

Missouri

Assessment Program

Technical Report 2008

Submitted to
Missouri Department of Elementary and Secondary Education
December 2008



Developed and published by CTB/McGraw-Hill LLC, a subsidiary of The McGraw-Hill Companies, Inc., 20 Ryan Ranch Road, Monterey, California 93940-5703. Copyright © 2008 by Missouri Department of Elementary and Secondary Education. All rights reserved. Only Missouri State educators and citizens may copy and/or download and print the document, located online at <http://dese.mo.gov/divimprov/assess/index.html>.

Table of Contents

Table of Tables	iv
Table of Figures	viii
Table of Figures	viii
EXECUTIVE SUMMARY.....	1
E.1 Background.....	1
E.2 Administration	1
E.3 Student Performance.....	1
E.4 Validity and Test Scores	2
CHAPTER 1: INTRODUCTION.....	4
1.1 Background of the Missouri Assessment Program.....	4
1.2 Purpose of the Missouri Assessment Program.....	5
1.3 Design of the Missouri Assessment Program	5
1.4 Overview of this Report.....	5
CHAPTER 2: VALIDITY AND THE USES OF TEST SCORES.....	10
2.1 Uses of Test Scores.....	10
2.2 Test-Level Scores.....	11
2.2.1 Use of Test-Level Scores	12
2.3 Content Standard Subscores	12
2.3.1 Use of the Content Standard Subscores	12
2.4 Process Standard Subscores.....	13
2.4.1 Use of the Process Standard Subscores	13
CHAPTER 3: TEST CONTENT DEVELOPMENT	14
3.1 Test Specifications.....	14
3.2 Item Development.....	23
3.3 Item Writing.....	23
3.4 Local Pilot Test.....	24
3.5 Score, Revise, Rewrite Workshop	24
3.6 Content and Bias Review Workshop.....	24
3.7 Field Test Selection and Administration.....	25
3.8 Operational Test Selection.....	25
3.9 Accommodations and Universal Design.....	27
CHAPTER 4: TEST ADMINISTRATION	39
4.1 Training of Districts.....	39
4.2 Ancillary Materials	40
4.2.1 Return Material Forms and Guidelines.....	43
4.2.2 Security Forms	43
4.2.3 Interpretive Guides.....	43
4.3 Test security measures	44
4.4 Test Administration	44
4.4.1 Time	44
4.4.2 Accommodations	44
4.5 Summary	45

CHAPTER 5: CONSTRUCTED-RESPONSE SCORING	61
5.1 MAP Scoring Process	61
5.1.1 Handscoring Process Used for MAP	61
5.1.2 Selection of Scoring Evaluators.....	61
5.1.3 Handscoring Process.....	62
5.1.4 Monitoring the Scoring Process.....	63
5.1.5 Security	64
5.2 Inter-Rater Reliability	64
CHAPTER 6: OPERATIONAL DATA ANALYSES	72
6.1 Calibration Sample.....	72
6.2 Classical Item Statistics	72
6.2.1 Test-Level Statistics.....	73
6.2.2 Item-Level Statistics	73
6.3 Item Response Theory	74
6.3.1 Model Fit.....	75
6.4 Scaling.....	79
6.4.1 Linking Method and Results.....	79
6.4.2 Vertical Scale.....	80
6.4.3 Lowest and Highest Obtainable Scale Scores.....	81
6.5 Item-Pattern Scoring	81
CHAPTER 7. TEST RESULTS.....	123
7.1 Student Participation.....	123
7.2 Current Administration Data.....	124
7.3 Cross-year, Cross-sectional Comparisons	124
7.4 Reports	125
7.4.1 Description of Each Type of Report	125
7.5 Data Structures.....	130
7.5.1 General Research File.....	130
7.6 Interpreting Test Results	130
CHAPTER 8: ACHIEVEMENT-LEVEL SETTING	139
8.1 Legislation Affecting MAP Standard Setting	139
8.2 Bookmark Standard Setting Procedure.....	140
8.3 Cut Scores	140
8.4 Achievement-Level Descriptors	141
CHAPTER 9: EVIDENCE OF CONSTRUCT-RELATED VALIDITY	159
9.1 Minimization of Construct-Irrelevant Variance and Under-Representation ...	159
9.2 Reliability.....	159
9.2.1 Test Reliability.....	160
9.2.2 Standard Error of Measurement.....	161
9.2.3 Conditional Standard Error of Measurement.....	161
9.2.4 Decision Accuracy and Consistency.....	162
9.2.5 Convergent Validity.....	163
9.3 Principal Components Analysis.....	164
9.4 Analyses by Content Standard.....	165
9.5 Reliability of Content Standards.....	165
9.6 Correlations among Content Standard Subscores.....	165

9.7	Divergent (Discriminant) Validity.....	166
CHAPTER 10: FAIRNESS		199
10.1	Minimizing Bias through Careful Test Development.....	199
10.2	Evaluating Bias through Differential Item Functioning Statistics.....	200
10.3	Evaluating Bias through Impact Analysis.....	202
10.3.1	Reliability.....	202
10.3.2	Effect Size.....	202
REFERENCES		212

