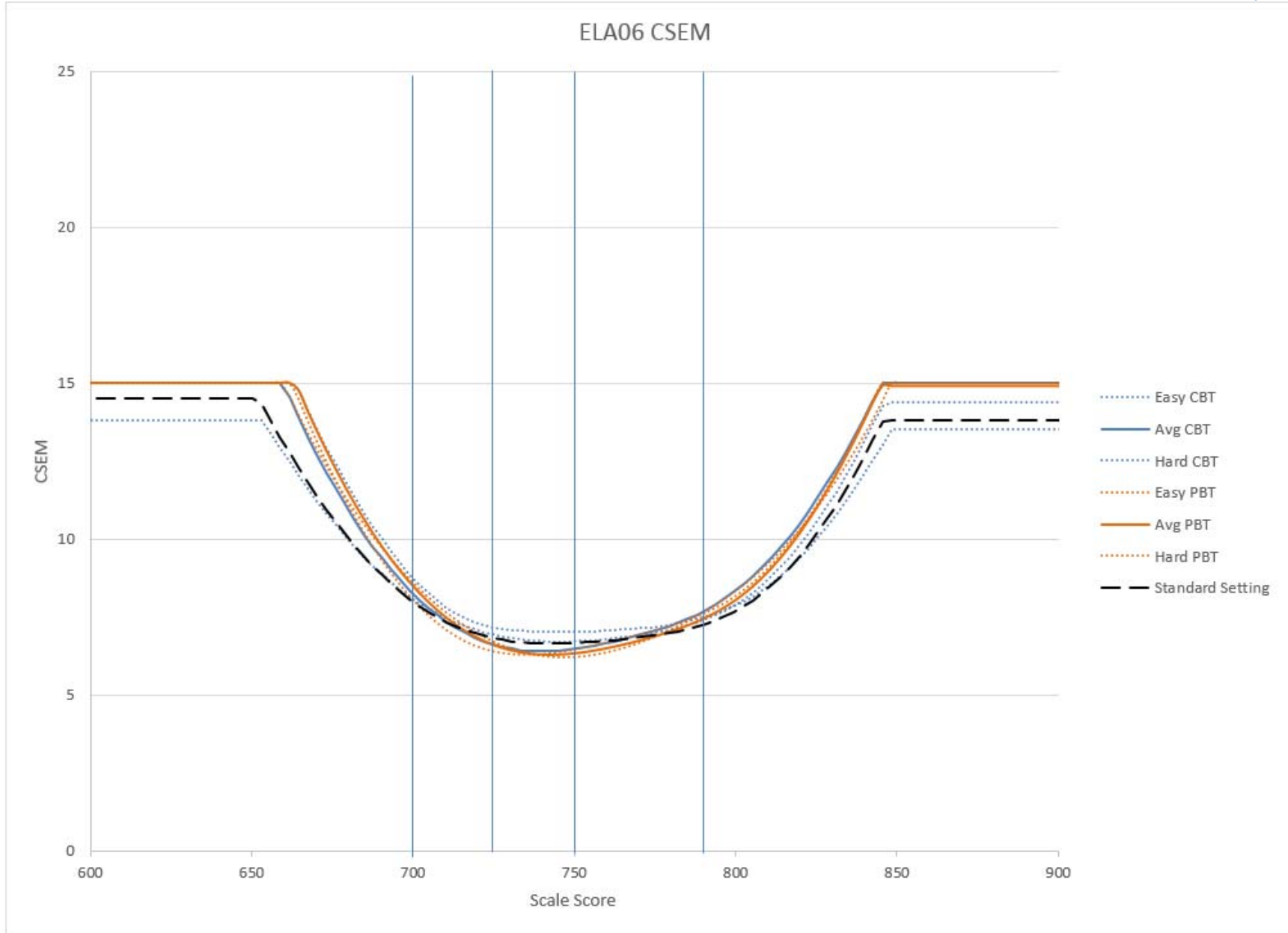


Figure A.12.46 CSEM Curves ELA/L Grade 6

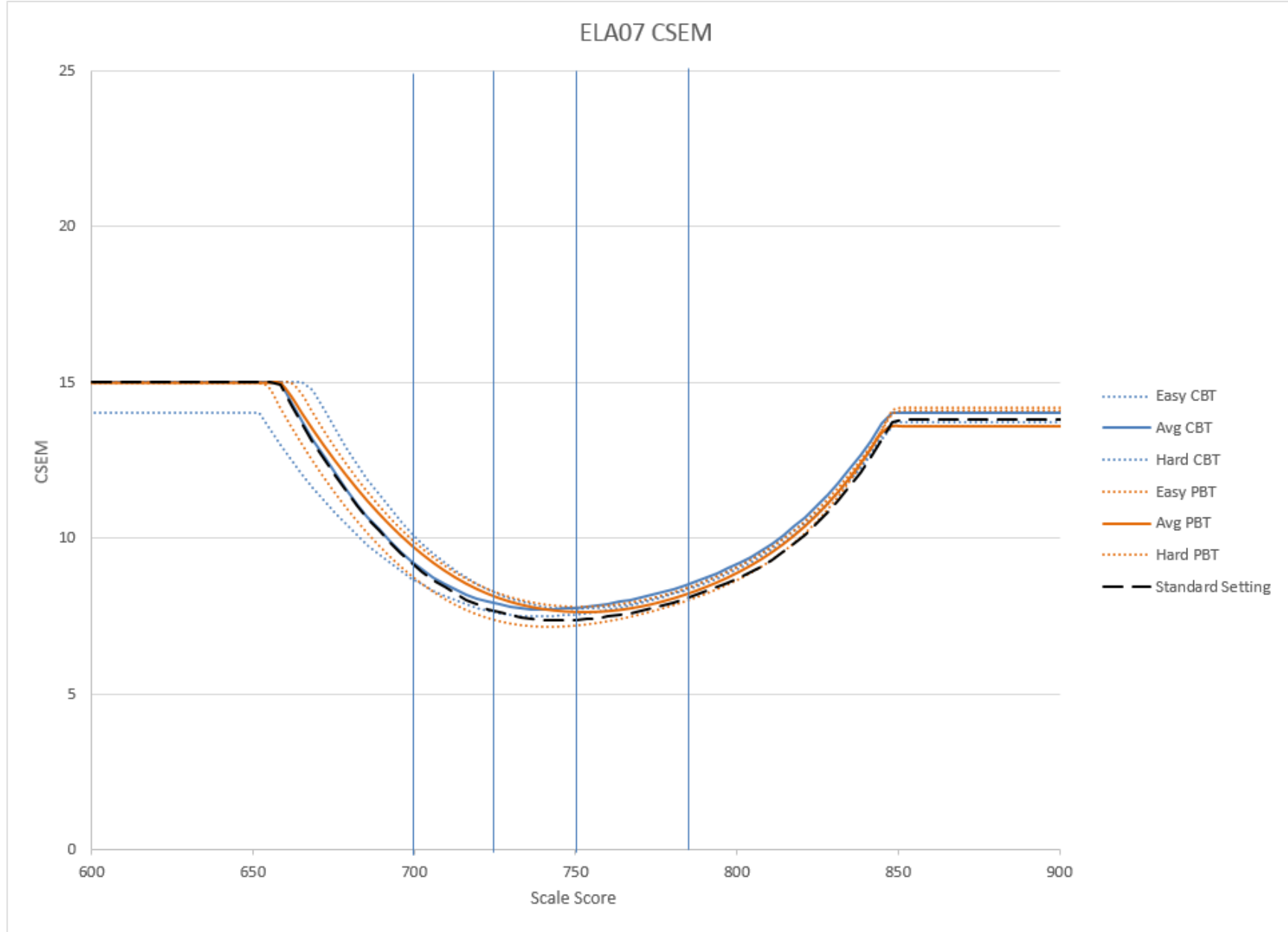


Figure A.12.47 CSEM Curves ELA/L Grade 7

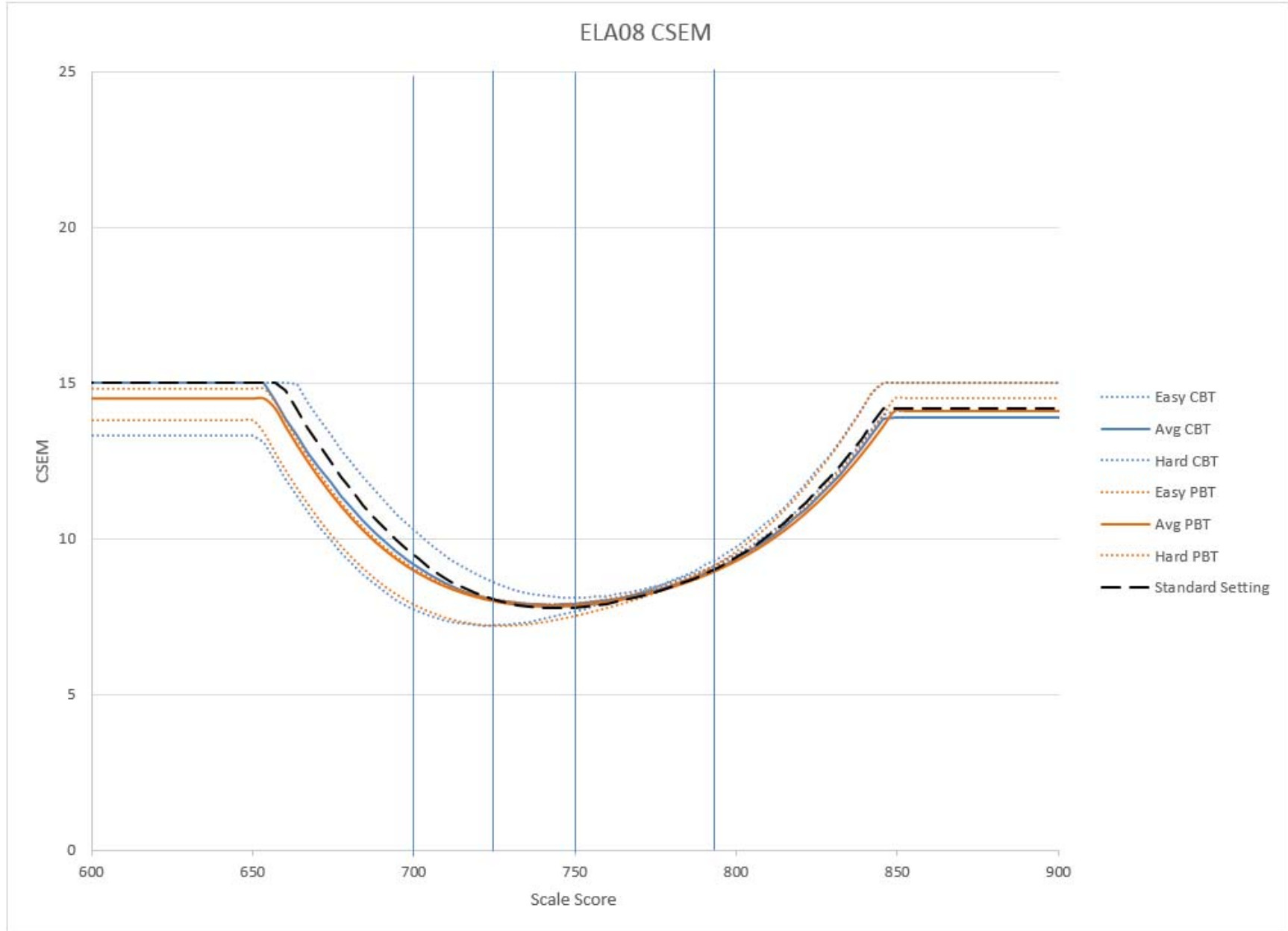


Figure A.12.48 CSEM Curves ELA/L Grade 8

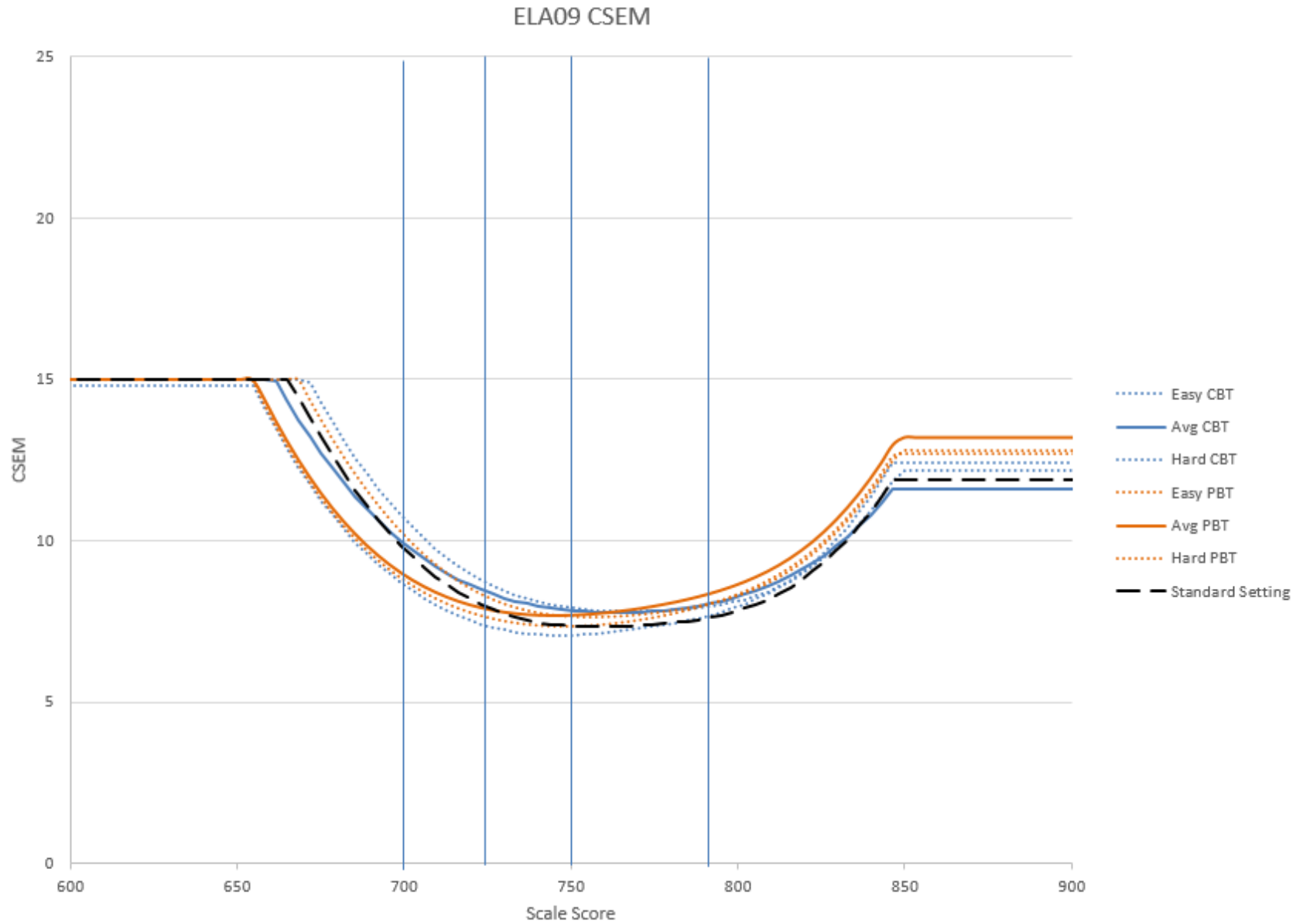


Figure A.12.49 CSEM Curves ELA/L Grade 9

April 18, 2016

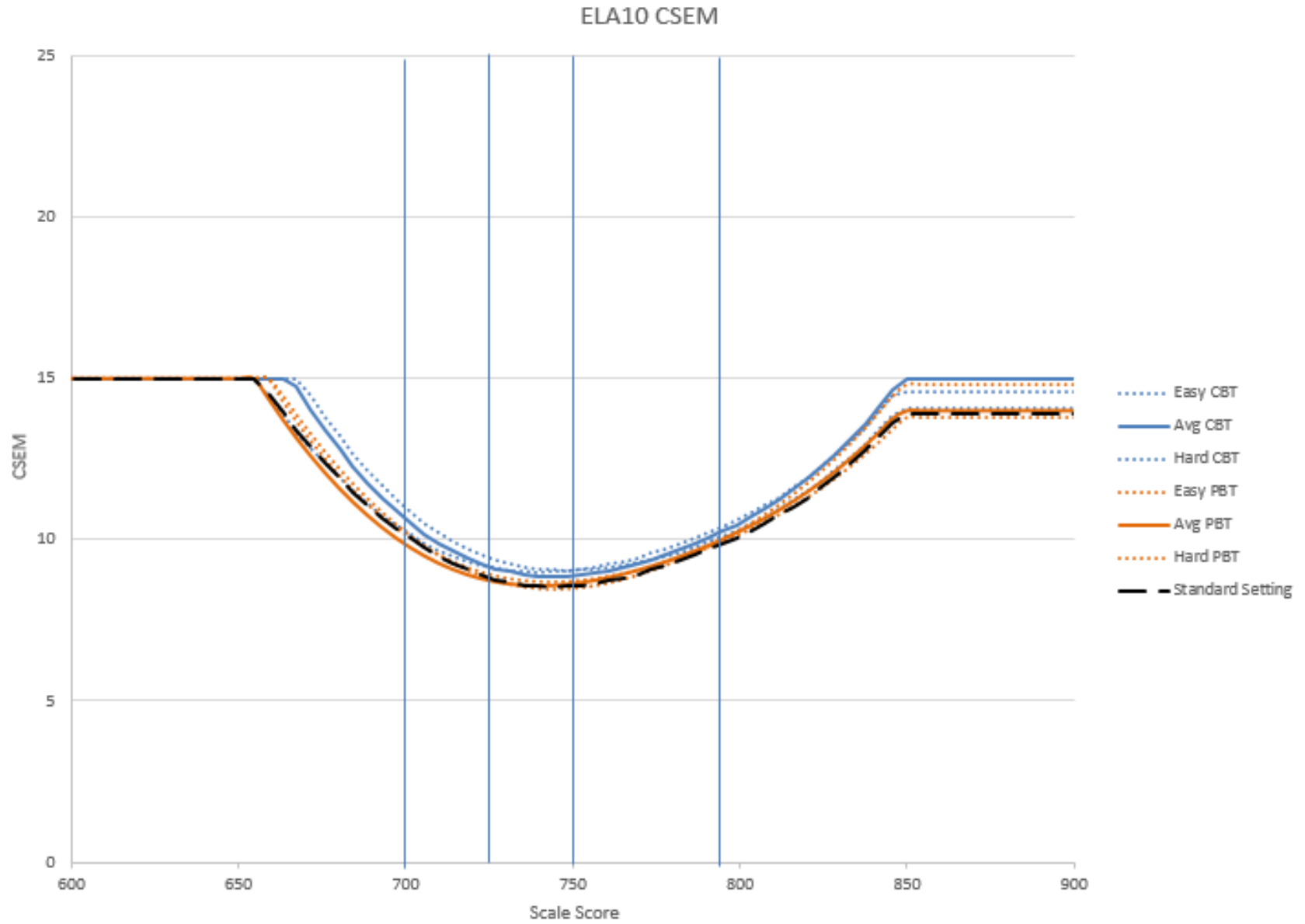


Figure A.12.50 CSEM Curves ELA/L Grade 10

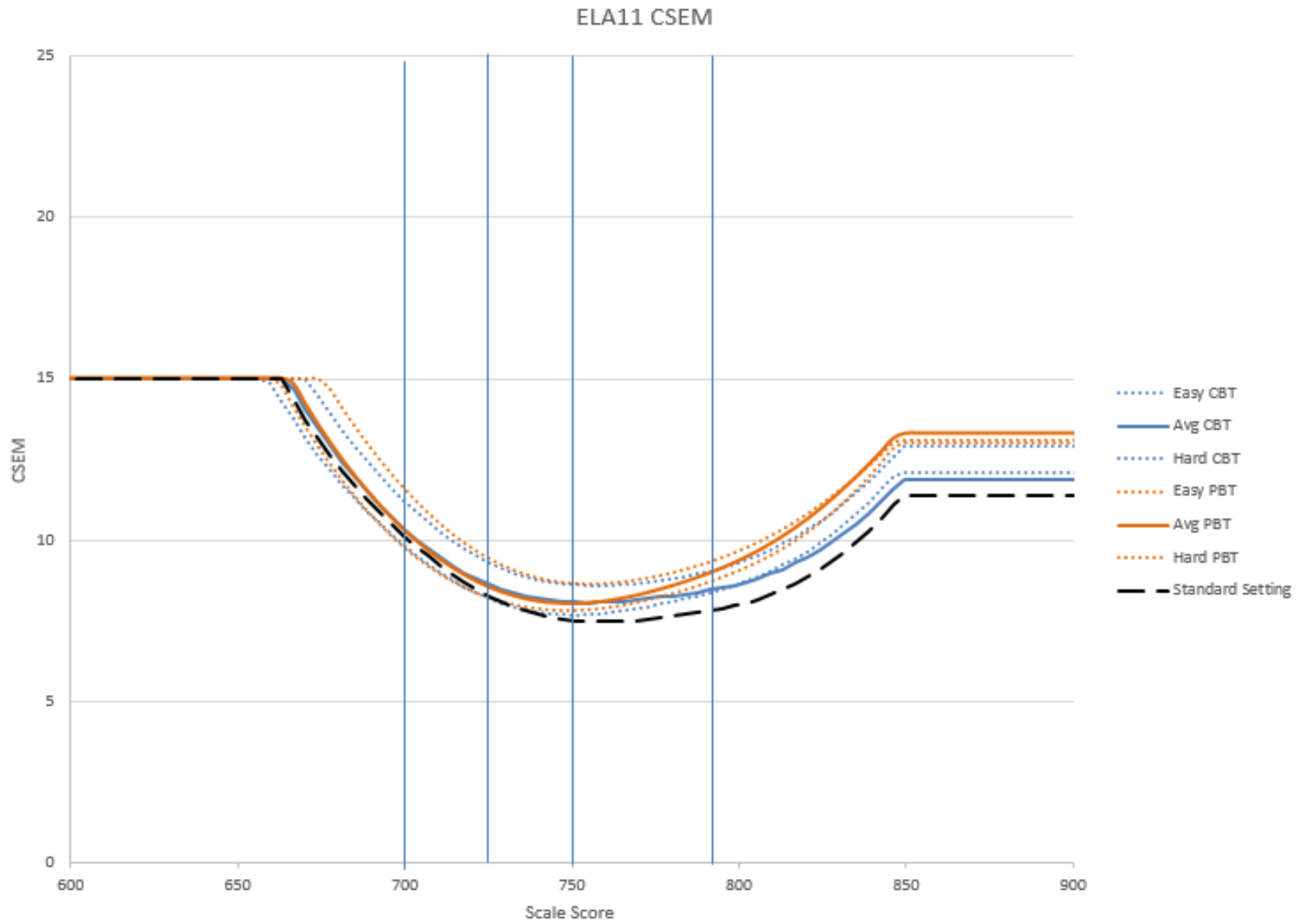


Figure A.12.51 CSEM Curves ELA/L Grade 11

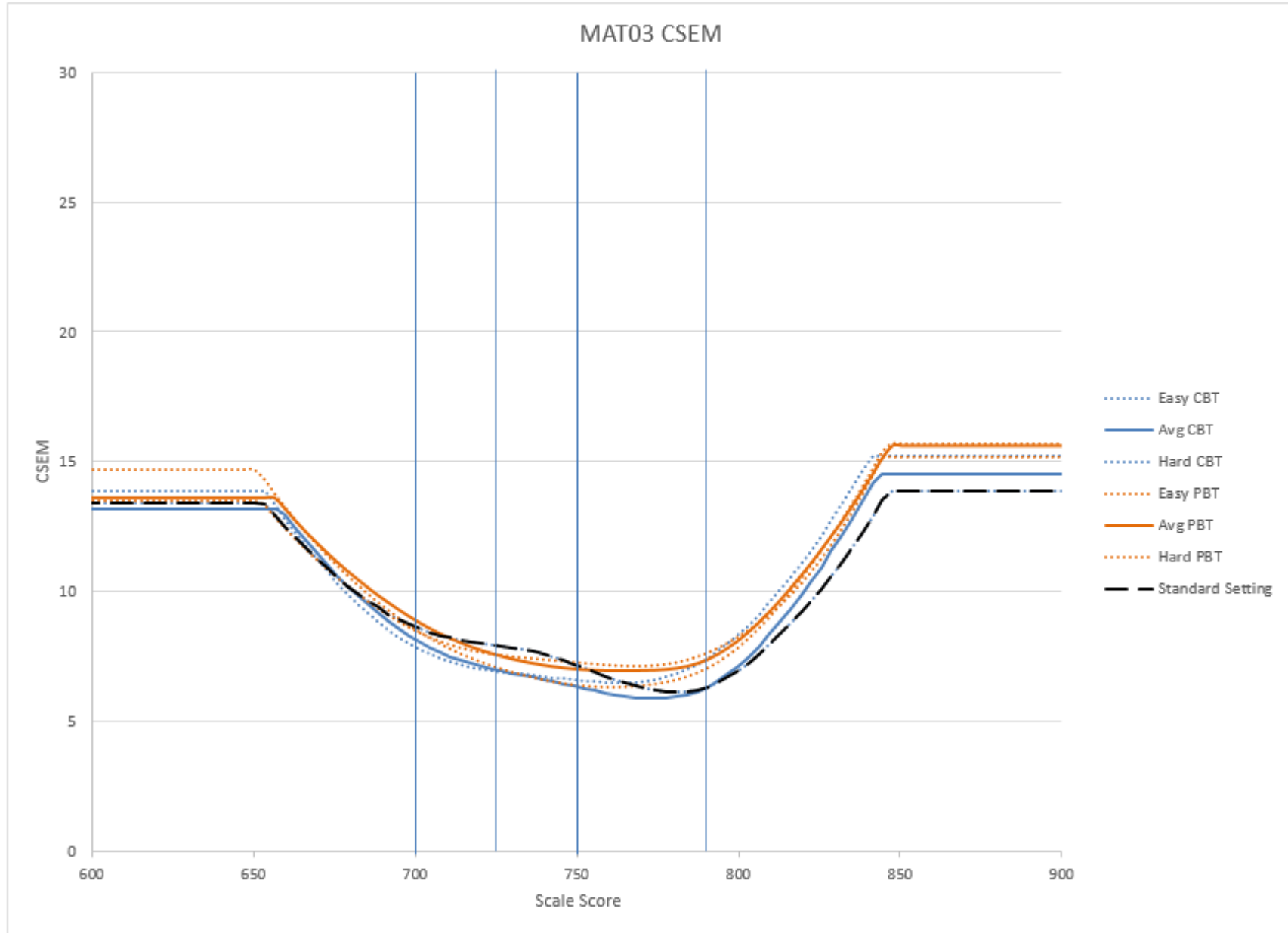


Figure A.12.52 CSEM Curves Mathematics Grade 3

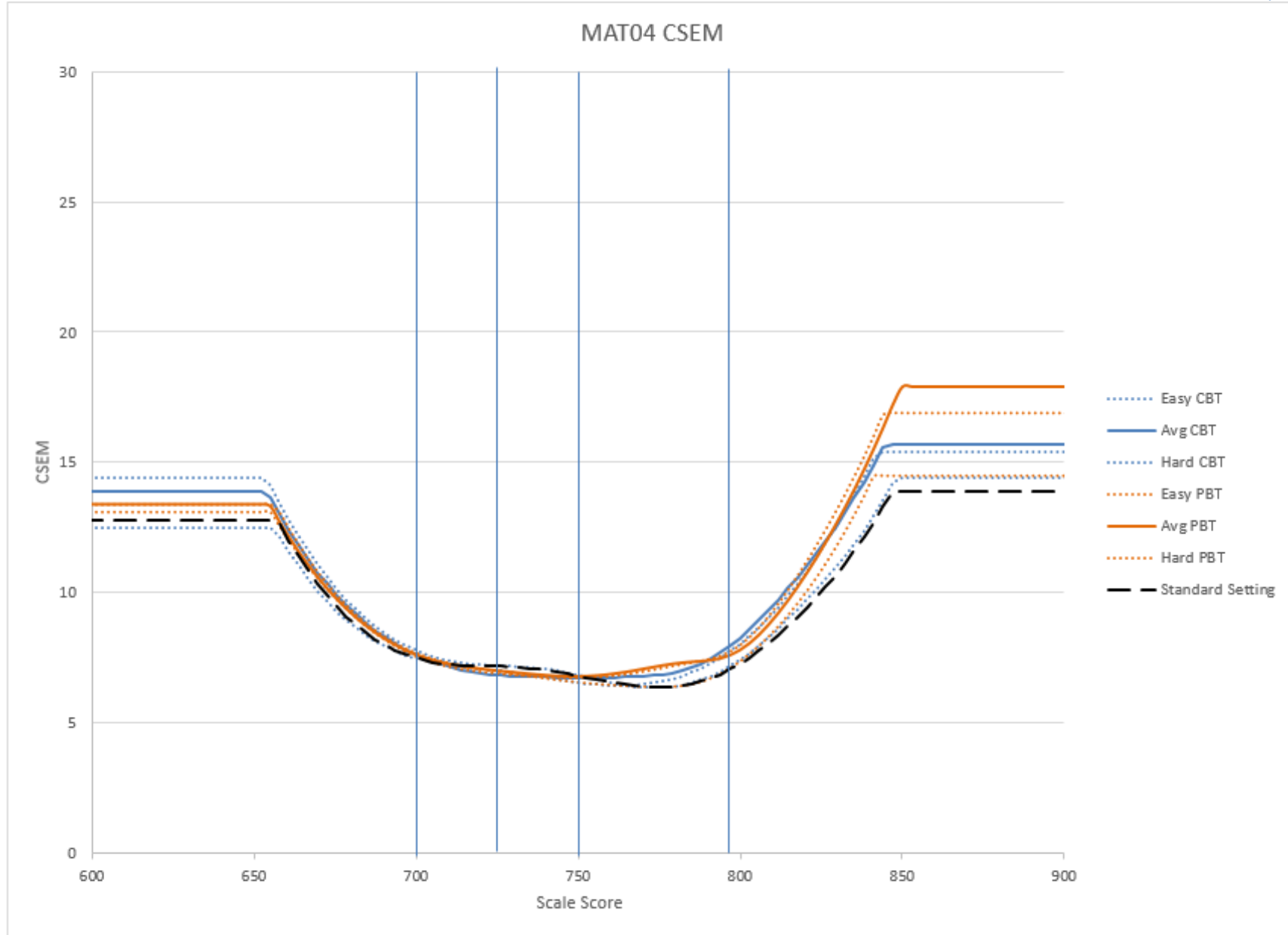


Figure A.12.53 CSEM Curves Mathematics Grade 4

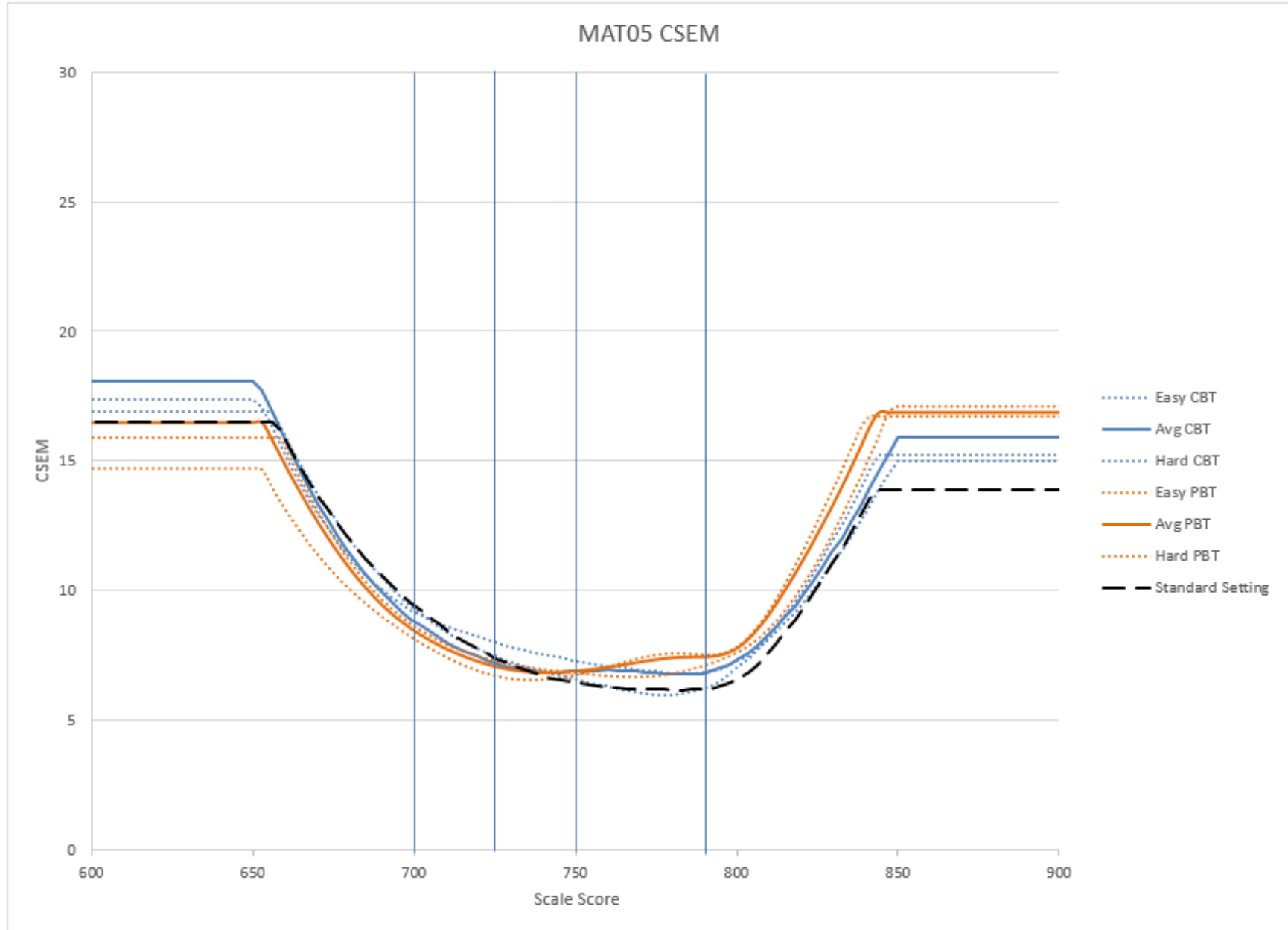


Figure A.12.54 CSEM Curves Mathematics Grade 5

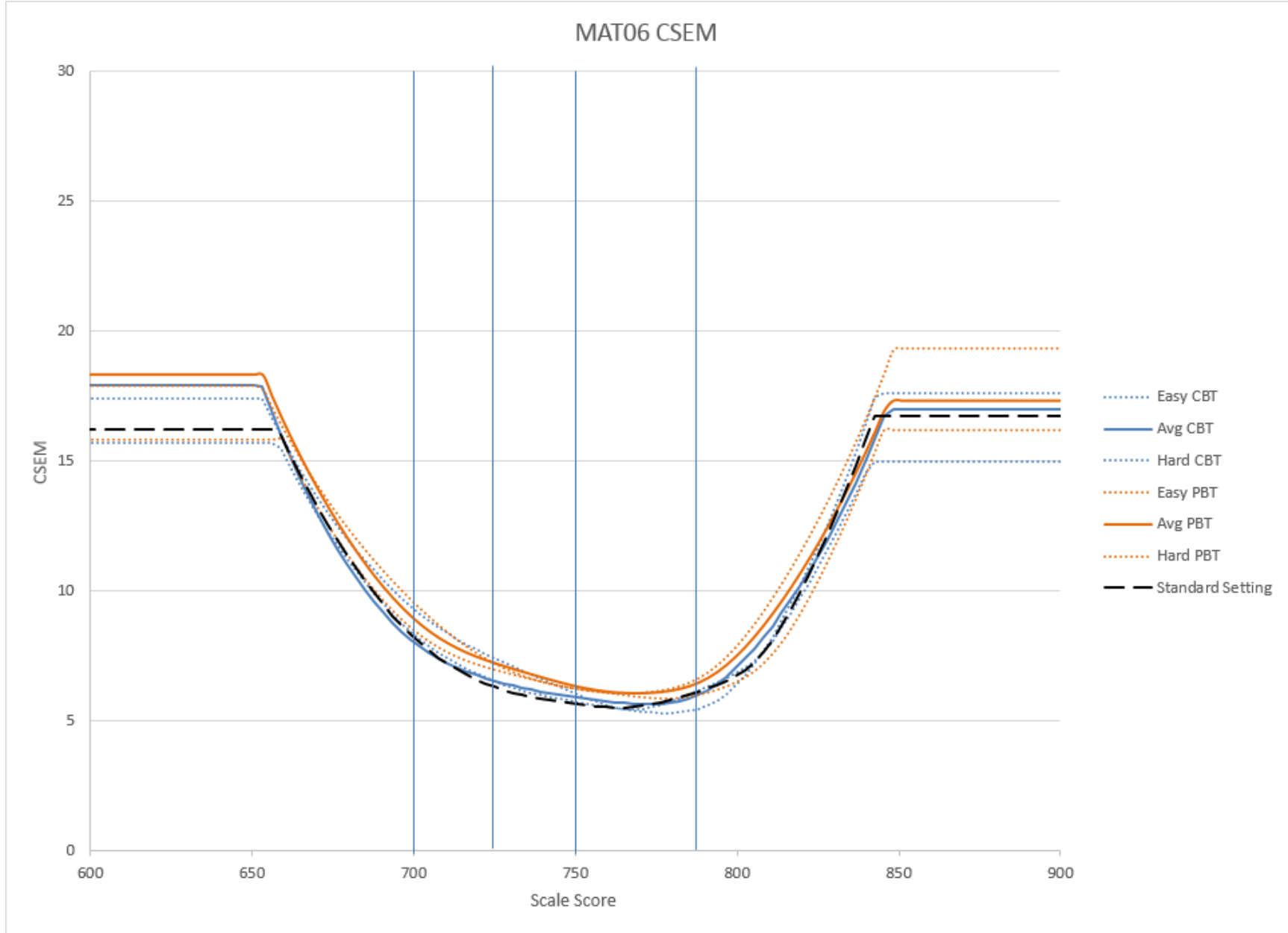


Figure A.12.55 CSEM Curves Mathematics Grade 6

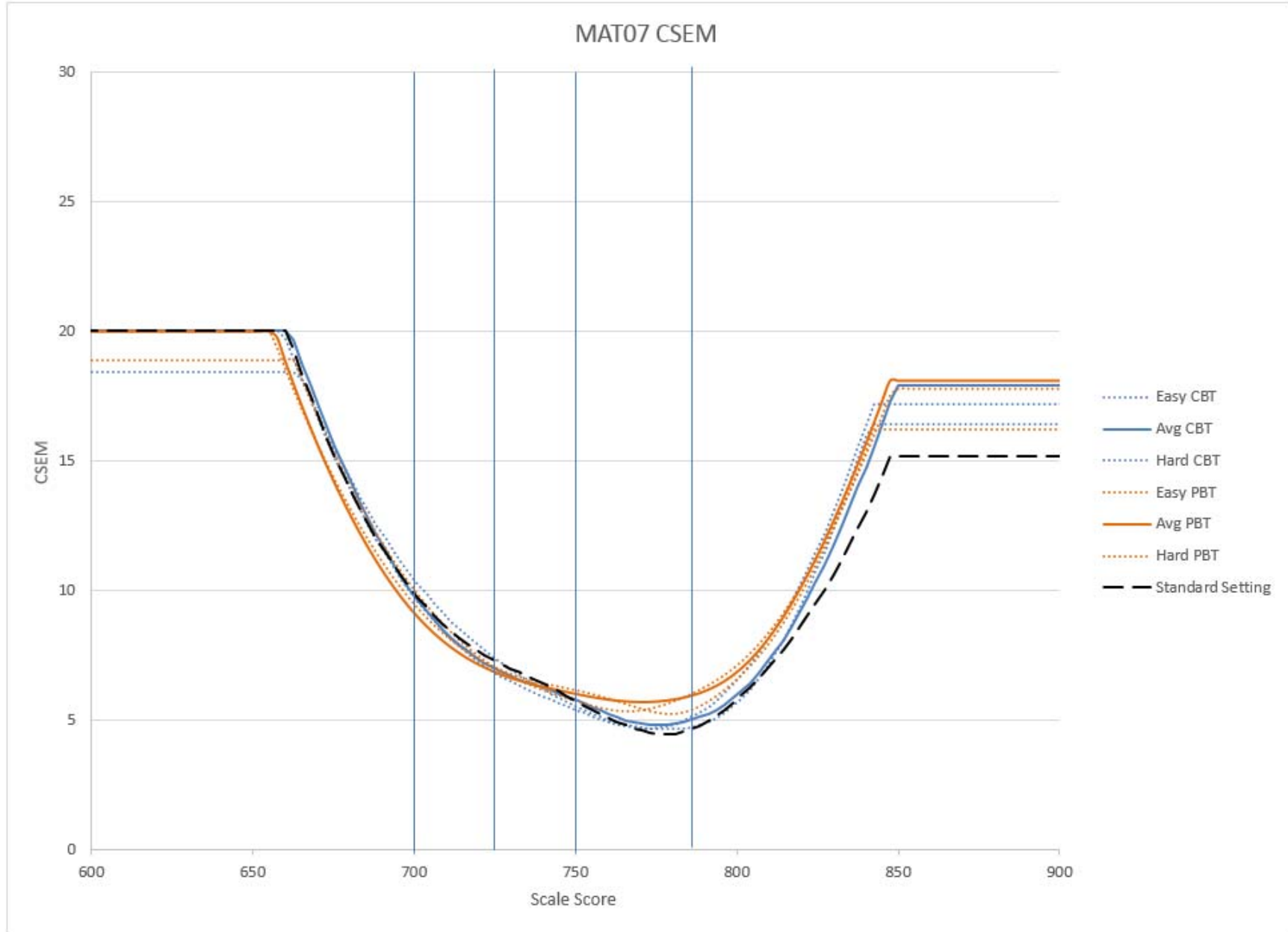


Figure A.12.56 CSEM Curves Mathematics Grade 7

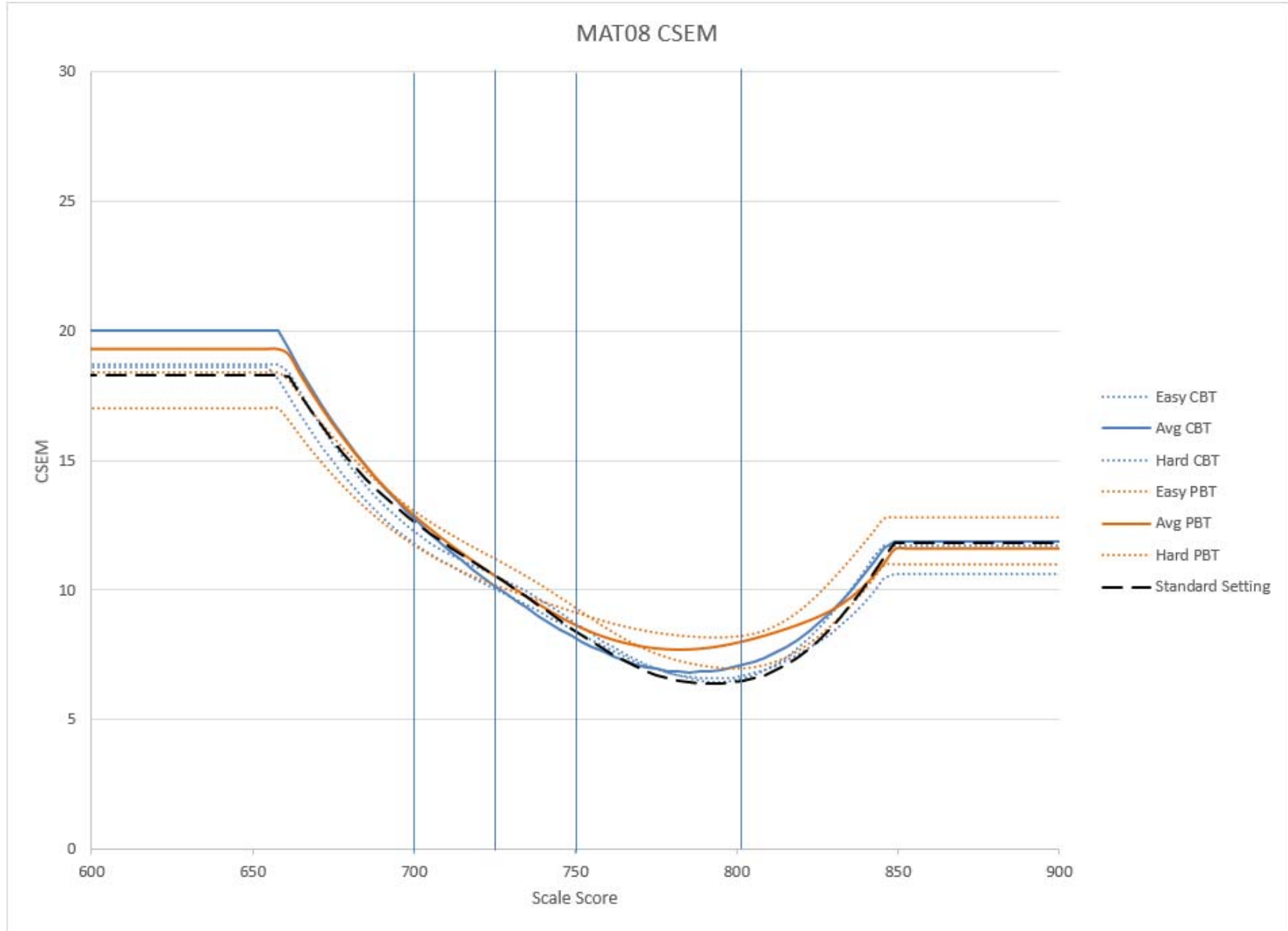


Figure A.12.57 CSEM Curves Mathematics Grade 8

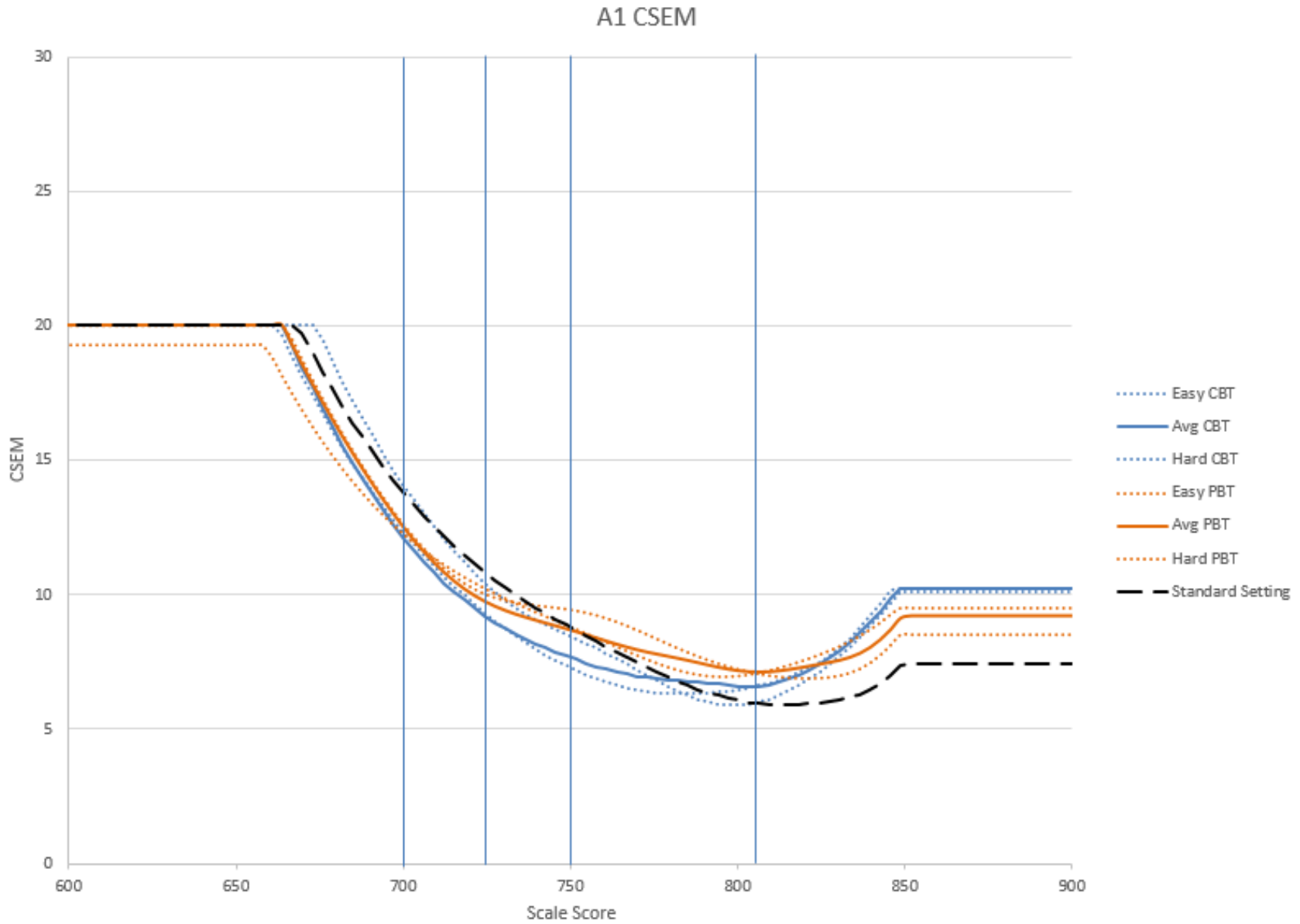


Figure A.12.58 CSEM Curves Algebra I

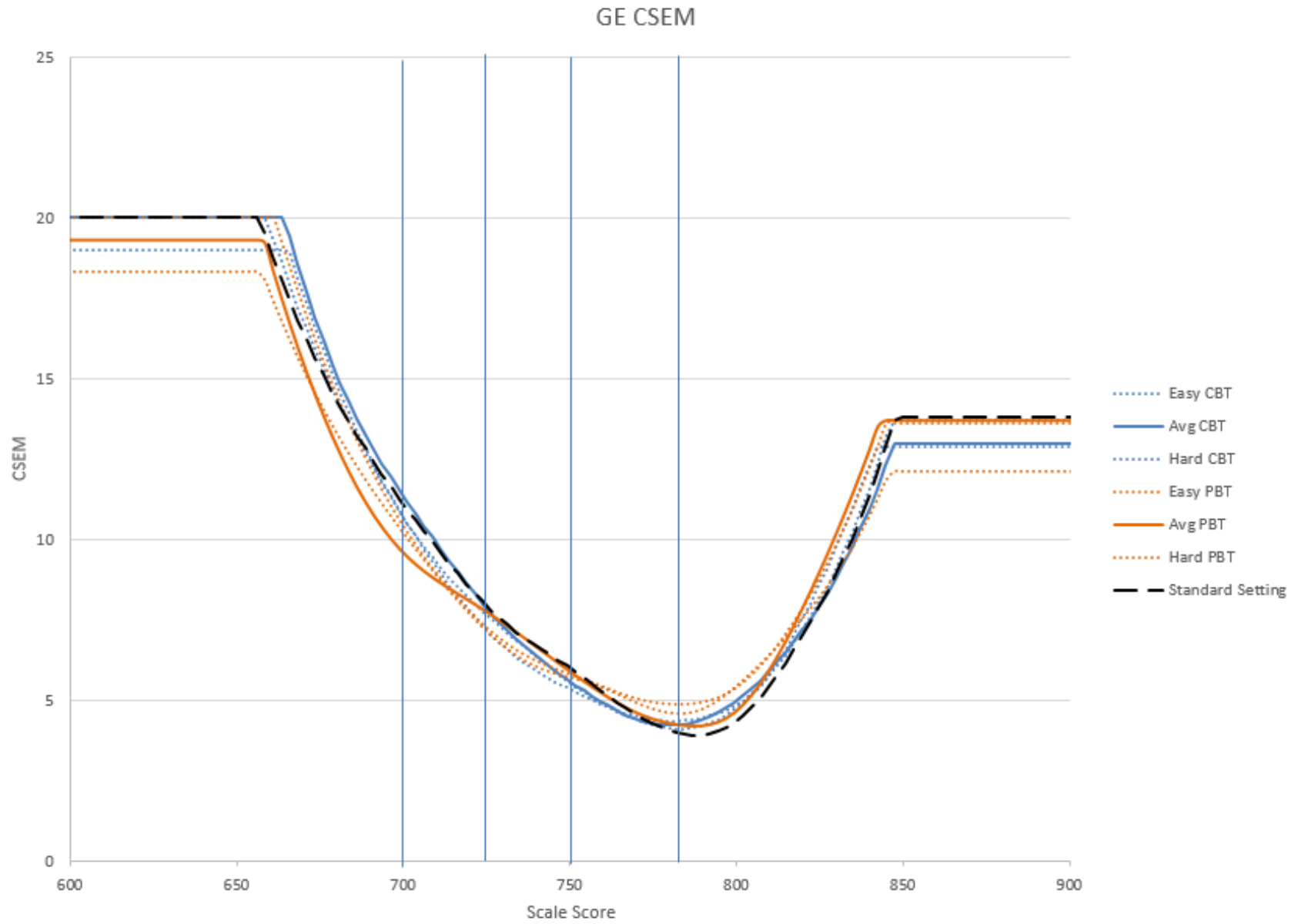


Figure A.12.59 CSEM Curves Geometry

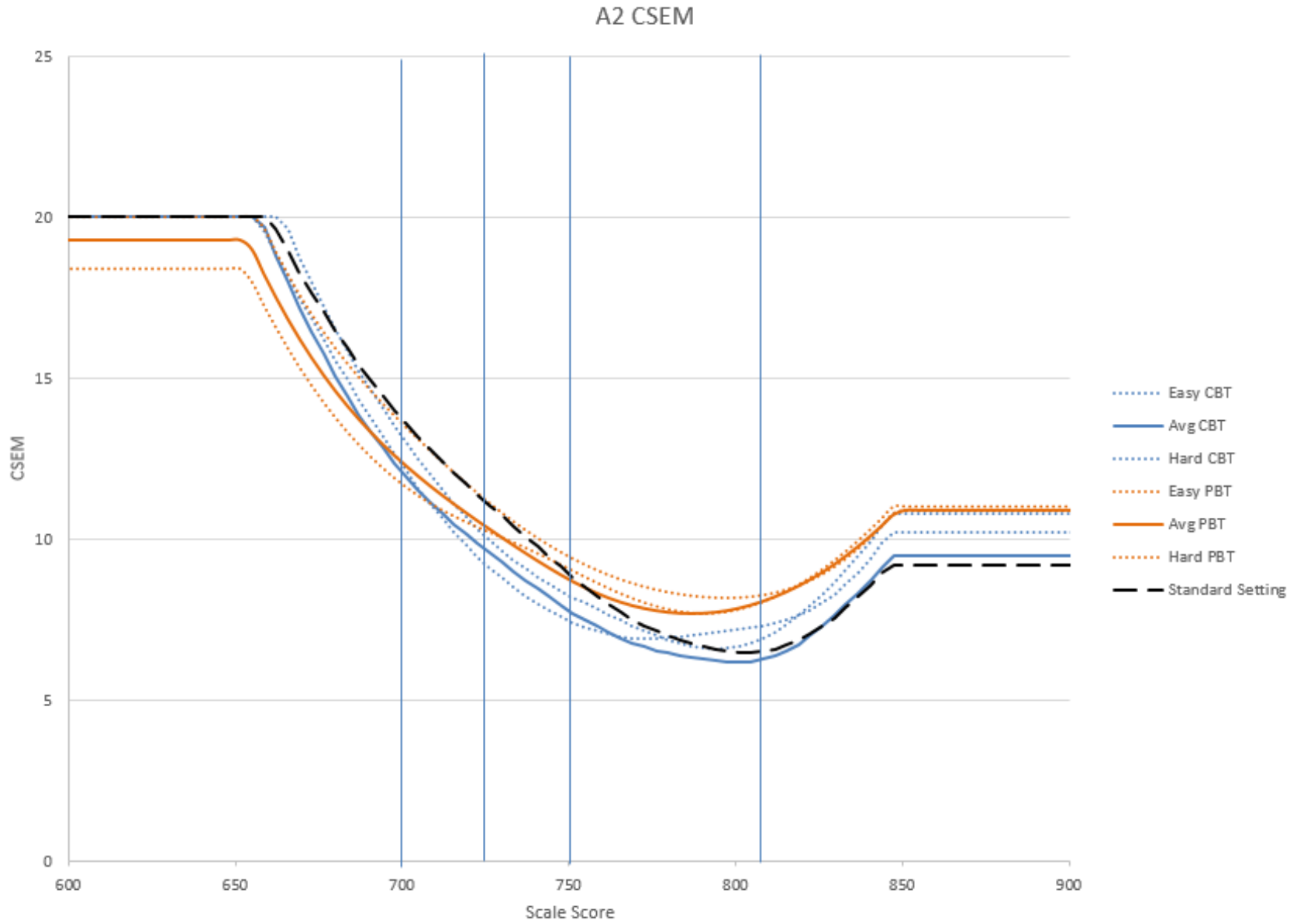


Figure A.12.60 CSEM Curves Algebra II

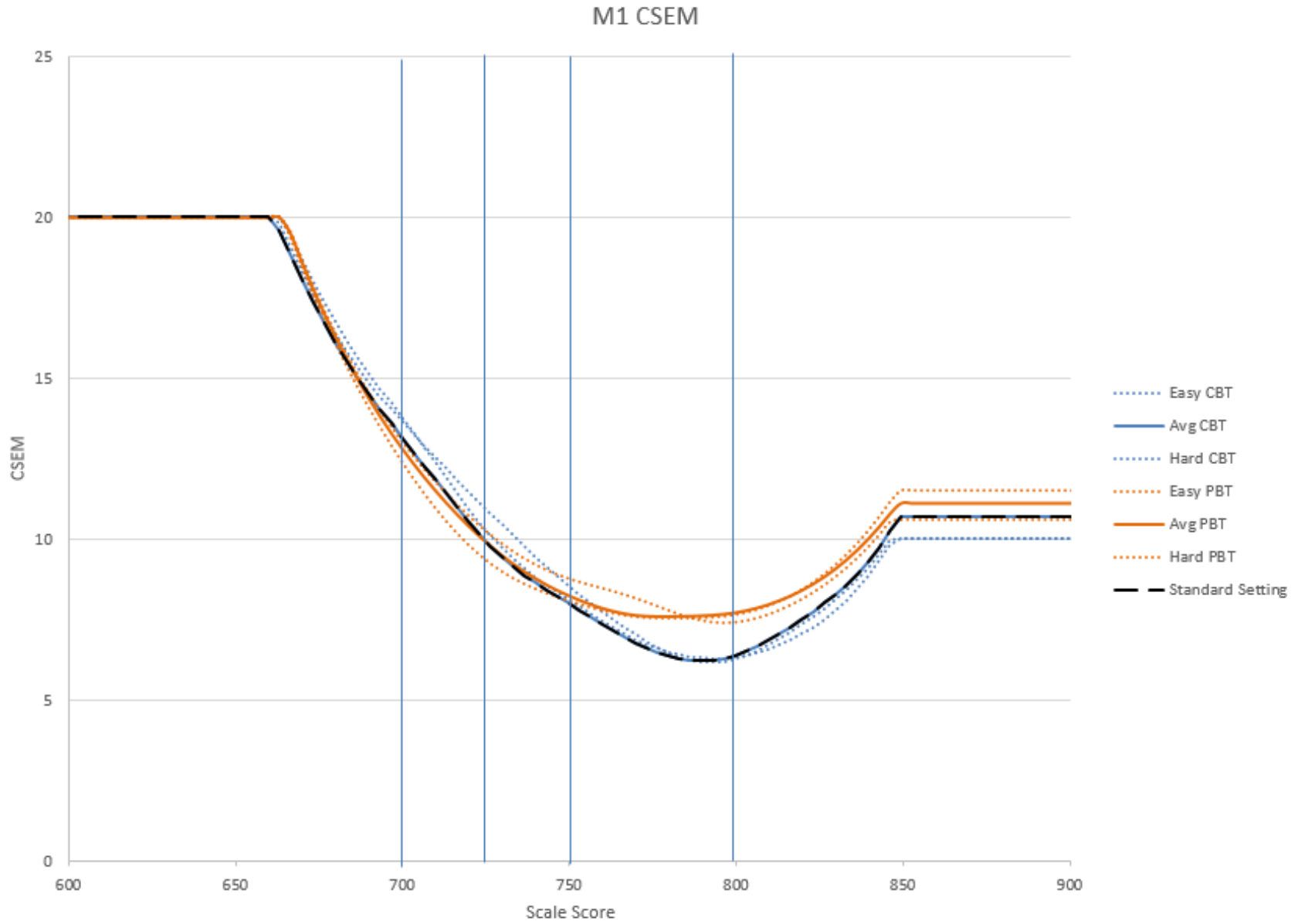


Figure A.12.61 CSEM Curves Integrated Mathematics I

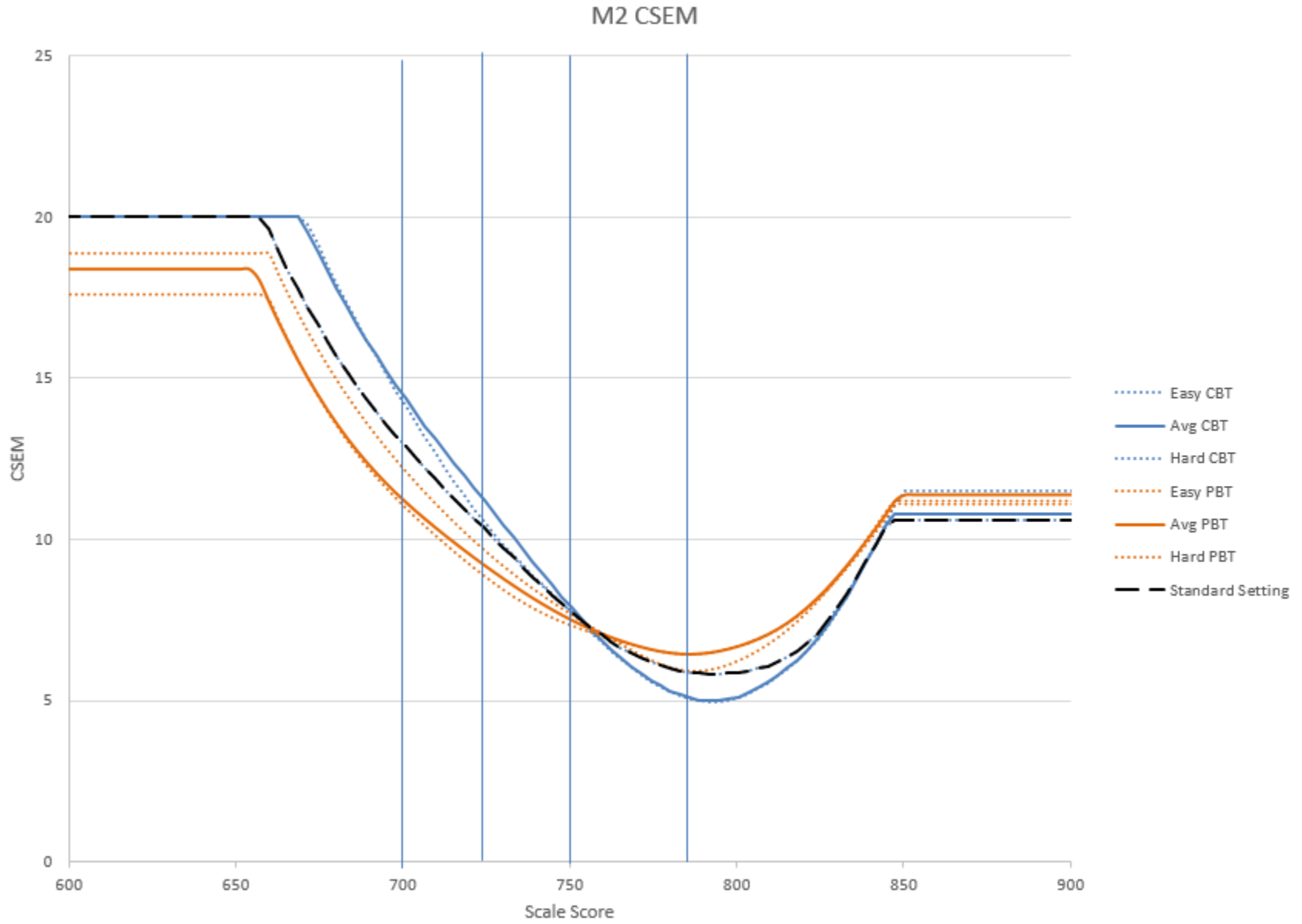