Table of Tables

Table E.1: Participation Rates: All Students.....	3
Table E.2: Percentage of Students Classified as <i>Proficient</i> or <i>Advanced</i> in 2006, 2007, and 2008 using Census Data: Communication Arts, Mathematics, and Science	3
Table 1.1: Timeline of Grade-Span MAP.....	8
Table 1.2: Timeline of Grade-Level MAP.....	8
Table 1.3: Number of Items that did not Map to a Missouri Grade-Level Expectation	8
Table 1.4: Spring 2008 MAP Test Design.....	9
Table 1.5: Spring 2008 Items Removed from Braille Forms.....	9
Table 3. 1: MAP Test Blueprint: Target Score Points by Content Standard (Communication Arts) or GLE Strand (Mathematics and Science)	15
Table 3. 2: Content Coverage: MAP 2008 Item Maps	15
Table 3. 3: Elements of Universal Design	28
Table 3. 4: Items Omitted from the MAP Spring 2008 Braille Version.....	30
Table 3. 5: MAP 2008 Content Standard Item/Point Distributions, Communication Arts	31
Table 3. 6: MAP 2008 GLE Strand Item/Point Distributions, Mathematics	32
Table 3. 7: MAP 2008 GLE Strand Item/Point Distributions, Science	33
Table 3. 8: MAP 2008 Number of Items/Points Measuring Process Standards, Communication Arts.....	34
Table 3. 9: MAP 2008 Number of Items/Points Measuring Process Standards, Mathematics.....	36
Table 3. 10: MAP 2008 Number of Items/Points Measuring Process Standards, Science	38
Table 4. 1: MAP Administration Schedule.....	47
Table 4. 2: MAP Accommodations for students who are English Language Learners	48
Table 4. 3: MAP Accommodations for Students with Disabilities.....	49
Table 4. 4: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Regular Edition.....	50
Table 4. 5: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Braille Edition	53
Table 4. 6: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Large Print Edition	56
Table 5. 1: Inter-rater Reliability, Communication Arts.....	65
Table 5. 2: Inter-rater Reliability, Mathematics.....	67
Table 5. 3: Inter-rater Reliability, Science.....	69
Table 6. 1: Large Districts that were Included in the 80% Calibration Sample	82
Table 6. 2: Summary of Calibration and Census Data: Communication Arts.....	83
Table 6. 3: Summary of Calibration and Census Data: Mathematics.....	86
Table 6. 4: Summary of Calibration and Census Data: Science	89

Table 6. 5: MAP Means, Standard Deviations for Raw Scores, P-Values, Item-Total Correlation (R_{it}): Communication Arts 2008.....	90
Table 6. 6: MAP Means, Standard Deviations for Raw Scores, P-Values, Item-Total Correlation (R_{it}): Mathematics 2008.....	90
Table 6. 7: MAP Means, Standard Deviations for Raw Scores, P-Values, Item-Total Correlation (R_{it}): Science 2008.....	90
Table 6. 8: Item Statistics: Grade 3.....	91
Table 6. 9: Item Statistics: Grade 4.....	93
Table 6. 10: Item Statistics: Grade 5.....	95
Table 6. 11: Item Statistics: Grade 6.....	97
Table 6. 12: Item Statistics: Grade 7.....	99
Table 6. 13: Item Statistics: Grade 8.....	101
Table 6. 14: Item Statistics: Grade 11 (Communication Arts, Science), Grade 10 (Mathematics).....	103
Table 6. 15: Item Statistics: Grade 11 Communication Arts Breach Form.....	105
Table 6. 16: Item Fit Statistics for Misfitting Items.....	106
Table 6. 17: LOSS and HOSS Values by Grade and Content Area.....	106
Table 7.1: Participation Rates: All Students.....	131
Table 7.2: Participation Rates: Males.....	131
Table 7.3: Participation Rates: Females.....	131
Table 7.4: Participation Rates: White.....	132
Table 7.5: Participation Rates: Black.....	132
Table 7.6: Participation Rates: Hispanic.....	132
Table 7.7: Participation Rates: Asian/Pacific Islander.....	133
Table 7.8: Participation Rates: Native American/Alaskan.....	133
Table 7.9: Participation Rates: Students Receiving Accommodations.....	133
Table 7.10: Summary Statistics for Communication Arts.....	134
Table 7.11: Summary Statistics for Mathematics.....	134
Table 7.12: Summary Statistics for Science.....	134
Table 7.13: Comparison of State-Level Means, 2006, 2007, and 2008 Census Data ...	135
Table 7.14: Comparison of Percent of Students in each Achievement Level, Communication Arts 2006, 2007, and 2008 Census Data.....	136
Table 7.15: Comparison of Percent of Students in each Achievement Level, Mathematics 2006, 2007, and 2008 Census Data.....	137
Table 7.16: Percent of Students in each Achievement Level, Science 2008 Census Data.....	137
Table 7.17: Summary of Score Reports for Spring 2008.....	138
Table 7.18: Types of Reports Available to Districts through Crystal Reports.....	138
Table 8. 1: Communication Arts Cut Scores.....	141
Table 8. 2: Mathematics Cut Scores.....	141
Table 8. 3: Science Cut Scores.....	141
Table 9. 1: Reliability in Communication Arts.....	167
Table 9. 2: Reliability in Mathematics.....	167
Table 9. 3: Reliability in Science.....	167

Table 9. 4: Decision Accuracy and Consistency Conditioned on Level of Achievement	168
Table 9. 5: Decision Accuracy and Consistency at Achievement Cut Points	168
Table 9. 6: SEM by Subgroup.....	169
Table 9. 7: Conditional Standard Error of Measurement at the Basic, Proficient, & Advanced Cut Scores.....	171
Table 9. 8: Principal Component Analysis for Communication Arts.....	172
Table 9. 9: Principal Component Analysis for Mathematics.....	173
Table 9. 10: Principal Component Analysis for Science	174
Table 9. 11: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 3.....	175
Table 9. 12: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 4.....	175
Table 9. 13: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 5.....	175
Table 9. 14: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 6.....	176
Table 9. 15: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 7.....	176
Table 9. 16: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 8.....	176
Table 9. 17: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 11.....	177
Table 9. 18: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 3.....	177
Table 9. 19: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards Mathematics, Grade 4.....	177
Table 9. 20: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 5.....	178
Table 9. 21: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 6.....	178
Table 9. 22: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 7.....	178