Figure A.12.62 CSEM Curves Integrated Mathematics II

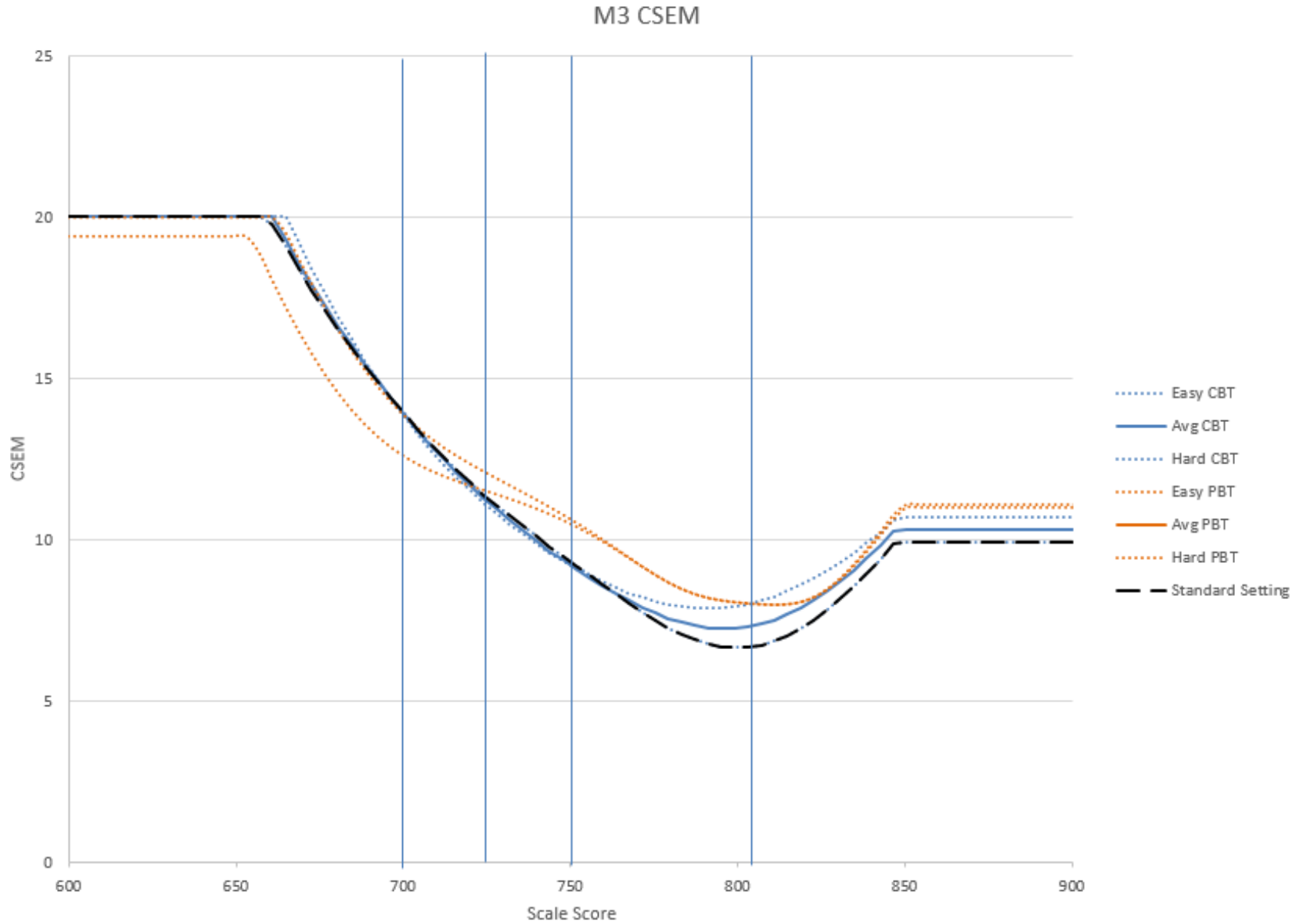


Figure A.12.63 CSEM Curves Integrated Mathematics III

April 18, 2016

Appendix 12.5: Subgroup Scale Score Performance

Table A.12.48 Subgroup Performance for ELA/L Scale Scores: Grade 3

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		518,857	736.20	39.74	650	850
Gender	Female	254,214	740.84	39.98	650	850
	Male	264,643	731.73	38.99	650	850
Ethnicity	American Indian/Alaska Native	5,088	715.25	33.61	650	850
	Asian	24,035	762.06	39.66	650	850
	Black or African American	95,098	719.84	35.79	650	850
	Hispanic/Latino	126,309	722.69	36.19	650	850
	Native Hawaiian or Pacific Islander	997	732.97	41.78	650	850
	Multiple Race Selected	43,004	750.47	39.45	650	850
	White	221,943	745.95	37.96	650	850
Economic Status*	Economically Disadvantaged	249,056	721.61	35.77	650	850
	Not Economically Disadvantaged	236,384	752.60	37.95	650	850
English Learner Status	English Learner (EL)	67,638	713.33	33.16	650	850
	Non English Learner	416,752	740.51	39.67	650	850
Disabilities	Students with Disabilities (SWD)	46,945	707.95	37.10	650	850
	Students without Disabilities	319,190	741.47	38.65	650	850
Reading Score		518,857	44.35	15.87	10	90
Gender	Female	254,214	45.71	15.89	10	90
	Male	264,643	43.05	15.74	10	90
Ethnicity	American Indian/Alaska Native	5,088	35.87	13.12	10	90
	Asian	24,035	53.62	15.91	10	90
	Black or African American	95,098	37.81	13.91	10	90
	Hispanic/Latino	126,309	38.55	14.02	10	90