Table 9. 23: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 8.....	179
Table 9. 24: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 10.....	179
Table 9. 25: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Science, Grade 5.....	179
Table 9. 26: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Science, Grade 8.....	180
Table 9. 27: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Science, Grade 11	180
Table 9. 28: Inter-Correlation of Communication Arts, Mathematics, and Science Scale Scores.....	180
Table 10. 1: 2008 MAP SMD Statistics: Number of Flagged Items, Communication Arts	204
Table 10. 2: 2008 MAP SMD Statistics: Number of Flagged Items, Mathematics.....	205
Table 10. 3: 2008 MAP SMD Statistics: Number of Flagged Items, Science.....	206
Table 10. 4: Impact Analysis, Grade 3	207
Table 10. 5: Impact Analysis, Grade 4	207
Table 10. 6: Impact Analysis, Grade 5	208
Table 10. 7: Impact Analysis, Grade 6	209
Table 10. 8: Impact Analysis, Grade 7	209
Table 10. 9: Impact Analysis, Grade 8	210
Table 10. 10: Impact Analysis, Grade 11 (Communication Arts, Science), Grade 10 (Mathematics)	211
Table 10. 11: Impact Analysis, Grade 11 Breach	211

Table of Figures

Figure 4. 1: Sample Script of Examiner’s Manual	59
Figure 4. 2: District Report Form	60
Figure 4. 3: Test Book Accountability Form	60
Figure 6. 1: Item characteristic curve for Grade 3 Communication Arts, Session 4, Item 1 (Z = 342.93)	107
Figure 6. 2: Item characteristic curve for Grade 4 Communication Arts, Session 2, Item 17 (Z = 166.51)	107
Figure 6. 3: Item characteristic curve for Grade 5 Communication Arts, Session 1, Item 3 (Z = 400.72)	108
Figure 6. 4: Item characteristic curve for Grade 5 Communication Arts, Session 1, Item 4 (Z = 229.02)	108
Figure 6. 5: Item characteristic curve for Grade 6 Communication Arts, Session 1, Item 6A (Z = 175.21)	109
Figure 6. 6: Item characteristic curve for Grade 8 Communication Arts, Session 1, Item 4 (Z = 206.03)	109
Figure 6. 7: Item characteristic curve for Grade 8 Communication Arts, Session 2, Item 11 (Z = 205.58)	110
Figure 6. 8: Item characteristic curve for Grade 8 Communication Arts, Session 2, Item 38 (Z = 230.62)	110
Figure 6. 9: Item characteristic curve for Grade 11 Communication Arts, Session 3, Item 29 (Z = 164.84)	110
Figure 6. 10: Item characteristic curve for Grade 11 Communication Arts breach form, Session 3, Item 33 (Z = 3.60).....	110
Figure 6. 11: Item characteristic curve for Grade 4 Mathematics, Session 3, Item 8 (Z = 430.14)	111
Figure 6. 12: Item characteristic curve for Grade 5 Mathematics, Session 3, Item 3 (Z = 228.08)	111
Figure 6. 13: Item characteristic curve for Grade 5 Mathematics, Session 3, Item 4 (Z = 151.06)	112
Figure 6. 14: Item characteristic curve for Grade 5 Mathematics, Session 3, Item 5 (Z = 384.17)	112
Figure 6. 15: Item characteristic curve for Grade 6 Mathematics, Session 3, Item 7 (Z = 294.91)	113
Figure 6. 16: Item characteristic curve for Grade 7 Mathematics, Session 3, Item 3 (Z = 173.79)	113
Figure 6. 17: Item characteristic curve for Grade 8 Mathematics, Session 2, Item 16 (Z = 259.40)	114
Figure 6. 18: Item characteristic curve for Grade 8 Mathematics, Session 3, Item 5 (Z = 253.38)	114
Figure 6. 19: Item characteristic curve for Grade 10 Mathematics, Session 2, Item 9 (Z = 312.37)	115
Figure 6. 20: Item characteristic curve for Grade 10 Mathematics, Session 3, Item 5 (Z = 287.86)	115