	Native Hawaiian or Pacific Islander	997	42.78	16.25	10	90
	Multiple Race Selected	43,004	50.02	16.06	10	90
	White	221,943	48.60	15.38	10	90
Economic Status*	Economically Disadvantaged	249,056	38.39	13.95	10	90
	Not Economically Disadvantaged	236,384	50.90	15.41	10	90
English Learner Status	English Learner (EL)	67,638	34.67	12.48	10	90
	Non English Learner	416,752	46.08	15.88	10	90
Disabilities	Students with Disabilities (SWD)	46,945	34.02	14.62	10	90
	Students without Disabilities	319,190	46.40	15.52	10	90
Writing Score		518,857	31.10	11.24	10	60
Gender	Female	254,214	32.76	10.99	10	60
	Male	264,643	29.50	11.24	10	60

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Ethnicity	American Indian/Alaska Native	5,088	26.52	10.63	10	60
	Asian	24,035	37.97	10.22	10	60
	Black or African American	95,098	27.34	11.03	10	60
	Hispanic/Latino	126,309	28.56	11.05	10	60
	Native Hawaiian or Pacific Islander	997	30.62	11.77	10	60
	Multiple Race Selected	43,004	34.43	10.65	10	60
	White	221,943	32.92	10.72	10	60
Economic Status*	Economically Disadvantaged	249,056	27.94	11.00	10	60
	Not Economically Disadvantaged	236,384	34.86	10.39	10	60
English Learner Status	English Learner (EL)	67,638	26.64	10.90	10	60
	Non English Learner	416,752	32.07	11.12	10	60
Disabilities	Students with Disabilities (SWD)	46,945	23.09	11.44	10	60
	Students without Disabilities	319,190	32.40	10.82	10	60

Note: This table is identical to Table 12.7 in Section 12.

*Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.49 Subgroup Performance for ELA/L Scale Scores: Grade 4

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		632,312	742.03	33.29	650	850
Gender	Female	309,674	746.68	33.18	650	850
	Male	322,638	737.57	32.79	650	850
Ethnicity	American Indian/Alaska Native	5,081	723.33	29.53	650	850
	Asian	26,985	764.24	33.25	650	850
	Black or African American	107,485	726.09	30.00	650	850
	Hispanic/Latino	128,961	730.63	30.91	650	850
	Native Hawaiian or Pacific Islander	1,050	741.13	34.09	650	848
	Multiple Race Selected	46,968	754.84	33.46	650	850
	White	305,206	749.03	31.33	650	850
Economic Status*	Economically Disadvantaged	293,807	729.45	30.14	650	850
	Not Economically Disadvantaged	299,170	755.32	31.40	650	850
English Learner Status	English Learner (EL)	46,571	714.20	27.79	650	843
	Non English Learner	544,332	745.05	32.69	650	850
Disabilities	Students with Disabilities (SWD)	62,951	714.04	31.47	650	850
	Students without Disabilities	409,507	747.15	31.61	650	850
Reading Score		632,312	46.71	13.32	10	90
Gender	Female	309,674	47.98	13.26	10	90
	Male	322,638	45.50	13.27	10	90
Ethnicity	American Indian/Alaska Native	5,081	39.30	11.72	10	89
	Asian	26,985	54.85	13.43	10	90
	Black or African American	107,485	40.42	11.75	10	90
	Hispanic/Latino	128,961	41.78	12.12	10	90
	Native Hawaiian or Pacific Islander	1,050	45.83	13.41	10	85
	Multiple Race Selected	46,968	51.41	13.45	10	90
	White	305,206	49.77	12.68	10	90
Economic Status*	Economically Disadvantaged	293,807	41.66	11.90	10	90
	Not Economically Disadvantaged	299,170	51.95	12.73	10	90
English Learner Status	English Learner (EL)	46,571	35.36	10.51	10	90
	Non English Learner	544,332	47.89	13.11	10	90
Disabilities	Students with Disabilities (SWD)	62,951	36.42	12.31	10	90
	Students without Disabilities	409,507	48.64	12.78	10	90
Writing Score		632,312	33.00	9.11	10	60
Gender	Female	309,674	34.65	8.82	10	60
	Male	322,638	31.42	9.10	10	60
Ethnicity	American Indian/Alaska Native	5,081	28.74	8.76	10	60

Group Type	Group	N	Mean	SD	Min	Max
	Asian	26,985	38.64	8.69	10	60
	Black or African American	107,485	29.27	8.98	10	60
	Hispanic/Latino	128,961	30.81	8.96	10	60
	Native Hawaiian or Pacific Islander	1,050	33.35	9.34	10	60
	Multiple Race Selected	46,968	36.26	8.85	10	60
	White	305,206	34.36	8.52	10	60
Economic Status*	Economically Disadvantaged	293,807	30.17	8.86	10	60
	Not Economically Disadvantaged	299,170	36.10	8.37	10	60
English Learner Status	English Learner (EL)	46,571	26.79	9.12	10	60
	Non English Learner	544,332	33.73	8.90	10	60
Disabilities	Students with Disabilities (SWD)	62,951	25.28	9.95	10	60
	Students without Disabilities	409,507	34.34	8.46	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.50 Subgroup Performance for ELA/L Scale Scores: Grade 5

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		638,239	741.08	31.85	650	850
Gender	Female	312,329	745.75	31.71	650	850
	Male	325,910	736.60	31.33	650	850
Ethnicity	American Indian/Alaska Native	5,033	723.47	28.07	650	831
	Asian	27,199	763.19	31.93	650	850
	Black or African American	107,694	725.77	29.05	650	849
	Hispanic/Latino	126,349	730.86	29.65	650	850
	Native Hawaiian or Pacific Islander	964	740.76	33.59	650	850
	Multiple Race Selected	46,382	752.54	31.65	650	850
	White	304,091	747.30	30.15	650	850
Economic Status*	Economically Disadvantaged	290,331	729.03	28.99	650	850
	Not Economically Disadvantaged	308,141	753.47	29.99	650	850
English Learner Status	English Learner (EL)	36,102	712.50	25.72	650	850
	Non English Learner	560,196	743.61	31.36	650	850
Disabilities	Students with Disabilities (SWD)	64,260	713.36	28.42	650	850
	Students without Disabilities	411,603	746.28	30.24	650	850
Reading Score		638,239	46.38	12.81	10	90
Gender	Female	312,329	47.59	12.69	10	90
	Male	325,910	45.23	12.83	10	90
Ethnicity	American Indian/Alaska Native	5,033	39.27	11.40	10	90
	Asian	27,199	54.42	12.89	10	90
	Black or African American	107,694	40.49	11.58	10	90
	Hispanic/Latino	126,349	42.00	11.73	10	90
	Native Hawaiian or Pacific Islander	964	45.58	13.21	10	90
	Multiple Race Selected	46,382	50.58	12.82	10	90
	White	304,091	49.01	12.26	10	90
Economic Status*	Economically Disadvantaged	290,331	41.60	11.58	10	90
	Not Economically Disadvantaged	308,141	51.18	12.21	10	90
English Learner Status	English Learner (EL)	36,102	34.98	9.89	10	87
	Non English Learner	560,196	47.33	12.64	10	90
Disabilities	Students with Disabilities (SWD)	64,260	36.00	11.35	10	90
	Students without Disabilities	411,603	48.37	12.27	10	90
Writing Score		638,239	31.68	10.09	10	60
Gender	Female	312,329	33.65	9.49	10	60
	Male	325,910	29.78	10.28	10	60
Ethnicity	American Indian/Alaska Native	5,033	27.45	9.88	10	57

Group Type	Group	N	Mean	SD	Min	Max
	Asian	27,199	37.84	8.84	10	60
	Black or African American	107,694	27.38	10.35	10	60
	Hispanic/Latino	126,349	29.39	10.19	10	60
	Native Hawaiian or Pacific Islander	964	32.31	10.33	10	60
	Multiple Race Selected	46,382	34.93	9.35	10	60
	White	304,091	33.18	9.38	10	60
Economic Status*	Economically Disadvantaged	290,331	28.52	10.17	10	60
	Not Economically Disadvantaged	308,141	35.03	8.89	10	60
English Learner Status	English Learner (EL)	36,102	23.82	10.49	10	60
	Non English Learner	560,196	32.42	9.81	10	60
Disabilities	Students with Disabilities (SWD)	64,260	22.77	10.88	10	60
	Students without Disabilities	411,603	33.22	9.29	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.51 Subgroup Performance for ELA/L Scale Scores: Grade 6

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		632,506	739.98	30.93	650	850
Gender	Female	308,806	745.69	30.29	650	850
	Male	323,700	734.53	30.56	650	850
Ethnicity	American Indian/Alaska Native	4,857	723.79	27.70	650	826
	Asian	26,837	761.67	31.16	650	850
	Black or African American	107,714	724.80	28.24	650	850
	Hispanic/Latino	123,646	729.82	28.95	650	850
	Native Hawaiian or Pacific Islander	956	739.27	31.79	654	839
	Multiple Race Selected	45,845	751.47	30.56	650	850
	White	301,694	745.98	29.14	650	850
Economic Status*	Economically Disadvantaged	283,067	728.19	28.42	650	850
	Not Economically Disadvantaged	310,693	751.35	29.20	650	850
English Learner Status	English Learner (EL)	29,835	709.56	26.02	650	839
	Non English Learner	561,409	742.06	30.44	650	850
Disabilities	Students with Disabilities (SWD)	62,028	711.00	27.30	650	850
	Students without Disabilities	409,302	745.46	29.07	650	850
Reading Score		632,506	46.05	12.26	10	90
Gender	Female	308,806	47.51	12.00	10	90
	Male	323,700	44.66	12.34	10	90
Ethnicity	American Indian/Alaska Native	4,857	39.44	10.88	10	79
	Asian	26,837	53.70	12.40	10	90
	Black or African American	107,714	40.22	10.97	10	90
	Hispanic/Latino	123,646	41.61	11.18	10	90
	Native Hawaiian or Pacific Islander	956	44.93	12.33	13	81
	Multiple Race Selected	45,845	50.31	12.27	10	90
	White	301,694	48.64	11.69	10	90
Economic Status*	Economically Disadvantaged	283,067	41.37	11.10	10	90
	Not Economically Disadvantaged	310,693	50.48	11.71	10	90
English Learner Status	English Learner (EL)	29,835	33.83	9.63	10	90
	Non English Learner	561,409	46.84	12.07	10	90
Disabilities	Students with Disabilities (SWD)	62,028	35.24	10.73	10	90
	Students without Disabilities	409,302	48.10	11.65	10	90
Writing Score		632,506	31.67	9.75	10	60
Gender	Female	308,806	34.09	8.87	10	60
	Male	323,700	29.36	9.98	10	60
Ethnicity	American Indian/Alaska Native	4,857	27.97	9.57	10	60

Group Type	Group	N	Mean	SD	Min	Max
	Asian	26,837	37.91	8.82	10	60
	Black or African American	107,714	27.52	10.00	10	60
	Hispanic/Latino	123,646	29.62	9.79	10	60
	Native Hawaiian or Pacific Islander	956	32.41	9.99	10	60
	Multiple Race Selected	45,845	34.85	8.98	10	60
	White	301,694	33.00	9.10	10	60
Economic Status*	Economically Disadvantaged	283,067	28.67	9.85	10	60
	Not Economically Disadvantaged	310,693	34.65	8.76	10	60
English Learner Status	English Learner (EL)	29,835	23.86	10.30	10	60
	Non English Learner	561,409	32.25	9.54	10	60
Disabilities	Students with Disabilities (SWD)	62,028	22.76	10.39	10	60
	Students without Disabilities	409,302	33.30	8.90	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.52 Subgroup Performance for ELA/L Scale Scores: Grade 7

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		624,008	740.56	36.13	650	850
Gender	Female	304,602	747.84	35.23	650	850
	Male	319,406	733.63	35.61	650	850
Ethnicity	American Indian/Alaska Native	4,901	718.60	31.97	650	850
	Asian	25,556	766.98	36.66	650	850
	Black or African American	106,713	722.38	32.67	650	850
	Hispanic/Latino	119,045	728.69	33.89	650	850
	Native Hawaiian or Pacific Islander	913	739.97	38.17	650	850
	Multiple Race Selected	45,016	754.39	35.26	650	850
	White	301,479	747.59	33.96	650	850
Economic Status*	Economically Disadvantaged	273,549	726.66	33.16	650	850
	Not Economically Disadvantaged	311,611	753.65	34.13	650	850
English Learner Status	English Learner (EL)	29,255	704.23	29.35	650	850
	Non English Learner	553,384	743.13	35.54	650	850
Disabilities	Students with Disabilities (SWD)	60,194	706.04	31.23	650	850
	Students without Disabilities	405,022	746.52	34.21	650	850
Reading Score		624,008	46.23	14.35	10	90
Gender	Female	304,602	48.34	13.99	10	90
	Male	319,406	44.22	14.39	10	90
Ethnicity	American Indian/Alaska Native	4,901	37.55	12.72	10	90
	Asian	25,556	55.57	14.64	10	90
	Black or African American	106,713	39.30	12.82	10	90
	Hispanic/Latino	119,045	41.12	13.18	10	90
	Native Hawaiian or Pacific Islander	913	45.19	14.93	10	90
	Multiple Race Selected	45,016	51.05	14.11	10	90
	White	301,479	49.24	13.64	10	90
Economic Status*	Economically Disadvantaged	273,549	40.77	13.06	10	90
	Not Economically Disadvantaged	311,611	51.28	13.69	10	90
English Learner Status	English Learner (EL)	29,255	31.60	10.95	10	90
	Non English Learner	553,384	47.21	14.12	10	90
Disabilities	Students with Disabilities (SWD)	60,194	33.10	12.30	10	90
	Students without Disabilities	405,022	48.49	13.68	10	90
Writing Score		624,008	31.83	10.68	10	60
Gender	Female	304,602	34.49	9.86	10	60
	Male	319,406	29.29	10.82	10	60
Ethnicity	American Indian/Alaska Native	4,901	26.60	10.30	10	56

Group Type	Group	N	Mean	SD	Min	Max
	Asian	25,556	39.14	9.65	10	60
	Black or African American	106,713	26.97	10.67	10	60
	Hispanic/Latino	119,045	29.45	10.61	10	60
	Native Hawaiian or Pacific Islander	913	32.46	11.19	10	60
	Multiple Race Selected	45,016	35.77	9.88	10	60
	White	301,479	33.37	9.99	10	60
Economic Status*	Economically Disadvantaged	273,549	28.35	10.61	10	60
	Not Economically Disadvantaged	311,611	35.19	9.72	10	60
English Learner Status	English Learner (EL)	29,255	23.05	10.58	10	60
	Non English Learner	553,384	32.50	10.49	10	60
Disabilities	Students with Disabilities (SWD)	60,194	22.10	10.77	10	60
	Students without Disabilities	405,022	33.47	9.97	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.53 Subgroup Performance for ELA/L Scale Scores: Grade 8

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		617,564	740.62	36.90	650	850
Gender	Female	300,713	748.42	35.61	650	850
	Male	316,851	733.22	36.58	650	850
Ethnicity	American Indian/Alaska Native	4,626	721.39	32.49	650	850
	Asian	24,982	768.07	36.73	650	850
	Black or African American	107,133	722.37	33.65	650	850
	Hispanic/Latino	114,806	728.87	34.74	650	850
	Native Hawaiian or Pacific Islander	833	737.85	40.62	650	850
	Multiple Race Selected	44,473	754.84	36.00	650	850
	White	299,203	747.35	34.81	650	850
Economic Status*	Economically Disadvantaged	265,283	726.60	34.06	650	850
	Not Economically Disadvantaged	313,632	753.30	35.02	650	850
English Learner Status	English Learner (EL)	28,436	704.00	29.82	650	842
	Non English Learner	545,424	743.22	36.36	650	850
Disabilities	Students with Disabilities (SWD)	58,679	705.09	31.76	650	850
	Students without Disabilities	397,692	746.54	35.01	650	850
Reading Score		617,564	46.21	14.66	10	90
Gender	Female	300,713	48.42	14.25	10	90
	Male	316,851	44.12	14.74	10	90
Ethnicity	American Indian/Alaska Native	4,626	38.42	13.05	10	90
	Asian	24,982	55.84	14.64	10	90
	Black or African American	107,133	39.28	13.21	10	90
	Hispanic/Latino	114,806	41.08	13.58	10	90
	Native Hawaiian or Pacific Islander	833	44.67	15.71	10	90
	Multiple Race Selected	44,473	51.19	14.41	10	90
	White	299,203	49.11	13.98	10	90
Economic Status*	Economically Disadvantaged	265,283	40.68	13.43	10	90
	Not Economically Disadvantaged	313,632	51.13	14.01	10	90
English Learner Status	English Learner (EL)	28,436	31.30	11.03	10	88
	Non English Learner	545,424	47.22	14.44	10	90
Disabilities	Students with Disabilities (SWD)	58,679	32.73	12.39	10	90
	Students without Disabilities	397,692	48.48	14.00	10	90
Writing Score		617,564	32.53	10.10	10	60
Gender	Female	300,713	35.15	9.36	10	60
	Male	316,851	30.05	10.16	10	60
Ethnicity	American Indian/Alaska Native	4,626	28.34	9.31	10	60

Group Type	Group	N	Mean	SD	Min	Max
	Asian	24,982	39.83	9.64	10	60
	Black or African American	107,133	27.96	9.92	10	60
	Hispanic/Latino	114,806	30.24	9.90	10	60
	Native Hawaiian or Pacific Islander	833	32.10	11.34	10	60
	Multiple Race Selected	44,473	36.35	9.54	10	60
	White	299,203	33.91	9.51	10	60
Economic Status*	Economically Disadvantaged	265,283	29.23	9.88	10	60
	Not Economically Disadvantaged	313,632	35.57	9.43	10	60
English Learner Status	English Learner (EL)	28,436	24.28	9.86	10	60
	Non English Learner	545,424	33.15	9.97	10	60
Disabilities	Students with Disabilities (SWD)	58,679	23.29	10.15	10	60
	Students without Disabilities	397,692	34.01	9.51	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.54 Subgroup Performance for ELA/L Scale Scores: Grade 9

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		413,098	739.45	36.55	650	850
Gender	Female	201,280	746.57	35.57	650	850
	Male	211,818	732.68	36.18	650	850
Ethnicity	American Indian/Alaska Native	3,982	721.02	31.41	650	836
	Asian	17,616	762.36	37.86	650	850
	Black or African American	56,553	720.83	32.84	650	850
	Hispanic/Latino	85,733	725.20	33.80	650	850
	Native Hawaiian or Pacific Islander	702	734.33	37.76	650	850
	Multiple Race Selected	15,568	741.13	36.50	650	850
	White	211,054	747.33	34.58	650	850
Economic Status*	Economically Disadvantaged	175,922	725.87	33.29	650	850
	Not Economically Disadvantaged	228,674	749.50	35.49	650	850
English Learner Status	English Learner (EL)	20,828	703.19	28.74	650	843
	Non English Learner	377,194	741.60	35.85	650	850
Disabilities	Students with Disabilities (SWD)	40,963	705.72	29.15	650	850
	Students without Disabilities	278,126	745.43	34.73	650	850
Reading Score		413,098	45.78	14.58	10	90
Gender	Female	201,280	47.80	14.26	10	90
	Male	211,818	43.86	14.62	10	90
Ethnicity	American Indian/Alaska Native	3,982	38.32	12.45	10	90
	Asian	17,616	53.95	15.25	10	90
	Black or African American	56,553	38.58	12.80	10	90
	Hispanic/Latino	85,733	39.93	13.12	10	90
	Native Hawaiian or Pacific Islander	702	43.06	14.60	10	89
	Multiple Race Selected	15,568	46.47	14.55	10	90
	White	211,054	49.04	13.98	10	90
Economic Status*	Economically Disadvantaged	175,922	40.45	13.11	10	90
	Not Economically Disadvantaged	228,674	49.75	14.30	10	90
English Learner Status	English Learner (EL)	20,828	31.41	10.77	10	90
	Non English Learner	377,194	46.64	14.33	10	90
Disabilities	Students with Disabilities (SWD)	40,963	33.04	11.52	10	90
	Students without Disabilities	278,126	48.07	13.99	10	90
Writing Score		413,098	31.44	10.83	10	60
Gender	Female	201,280	34.19	9.89	10	60
	Male	211,818	28.83	11.04	10	60
Ethnicity	American Indian/Alaska Native	3,982	27.33	10.26	10	57

Group Type	Group	N	Mean	SD	Min	Max
	Asian	17,616	37.77	10.01	10	60
	Black or African American	56,553	26.50	10.96	10	60
	Hispanic/Latino	85,733	28.18	10.93	10	60
	Native Hawaiian or Pacific Islander	702	30.95	11.31	10	60
	Multiple Race Selected	15,568	31.77	10.89	10	60
	White	211,054	33.27	10.08	10	60
Economic Status*	Economically Disadvantaged	175,922	28.05	10.76	10	60
	Not Economically Disadvantaged	228,674	33.92	10.18	10	60
English Learner Status	English Learner (EL)	20,828	22.36	10.62	10	60
	Non English Learner	377,194	31.96	10.63	10	60
Disabilities	Students with Disabilities (SWD)	40,963	21.84	10.53	10	60
	Students without Disabilities	278,126	33.08	10.09	10	60

Note: This table is identical to Table 12.8 in Section 12.

*Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.55 Subgroup Performance for ELA/L Scale Scores: Grade 10

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		267,159	735.21	44.57	650	850
Gender	Female	130,374	743.60	43.41	650	850
	Male	136,785	727.20	44.19	650	850
Ethnicity	American Indian/Alaska Native	3,432	717.84	36.87	650	850
	Asian	13,966	760.59	47.49	650	850
	Black or African American	57,287	721.51	39.65	650	850
	Hispanic/Latino	58,093	721.68	40.78	650	850
	Native Hawaiian or Pacific Islander	576	728.84	46.05	650	850
	Multiple Race Selected	4,668	742.57	45.24	650	850
	White	127,060	745.03	44.26	650	850
Economic Status*	Economically Disadvantaged	103,720	720.02	40.17	650	850
	Not Economically Disadvantaged	142,189	745.06	45.17	650	850
English Learner Status	English Learner (EL)	11,858	694.40	33.55	650	850
	Non English Learner	249,876	737.36	44.09	650	850
Disabilities	Students with Disabilities (SWD)	27,058	694.74	33.93	650	850
	Students without Disabilities	201,777	739.52	43.72	650	850
Reading Score		267,159	44.26	17.48	10	90
Gender	Female	130,374	46.60	17.10	10	90
	Male	136,785	42.04	17.55	10	90
Ethnicity	American Indian/Alaska Native	3,432	37.16	14.41	10	90
	Asian	13,966	53.68	18.83	10	90
	Black or African American	57,287	39.15	15.36	10	90
	Hispanic/Latino	58,093	38.76	15.65	10	90
	Native Hawaiian or Pacific Islander	576	40.87	17.26	10	90
	Multiple Race Selected	4,668	47.55	18.01	10	90
	White	127,060	48.15	17.58	10	90
Economic Status*	Economically Disadvantaged	103,720	38.44	15.49	10	90
	Not Economically Disadvantaged	142,189	48.10	17.88	10	90
English Learner Status	English Learner (EL)	11,858	28.95	12.36	10	90
	Non English Learner	249,876	45.08	17.36	10	90
Disabilities	Students with Disabilities (SWD)	27,058	29.45	13.29	10	90
	Students without Disabilities	201,777	45.87	17.25	10	90
Writing Score		267,159	30.85	12.30	10	60
Gender	Female	130,374	33.77	11.73	10	60
	Male	136,785	28.06	12.18	10	60
Ethnicity	American Indian/Alaska Native	3,432	27.22	10.66	10	60

Group Type	Group	N	Mean	SD	Min	Max
	Asian	13,966	37.28	12.57	10	60
	Black or African American	57,287	27.41	11.44	10	60
	Hispanic/Latino	58,093	27.87	11.80	10	60
	Native Hawaiian or Pacific Islander	576	30.09	13.14	10	60
	Multiple Race Selected	4,668	32.16	12.31	10	60
	White	127,060	33.11	12.10	10	60
Economic Status*	Economically Disadvantaged	103,720	27.15	11.67	10	60
	Not Economically Disadvantaged	142,189	33.17	12.32	10	60
English Learner Status	English Learner (EL)	11,858	20.48	10.55	10	60
	Non English Learner	249,876	31.39	12.16	10	60
Disabilities	Students with Disabilities (SWD)	27,058	20.03	10.38	10	60
	Students without Disabilities	201,777	31.96	12.04	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.56 Subgroup Performance for ELA/L Scale Scores: Grade 11

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		171,171	738.82	39.11	650	850
Gender	Female	83,114	746.32	37.84	650	850
	Male	88,057	731.75	38.97	650	850
Ethnicity	American Indian/Alaska Native	2,844	734.08	31.32	650	850
	Asian	9,069	760.09	42.60	650	850
	Black or African American	19,151	723.77	34.99	650	850
	Hispanic/Latino	45,363	730.59	35.54	650	850
	Native Hawaiian or Pacific Islander	344	748.73	39.35	650	850
	Multiple Race Selected	8,283	738.91	41.11	650	850
	White	85,713	744.47	39.29	650	850
Economic Status*	Economically Disadvantaged	68,337	728.67	35.13	650	850
	Not Economically Disadvantaged	102,591	745.61	40.15	650	850
English Learner Status	English Learner (EL)	7,144	706.36	29.20	650	846
	Non English Learner	161,653	740.49	38.84	650	850
Disabilities	Students with Disabilities (SWD)	15,989	708.85	30.73	650	850
	Students without Disabilities	112,433	745.80	38.23	650	850
Reading Score		171,171	45.76	15.28	10	90
Gender	Female	83,114	47.85	14.87	10	90
	Male	88,057	43.79	15.40	10	90
Ethnicity	American Indian/Alaska Native	2,844	42.88	12.28	10	90
	Asian	9,069	53.54	17.05	10	90
	Black or African American	19,151	40.01	13.28	10	90
	Hispanic/Latino	45,363	42.18	13.57	10	90
	Native Hawaiian or Pacific Islander	344	48.62	15.48	10	90
	Multiple Race Selected	8,283	46.12	16.00	10	90
	White	85,713	48.19	15.49	10	90
Economic Status*	Economically Disadvantaged	68,337	41.65	13.46	10	90
	Not Economically Disadvantaged	102,591	48.51	15.80	10	90
English Learner Status	English Learner (EL)	7,144	33.11	10.72	10	90
	Non English Learner	161,653	46.41	15.20	10	90
Disabilities	Students with Disabilities (SWD)	15,989	34.59	11.80	10	90
	Students without Disabilities	112,433	48.19	15.07	10	90
Writing Score		171,171	30.93	11.86	10	60
Gender	Female	83,114	33.69	11.09	10	60
	Male	88,057	28.32	11.97	10	60
Ethnicity	American Indian/Alaska Native	2,844	31.21	9.61	10	60

Group Type	Group	N	Mean	SD	Min	Max
	Asian	9,069	36.65	11.72	10	60
	Black or African American	19,151	26.86	11.64	10	60
	Hispanic/Latino	45,363	29.25	11.36	10	60
	Native Hawaiian or Pacific Islander	344	34.34	11.36	10	60
	Multiple Race Selected	8,283	30.58	12.40	10	60
	White	85,713	32.14	11.76	10	60
Economic Status*	Economically Disadvantaged	68,337	28.53	11.37	10	60
	Not Economically Disadvantaged	102,591	32.52	11.91	10	60
English Learner Status	English Learner (EL)	7,144	22.47	10.74	10	60
	Non English Learner	161,653	31.35	11.76	10	60
Disabilities	Students with Disabilities (SWD)	15,989	22.46	10.92	10	60
	Students without Disabilities	112,433	33.00	11.37	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.57 Subgroup Performance for Mathematics Scale Scores: Grade 3

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		650,262	738.42	33.12	650	850
Gender	Female	318,517	738.86	32.04	650	850
	Male	331,745	737.99	34.11	650	850
Ethnicity	American Indian/Alaska Native	5,270	722.18	28.62	650	836
	Asian	26,866	765.78	33.45	650	850
	Black or African American	114,909	721.97	29.90	650	850
	Hispanic/Latino	136,010	727.57	30.03	650	850
	Native Hawaiian or Pacific Islander	1,149	736.09	35.00	650	845
	Multiple Race Selected	48,249	748.37	33.56	650	850
	White	307,191	745.89	31.05	650	850
Economic Status*	Economically Disadvantaged	311,763	726.10	30.05	650	850
	Not Economically Disadvantaged	296,843	751.70	31.42	650	850
English Learner Status	English Learner (EL)	77,582	723.18	29.75	650	850
	Non English Learner	528,210	741.04	33.13	650	850
Disabilities	Students with Disabilities (SWD)	61,364	716.60	32.90	650	850
	Students without Disabilities	424,037	742.84	31.73	650	850

Note: This table is identical to Table 12.9 in Section 12.

*Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.58 Subgroup Performance for Mathematics Scale Scores: Grade 4

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		635,075	735.70	30.47	650	850
Gender	Female	310,944	736.38	29.53	650	850
	Male	324,131	735.06	31.33	650	850
Ethnicity	American Indian/Alaska Native	5,070	720.31	25.62	650	843
	Asian	27,316	762.14	31.91	650	850
	Black or African American	107,759	720.25	26.13	650	850
	Hispanic/Latino	130,918	725.27	26.70	650	850
	Native Hawaiian or Pacific Islander	1,063	734.45	31.33	655	848
	Multiple Race Selected	47,058	746.59	31.46	650	850
	White	305,351	742.03	29.00	650	850
Economic Status*	Economically Disadvantaged	295,888	723.84	26.69	650	850
	Not Economically Disadvantaged	299,855	747.74	29.64	650	850
English Learner Status	English Learner (EL)	49,545	715.32	25.17	650	850
	Non English Learner	544,120	737.89	30.39	650	850
Disabilities	Students with Disabilities (SWD)	63,085	714.31	27.98	650	850
	Students without Disabilities	411,533	740.51	29.32	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.59 Subgroup Performance for Mathematics Scale Scores: Grade 5

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		640,102	735.54	30.15	650	850
Gender	Female	313,222	736.66	28.94	650	850
	Male	326,880	734.47	31.23	650	850
Ethnicity	American Indian/Alaska Native	5,036	721.65	26.05	650	828
	Asian	27,424	761.77	31.45	650	850
	Black or African American	107,918	720.03	25.75	650	850
	Hispanic/Latino	128,039	725.56	26.82	650	850
	Native Hawaiian or Pacific Islander	985	735.43	32.00	650	834
	Multiple Race Selected	46,414	745.52	31.55	650	850
	White	303,874	741.46	28.63	650	850
Economic Status*	Economically Disadvantaged	291,979	723.97	26.65	650	850
	Not Economically Disadvantaged	308,277	746.95	29.34	650	850
English Learner Status	English Learner (EL)	38,610	713.90	25.46	650	849
	Non English Learner	559,517	737.43	30.04	650	850
Disabilities	Students with Disabilities (SWD)	64,295	713.79	27.22	650	850
	Students without Disabilities	412,825	740.62	29.05	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.60 Subgroup Performance for Mathematics Scale Scores: Grade 6

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		632,127	735.11	29.74	650	850
Gender	Female	308,782	736.32	28.80	650	850
	Male	323,345	733.95	30.56	650	850
Ethnicity	American Indian/Alaska Native	4,851	719.93	26.10	650	824
	Asian	26,768	760.70	30.54	650	850
	Black or African American	107,385	718.31	25.64	650	850
	Hispanic/Latino	124,746	724.74	26.64	650	850
	Native Hawaiian or Pacific Islander	968	733.46	30.83	650	829
	Multiple Race Selected	45,820	745.75	30.37	650	850
	White	300,790	741.56	27.72	650	850
Economic Status*	Economically Disadvantaged	283,852	723.20	26.42	650	850
	Not Economically Disadvantaged	309,947	746.56	28.38	650	850
English Learner Status	English Learner (EL)	31,926	710.27	25.50	650	850
	Non English Learner	559,399	736.98	29.40	650	850
Disabilities	Students with Disabilities (SWD)	61,813	711.09	26.35	650	850
	Students without Disabilities	409,292	740.24	28.44	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.61 Subgroup Performance for Mathematics Scale Scores: Grade 7

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		608,990	733.95	27.15	650	850
Gender	Female	297,392	735.11	26.13	650	850
	Male	311,598	732.85	28.04	650	850
Ethnicity	American Indian/Alaska Native	4,884	718.72	23.52	650	805
	Asian	22,565	755.16	28.26	650	850
	Black or African American	105,246	719.26	23.53	650	850
	Hispanic/Latino	119,429	725.25	24.73	650	850
	Native Hawaiian or Pacific Islander	911	732.58	29.49	650	818
	Multiple Race Selected	44,191	744.16	28.30	650	850
	White	291,912	739.75	25.44	650	850
Economic Status*	Economically Disadvantaged	272,126	723.92	24.49	650	850
	Not Economically Disadvantaged	298,096	743.76	26.14	650	850
English Learner Status	English Learner (EL)	31,366	711.13	23.42	650	839
	Non English Learner	536,440	735.79	26.81	650	850
Disabilities	Students with Disabilities (SWD)	59,870	710.27	24.47	650	850
	Students without Disabilities	390,280	738.10	25.65	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.62 Subgroup Performance for Mathematics Scale Scores: Grade 8

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		504,561	728.21	35.33	650	850
Gender	Female	243,046	730.73	34.21	650	850
	Male	261,515	725.86	36.18	650	850
Ethnicity	American Indian/Alaska Native	4,245	711.00	28.71	650	827
	Asian	15,120	756.32	40.10	650	850
	Black or African American	96,669	712.02	30.65	650	850
	Hispanic/Latino	103,379	718.40	32.34	650	850
	Native Hawaiian or Pacific Islander	688	722.78	37.15	650	850
	Multiple Race Selected	38,270	745.02	38.77	650	850
	White	231,454	734.91	33.47	650	850
Economic Status*	Economically Disadvantaged	239,097	717.13	31.82	650	850
	Not Economically Disadvantaged	230,003	739.58	35.41	650	850
English Learner Status	English Learner (EL)	29,346	702.61	28.92	650	850
	Non English Learner	435,423	730.14	35.16	650	850
Disabilities	Students with Disabilities (SWD)	56,246	701.34	29.32	650	850
	Students without Disabilities	292,146	731.67	33.33	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.63 Subgroup Performance for Mathematics Scale Scores: Algebra I

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		480,604	733.96	32.23	650	850
Gender	Female	233,330	735.72	31.13	650	850
	Male	247,274	732.29	33.15	650	850
Ethnicity	American Indian/Alaska Native	4,024	718.89	25.82	650	822
	Asian	20,209	761.40	35.88	650	850
	Black or African American	93,290	718.84	26.79	650	850
	Hispanic/Latino	91,047	722.22	28.10	650	850
	Native Hawaiian or Pacific Islander	775	730.50	32.40	650	839
	Multiple Race Selected	18,497	740.96	35.21	650	850
	White	230,890	741.54	30.97	650	850
Economic Status*	Economically Disadvantaged	194,799	722.37	27.77	650	850
	Not Economically Disadvantaged	249,242	743.28	32.68	650	850
English Learner Status	English Learner (EL)	25,094	711.30	25.68	650	850
	Non English Learner	438,285	735.42	32.05	650	850
Disabilities	Students with Disabilities (SWD)	45,685	709.45	25.77	650	850
	Students without Disabilities	333,456	739.43	32.28	650	850

Note: This table is identical to Table 12.10 in Section 12.

*Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.64 Subgroup Performance for Mathematics Scale Scores: Geometry

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		205,219	732.35	26.75	650	850
Gender	Female	100,927	733.14	25.62	650	848
	Male	104,292	731.58	27.78	650	850
Ethnicity	American Indian/Alaska Native	2,886	719.85	20.09	650	793
	Asian	11,143	752.18	28.51	650	850
	Black or African American	23,874	716.81	21.57	650	809
	Hispanic/Latino	46,897	719.80	21.96	650	817
	Native Hawaiian or Pacific Islander	538	726.32	24.49	650	809
	Multiple Race Selected	4,829	737.25	25.87	650	849
	White	108,072	738.31	25.65	650	850
Economic Status*	Economically Disadvantaged	79,099	721.44	22.57	650	826
	Not Economically Disadvantaged	123,268	739.13	26.82	650	850
English Learner Status	English Learner (EL)	9,850	712.33	21.54	650	812
	Non English Learner	190,907	733.38	26.58	650	850
Disabilities	Students with Disabilities (SWD)	17,553	709.10	21.07	650	850
	Students without Disabilities	175,544	734.83	26.20	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.65 Subgroup Performance for Mathematics Scale Scores: Algebra II

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		186,890	719.06	36.79	650	850
Gender	Female	93,563	719.94	34.95	650	850
	Male	93,327	718.19	38.53	650	850
Ethnicity	American Indian/Alaska Native	2,328	710.09	28.39	650	810
	Asian	13,174	748.73	39.70	650	850
	Black or African American	27,655	700.86	29.87	650	848
	Hispanic/Latino	40,830	707.97	31.65	650	850
	Native Hawaiian or Pacific Islander	331	723.55	35.38	650	822
	Multiple Race Selected	7,547	715.29	38.61	650	850
	White	94,500	725.53	35.97	650	850
Economic Status*	Economically Disadvantaged	65,527	705.85	31.05	650	850
	Not Economically Disadvantaged	120,973	726.19	37.67	650	850
English Learner Status	English Learner (EL)	5,915	695.33	31.21	650	830
	Non English Learner	179,177	719.85	36.65	650	850
Disabilities	Students with Disabilities (SWD)	11,170	690.58	29.62	650	850
	Students without Disabilities	138,629	721.95	36.58	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.66 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics I

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		30,217	730.86	35.11	650	850
Gender	Female	14,546	732.49	34.06	650	850
	Male	15,671	729.34	35.99	650	850
Ethnicity	American Indian/Alaska Native	341	712.77	29.53	650	799
	Asian	826	746.95	38.94	650	850
	Black or African American	4,147	716.88	29.20	650	825
	Hispanic/Latino	9,212	718.92	30.37	650	842
	Native Hawaiian or Pacific Islander	48	730.42	31.81	668	800
	Multiple Race Selected	1,283	738.76	38.58	650	850
	White	13,336	741.21	34.67	650	850
Economic Status*	Economically Disadvantaged	15,591	719.96	30.39	650	850
	Not Economically Disadvantaged	14,424	742.42	35.80	650	850
English Learner Status	English Learner (EL)	2,670	705.67	25.54	650	842
	Non English Learner	25,207	734.75	34.84	650	850
Disabilities	Students with Disabilities (SWD)	3,515	701.55	26.69	650	821
	Students without Disabilities	18,880	737.05	34.79	650	850

Note: This table is identical to Table 12.11 in Section 12.

*Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.67 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics II

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		12,282	729.17	31.53	650	850
Gender	Female	6,004	729.00	30.71	650	850
	Male	6,278	729.32	32.30	650	850
Ethnicity	American Indian/Alaska Native	496	717.38	21.38	650	815
	Asian	441	745.13	36.16	664	850
	Black or African American	1,081	715.85	26.42	650	810
	Hispanic/Latino	4,108	715.71	24.45	650	827
	Native Hawaiian or Pacific Islander	18	714.17	37.40	650	775
	Multiple Race Selected	259	730.41	32.08	650	837
	White	5,654	740.53	31.79	650	850
Economic Status*	Economically Disadvantaged	5,455	717.52	25.67	650	841
	Not Economically Disadvantaged	6,678	737.94	32.07	650	850
English Learner Status	English Learner (EL)	801	707.06	21.76	650	826
	Non English Learner	9,598	733.99	31.73	650	850
Disabilities	Students with Disabilities (SWD)	947	705.30	23.43	650	814
	Students without Disabilities	8,920	735.55	31.62	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table A.12.68 Subgroup Performance for Mathematics Scale Scores: Integrated Mathematics III

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		8,396	717.54	38.73	650	850
Gender	Female	4,186	719.01	37.18	650	841
	Male	4,210	716.09	40.17	650	850
Ethnicity	American Indian/Alaska Native	183	709.03	30.02	650	792
	Asian	393	733.62	40.55	650	836
	Black or African American	671	705.78	35.70	650	811
	Hispanic/Latino	3,018	707.35	34.58	650	823
	Native Hawaiian or Pacific Islander	19	710.32	38.54	650	775
	Multiple Race Selected	380	713.37	39.91	650	831
	White	3,729	727.08	39.53	650	850
Economic Status*	Economically Disadvantaged	3,877	709.89	34.98	650	829
	Not Economically Disadvantaged	4,519	724.11	40.55	650	850
English Learner Status	English Learner (EL)	373	686.03	27.14	650	791
	Non English Learner	6,526	721.38	38.89	650	850
Disabilities	Students with Disabilities (SWD)	432	680.39	28.56	650	819
	Students without Disabilities	4,762	720.55	39.22	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Addendum: Statistical Summary of the Fall 2014 Administration

The addendum presents the results of analyses for the Fall 2014 operational administration. These results are reported separately from the Spring 2015 results because fall testing involved a nonrepresentative subset of students testing only for ELA/L grades 9, 10, and 11, as well as Algebra I, Geometry, and Algebra II. Only one paper test form was administered for each test.

To organize the addendum, tables are numbered sequentially according to the section represented by the tables. The reader can refer back to the corresponding section in the technical report for related information on the topic. For example, the first addendum table for Section 5 is numbered ADD.5.1, the second addendum table for Section 5 is numbered ADD.5.2, and so on.

Addendum 5: Test Taker Characteristics

Table ADD.5.1 State Participation in ELA/L Fall 2014 Operational Tests, by Grade

State	Category	English Language Arts/Literacy			
		Total	Grade 9	Grade 10	Grade 11
PARCC	N of Students	13,295	4,184	7,910	1,201
AR	N of Students	808	310	336	162
AR	% of PARCC Data	6.1	7.4	4.2	13.5
IL	N of Students	1,035	0	0	1,035
IL	% of PARCC Data	7.8	0.0	0.0	86.2
MD	N of Students	3,498	10	3,484	4
MD	% of PARCC Data	26.3	0.2	44.0	0.3
MS	N of Students	3,874	0	3,874	0
MS	% of PARCC Data	29.1	0.0	49.0	0.0
OH	N of Students	3,928	3,850	78	0
OH	% of PARCC Data	29.5	92.0	1.0	0.0
RI	N of Students	152	14	138	0
RI	% of PARCC Data	1.1	0.3	1.7	0.0

Table ADD.5.2 State Participation in Mathematics Fall 2014 Operational Tests, by Grade

State	Category	Mathematics			
		Total	A1	GO	A2
PARCC	N of Students	13,202	8,308	1,294	3,600
AR	N of Students	911	247	492	172
AR	% of PARCC Data	6.9	3.0	38.0	4.8
IL	N of Students	665	0	0	665
IL	% of PARCC Data	5.0	0.0	0.0	18.5
MD	N of Students	5,723	2,984	0	2,739
MD	% of PARCC Data	43.3	35.9	0.0	76.1
MS	N of Students	2,201	2,201	0	0
MS	% of PARCC Data	16.7	26.5	0.0	0.0
OH	N of Students	3,502	2,792	686	24
OH	% of PARCC Data	26.5	33.6	53.0	0.7
RI	N of Students	200	84	116	0
RI	% of PARCC Data	1.5	1.0	9.0	0.0

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra.

Table ADD.5.3 All States Combined: Fall 2014 ELA/L Test Takers by Grade and Gender

Grade	Mode	Valid Cases	Gender			
			Female		Male	
		N	N	%	N	%
9	PBT	4,184	2,203	52.7	1,981	47.3
10	PBT	7,910	4,078	51.6	3,832	48.4
11	PBT	1,201	572	47.6	629	52.4

Table ADD.5.4 All States Combined: Fall 2014 Mathematics Test Takers by Grade and Gender

Grade	Mode	Valid Cases	Gender			
			Female		Male	
		N	N	%	N	%
A1	PBT	8,308	4,152	50.0	4,156	50.0
GO	PBT	1,294	650	50.2	644	49.8
A2	PBT	3,600	1,801	50.0	1,799	50.0

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra.

Table ADD.5.5 Fall 2014 ELA/L Test Takers by State, Grade, and Gender

State	Grade	Mode	Valid Cases	Gender			
				Female		Male	
				N	%	N	%
AR	9	PBT	310	159	51.3	151	48.7
AR	10	PBT	336	167	49.7	169	50.3
AR	11	PBT	162	84	51.9	78	48.1
IL	11	PBT	1,035	488	47.1	547	52.9
MD	9	PBT	10	n/a	n/a	10	100.0
MD	10	PBT	3,484	1,695	48.7	1,789	51.3
MD	11	PBT	4	n/a	n/a	4	100.0
MS	10	PBT	3,874	2,100	54.2	1,774	45.8
OH	9	PBT	3,850	2,036	52.9	1,814	47.1
OH	10	PBT	78	49	62.8	29	37.2
RI	9	PBT	14	8	57.1	6	42.9
RI	10	PBT	138	67	48.6	71	51.4

Note: n/a = not applicable.

Table ADD.5.6 Fall 2014 Mathematics Test Takers by State, Content Area, and Gender

State	Grade	Mode	Valid Cases	Gender			
				Female		Male	
				N	%	N	%
AR	A1	PBT	247	124	50.2	123	49.8
AR	GO	PBT	492	239	48.6	253	51.4
AR	A2	PBT	172	83	48.3	89	51.7
IL	A2	PBT	665	330	49.6	335	50.4
MD	A1	PBT	2,984	1,463	49.0	1,521	51.0
MD	A2	PBT	2,739	1,378	50.3	1,361	49.7
MS	A1	PBT	2,201	1,106	50.2	1,095	49.8
OH	A1	PBT	2,792	1,419	50.8	1,373	49.2
OH	GO	PBT	686	351	51.2	335	48.8
OH	A2	PBT	24	10	41.7	14	58.3
RI	A1	PBT	84	40	47.6	44	52.4
RI	GO	PBT	116	60	51.7	56	48.3

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra.

Table ADD.5.7 Demographic Information for Fall 2014 Grade 9 ELA/L, Overall and by State

Demographic	PARCC	AR	IL	MD	MS	OH	RI
Econ Dis (%)	34.3	34.8	n/a	0.0	n/a	34.2	71.4
SWD (%)	7.8	1.3	n/a	0.0	n/a	8.3	21.4
EL (%)	1.4	0.3	n/a	0.0	n/a	1.5	7.1
Male (%)	47.3	48.7	n/a	100.0	n/a	47.1	42.9
Female (%)	52.7	51.3	n/a	0.0	n/a	52.9	57.1
AmInd/ANat (%)	0.2	0.6	n/a	0.0	n/a	0.2	7.1
Asian (%)	1.1	0.0	n/a	0.0	n/a	1.1	0.0
Black/AA (%)	15.5	10.6	n/a	70.0	n/a	15.8	21.4
Hisp/Lat (%)	3.2	2.9	n/a	10.0	n/a	3.1	35.7
Wh/Caus (%)	43.5	60.6	n/a	20.0	n/a	42.2	28.6
NtvHawaii/Pacific (%)	0.0	0.0	n/a	0.0	n/a	0.1	0.0
Two Or More (%)	1.0	0.0	n/a	0.0	n/a	1.1	7.1
Unknown (%)	35.5	25.2	n/a	0.0	n/a	36.5	0.0

Note: Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table ADD.5.8 Demographic Information for Fall 2014 Grade 10 ELA/L, Overall and by State

Demographic	PARCC	AR	IL	MD	MS	OH	RI
Econ Dis (%)	20.8	38.1	n/a	28.2	12.7	24.4	15.9
SWD (%)	2.2	5.7	n/a	1.6	2.1	0.0	10.9
EL (%)	0.6	2.7	n/a	0.3	0.5	0.0	2.9
Male (%)	48.4	50.3	n/a	51.3	45.8	37.2	51.4
Female (%)	51.6	49.7	n/a	48.7	54.2	62.8	48.6
AmInd/ANat (%)	0.2	0.6	n/a	0.3	0.2	0.0	0.0
Asian (%)	1.7	0.3	n/a	2.6	1.2	1.3	0.0
Black/AA (%)	13.7	13.1	n/a	12.9	15.2	3.8	0.7
Hisp/Lat (%)	4.3	7.7	n/a	6.4	1.9	7.7	8.7
Wh/Caus (%)	64.0	68.8	n/a	74.1	53.7	55.1	90.6
NtvHawaii/Pacific (%)	0.1	0.0	n/a	0.1	0.0	0.0	0.0
Two Or More (%)	1.5	0.3	n/a	3.4	0.0	0.0	0.0
Unknown (%)	14.4	9.2	n/a	0.1	27.8	32.1	0.0

Note: Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawaii/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table ADD.5.9 Demographic Information for Fall 2014 Grade 11 ELA/L, Overall and by State

Demographic	PARCC	AR	IL	MD	MS	OH	RI
Econ Dis (%)	39.0	36.4	39.5	0.0	n/a	n/a	n/a
SWD (%)	10.6	16.7	9.7	0.0	n/a	n/a	n/a
EL (%)	0.4	3.1	0.0	0.0	n/a	n/a	n/a
Male (%)	52.4	48.1	52.9	100.0	n/a	n/a	n/a
Female (%)	47.6	51.9	47.1	0.0	n/a	n/a	n/a
AmInd/ANat (%)	0.5	0.0	0.6	0.0	n/a	n/a	n/a
Asian (%)	2.3	0.6	2.6	0.0	n/a	n/a	n/a
Black/AA (%)	5.9	4.9	5.9	50.0	n/a	n/a	n/a
Hisp/Lat (%)	18.2	6.2	20.1	25.0	n/a	n/a	n/a
Wh/Caus (%)	69.4	85.8	67.0	25.0	n/a	n/a	n/a
NtvHawaii/Pacific (%)	0.0	0.0	0.0	0.0	n/a	n/a	n/a
Two Or More (%)	3.4	0.6	3.9	0.0	n/a	n/a	n/a
Unknown (%)	0.2	1.9	0.0	0.0	n/a	n/a	n/a

Note: Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawaii/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table ADD.5.10 Demographic Information for Fall 2014 Algebra I, Overall and by State

Demographic	PARCC	AR	IL	MD	MS	OH	RI
Econ Dis (%)	28.4	49.4	n/a	27.2	19.0	35.4	25.0
SWD (%)	3.4	0.8	n/a	1.4	3.3	5.9	3.6
EL (%)	1.1	5.7	n/a	0.7	0.4	1.6	1.2
Male (%)	50.0	49.8	n/a	51.0	49.8	49.2	52.4
Female (%)	50.0	50.2	n/a	49.0	50.2	50.8	47.6
AmInd/ANat (%)	0.3	3.6	n/a	0.4	0.1	0.2	0.0
Asian (%)	1.9	3.2	n/a	3.0	1.1	1.1	0.0
Black/AA (%)	15.9	11.3	n/a	12.3	18.9	18.2	4.8
Hisp/Lat (%)	4.3	13.0	n/a	6.9	2.4	1.9	13.1
Wh/Caus (%)	60.0	49.0	n/a	74.1	65.5	41.4	69.0
NtvHawaii/Pacific (%)	0.1	0.0	n/a	0.1	0.0	0.1	0.0
Two Or More (%)	1.4	0.0	n/a	3.2	0.2	0.6	3.6
Unknown (%)	16.0	19.8	n/a	0.0	11.8	36.4	9.5

Note: Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table ADD.5.11 Demographic Information for Fall 2014 Geometry, Overall and by State

Demographic	PARCC	AR	IL	MD	MS	OH	RI
Econ Dis (%)	29.7	31.9	n/a	n/a	n/a	26.1	41.4
SWD (%)	2.2	1.2	n/a	n/a	n/a	0.6	16.4
EL (%)	2.3	4.3	n/a	n/a	n/a	1.2	0.9
Male (%)	49.8	51.4	n/a	n/a	n/a	48.8	48.3
Female (%)	50.2	48.6	n/a	n/a	n/a	51.2	51.7
AmInd/ANat (%)	0.9	2.0	n/a	n/a	n/a	0.1	0.9
Asian (%)	1.7	2.2	n/a	n/a	n/a	1.0	3.4
Black/AA (%)	6.0	6.5	n/a	n/a	n/a	5.7	5.2
Hisp/Lat (%)	8.4	11.2	n/a	n/a	n/a	2.8	30.2
Wh/Caus (%)	56.6	64.8	n/a	n/a	n/a	50.1	59.5
NtvHawaii/Pacific (%)	0.1	0.2	n/a	n/a	n/a	0.0	0.0
Two Or More (%)	0.5	0.0	n/a	n/a	n/a	1.0	0.0
Unknown (%)	25.8	13.0	n/a	n/a	n/a	39.2	0.9

Note: Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Table ADD.5.12 Demographic Information for Fall 2014 Algebra II, Overall and by State

Demographic	PARCC	AR	IL	MD	MS	OH	RI
Econ Dis (%)	27.5	33.7	36.1	25.2	n/a	8.3	n/a
SWD (%)	3.9	5.2	6.5	3.2	n/a	0.0	n/a
EL (%)	0.2	1.2	0.0	0.1	n/a	0.0	n/a
Male (%)	50.0	51.7	50.4	49.7	n/a	58.3	n/a
Female (%)	50.0	48.3	49.6	50.3	n/a	41.7	n/a
AmInd/ANat (%)	0.4	1.7	0.3	0.4	n/a	0.0	n/a
Asian (%)	2.9	1.7	3.2	2.7	n/a	25.0	n/a
Black/AA (%)	13.8	2.9	7.8	16.1	n/a	4.2	n/a
Hisp/Lat (%)	7.6	3.5	15.6	6.0	n/a	8.3	n/a
Wh/Caus (%)	71.7	86.6	69.0	71.5	n/a	58.3	n/a
NtvHawaii/Pacific (%)	0.1	0.6	0.0	0.1	n/a	0.0	n/a
Two Or More (%)	2.6	0.6	4.1	2.3	n/a	0.0	n/a
Unknown (%)	0.9	2.3	0.0	1.0	n/a	4.2	n/a

Note: Econ Dis = Economically Disadvantaged; SWD = Student with Disabilities; EL = English learner; AmInd/ANat = American Indian/Alaska Native; Black/AA = Black/African American; Hisp/Lat = Hispanic/Latino; Wh/Caus = White/Caucasian; NtvHawai/Pacific = Native Hawaiian or Other Pacific Islander; Two or More = two or more races reported; Unknown = not reported. n/a = not applicable.

Addendum 6: Classical Item Analysis

Table ADD.6.1 Summary of p Values for Fall 2014 ELA/L and Mathematics Operational Items, by Grade and Component

Grade	Component	N of Items	Mean p Value	SD p Value	Min p Value	Max p Value	Median p Value
9	PBA	22	0.43	0.17	0.17	0.73	0.41
9	EOY	22	0.48	0.18	0.16	0.79	0.47
10	PBA	23	0.52	0.15	0.17	0.80	0.49
10	EOY	22	0.62	0.13	0.34	0.79	0.63
11	PBA	23	0.45	0.14	0.16	0.68	0.43
11	EOY	22	0.41	0.09	0.22	0.65	0.41
A1	PBA	17	0.27	0.14	0.03	0.64	0.25
A1	EOY	35	0.28	0.19	0.04	0.69	0.24
GO	PBA	18	0.35	0.18	0.09	0.61	0.35
GO	EOY	35	0.42	0.20	0.08	0.83	0.42
A2	PBA	19	0.22	0.16	0.06	0.67	0.17
A2	EOY	32	0.30	0.18	0.05	0.74	0.26

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Table ADD.6.2 Summary of Polyserial Correlations for Fall 2014 ELA/L and Mathematics Operational Items, by Grade and Test

Grade	Component	N of Items	Mean Polyserial	SD Polyserial	Min Polyserial	Max Polyserial	Median Polyserial
9	PBA	22	0.19	0.09	0.04	0.31	0.21
9	EOY	22	0.25	0.07	0.07	0.34	0.28
10	PBA	23	0.22	0.08	0.05	0.31	0.24
10	EOY	22	0.29	0.05	0.16	0.35	0.30
11	PBA	23	0.21	0.09	0.05	0.31	0.24
11	EOY	22	0.28	0.05	0.12	0.38	0.27
A1	PBA	17	0.30	0.16	0.12	0.61	0.21
A1	EOY	35	0.37	0.15	0.04	0.64	0.36
GO	PBA	18	0.37	0.19	0.13	0.70	0.38
GO	EOY	35	0.46	0.19	0.16	0.83	0.44
A2	PBA	19	0.33	0.21	0.09	0.66	0.23
A2	EOY	32	0.37	0.16	0.11	0.65	0.34

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Addendum 7: Differential Item Functioning

Table ADD.7.1 Differential Item Functioning for Fall 2014 High School ELA/L

Grade	DIF Comparisons	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF		
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	
9	Male vs Female	45	45	45	2	4	3	7	39	87	0	0	1	2	
	White vs AmerIndian	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Asian	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Black	45	45	45	0	0	1	2	44	98	0	0	0	0	
	White vs Hispanic	45	45	45	0	0	2	4	43	96	0	0	0	0	
	White vs Pacific Islander	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Multiracial	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	NoEcnDis vs EcnDis	45	45	45	0	0	0	0	45	100	0	0	0	0	
	ELN vs ELY	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	SWDN vs SWDY	45	45	45	0	0	2	4	43	96	0	0	0	0	
10	Male vs Female	45	45	45	0	0	2	4	41	91	2	4	0	0	
	White vs AmerIndian	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Asian	45	45	45	0	0	4	9	39	87	2	4	0	0	
	White vs Black	45	45	45	0	0	1	2	44	98	0	0	0	0	
	White vs Hispanic	45	45	45	0	0	0	0	45	100	0	0	0	0	
	White vs Pacific Islander	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	

Grade	DIF Comparisons	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF	
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF
	White vs Multiracial	45	45	45	0	0	1	2	43	96	0	0	1	2
	NoEcnDis vs EcnDis	45	45	45	0	0	0	0	45	100	0	0	0	0
	ELN vs ELY	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	SWDN vs SWDY	45	45	45	0	0	1	2	44	98	0	0	0	0
	Male vs Female	45	45	45	0	0	3	7	42	93	0	0	0	0
	White vs AmerIndian	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	White vs Asian	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	White vs Black	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	White vs Hispanic	45	45	45	2	4	5	11	38	84	0	0	0	0
	White vs Pacific Islander	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
11	White vs Multiracial	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	NoEcnDis vs EcnDis	45	45	45	0	0	1	2	44	98	0	0	0	0
	ELN vs ELY	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	SWDN vs SWDY	45	45	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note: AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability(ies).

Table ADD.7.2 Differential Item Functioning for Fall 2014 High School Mathematics

Grade	DIF Comparisons	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF		
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	
A1	Male vs Female	53	53	53	0	0	0	0	52	98	1	2	0	0	
	White vs AmerIndian	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Asian	53	53	53	0	0	0	0	51	96	2	4	0	0	
	White vs Black	53	53	53	0	0	1	2	50	94	2	4	0	0	
	White vs Hispanic	53	53	53	0	0	0	0	52	98	1	2	0	0	
	White vs Pacific Islander	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Multiracial	53	53	52	0	0	0	0	50	96	2	4	0	0	
	NoEcnDis vs EcnDis	53	53	53	0	0	0	0	53	100	0	0	0	0	
	ELN vs ELY	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	SWDN vs SWDY	53	53	53	0	0	4	8	47	89	2	4	0	0	
GO	Male vs Female	53	53	53	0	0	3	6	48	91	2	4	0	0	
	White vs AmerIndian	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Asian	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Black	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Hispanic	53	53	52	0	0	1	2	51	98	0	0	0	0	
	White vs Pacific Islander	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Multiracial	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	

Grade	DIF Comparisons	Total N of Unique Items	Total N of Item Occurrences	Total N of Item Occurrences Included in DIF Analysis	C- DIF		B- DIF		A DIF		B+ DIF		C+ DIF		
					N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	N of Occurrences	% of Total Occurrences in DIF	
	NoEcnDis vs EcnDis	53	53	53	0	0	3	6	50	94	0	0	0	0	
	ELN vs ELY	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	SWDN vs SWDY	53	53	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
A2	Male vs Female	54	54	54	0	0	1	2	53	98	0	0	0	0	
	White vs AmerIndian	54	54	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Asian	54	54	52	0	0	1	2	51	98	0	0	0	0	
	White vs Black	54	54	54	1	2	1	2	52	96	0	0	0	0	
	White vs Hispanic	54	54	54	0	0	0	0	53	98	1	2	0	0	
	White vs Pacific Islander	54	54	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	White vs Multiracial	54	54	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
	NoEcnDis vs EcnDis	54	54	54	0	0	0	0	54	100	0	0	0	0	
	ELN vs ELY	54	54	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	SWDN vs SWDY	54	54	40	0	0	1	3	37	93	0	0	2	5	

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, AmerIndian = American Indian/Alaska Native, Black = Black/African American, Hispanic = Hispanic/Latino, Pacific Islander = Native Hawaiian or Pacific Islander, Multiracial = Multiple Race Selected, NoEcnDis = not economically disadvantaged, EcnDis = economically disadvantaged, ELN = not an English learner, ELY = English learner, SWDN = not student with disability(ies), SWDY = student with disability(ies).

Addendum 8: Reliability

Table ADD.8.1 shows that at the total group level, the Fall 2014 forms have reliability estimates within .01 of their corresponding Spring 2015 PBT forms. The Fall 2014 scale score SEM estimates are at least 10% lower for the English Language/Literacy Grade 10, Grade 11, and Mathematics Geometry than for the Spring 2015 PBT SEM estimates.

Even though considerably fewer students took the Fall 2014 test form, the reliability and SEM estimates for subgroups found in Tables ADD.8.2-ADD.8.7 for the Fall 2014 test forms show nothing unusual given the lower scale score SEMs indicated above and the Spring 2015 subgroup estimates. Table ADD.8.8 and ADD.8.9 provide the claim and subclaim reliability and raw score SEM estimates for the Fall 2014 forms. Given the sample sizes and the number of items for each claim and subclaim, the reliability and SEM estimates are not unusual given the Spring 2015 results.

Table ADD.8.1 Summary of Test Reliability Estimates for Fall 2014 Total Group

Content	Grade/ Course	Sample Size	Maximum Possible Score	Reliability	Raw Score SEM	Scale Score SEM
ELA/L	9	4,034	135	0.91	6.59	9.86
ELA/L	10	7,609	137	0.93	6.42	9.74
ELA/L	11	1,093	137	0.92	7.12	9.99
Mathematics	A1	6,733	96	0.90	4.25	9.19
Mathematics	GO	1,184	97	0.94	4.42	5.68
Mathematics	A2	2,640	104	0.92	4.39	10.24

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra.

Table ADD.8.2 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Grade 9 ELA/L

	PBT			
	Sample Size	Maximum Possible Raw Score	Reliability	Scale Score Standard Error of Measurement
Total Group	4,034	135	0.91	9.86
Gender				
Male	1,877	135	0.91	9.70
Female	2,157	135	0.90	9.99
Unknown/Missing	-	-	-	-
Ethnicity				
White	1,774	135	0.92	9.81
African American	594	135	0.89	9.89
Asian/Pacific Islander	-	-	-	-
American Indian/Alaska Native	-	-	-	-
Hispanic	131	135	0.90	10.36
Multiple	-	-	-	-
Special Instructional Needs				
Economically Disadvantaged	1,356	135	0.90	9.76
Not-economically Disadvantaged	1,734	135	0.91	10.00
English Learner (EL)	-	-	-	-
Non English Learner	2,835	135	0.91	9.88
Students with Disabilities (SWD)	329	135	0.89	9.26
Students without Disabilities	2,375	135	0.90	9.87
Students Taking Accommodated Forms				
A: ASL	-	-	-	-
C: Closed-Caption	-	-	-	-
R: Screen Reader	-	-	-	-
T: Text-to-Speech	-	-	-	-

Table ADD.8.3 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Grade 10 ELA/L

	PBT			
	Sample Size	Maximum Possible Raw Score	Reliability	Scale Score Standard Error of Measurement
Total Group	7,609	137	0.93	9.74
Gender				
Male	3,633	137	0.94	9.60
Female	3,976	137	0.93	9.86
Unknown/Missing	-	-	-	-
Ethnicity				
White	4,883	137	0.93	9.80
African American	1,011	137	0.93	9.44
Asian/Pacific Islander	139	137	0.94	9.87
American Indian/Alaska Native	-	-	-	-
Hispanic	325	137	0.94	9.50
Multiple	109	137	0.94	10.29
Special Instructional Needs				
Economically Disadvantaged	1,523	137	0.93	9.53
Not-economically Disadvantaged	4,582	137	0.93	9.82
English Learner (EL)	-	-	-	-
Non English Learner	6,252	137	0.94	9.74
Students with Disabilities (SWD)	147	137	0.93	9.15
Students without Disabilities	3,036	137	0.92	9.86
Students Taking Accommodated Forms				
A: ASL	-	-	-	-
C: Closed-Caption	-	-	-	-
R: Screen Reader	-	-	-	-
T: Text-to-Speech	-	-	-	-

Table ADD.8.4 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Grade 11 ELA/L

	PBT			
	Sample Size	Maximum Possible Raw Score	Reliability	Scale Score Standard Error of Measurement
Total Group	1,093	137	0.92	9.99
Gender				
Male	564	137	0.92	9.91
Female	529	137	0.91	10.02
Unknown/Missing	-	-	-	-
Ethnicity				
White	774	137	0.91	9.88
African American	-	-	-	-
Asian/Pacific Islander	-	-	-	-
American Indian/Alaska Native	-	-	-	-
Hispanic	196	137	0.91	10.17
Multiple	-	-	-	-
Special Instructional Needs				
Economically Disadvantaged	418	137	0.90	9.95
Not-economically Disadvantaged	666	137	0.91	10.01
English Learner (EL)	-	-	-	-
Non English Learner	155	137	0.94	10.01
Students with Disabilities (SWD)	-	-	-	-
Students without Disabilities	867	137	0.91	10.03
Students Taking Accommodated Forms				
A: ASL	-	-	-	-
C: Closed-Caption	-	-	-	-
R: Screen Reader	-	-	-	-
T: Text-to-Speech	-	-	-	-

Table ADD.8.5 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Algebra I

	PBT			
	Sample Size	Maximum Possible Raw Score	Reliability	Scale Score Standard Error of Measurement
Total Group	6,733	96	0.90	9.19
Gender				
Male	3,198	96	0.90	9.23
Female	3,535	96	0.89	9.13
Unknown/Missing	-	-	-	-
Ethnicity				
White	4,158	96	0.89	9.18
African American	905	96	0.86	9.75
Asian/Pacific Islander	131	96	0.91	8.39
American Indian/Alaska Native	-	-	-	-
Hispanic	272	96	0.88	9.52
Multiple	-	-	-	-
Special Instructional Needs				
Economically Disadvantaged	1,767	96	0.86	9.55
Not-economically Disadvantaged	3,796	96	0.90	9.14
English Learner (EL)	-	-	-	-
Non English Learner	5,501	96	0.90	9.19
Students with Disabilities (SWD)	245	96	0.87	8.98
Students without Disabilities	3,130	96	0.91	9.00
Students Taking Accommodated Forms				
A: ASL	-	-	-	-
C: Closed-Caption	-	-	-	-
R: Screen Reader	-	-	-	-
T: Text-to-Speech	-	-	-	-
Students Taking Translated Forms				
Spanish Language Form	-	-	-	-

Table ADD.8.6 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Geometry

	PBT			
	Sample Size	Maximum Possible Raw Score	Reliability	Scale Score Standard Error of Measurement
Total Group	1,184	97	0.94	5.68
Gender				
Male	575	97	0.94	5.72
Female	609	97	0.93	5.63
Unknown/Missing	-	-	-	-
Ethnicity				
White	680	97	0.93	5.54
African American	-	-	-	-
Asian/Pacific Islander	-	-	-	-
American Indian/Alaska Native	-	-	-	-
Hispanic	-	-	-	-
Multiple	-	-	-	-
Special Instructional Needs				
Economically Disadvantaged	329	97	0.91	5.80
Not-economically Disadvantaged	585	97	0.93	5.55
English Learner (EL)	-	-	-	-
Non English Learner	826	97	0.94	5.60
Students with Disabilities (SWD)	-	-	-	-
Students without Disabilities	542	97	0.93	5.54
Students Taking Accommodated Forms				
A: ASL	-	-	-	-
C: Closed-Caption	-	-	-	-
R: Screen Reader	-	-	-	-
T: Text-to-Speech	-	-	-	-
Students Taking Translated Forms				
Spanish Language Form	-	-	-	-

Table ADD.8.7 Summary of Test Reliability Estimates for Fall 2014 Subgroups: Algebra II

	PBT			
	Sample Size	Maximum Possible Raw Score	Reliability	Scale Score Standard Error of Measurement
Total Group	2,640	104	0.92	10.24
Gender				
Male	1,266	104	0.93	10.29
Female	1,374	104	0.91	10.17
Unknown/Missing	-	-	-	-
Ethnicity				
White	2,007	104	0.92	10.15
African American	244	104	0.85	10.93
Asian/Pacific Islander	-	-	-	-
American Indian/Alaska Native	-	-	-	-
Hispanic	201	104	0.91	10.53
Multiple	-	-	-	-
Special Instructional Needs				
Economically Disadvantaged	614	104	0.89	10.84
Not-economically Disadvantaged	2,014	104	0.92	10.06
English Learner (EL)	-	-	-	-
Non English Learner	2,126	104	0.92	10.19
Students with Disabilities (SWD)	-	-	-	-
Students without Disabilities	983	104	0.91	10.43
Students Taking Accommodated Forms				
A: ASL	-	-	-	-
C: Closed-Caption	-	-	-	-
R: Screen Reader	-	-	-	-
T: Text-to-Speech	-	-	-	-
Students Taking Translated Forms				
Spanish Language Form	-	-	-	-

Table ADD.8.8 Fall 2014 Average ELA/L Reliability Estimates for Reading Total and Subscores

Grade Level	Reading: Total		Reading: Literature		Reading: Information		Reading: Vocabulary	
	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability
9	90	0.88	36	0.77	32	0.75	22	0.59
10	92	0.92	28	0.73	44	0.87	20	0.70
11	92	0.91	34	0.78	32	0.81	26	0.68

Table ADD.8.9 Fall 2014 Average ELA/L Reliability Estimates for Writing Total and Subscores

Grade Level	Writing: Total		Writing: Written Expression		Writing: Knowledge Language and Conventions	
	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability
9	45	0.83	36	0.78	9	0.80
10	45	0.84	36	0.82	9	0.83
11	45	0.83	36	0.80	9	0.80

Table ADD.8.10 Fall 2014 Average Mathematics Reliability Estimates for Total Test and Subscores

Grade Level	Major Content		Additional & Supporting Content		Mathematics Reasoning		Modeling Practice	
	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability	Max Possible Raw Score	Reliability
A1	38	0.76	26	0.76	14	0.51	18	0.69
GO	39	0.88	26	0.80	14	0.62	18	0.72
A2	30	0.81	32	0.81	18	0.56	24	0.72

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra.