Figure 6. 21: Item characteristic curve for Grade 5 Science, Session 3, Item 2 (Z = 234.12)	116
Figure 6. 22: Item characteristic curve for Grade 8 Science, Session 2, Item 32 (Z = 184.31)	117
Figure 6. 23: Item characteristic curve for Grade 11 Science, Session 3, Item 15 (Z = 146.34)	117
Figure 6. 24: Cross-Grade Articulation of Scale Scores at Selected Percentiles, Communication Arts MAP	118
Figure 6. 25: Cross-Grade Articulation of Scale Scores at Selected Percentiles, Mathematics MAP	119
Figure 6. 26: Cross-Grade Articulation of Scale Scores at Selected Percentiles, Science MAP	120
Figure 6. 27: Communication Arts Test Characteristic Curves (TCC) by grade, 2008..	121
Figure 6. 28: Mathematics Test Characteristic Curves (TCC) by grade, 2008.....	121
Figure 6. 29: Science Test Characteristic Curves (TCC) by grade, 2008.....	122
Figure 9. 1: SEM Plot Communication Arts, Grade 3	181
Figure 9. 2: SEM Plot Communication Arts, Grade 4	181
Figure 9. 3: SEM Plot Communication Arts, Grade 5	182
Figure 9. 4: SEM Plot Communication Arts, Grade 6	182
Figure 9. 5: SEM Plot Communication Arts, Grade 7	183
Figure 9. 6: SEM Plot Communication Arts, Grade 8	183
Figure 9. 7: SEM Plot Communication Arts, Grade 11	184
Figure 9. 8: SEM Plot, Communication Arts Grade 11 Breach.....	184
Figure 9. 9: SEM Plot Mathematics, Grade 3	185
Figure 9. 10: SEM Plot Mathematics, Grade 4	185
Figure 9. 11: SEM Plot Mathematics, Grade 5	186
Figure 9. 12: SEM Plot Mathematics, Grade 6	186
Figure 9. 13: SEM Plot Mathematics, Grade 7	187
Figure 9. 14: SEM Plot Mathematics Grade, 10	188
Figure 9. 15: SEM Plot Science, Grade 5	188
Figure 9. 16: SEM Plot Science, Grade 8	189
Figure 9. 17: SEM Plot Science, Grade 11	189
Figure 9. 18: Scree Plot Communication Arts, Grade 3	190
Figure 9. 19: Scree Plot Communication Arts, Grade 4	190
Figure 9. 20: Scree Plot Communication Arts, Grade 5	191
Figure 9. 21: Scree Plot Communication Arts, Grade 6	191
Figure 9. 22: Scree Plot Communication Arts, Grade 7	192
Figure 9. 23: Scree Plot Communication Arts, Grade 8	192
Figure 9. 24: Scree Plot Communication Arts, Grade 11	193
Figure 9. 25: Scree Plot Communication Arts, Grade 11 Breach.....	193
Figure 9. 26: Scree Plot Mathematics, Grade 3	194
Figure 9. 27: Scree Plot Mathematics, Grade 4	194
Figure 9. 28: Scree Plot Mathematics, Grade 5	195
Figure 9. 29: Scree Plot Mathematics, Grade 6	195
Figure 9. 30: Scree Plot Mathematics, Grade 7	196

Figure 9. 31: Scree Plot Mathematics, Grade 8	196
Figure 9. 32: Scree Plot Mathematics, Grade 10	197
Figure 9. 33: Scree Plot Science, Grade 5	197
Figure 9. 34: Scree Plot Science, Grade 8	198
Figure 9. 35: Scree Plot Science, Grade 11	198

EXECUTIVE SUMMARY

This report is a technical summary of the 2008 operational administration of the Missouri Assessment Program (MAP). The MAP is a grade-level test in Communication Arts administered in Grades 3 through 8 and 11. It is also a grade-level test in Mathematics administered in Grade 3 through 8 and 10. The MAP is a grade-span test in Science administered in Grades 5, 8, and 11. The MAP is designed to measure students' knowledge of Communication Arts, Mathematics, and Science. This section provides a summary of the 2008 Technical Report.

E.1 Background

The MAP was originally designed as grade-span tests to measure Missouri's Show-Me Standards. These standards were adopted by the Missouri State Board of Education in 1996. Since their inception, Missouri's Show-Me Standards have been further refined to better delineate Content Standards, Process Standards, and Content Strands/Grade-Level Expectations as Missouri changed their testing program to comply with the requirements of No Child Left Behind. Starting in 2006, grade-level tests were administered in Communication Arts and Mathematics. In 2008, grade-span tests were administered in Science. The MAP tests have therefore undergone multiple alignment analyses to ensure that MAP content reflects these refinements. Further details of the development of the 2008 MAP may be found in Chapter 3 of this report.

E.2 Administration

In the spring of 2008, Missouri administered assessments in Communication Arts to students in Grades 3 through 8 and 11, Mathematics to students in Grades 3 through 8 and 10, and Science to students in Grades 5, 8, and 11. The MAP was administered from March 31 to May 2, 2008. Test administration is discussed in Chapter 4 of this report.

Approximately 550 districts administered Communication Arts and/or Mathematics MAP tests in Grades 3 through 8 and/or high school. These districts also administered Science MAP tests in Grades 5, 8, and 11. Table E.1 shows participation rates based on the census data.¹ For the purposes of this report, participation rate is defined as the percent of students who received a valid scale score given the total number of students who received a test book. The "accountable" column shows the total number of students who received a test book. The reportable students are the number of students who received a scale score on MAP. Further analysis of participation rates is provided in Chapter 7 of this report.

E.3 Student Performance

This is the third year of the grade-level MAP testing programs in Communication Arts and Mathematics and the first year for the grade-span tests in Science. Table E.2 presents

¹ The census data used in this report does not reflect additional cleaning steps that DESE staff implements once CTB releases data to DESE; therefore, the numbers in this report may differ from numbers that DESE reports using their cleaned data.

the percentage of students classified as *Proficient* or *Advanced* in 2006, 2007, and 2008 in both Communication Arts and Mathematics. Table E.2 also shows the percentage of students classified as *Proficient* or *Advanced* in 2008 on the Science MAP.

For Communication Arts and Mathematics, small decreases in performance were seen in Grades 3 through 5 in both content areas, and increases were observed in Grades 6 through 8 in both content areas. On the high school assessments, there was a small decline in performance on the Communication Arts test, and there was an increase in performance on the Mathematics test. This was the first operational year for the Science MAP. More information on student performance may be found in Chapter 7 of this report.