Addendum 9: Validity

The intercorrelations for the Fall 2014 tests are presented in Tables ADD.9.1 through ADD.9.3 for ELA/L grades 9, 10, and 11 and Tables ADD.9.4 through ADD.9.6 for the traditional mathematics courses (A1, GO, A2). Like the spring intercorrelations, the ELA/L all have moderate to high values with the writing subclaims being highly intercorrelated. The mathematics intercorrelations have moderate values. Tables ADD.9.7 through ADD.9.9 are the correlations between ELA/L and mathematics from the fall block.

Table ADD.9.1 Average Intercorrelations and Reliability between Grade 9 ELA/L Reading Subclaims

	PBT						
	RD	RL	RI	RV	WR	WE	WKL
RD	0.88	4,175	4,175	4,175	4,175	4,175	4,175
RL	0.92	0.77	4,175	4,175	4,175	4,175	4,175
RI	0.91	0.73	0.75	4,175	4,175	4,175	4,175
RV	0.83	0.65	0.66	0.59	4,175	4,175	4,175
WR	0.77	0.71	0.73	0.58	0.83	4,174	4,174
WE	0.76	0.70	0.72	0.58	1.00	0.78	4,174
WKL	0.76	0.70	0.72	0.58	0.96	0.93	0.80

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table ADD.9.2 Average Intercorrelations and Reliability between Grade 10 ELA/L Reading Subclaims

	PBT						
	RD	RL	RI	RV	WR	WE	WKL
RD	0.92	7,897	7,897	7,897	7,897	7,897	7,897
RL	0.90	0.73	7,897	7,897	7,897	7,897	7,897
RI	0.96	0.79	0.87	7,897	7,897	7,897	7,897
RV	0.87	0.71	0.78	0.70	7,897	7,897	7,897
WR	0.79	0.75	0.76	0.65	0.84	7,896	7,896
WE	0.78	0.74	0.75	0.64	1.00	0.82	7,896
WKL	0.78	0.73	0.75	0.64	0.96	0.93	0.83

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table ADD.9.3 Average Intercorrelations and Reliability between Grade 11 ELA/L Reading Subclaims

	PBT						
	RD	RL	RI	RV	WR	WE	WKL
RD	0.91	1,199	1,199	1,199	1,199	1,199	1,199
RL	0.93	0.78	1,199	1,199	1,199	1,199	1,199
RI	0.93	0.78	0.81	1,199	1,199	1,199	1,199
RV	0.88	0.75	0.71	0.68	1,199	1,199	1,199
WR	0.77	0.73	0.76	0.61	0.83	1,199	1,199
WE	0.77	0.72	0.75	0.61	1.00	0.80	1,199
WKL	0.77	0.72	0.75	0.62	0.98	0.96	0.80

Note: RD = Reading, RL = Reading Literature, RI = Reading Information, RV = Reading Vocabulary WR = Writing, WE = Written Expression, and WKL = Writing Knowledge and Conventions. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table ADD.9.4 Average Intercorrelations and Reliability between Algebra I Subclaims

	PBT			
	MC	ASC	MR	MP
MC	0.76	8,248	8,248	8,248
ASC	0.73	0.76	8,248	8,248
MR	0.60	0.56	0.51	8,146
MP	0.64	0.64	0.54	0.69

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table ADD.9.5 Average Intercorrelations and Reliability between Geometry Subclaims

	PBT			
	MC	ASC	MR	MP
MC	0.88	1,291	1,291	1,291
ASC	0.82	0.80	1,291	1,291
MR	0.73	0.69	0.62	1,289
MP	0.79	0.73	0.67	0.72

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table ADD.9.6 Average Intercorrelations and Reliability between Algebra II Subclaims

	PBT			
	MC	ASC	MR	MP
MC	0.81	3,573	3,573	3,573
ASC	0.82	0.81	3,573	3,573
MR	0.70	0.67	0.56	3,559
MP	0.73	0.73	0.69	0.72

Note: MC = Major Content, ASC = Additional and Supporting Content, MR = Mathematical Reasoning, and MP = Modeling Practice. The shaded values along the diagonal are the reliabilities as reported in Section 8. The average intercorrelations are provided in the lower portion of the table and the average sample sizes are provided in the upper portion of the table.

Table ADD.9.7 Average Correlations between ELA/L and Mathematics for High School

WR	PBT		
	A1	GO	A2
9	0.72 (1,472)	0.58 (308)	n/a
10	0.65 (542)	0.63 (184)	0.57 (314)
11	n/a	n/a	0.7 (244)

Note: ELA/L = English Language Arts/Literacy, A1 = Algebra I, GO = Geometry, A2 = Algebra II. The average correlations are provided with the average sample sizes are provided below in the parentheses. n/a = n-counts < 100; correlations not calculated.

Table ADD.9.8 Average Correlations between Reading and Mathematics for High School

WR	PBT		
	A1	GO	A2
9	0.69 (1,472)	0.62 (308)	n/a
10	0.64 (542)	0.65 (184)	0.54 (314)
11	n/a	n/a	0.68 (244)

Note: ELA/L = English Language Arts/Literacy, A1 = Algebra I, GO = Geometry, A2 = Algebra II. The average correlations are provided with the average sample sizes are provided below in the parentheses. n/a = n-counts < 100; correlations not calculated.

Table ADD.9.9 Average Correlations between Writing and Mathematics for High School

WR	PBT		
	A1	GO	A2
9	0.65 (1,472)	0.41 (308)	n/a
10	0.57 (542)	0.49 (184)	0.51 (314)
11	n/a	n/a	0.62 (244)

Note: ELA/L = English Language Arts/Literacy, A1 = Algebra I, GO = Geometry, A2 = Algebra II. The average correlations are provided with the average sample sizes are provided below in the parentheses. n/a = n-counts < 100; correlations not calculated.

Addendum 10: IRT Calibration and Scaling for the Fall 2014 Tests

In addition to the Spring 2015 administrations, operational administrations for high school ELA/L and mathematics tests were administered in Fall 2014. The six tests included ELA/L Grades 9, 10, and 11, and mathematics Algebra I, Geometry, and Algebra II. Integrated Mathematics tests were not administered as part of the Fall 2014 administrations. The operational tests administered in Fall 2014 will be referred to as the 2014 Fall block tests. Only one version of each Performance Based Assessment (PBA) and End-of-Year assessment (EOY) was administered per grade/subject and all versions were delivered on paper only (PBT). Item Response Theory (IRT) calibration and scaling was completed to place the Fall block operational item data for each grade/subject on the corresponding Spring 2015 base scale (i.e., the CBT scale).

The calibration and scaling procedures for 2014 Fall block tests were very similar to those used for the spring calibration and scaling. This section will address any differences in procedures between Fall block and Spring 2015 and will present the 2014 Fall block results.

The 2014 Fall block data were calibrated concurrently across the PBA and EOY components. Since the 2014 Fall block data were only from paper forms, only one calibration was needed per grade/subject for ELA/L and mathematics.

The 2014 Fall block IRT data files were prepared using the same specifications as those used in spring. Also, the same IRT software, software specifications, and IRT models (2PL/PCM) were used for Fall block. The 2014 Fall block paper-only IRT parameter estimates were transformed onto the 2015 spring base scales using the Stocking and Lord (1983) procedure after the spring PBT parameter estimates were transformed onto the spring CBT scales. Refer to Section 10 of the 2014-2015 PARCC Technical Report for more information on the following topics:

Calibration:

- 10.2 IRT Sparse Data File Preparation
- 10.3 Description of the calibration process
- 10.4 Model fit evaluation criteria
- 10.5 Items excluded from score reporting

Scaling:

- 10.6 Description of scaling process (Paper to Online)
- 10.7 Items Excluded from Spring 2015 Paper to Online Linking Sets
- 10.8 Correlations and Plots of Parameter Estimates
- 10.9 Scaling constants
- 10.10 Summary Statistics and Distributions from IRT Analyses
- 10.11 Effect Sizes of Linking Average Item Score Differences versus Scale Score Differences
- 10.12 Local Item Independence/Minimal Testlet Effect

The main difference between the 2014 Fall block to Spring 2015 scaling was the selection of reference item parameter estimates for the Spring 2015 scale. Although only PBT parameter estimates were

available for the 2014 Fall block common items, Spring 2015 linking items were sometimes delivered on both CBT and PBT forms. In all cases, when a Fall block linking item was delivered via both PBT and CBT modes in Spring 2015, the reference parameters were based on the “parameters estimates of record” used in the score reporting tables. The parameters of record were always the CBT parameter estimates when an item was delivered in both modes. In this case, the item was kept in the CBT/PBT Spring 2015 linking set. If a linking item was delivered on a PBT form only, or was not treated as a linking item across modes in Spring 2015, the parameter estimates of record were the PBT parameter estimates.

Also worth noting, most 2014 Fall block linking items were operational items on one of more of the Spring 2015 forms. However, there were a couple of exceptions. All 2014 Fall block operational items that were administered in Spring 2015 field-test positions were included in the linking sets.

The following tables (Tables ADD.10.1 through ADD.10.7) pertain to the calibration and scaling of the 2014 Fall block item response data.

The number of test takers per grade and subject for the 2014 Fall block tests was much smaller than the Spring 2015 calibration. In addition, there was only one PBA and one EOY form for the 2014 Fall block administration. Table ADD.10.1 presents the number of students, as well as the number and percent of items excluded from score reporting for each grade/subject for the 2014 Fall block.

Table ADD.10.1 Number of Students and Number and Percent of Fall block Items Excluded from Calibration

Content	Grade	Total N of Students	Total <i>n</i> of Items	No. of Items Excluded	Percent Excluded	Reason Excluded			
						Small Sample Size	Poor IA Stats.	Did Not Calibrate	Other*
ELA/L	9	4,277	50	1	2.0				1
ELA/L	10	8,058	50	0	0.0				
ELA/L	11	1,308	50	0	0.0				
Math	A1	8,525	53	1	1.9		1		
Math	GO	1,309	53	0	0.0				
Math	A2	3,980	54	3	5.6		1		2

Note: *ELA/L Grade 9 Other: mean scores similar for students scoring 0 and 1, did not calibrate unless collapsed. Algebra II dropped two items due to poor IRT fit. A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Table ADD.10.2 presents the total number of common items, items excluded from the 2014 Fall block/Spring 2015 linking sets, and items kept in the linking sets for each grade/subject. Items were removed from the linking sets due only to high weighted root mean differences, which indicate the item parameter estimates were not similar across the 2014 Fall block and Spring 2015 administrations.

Table ADD.10.2 Number of Fall Block Items Excluded from the Fall Block/Spring Linking Sets

Content	Grade/ Subject	Total <i>n</i> of Common Items	Number Excluded	Final Number in Linking Set	Reason Excluded				
					Low Polyserial	Diff. No. of Cat.	Mode C-DIF	Other	High WRMSD
ELA/L	9	38	4	34			n/a		4
ELA/L	10	46	1	45			n/a		1
ELA/L	11	31	4	27			n/a		4
Math	A1	40	6	34			n/a		6
Math	GO	33	5	28			n/a		5
Math	A2	32	4	28			n/a		4

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Once the final group of items for each linking set was determined, and the 2014 Fall block item parameter estimates were transformed onto the Spring 2015 scales, the a - and b -parameter estimates across administrations were plotted and the correlation between the a -parameter estimates and the b -parameter estimates were calculated. Tables ADD.10.3 shows the number of: linking items, score points of the linking items, and the correlation of the a - and b -parameter estimates across modes.

Table ADD.10.3 Number of Items, Points, and Correlations for Fall Block/Spring Linking Items

Content	Grade	Number		Parameter Correlations	
		Items	Points	a -	b -
ELA/L	9	34	80	0.94	0.97
ELA/L	10	45	114	0.95	0.97
ELA/L	11	27	65	0.94	0.94
Math	A1	34	61	0.90	0.89
Math	GO	28	54	0.88	0.96
Math	A2	28	57	0.92	0.98

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Table ADD.10.4 presents the slope and intercept scaling constants for the 2014 Fall block tests, derived from **STUIRT** (Kim & Kolen, 2004) using the Stocking and Lord (1983) test characteristic curve procedure. Of interest are the intercept values which provide an indication of the difference in ability/performance between the students in the 2014 Fall block and Spring 2015 groups on the common items. If the two groups are similar in ability, the intercept would be close to 0.00.

In all cases, the intercept is fairly large and positive indicating that the 2014 Fall block test takers performed better on the common items than Spring 2015 test takers. For ELA/L, at grades 9 and 10, the

intercepts are fairly large (i.e., about 0.56 and 0.55), with a smaller difference at Grade 11 (0.22). For mathematics, the largest intercept was for Geometry (0.88) and the smallest was for Algebra II (0.28).

Table ADD.10.4 Scaling Constants 2014 Fall Block to Spring 2015

Content	Grade/Subject	FB to Spring 2015	
		Slope	Intercept
ELA/L	9	0.928454	0.55893
ELA/L	10	0.876710	0.55163
ELA/L	11	0.947775	0.21618
Math	A1	0.913549	0.54747
Math	GO	0.912718	0.88536
Math	A2	1.016200	0.27998

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Tables ADD.10.5 through ADD.10.7 show the b - and a -parameter estimates, the standard errors of estimate, and the model fit statistics for the Fall block tests.

Table ADD.10.5 2014 Fall Block IRT Summary Parameter Estimates by Grade/Subject

Content	Grade/Subject	No. of Score Points	No. of Items	b Estimates Summary				a Estimates Summary			
				Mean	SD	Min	Max	Mean	SD	Min	Max
ELA/L	9	135	49	1.26	1.54	-0.73	6.87	0.48	0.26	0.10	1.13
ELA/L	10	137	50	0.40	0.75	-0.69	2.68	0.59	0.23	0.15	1.10
ELA/L	11	137	50	0.72	0.81	-0.62	4.01	0.51	0.21	0.15	0.98
Math	A1	96	52	2.21	1.91	-0.19	13.04	0.54	0.23	0.08	1.00
Math	GO	97	53	1.31	0.92	-1.17	3.41	0.75	0.29	0.22	1.68
Math	A2	104	51	1.92	1.33	-0.63	4.88	0.56	0.24	0.12	1.15

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Table ADD.10.6 2014 Fall Block IRT Standard Errors of Parameter Estimates by Grade/Subject

Content	Grade/Subject	No. of Score Points	No. of Items	SE of b Estimates				SE of a Estimates			
				Mean	SD	Min	Max	Mean	SD	Min	Max
ELA/L	9	135	49	0.067	0.105	0.013	0.614	0.015	0.006	0.005	0.029
ELA/L	10	137	50	0.022	0.012	0.010	0.073	0.012	0.004	0.004	0.020
ELA/L	11	137	50	0.069	0.053	0.026	0.268	0.028	0.009	0.011	0.045
Math	A1	96	52	0.063	0.066	0.010	0.372	0.018	0.008	0.003	0.038
Math	GE	97	53	0.069	0.057	0.023	0.341	0.051	0.022	0.019	0.136
Math	A2	104	51	0.079	0.089	0.017	0.560	0.024	0.011	0.004	0.066

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Table ADD.10.7 2014 Fall Block IRT Model Fit by Grade/Subject

Content	Grade/ Subject	No. of Score Points	No. of Items	G^2				Adjusted Fit			
				Mean	SD	Min	Max	Mean	SD	Min	Max
ELA/L	9	135	49	134	54	40	302	0.17	0.03	0.10	0.26
ELA/L	10	137	50	191	99	61	599	0.15	0.04	0.09	0.26
ELA/L	11	137	50	69	20	37	127	0.23	0.03	0.17	0.30
Math	A1	96	52	124	107	27	451	0.11	0.05	0.06	0.23
Math	GE	97	53	57	26	19	130	0.20	0.04	0.12	0.30
Math	A2	104	51	105	63	28	301	0.16	0.05	0.09	0.28

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II.

Tables ADD.10.8 to ADD.10.10 present IRT parameter estimates, standard errors of estimate and model fit for ELA/L Fall block data for grades 9, 10, and 11. This information is presented for all items, reading items, writing items, PBA-only, and EOY-only. Similar information is provided in Tables ADD.10.11 through ADD.10.13 for the Fall block administration of mathematics (Algebra I, Geometry, and Algebra II). For mathematics, the results are presented for all items, single select multiple choice items, CR items, PBA-only, EOY-only, and subclaims.

Table ADD.10.8 PBT IRT Summary Parameter Estimates for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E09	All Items	135	49	1.26	1.54	-0.73	6.87	0.48	0.26	0.10	1.13
PBT	E09	Reading	90	43	1.29	1.63	-0.73	6.87	0.41	0.20	0.10	0.84
PBT	E09	Writing	45	6	1.07	0.57	0.45	1.89	0.97	0.11	0.80	1.13
PBT	E09	All PBA	91	27	1.34	1.40	-0.57	6.18	0.53	0.31	0.10	1.13
PBT	E09	All EOY	44	22	1.17	1.72	-0.73	6.87	0.41	0.19	0.10	0.84
PBT	E10	All Items	137	50	0.40	0.75	-0.69	2.68	0.59	0.23	0.15	1.10
PBT	E10	Reading	92	44	0.39	0.79	-0.69	2.68	0.53	0.18	0.15	0.99
PBT	E10	Writing	45	6	0.51	0.43	-0.02	1.09	0.98	0.10	0.83	1.10
PBT	E10	All PBA	93	28	0.62	0.77	-0.57	2.68	0.62	0.25	0.23	1.10
PBT	E10	All EOY	44	22	0.13	0.64	-0.69	2.12	0.55	0.19	0.15	0.92
PBT	E11	All Items	137	50	0.72	0.81	-0.62	4.01	0.51	0.21	0.15	0.98
PBT	E11	Reading	92	44	0.75	0.85	-0.62	4.00	0.46	0.18	0.15	0.96
PBT	E11	Writing	45	6	0.51	0.37	0.08	1.03	0.83	0.09	0.72	0.98
PBT	E11	All PBA	93	28	0.63	0.81	-0.62	2.40	0.53	0.24	0.15	0.98
PBT	E11	All EOY	44	22	0.84	0.81	-0.35	4.00	0.48	0.16	0.21	0.96

Table ADD.10.9 PBT IRT Standard Errors of Parameter Estimates for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>SE of b Estimates</i>				<i>SE of a Estimates</i>			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E09	All Items	135	49	0.067	0.105	0.013	0.614	0.015	0.006	0.005	0.029
PBT	E09	Reading	90	43	0.074	0.110	0.018	0.614	0.014	0.005	0.005	0.026
PBT	E09	Writing	45	6	0.015	0.002	0.013	0.017	0.024	0.003	0.020	0.029
PBT	E09	All PBA	91	27	0.071	0.123	0.013	0.614	0.016	0.006	0.007	0.029
PBT	E09	All EOY	44	22	0.062	0.081	0.022	0.402	0.013	0.005	0.005	0.026
PBT	E10	All Items	137	50	0.022	0.012	0.010	0.073	0.012	0.004	0.004	0.020
PBT	E10	Reading	92	44	0.024	0.012	0.010	0.073	0.011	0.003	0.004	0.017
PBT	E10	Writing	45	6	0.010	0.000	0.010	0.011	0.018	0.002	0.015	0.020
PBT	E10	All PBA	93	28	0.021	0.012	0.010	0.073	0.012	0.004	0.005	0.020
PBT	E10	All EOY	44	22	0.024	0.012	0.013	0.063	0.012	0.003	0.004	0.017
PBT	E11	All Items	137	50	0.069	0.053	0.026	0.268	0.028	0.009	0.011	0.045
PBT	E11	Reading	92	44	0.074	0.054	0.028	0.268	0.026	0.008	0.011	0.043
PBT	E11	Writing	45	6	0.028	0.003	0.026	0.032	0.039	0.004	0.034	0.045
PBT	E11	All PBA	93	28	0.070	0.057	0.026	0.255	0.028	0.010	0.011	0.045
PBT	E11	All EOY	44	22	0.067	0.047	0.034	0.268	0.028	0.007	0.016	0.043

Table ADD.10.10 PBT IRT Model Fit for All Items for ELA/L by Grade

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	E09	All Items	135	49	134	54	40	302	0.17	0.03	0.10	0.26
PBT	E09	Reading	90	43	134	56	40	302	0.17	0.03	0.10	0.26
PBT	E09	Writing	45	6	139	27	106	183	0.18	0.02	0.16	0.21
PBT	E09	All PBA	91	27	142	62	40	302	0.18	0.04	0.10	0.26
PBT	E09	All EOY	44	22	125	40	63	208	0.17	0.03	0.12	0.22
PBT	E10	All Items	137	50	191	99	61	599	0.15	0.04	0.09	0.26
PBT	E10	Reading	92	44	193	106	61	599	0.15	0.04	0.09	0.26
PBT	E10	Writing	45	6	178	24	135	203	0.15	0.01	0.13	0.16
PBT	E10	All PBA	93	28	209	113	61	599	0.15	0.04	0.09	0.26
PBT	E10	All EOY	44	22	167	74	84	308	0.14	0.03	0.10	0.19
PBT	E11	All Items	137	50	69	20	37	127	0.23	0.03	0.17	0.30
PBT	E11	Reading	92	44	66	19	37	127	0.22	0.03	0.17	0.30
PBT	E11	Writing	45	6	89	15	69	104	0.26	0.02	0.23	0.29
PBT	E11	All PBA	93	28	71	24	37	127	0.23	0.04	0.17	0.30
PBT	E11	All EOY	44	22	65	13	46	88	0.22	0.02	0.19	0.26

Table ADD.10.11 PBT IRT Summary Parameter Estimates for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>b</i> Estimates Summary				<i>a</i> Estimates Summary			
					Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
PBT	A1	All Items	96	52	2.21	1.91	-0.19	13.04	0.54	0.23	0.08	1.00
PBT	A1	SSMC	17	17	1.56	1.16	-0.06	3.98	0.43	0.17	0.13	0.70
PBT	A1	CR	79	35	2.52	2.13	-0.19	13.04	0.59	0.23	0.08	1.00
PBT	A1	All PBA	41	17	2.09	1.01	-0.06	4.02	0.46	0.19	0.13	0.92
PBT	A1	PBA Type I	9	9	1.84	1.21	-0.06	3.98	0.50	0.23	0.13	0.92
PBT	A1	PBA Type II	14	4	2.69	0.89	2.19	4.02	0.41	0.15	0.21	0.54
PBT	A1	PBA Type III	18	4	2.08	0.37	1.55	2.43	0.42	0.08	0.33	0.52
PBT	A1	All EOY	55	35	2.27	2.23	-0.19	13.04	0.58	0.23	0.08	1.00
PBT	GO	All Items	97	53	1.31	0.92	-1.17	3.41	0.75	0.29	0.22	1.68
PBT	GO	SSMC	11	11	0.70	1.00	-1.17	2.07	0.54	0.19	0.22	0.84
PBT	GO	CR	86	42	1.47	0.84	-0.14	3.41	0.80	0.29	0.43	1.68
PBT	GO	All PBA	42	18	1.56	0.78	0.45	2.99	0.65	0.17	0.36	1.03
PBT	GO	PBA Type I	10	10	1.37	0.85	0.45	2.99	0.68	0.19	0.36	1.03
PBT	GO	PBA Type II	14	4	2.12	0.31	1.88	2.54	0.68	0.16	0.53	0.91
PBT	GO	PBA Type III	18	4	1.47	0.81	0.51	2.50	0.56	0.10	0.43	0.66
PBT	GO	All EOY	55	35	1.18	0.96	-1.17	3.41	0.80	0.33	0.22	1.68
PBT	A2	All Items	104	51	1.92	1.33	-0.63	4.88	0.56	0.24	0.12	1.15
PBT	A2	SSMC	14	14	1.39	1.57	-0.50	4.71	0.41	0.18	0.15	0.72
PBT	A2	CR	90	37	2.12	1.19	-0.63	4.88	0.61	0.24	0.12	1.15
PBT	A2	All PBA	51	19	2.11	1.25	-0.50	4.88	0.58	0.25	0.25	1.05
PBT	A2	PBA Type I	9	9	1.52	0.86	-0.50	2.42	0.65	0.30	0.25	1.05
PBT	A2	PBA Type II	18	5	3.12	1.69	0.73	4.88	0.46	0.18	0.27	0.66
PBT	A2	PBA Type III	24	5	2.15	0.78	0.78	2.65	0.58	0.22	0.38	0.91
PBT	A2	All EOY	53	32	1.81	1.38	-0.63	4.71	0.54	0.24	0.12	1.15

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, SSMC = single select multiple choice items, CR = constructed response items.

Table ADD.10.12 PBT IRT Standard Errors of Parameter Estimates for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	<i>SE of b Estimates</i>				<i>SE of a Estimates</i>			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	A1	All Items	96	52	0.063	0.066	0.010	0.372	0.018	0.008	0.003	0.038
PBT	A1	SSMC	17	17	0.082	0.091	0.020	0.372	0.017	0.002	0.014	0.020
PBT	A1	CR	79	35	0.053	0.047	0.010	0.204	0.018	0.010	0.003	0.038
PBT	A1	All PBA	41	17	0.063	0.084	0.017	0.372	0.014	0.006	0.006	0.024
PBT	A1	PBA Type I	9	9	0.088	0.110	0.022	0.372	0.019	0.003	0.014	0.024
PBT	A1	PBA Type II	14	4	0.045	0.032	0.019	0.090	0.010	0.003	0.006	0.014
PBT	A1	PBA Type III	18	4	0.024	0.010	0.017	0.038	0.009	0.003	0.006	0.012
PBT	A1	All EOY	55	35	0.062	0.056	0.010	0.241	0.020	0.009	0.003	0.038
PBT	GO	All Items	97	53	0.069	0.057	0.023	0.341	0.051	0.022	0.019	0.136
PBT	GO	SSMC	11	11	0.109	0.083	0.046	0.341	0.046	0.007	0.035	0.058
PBT	GO	CR	86	42	0.059	0.043	0.023	0.260	0.052	0.024	0.019	0.136
PBT	GO	All PBA	42	18	0.066	0.045	0.023	0.205	0.043	0.013	0.019	0.060
PBT	GO	PBA Type I	10	10	0.086	0.052	0.041	0.205	0.051	0.007	0.040	0.060
PBT	GO	PBA Type II	14	4	0.047	0.016	0.034	0.070	0.036	0.009	0.025	0.046
PBT	GO	PBA Type III	18	4	0.038	0.018	0.023	0.064	0.028	0.008	0.019	0.039
PBT	GO	All EOY	55	35	0.071	0.063	0.026	0.341	0.055	0.024	0.024	0.136
PBT	A2	All Items	104	51	0.079	0.089	0.017	0.560	0.024	0.011	0.004	0.066
PBT	A2	SSMC	14	14	0.130	0.147	0.035	0.560	0.023	0.003	0.019	0.028
PBT	A2	CR	90	37	0.059	0.041	0.017	0.200	0.025	0.013	0.004	0.066
PBT	A2	All PBA	51	19	0.068	0.054	0.017	0.200	0.023	0.009	0.010	0.040
PBT	A2	PBA Type I	9	9	0.072	0.047	0.035	0.165	0.029	0.008	0.019	0.040
PBT	A2	PBA Type II	18	5	0.097	0.078	0.017	0.200	0.018	0.003	0.015	0.023
PBT	A2	PBA Type III	24	5	0.032	0.008	0.018	0.037	0.019	0.010	0.010	0.034
PBT	A2	All EOY	53	32	0.085	0.104	0.022	0.560	0.025	0.012	0.004	0.066

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, SSMC = single select multiple choice items, CR = constructed response items.

Table ADD.10.13 PBT IRT Model Fit for All Items for Mathematics by Grade/Subject

Mode	Grade	Item Grouping	No. of Score Points	No. of Items	G^2				Adjusted Fit			
					Mean	SD	Min	Max	Mean	SD	Min	Max
PBT	A1	All Items	96	52	124	107	27	451	0.11	0.05	0.06	0.23
PBT	A1	SSMC	17	17	65	32	32	122	0.09	0.02	0.06	0.12
PBT	A1	CR	79	35	153	118	27	451	0.13	0.05	0.06	0.23
PBT	A1	All PBA	41	17	162	141	32	451	0.13	0.06	0.06	0.23
PBT	A1	PBA Type I	9	9	59	26	32	100	0.08	0.02	0.06	0.11
PBT	A1	PBA Type II	79	35	153	118	27	451	0.13	0.05	0.06	0.23
PBT	A1	PBA Type III	55	35	106	81	27	405	0.11	0.04	0.06	0.22
PBT	A1	All EOY	17	17	65	32	32	122	0.09	0.02	0.06	0.12
PBT	GO	All Items	97	53	57	26	19	130	0.20	0.04	0.12	0.30
PBT	GO	SSMC	11	11	42	9	21	52	0.18	0.02	0.13	0.20
PBT	GO	CR	86	42	61	27	19	130	0.21	0.04	0.12	0.30
PBT	GO	All PBA	42	18	65	34	21	130	0.21	0.06	0.13	0.30
PBT	GO	PBA Type I	10	10	39	12	21	65	0.17	0.03	0.13	0.22
PBT	GO	PBA Type II	86	42	61	27	19	130	0.21	0.04	0.12	0.30
PBT	GO	PBA Type III	55	35	53	20	19	100	0.20	0.03	0.12	0.27
PBT	GO	All EOY	11	11	42	9	21	52	0.18	0.02	0.13	0.20
PBT	A2	All Items	104	51	105	63	28	301	0.16	0.05	0.09	0.28
PBT	A2	SSMC	14	14	68	21	45	122	0.13	0.02	0.11	0.18
PBT	A2	CR	90	37	119	67	28	301	0.17	0.05	0.09	0.28
PBT	A2	All PBA	51	19	119	72	36	301	0.17	0.05	0.10	0.28
PBT	A2	PBA Type I	9	9	68	25	36	112	0.13	0.02	0.10	0.17
PBT	A2	PBA Type II	90	37	119	67	28	301	0.17	0.05	0.09	0.28
PBT	A2	PBA Type III	53	32	97	56	28	276	0.16	0.04	0.09	0.27
PBT	A2	All EOY	14	14	68	21	45	122	0.13	0.02	0.11	0.18

Note: A1 = Algebra I, GO = Geometry, A2 = Algebra II, SSMC = single select multiple choice items, CR = constructed response items.

Addendum 12: Scale Scores

Table ADD.12.1 Fall 2014 Subgroup Performance for ELA/L Scale Scores: Grade 9

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		4,184	756.43	33.76	650	850
Gender	Female	2,203	762.58	32.23	667	850
	Male	1,981	749.60	34.11	650	850
Ethnicity	American Indian/Alaska Native	9	733.67	33.48	698	794
	Asian	44	764.89	36.88	679	834
	Black or African American	650	742.39	32.49	650	836
	Hispanic/Latino	135	756.30	32.31	671	832
	Native Hawaiian or Pacific Islander	2	745.50	21.92	730	761
	Multiple Race Selected	42	741.81	29.69	667	816
	White	1,818	757.65	34.51	650	850
Economic Status*	Economically Disadvantaged	1,434	745.21	31.91	650	846
	Not Economically Disadvantaged	1,784	765.32	35.02	650	850
English Learner Status	English Learner (EL)	58	741.40	33.74	671	828
	Non English Learner	2,949	758.91	34.58	650	850
Disabilities	Students with Disabilities (SWD)	325	718.59	30.47	650	825
	Students without Disabilities	2,489	759.92	32.57	667	850
Delivery Mode	PBT	4,184	756.43	33.76	650	850
Reading Score		4,184	52.89	13.64	10	90
Gender	Female	2,203	54.44	13.08	17	90
	Male	1,981	51.16	14.03	10	90
Ethnicity	American Indian/Alaska Native	9	44.00	13.94	29	67
	Asian	44	55.50	14.66	25	85
	Black or African American	650	47.48	12.91	10	90
	Hispanic/Latino	135	52.84	12.94	21	82
	Native Hawaiian or Pacific Islander	2	45.50	2.12	44	47
	Multiple Race Selected	42	47.12	11.24	19	71
	White	1,818	53.34	14.00	10	90
Economic Status*	Economically Disadvantaged	1,434	48.77	12.71	10	90
	Not Economically Disadvantaged	1,784	56.15	14.23	10	90
English Learner Status	English Learner (EL)	58	47.00	13.00	21	81
	Non English Learner	2,949	53.70	13.90	10	90
Disabilities	Students with Disabilities (SWD)	325	38.62	12.10	10	81
	Students without Disabilities	2,489	54.17	13.11	14	90
Delivery Mode	PBT	4,184	52.89	13.64	10	90
Writing Score		4,184	35.55	9.19	10	60

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Gender	Female	2,203	37.84	8.26	10	60
	Male	1,981	33.00	9.50	10	60
Ethnicity	American Indian/Alaska Native	9	30.11	9.28	10	44
	Asian	44	37.86	9.97	10	52
	Black or African American	650	31.90	9.93	10	55
	Hispanic/Latino	135	35.55	8.77	10	52
	Native Hawaiian or Pacific Islander	2	36.50	10.61	29	44
	Multiple Race Selected	42	31.26	10.78	10	52
	White	1,818	35.85	9.23	10	60
Economic Status*	Economically Disadvantaged	1,434	32.55	9.37	10	55
	Not Economically Disadvantaged	1,784	37.73	9.28	10	60
English Learner Status	English Learner (EL)	58	31.47	10.73	10	51
	Non English Learner	2,949	36.19	9.41	10	60
Disabilities	Students with Disabilities (SWD)	325	25.06	10.59	10	51
	Students without Disabilities	2,489	36.45	8.75	10	60
Delivery Mode	PBT	4,184	35.55	9.19	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table ADD.12.2 Fall 2014 Subgroup Performance for ELA/L Scale Scores: Grade 10

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		7,910	759.76	39.26	650	850
Gender	Female	4,078	767.03	37.42	650	850
	Male	3,832	752.02	39.70	650	850
Ethnicity	American Indian/Alaska Native	19	749.37	33.08	697	814
	Asian	137	778.33	39.16	667	850
	Black or African American	1,087	740.08	38.08	650	850
	Hispanic/Latino	339	746.12	38.40	650	850
	Native Hawaiian or Pacific Islander	5	722.60	32.89	680	756
	Multiple Race Selected	122	743.20	43.35	650	847
	White	5,064	762.70	38.74	650	850
Economic Status*	Economically Disadvantaged	1,644	739.15	38.23	650	850
	Not Economically Disadvantaged	4,751	762.90	38.73	650	850
English Learner Status	English Learner (EL)	44	720.89	31.17	657	791
	Non English Learner	6,540	757.49	39.91	650	850
Disabilities	Students with Disabilities (SWD)	172	716.55	35.22	657	817
	Students without Disabilities	3,113	764.23	36.00	650	850
Delivery Mode	PBT	7,910	759.76	39.26	650	850
Reading Score		7,910	54.02	15.59	10	90
Gender	Female	4,078	56.16	15.04	10	90
	Male	3,832	51.74	15.85	10	90
Ethnicity	American Indian/Alaska Native	19	50.11	13.03	26	77
	Asian	137	60.78	15.54	19	90
	Black or African American	1,087	46.50	15.48	10	90
	Hispanic/Latino	339	49.05	15.56	10	90
	Native Hawaiian or Pacific Islander	5	38.20	14.32	22	53
	Multiple Race Selected	122	47.84	16.68	10	89
	White	5,064	55.11	15.27	10	90
Economic Status*	Economically Disadvantaged	1,644	46.34	15.29	10	90
	Not Economically Disadvantaged	4,751	55.10	15.28	10	90
English Learner Status	English Learner (EL)	44	39.27	14.24	10	70
	Non English Learner	6,540	53.09	15.72	10	90
Disabilities	Students with Disabilities (SWD)	172	37.64	14.16	12	79
	Students without Disabilities	3,113	55.67	14.43	10	90
Delivery Mode	PBT	7,910	54.02	15.59	10	90
Writing Score		7,910	36.88	10.17	10	60
Gender	Female	4,078	39.29	9.22	10	60

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
	Male	3,832	34.31	10.51	10	60
Ethnicity	American Indian/Alaska Native	19	34.74	8.97	21	49
	Asian	137	41.59	9.81	10	60
	Black or African American	1,087	32.24	10.09	10	60
	Hispanic/Latino	339	33.38	10.30	10	60
	Native Hawaiian or Pacific Islander	5	30.40	6.47	21	37
	Multiple Race Selected	122	32.57	11.98	10	60
	White	5,064	37.58	10.14	10	60
Economic Status*	Economically Disadvantaged	1,644	31.74	10.38	10	60
	Not Economically Disadvantaged	4,751	37.73	10.13	10	60
English Learner Status	English Learner (EL)	44	28.07	7.25	10	46
	Non English Learner	6,540	36.37	10.49	10	60
Disabilities	Students with Disabilities (SWD)	172	25.94	10.36	10	50
	Students without Disabilities	3,113	38.08	9.25	10	60
Delivery Mode	PBT	7,910	36.88	10.17	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table ADD.12.3 Fall 2014 Subgroup Performance for ELA/L Scale Scores: Grade 11

Group Type	Group	N	Mean	SD	Min	Max
Full Summative Score		1,201	748.59	36.05	650	850
Gender	Female	572	752.92	35.19	654	835
	Male	629	744.65	36.41	650	850
Ethnicity	American Indian/Alaska Native	6	761.83	20.61	722	783
	Asian	28	751.89	30.52	689	811
	Black or African American	71	713.38	31.22	659	779
	Hispanic/Latino	219	734.96	34.34	654	835
	Native Hawaiian or Pacific Islander	0	-	-	-	-
	Multiple Race Selected	41	746.63	38.90	675	820
	White	833	755.00	34.25	650	850
Economic Status*	Economically Disadvantaged	468	735.84	32.90	654	840
	Not Economically Disadvantaged	722	757.04	35.46	650	850
English Learner Status	English Learner (EL)	5	714.60	34.02	668	763
	Non English Learner	165	748.81	42.33	654	850
Disabilities	Students with Disabilities (SWD)	127	724.11	35.35	654	850
	Students without Disabilities	960	750.08	34.75	650	850
Delivery Mode	PBT	1,201	748.59	36.05	650	850
Reading Score		1,201	49.60	14.08	13	90
Gender	Female	572	50.52	13.64	15	89
	Male	629	48.77	14.43	13	90
Ethnicity	American Indian/Alaska Native	6	52.33	9.56	34	62
	Asian	28	51.71	13.06	26	83
	Black or African American	71	37.06	12.17	13	64
	Hispanic/Latino	219	43.46	13.19	15	80
	Native Hawaiian or Pacific Islander	0	-	-	-	-
	Multiple Race Selected	41	49.63	14.77	23	76
	White	833	52.19	13.43	13	90
Economic Status*	Economically Disadvantaged	468	44.79	12.88	13	87
	Not Economically Disadvantaged	722	52.81	13.92	13	90
English Learner Status	English Learner (EL)	5	34.20	14.79	20	57
	Non English Learner	165	50.05	16.10	15	90
Disabilities	Students with Disabilities (SWD)	127	40.27	13.59	13	90
	Students without Disabilities	960	50.19	13.64	13	90
Delivery Mode	PBT	1,201	49.60	14.08	13	90
Writing Score		1,201	34.06	10.01	10	60
Gender	Female	572	35.90	9.74	10	60

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
	Male	629	32.40	9.97	10	56
Ethnicity	American Indian/Alaska Native	6	39.67	6.68	33	51
	Asian	28	34.00	8.69	10	43
	Black or African American	71	24.03	10.23	10	48
	Hispanic/Latino	219	31.79	9.71	10	60
	Native Hawaiian or Pacific Islander	0	-	-	-	-
	Multiple Race Selected	41	32.68	11.16	10	52
	White	833	35.52	9.43	10	60
Economic Status*	Economically Disadvantaged	468	30.89	9.61	10	60
	Not Economically Disadvantaged	722	36.17	9.65	10	60
English Learner Status	English Learner (EL)	5	28.20	8.81	14	36
	Non English Learner	165	33.82	11.31	10	60
Disabilities	Students with Disabilities (SWD)	127	27.78	11.07	10	60
	Students without Disabilities	960	34.46	9.55	10	60
Delivery Mode	PBT	1,201	34.06	10.01	10	60

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table ADD.12.4 Fall 2014 Subgroup Performance for Mathematics Scale Scores: Algebra I

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		8,308	749.73	29.65	650	850
Gender	Female	4,152	751.85	28.46	650	850
	Male	4,156	747.62	30.65	650	850
Ethnicity	American Indian/Alaska Native	29	742.41	29.68	686	833
	Asian	155	763.91	28.26	678	850
	Black or African American	1,323	733.66	28.00	650	843
	Hispanic/Latino	355	741.36	27.16	650	824
	Native Hawaiian or Pacific Islander	5	743.60	32.02	694	770
	Multiple Race Selected	120	744.37	29.60	666	828
	White	4,988	752.23	28.31	650	850
Economic Status*	Economically Disadvantaged	2,362	737.83	27.07	650	833
	Not Economically Disadvantaged	4,621	753.13	29.47	651	850
English Learner Status	English Learner (EL)	89	740.15	30.47	650	815
	Non English Learner	6,897	749.56	29.82	650	850
Disabilities	Students with Disabilities (SWD)	284	724.77	25.98	666	850
	Students without Disabilities	3,798	750.60	30.09	651	850
Delivery Mode	PBT	8,308	749.73	29.65	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table ADD.12.5 Fall 2014 Subgroup Performance for Mathematics Scale Scores: Geometry

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		1,294	752.32	23.68	675	822
Gender	Female	650	751.51	22.24	691	808
	Male	644	753.14	25.05	675	822
Ethnicity	American Indian/Alaska Native	12	748.33	34.47	701	800
	Asian	22	761.95	30.40	697	801
	Black or African American	77	732.26	21.04	684	774
	Hispanic/Latino	109	736.20	25.51	675	791
	Native Hawaiian or Pacific Islander	1	752.00	0.00	752	752
	Multiple Race Selected	7	757.29	9.16	742	768
	White	732	754.47	21.55	675	822
Economic Status*	Economically Disadvantaged	384	739.08	21.29	675	801
	Not Economically Disadvantaged	620	758.51	22.13	675	822
English Learner Status	English Learner (EL)	30	741.43	25.03	701	790
	Non English Learner	915	751.20	24.06	675	822
Disabilities	Students with Disabilities (SWD)	29	718.52	26.92	675	776
	Students without Disabilities	598	751.69	21.31	684	822
Delivery Mode	PBT	1,294	752.32	23.68	675	822

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).

Table ADD.12.6 Fall 2014 Subgroup Performance for Mathematics Scale Scores: Algebra II

Group Type	Group	<i>N</i>	Mean	<i>SD</i>	Min	Max
Full Summative Score		3,600	730.89	38.53	650	850
Gender	Female	1,801	731.73	35.60	650	850
	Male	1,799	730.04	41.24	650	850
Ethnicity	American Indian/Alaska Native	15	748.40	39.59	690	828
	Asian	103	751.07	45.62	650	850
	Black or African American	498	694.60	28.47	650	794
	Hispanic/Latino	275	721.13	35.47	650	850
	Native Hawaiian or Pacific Islander	5	738.40	55.45	686	831
	Multiple Race Selected	92	726.58	31.27	650	838
	White	2,580	738.32	36.13	650	850
Economic Status*	Economically Disadvantaged	989	711.51	34.55	650	831
	Not Economically Disadvantaged	2,600	738.34	37.39	650	850
English Learner Status	English Learner (EL)	6	717.67	32.43	686	757
	Non English Learner	2,930	732.33	39.04	650	850
Disabilities	Students with Disabilities (SWD)	140	692.11	27.87	650	797
	Students without Disabilities	1,439	720.85	36.17	650	850
Delivery Mode	PBT	3,600	730.89	38.53	650	850

Note: *Economic status was based on participation in National School Lunch Program (NSLP): receipt of free or reduced-price lunch (FRL).