E.4 Validity and Test Scores

Most sections of this technical report are designed to provide validity evidence to support the use of MAP test scores. Chapter 2 discusses the concept of validity and the uses of MAP scores. Chapter 3 focuses on the test development process used to create MAP. The discussions in this section are important to the content-related validity of the MAP scores. Chapter 4 presents information on test administration. Chapter 5 discusses the scoring of constructed-response items, as well as the results of the inter-rater reliability studies. Chapter 6 presents the scaling and linking procedures as well as the results of other operational data analyses. Chapter 7 reviews the results of the 2008 operational administration and overviews the score reports sent to parents, schools, and districts. Chapter 8 highlights the standard setting procedures used in 2008. Chapter 9 discusses reliability and construct-related validity. In this section, we evaluate the assumption that the content-area MAPs are unidimensional. For example, the grade-level Mathematics MAP should measure one primary dimension (Mathematics). Chapter 10 overviews the statistical and development processes used to assure fairness of the MAP for all examinees. Some analyses in this document are based on the calibration sample while other are based on census data. The source of data used for particular analyses is indicated throughout the Technical Report.

Table E.1: Participation Rates: All Students

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	66,357	99.73	66,357	99.85	-	-
4	67,049	99.74	67,049	99.84	-	-
5	65,734	99.71	65,734	99.85	65,734	99.77
6	65,830	99.76	65,830	99.83	-	-
7	66,923	99.67	66,923	99.71	-	-
8	67,574	99.56	67,574	99.61	67,574	99.46
HS	61,512	99.23	69,220	99.36	62,700	99.10
11b	1,309	99.24				

Table E.2: Percentage of Students Classified as *Proficient* or *Advanced* in 2006, 2007, and 2008 using Census Data: Communication Arts, Mathematics, and Science

Grade	Communication Arts				Mathematics				Science
	2006	2007	2008	2008-2007	2006	2007	2008	2008-2007	2008
3	43.0	43.2	40.3	-2.9	43.7	45.4	43.8	-1.6	
4	44.3	45.6	45.0	-0.6	43.8	44.9	44.2	-0.7	
5	45.5	48.3	48.2	-0.1	43.8	47.0	45.8	-1.2	44.5
6	42.7	44.1	47.3	3.2	44.3	48.3	50.7	2.4	
7	43.5	45.2	49.0	3.8	43.5	45.5	49.5	4.0	
8	42.1	42.2	48.1	5.9	40.4	41.2	43.8	2.6	43.2
HS	42.5	41.5	39.2	-2.3	42.3	40.8	45.6	4.8	47.2

CHAPTER 1: INTRODUCTION

The 2008 Missouri Assessment Program (MAP) marked the third administration of grade-level Communication Arts and Mathematics MAP in Missouri. This was the first administration of the grade-span Science MAP at Grades 5, 8, and 11. The MAP is designed to measure students' knowledge of Communication Arts, Mathematics, and Science. This report provides a technical overview of the Communication Arts, Mathematics, and Science assessments of the 2008 MAP. As such, it presents evidence for the validity of the 2008 MAP scores.

This chapter of the Technical Report serves to describe the background, history, purpose, and design of the Missouri Assessment Program (MAP), followed by an overview of the major sections for the current report.

1.1 Background of the Missouri Assessment Program

The MAP traces its origin to the 1993 Outstanding Schools Acts. This act required that Missouri create a statewide assessment system that measured challenging academic standards. From this act, grade-span assessments were created that measured Missouri's Show-Me standards. Historically, MAP was designed to be a grade-span test: Grades 3, 7, and 11 in Communication Arts, Grades 4, 8, and 10 in Mathematics, and Grades 3, 7, and 10 in Science. Table 1.1 provides a brief timeline of the events of the grade-span MAP.

In 2001, the federal *No Child Left Behind* (NCLB) was enacted, which required states to develop grade-level tests to be administered in Grades 3 through 8 and once in Grades 10 through 12 in both Reading and Mathematics. It also required that states have in place Science assessments to be administered at least once in Grades 3 through 5, Grades 6 through 9, and Grades 10 through 12 by the 2007-2008 school year. Based on the NCLB legislation, student performance, reported in terms of proficiency categories, is used to determine the adequate yearly progress of students at the school, district, and state levels.

In response to NCLB, the Department of Elementary and Secondary Education (DESE) contracted with CTB/McGraw-Hill in 2003 to expand the testing program to grade-level testing for Communication Arts and Mathematics. This contract was renewed in 2007 and extends through 2013. In the spring of 2005, Missouri administered a field test in Communication Arts and Mathematics, which was the basis for the construction of the 2006 and 2007 operational test forms.

The construction of the new Science MAP has been on a different trajectory. In 2005 DESE contracted with CTB/McGraw-Hill to construct a grade-span Science assessment in order to comply with the requirements of No Child Left Behind. In Spring of 2006, Missouri administered a field test in Science, which was the basis for the construction of the 2008 and 2009 operational Science forms. The contract to create grade-span Science assessments was renewed in 2007. This contract extends through 2013.

Table 1.2 shows a timeline of the development history of the NCLB-compliant testing program.

1.2 Purpose of the Missouri Assessment Program

The MAP is designed to measure how well students acquire the skills and knowledge described in Missouri's Grade-Level Expectations. The assessments yield information on academic achievement at the student, class, school, district, and state levels. This information is used to diagnose individual student strengths and weaknesses in relation to the instruction of the GLEs and to gauge the overall quality of education throughout Missouri.

1.3 Design of the Missouri Assessment Program

The spring 2008 MAP administration consisted of 17 operational assessments. Within Grades 3 through 8, six versions of the operational form were administered in a grade/content area. These versions were spiraled within classrooms and differed only by a set of embedded field test items. Note that the field test items embedded in the MAP did not contribute to a student's scale score.

In high school, only one form of the test was administered in Mathematics and Science. In Communication Arts, a breach form was administered to four districts after it was discovered that these districts had access to the *TerraNova* form embedded in the 2008 MAP.

Each form contained a norm-referenced test form from which norm-referenced scores were derived. The norm-referenced items served as anchor items to link performance on the 2008 MAP to prior administrations. These counted toward the student scale score if they could be mapped to a Missouri Grade-Level Expectation (GLE). If an item could not be mapped to a Missouri GLE, then it did not count to the criterion-referenced score, nor was it used as an anchor item. Table 1.3 shows the number of items that could not be mapped to a Missouri GLE. Table 1.4 provides an overview of the 2008 MAP test design.

Braille and large print versions of each operational MAP form were constructed for each grade/content area to enable visually impaired students to participate in MAP testing. At some grade levels/content areas, it was necessary to drop items from the assessment due to difficulties associated with the Braille translation. Table 1.5 lists these items that were omitted from the Braille forms. Note that students taking the Braille form were given full credit for the omitted items.

1.4 Overview of this Report

This Technical Report documents the major activities of the testing cycle in the subsequent chapters. This report provides comprehensive detail that confirms the processes and procedures applied in the MAP adhered to appropriate professional standards and practices of educational assessment. Ultimately, this report serves to document evidence that valid inferences about Missouri student performance can be

derived from the MAP. An overview of major activities documented within this report is provided below:

Validity and the Use of Test Scores (Chapter 2)

Chapter 2 of the Technical Report discusses the concept of validity evidence. This Technical Report is comprised of evidence that supports the use of the MAP score. In Chapter 2, we discuss some of the uses of the MAP score.

Item and Test Development (Chapter 3)

Chapter 3 of the technical report provides a summary of the major test development activities that occurred to create the spring 2008 operational test forms, the embedded field test items, and the materials developed to inform the public about the testing program. As each major event is presented and discussed, the role of the event in contributing to evidence for validity of the use of test results is discussed.

Test Administration (Chapter 4)

Chapter 4 of the Technical Report serves to describe the processes and activities implemented and information disseminated to help ensure standardized test administration procedures and, thus, uniform test administration conditions for students.

Scoring Constructed-Response Items (Chapter 5)

Chapter 5 describes the processes and activities for scoring constructed-response items. This chapter discusses how raters are trained and the measures for assuring consistency among scorers. Finally, this chapter presents the results of the inter-rater reliability studies.

Operational Analyses (Chapter 6)

Chapter 6 of the Technical Report includes a detailed description of the operational analyses of the 2008 MAP, which are comprised of three major parts: the calibration sample, the classical item analysis and the calibration, scaling, and linking using IRT models. This chapter describes the demographics of the calibration sample and compares it to the state census data. It reports the results of the classical item analysis as well as the results of the calibration, scaling, and linking.

Test Results and Reporting (Chapter 7)

Chapter 7 of the Technical Report contains information on the results of the spring 2008 administration of the MAP. Detailed summary statistics based on scale scores and achievement level information are also provided. Finally, this chapter presents information on the score reports sent to parents, schools, and districts.

Standard Setting (Chapter 8)

Chapter 8 of the Technical Report briefly discusses standard setting. It provides an overview of the standard setting activities that occurred for the MAP in 2008.

Reliability and Validity Evidence (Chapter 9)

Chapter 9 of the Technical Report provides evidence of reliability and validity of MAP scores. This chapter provides detailed results of the reliability of the tests as well as information on the decision consistency of the cut scores. It also provides evidence of construct validity for MAP scores.

Fairness (Chapter 10)

Chapter 10 of the Technical Reports discusses fairness and how the MAP tests are constructed to be fair to all Missouri students. This chapter summarizes the results of the differential item (DIF) analysis. It also discusses the results of an impact analysis to determine if large differences exist between demographic groups in Missouri.

Table 1.1: Timeline of Grade-Span MAP

Year	Event
1996	Show-Me Standards Approved
1996	Frameworks for Curriculum Development published
1997	Annotations to the Curriculum Frameworks published
1998	First operational administration of Mathematics MAP (Grades 4, 8, and 10)
1999	First operational administration of Communication Arts MAP (Grades 3, 7, and 11) and Science MAP (Grades 4, 8, and 11)
2000	First operational administration of Social Studies MAP (Grades 4, 8, and 10)
2001	Mathematics Curriculum Supplement published
2005	Last year of grade-span MAP

Table 1.2: Timeline of Grade-Level MAP

Year	Event
2004	Grade-Level Expectations published
2005	Communication Arts and Mathematics Field Test
2005	Standard Setting for Communication Arts and Mathematics
2006	First Operational Communication Arts and Mathematics MAP
2007	Science Field Test
2008	First Operational Science MAP
2008	Standard Setting for Science

Table 1.3: Number of Items that did not Map to a Missouri Grade-Level Expectation

Content	Grade	Number of Items
Science	5	3
	8	1
	11	2

Table 1.4: Spring 2008 MAP Test Design

Content	Grade	Number of Forms	Operational Items	Anchor Items	Total Raw Score Points	Embedded Field Test Items	Total Number of OP Items
Communication Arts	3	6	28	30	68	3-5	58
	4	6	21	35	63	3-5	56
	5	6	24	32	63	3-5	56
	6	6	24	31	62	3-4	55
	7	6	29	33	73	3-5	62
	8	6	27	34	69	3-5	61
	11	2	29	34	73		63
Mathematics	3	6	30	30	67	9	60
	4	6	33	32	77	10	65
	5	6	30	32	69	9	62
	6	6	30	31	68	9	61
	7	6	30	32	69	9	62
	8	6	33	31	76	10	64
	10	1	33	25	70		58
Science	5	6	31	22	80	10-11	53
	8	6	34	24	87	10	58
	11	1	35	23	92		58

Table 1.5: Spring 2008 Items Removed from Braille Forms

Content	Grade	Session	Item
Mathematics	3	1	15
	3	2	26
	4	2	24
	5	3	7
	6	1	12
	6	1	22
	6	3	3
	7	1	23
	7	3	7
	8	1	17
	8	2	17
	10	3	4
	10	3	8
Science	5	2	8
	5	3	2
	5	3	4
	8	1	4
	8	2	4
	8	3	6
	11	2	3

CHAPTER 2: VALIDITY AND THE USES OF TEST SCORES

Validity is the overarching component of the MAP testing program. The following excerpt is from the *Standards for Educational and Psychological Testing* (AERA, 1999) [hereafter referred to as the *Standards*]:

Ultimately, the validity of an intended interpretation of test scores relies on all the available evidence relevant to the technical quality of a testing system. This includes evidence of careful test construction; adequate score reliability; appropriate test administration and scoring; accurate score scaling, equating, and standard setting; and careful attention to fairness for all examinees (17).

As stated by the *Standards*, the validity of a testing program hinges on the use of the test scores. *Validity evidence* that supports the uses of the MAP test scores is provided in this Technical Report. In this section, we examine some possible uses of the MAP test scores.

The following sections (Chapters 3 through 10) of this Technical Report provide additional evidence for these uses, as well as technical support for some of the interpretations and uses of test scores. The information in Chapters 3 through 10 also provides a firm foundation that the MAP tests measure what they are intended to measure. However, this Technical Report cannot anticipate all possible interpretations and uses of MAP scores. It is recommended that policy and program evaluation studies, in accordance with the *Standards*, be conducted to support some of the uses of the MAP scores. To this end, DESE is conducting a study on consequential validity that is being implemented by Assessment Resource Center.

2.1 Uses of Test Scores

The validity of a test score ultimately rests in how that test score is used. To understand whether a test score is being used properly, we must first understand the purpose of the test. The intended uses of MAP scores include:

- identifying students' strengths and weaknesses on Missouri's Grade-Level Expectations
- communicating expectations for all students
- evaluating school-, district-, and/or state-level programs
- informing stakeholders (teachers, school administrators, district administrators, DESE staff, parents, and the public) on the status of the progress toward meeting academic achievement standards of the state
- meeting the requirements to measure Adequate Yearly Progress by NCLB
- meeting the requirements of the state's accountability program, Missouri School Improvement Program (MSIP)

This Technical Report refers to the use of several kinds of scores: the test-level scores (scale scores and achievement levels), the content standard scores, and the process standard scores.

2.2 Test-Level Scores

At the test level, an overall scale score that is based on student performance on the entire test is reported. In addition, an associated level of achievement is reported. These scores indicate, in varying ways, a student's achievement in Communication Arts, Mathematics, or Science. Test-level scores are reported at four reporting levels: the state, the school district, the school, and the student.

Custom-written portions of the MAP were directly authored by Missouri educators, edited by both CTB and Missouri educators, and subsequently reviewed and approved for use by Missouri educators. This procedure fosters a close relationship between the items and the Missouri Show-Me Standards, from which the MAP was developed. Portions of the MAP from CTB's item pool were also aligned to Missouri Content Standards, Process Standards, and Grade-Level Expectations (GLEs) to further solidify the Show-Me Standards as the foundation of the MAP. As shown in Table 1.3 in the previous chapter, all *TerraNova* items in Communication Arts and Mathematics MAP align to Missouri standards. Only three Grade 5 Science items, one Grade 8 Science item, and two Grade 11 Science items did not map to Missouri standards. Item development is described in Chapter 3; however, detailed descriptions of processes used to delineate the knowledge, skills, abilities, including content limits and descriptions for each content area, are beyond the scope of this report.

At the test level, two types of scores are reported to indicate a student's achievement on the MAP: (1) a scale score and (2) its associated level of achievement.

1. Scale Score

A scale score indicating a student's total performance is determined for each content area on the MAP. The overall scale score for a content area quantifies the achievement being measured by the Communication Arts, Mathematics, or Science test. In other words, the scale score represents the students' level of achievement, where higher scale scores indicate higher levels of achievement on the test and lower scale scores indicate the opposite.

2. Level of Achievement

A student's performance on the on the Communication Arts, Mathematics, or Science MAP is reported in one of four levels of achievement: *Below Basic*, *Basic*, *Proficient*, or *Advanced*. The cut scores for the levels of achievement were recommended by Missouri educators and citizens at the Bookmark Standard Setting Workshop in December 2005 for Communication Arts and Mathematics and in July 2008 for Science. The cut scores reflect the expectations of Missouri educators and citizens of what Missouri students should know and be able to do in each grade/content area. The Missouri Show-Me Standards guided these recommendations, as did Senate Bill 1080. (See Chapter 8 of this

report for a discussion of MAP standard setting.) Thus, MAP achievement levels reflect the achievement standards and abilities intended by the Missouri legislature, Missouri teachers, Missouri citizens, and DESE. Descriptions of each level of achievement in terms of what a student should know and be able to do are provided with the *Guide to Interpreting Results* (see Chapter 3).

2.2.1 Use of Test-Level Scores

MAP scale scores and achievement levels provide summary evidence of student achievement in Communication Arts, Mathematics, or Science. Classroom teachers may use these scores as evidence of student achievement in these content areas. At the aggregate level, district and school administrators may use this information for activities such as planning curriculum. At the state level, the aggregate test-level scale scores are used for accountability programs associated with No Child Left Behind and the Missouri School Improvement Program. The results presented in this Technical Report provide evidence that the scale scores are a valid and reliable indicator of student performance in Communication Arts, Mathematics, and Science.

2.3 Content Standard Subscores

The Content Standard subscores indicate student performance in terms of the number- and percent-correct score for each Content Standard in Communication Arts and each GLE strand in Mathematics and Science. Starting in 2008, Content Standard subscores were reported only through DESE's Crystal Reporting system. These scores may be aggregated by the state, district, or schools to determine the mean Content Standard subscores. These means may be used as indicators of the performance of the school or district in teaching students the knowledge and skills defined for each subject area.

2.3.1 Use of the Content Standard Subscores

The purpose of reporting Content Standard subscores on MAP is to show for each student the relationship between the overall achievement being measured and the skills in each of the areas delimited by the Content Standards in Communication Arts and the GLE strands in Mathematics and Science. Teachers may use these subscores for individual students as indicators of strengths and weaknesses, but they are best corroborated by other evidence, such as homework, class participation, diagnostic test scores, or observation. Chapter 3 of this Technical Report provides content validity evidence that supports the use of the Content Standard subscores. Chapter 9 of this Technical Report provides evidence of construct validity that further supports the use of the Content Standard subscores.

District and school administrators may compare their aggregate results with the state mean to better understand their strengths and weaknesses within a content area. Caution should be exercised when comparing Content Standards subscores between students or across years. The user should be aware that different items will comprise the Content Standards across years and that these items may vary in difficulty.

2.4 Process Standard Subscores

For each MAP content area, Process Standard and Content Standard subscores are determined from the same pool of items. These items were classified by the particular underlying processes used to teach each item's content, and each item's assigned Process Standard was verified by Missouri teachers in a Content Review workshop specifically designed to fulfill that purpose. Content Standard and Process Standard subscores generally show a directly proportional relationship, because the same pool of items is used to measure both sets of standards. Process Standard subscores are only reported through DESE's Crystal Reporting system.

2.4.1. Use of the Process Standard Subscores

The purpose of reporting Process Standard subscores on MAP is to show the achievement of students in each of the areas delimited by the Process Standards in Communication Arts, Mathematics, or Science. When the Process Standard processes are used to teach the subject area content, the Process Standard scores can be said to reflect the strategies Missouri teachers want Missouri students to adopt in the learning and handling of "real world" activities.

Caution should be exercised when making comparisons of Process Standards subscores between students or across years. The user should be aware that different items will comprise the Process Standards across years and that these items may vary in difficulty.

CHAPTER 3: TEST CONTENT DEVELOPMENT

Content-related validity in achievement tests is evidenced by a correspondence between test content and a specification of the content domain. Content-related validity can be demonstrated through consistent adherence to test blueprints, through a high-quality test development process that includes review of items for accessibility to English Language Learners and students with disabilities, and through alignment studies performed by independent groups. In this section, we will provide a detailed discussion of the test development cycle, from aligning items with Missouri’s rigorous Show-Me Standards and GLE strands to selecting items for the final operational test form. In particular, this section will show how the Missouri Assessment Program (MAP) follows rigorous procedures to select tests that reflect the full range of content that MAP is expected to cover.

3.1 Test Specifications

Evidence of validity based on test content includes information about the test specifications, including the test design and test blueprint. Test development involves creating a design framework from the statement of the construct to be measured. The MAP test specifications evolve from the tension between the constraints of the assessment program and the benefits sought from the examination of students. Many of the benefits sought are not scientific in nature, nor are many of the constraints; rather, they are policy considerations. The 2008–2009 MAP specifications were finalized in August 2007 prior to item selection for the operational forms.

The MAP test specifications consist of a test blueprint and a test design for each grade level/content area. The key structural aspect of the MAP tests is the test blueprint, which specifies the target score points for each Content Standard (Table 3.1). The blueprint represents a compromise between many constraints, including the target weights for each Content Standard recommended by Missouri teachers, availability of items from field testing, and results of multiple reviews by content specialists. Test design elements include such elements as number and types of items/tasks for each of the scores reported (tasks are measured by constructed-response items in MAP). The MAP test design is documented in this chapter with *item maps* (Table 3.2), which show the distribution of items/tasks by Content Standards (Communication Arts) and GLE Strands (Mathematics and Science). The item maps show the design of the test administration by representing the sessions into which the test is divided (*session* assignments determine which items will be taken together). The degree to which the 2008 MAP operational forms matched the test blueprint can be assessed by comparing the targeted score point distributions defined in the test blueprint with the actual point distributions in Tables 3.5–3.7. Actual point distributions on the 2008 MAP operational forms matched blueprint targets within 10% of the blueprint targets, which was the tolerance for variation approved by DESE.

Table 3. 1: MAP Test Blueprint: Target Score Points by Content Standard (Communication Arts) or GLE Strand (Mathematics and Science)

Content Area Content Standard/ GLE Strand	Grade							
	3	4	5	6	7	8	10	11
Reading								
Speaking/Writing Standard English	15	10	12	13	16	15		15
Reading—Fiction & Nonfiction	48	54	52	51	50	53		52
Writing Formally & Informally	6	2	2	1	7	1		6
Mathematics								
Number and Operations	25	19	14	14	14	12	11	
Algebraic Relationships	14	16	14	14	14	23	23	
Geometric and Spatial Relationships	14	16	14	14	14	15	15	
Measurement	10	16	14	14	14	12	11	
Data and Probability	7	11	14	14	14	15	15	
Science								
Matter and Energy			11			11		12
Force and Motion			8			7		10
Living Organisms			8			10		11
Ecology			9			8		8
Earth Systems			10			11		8
Universe			9			9		8
Scientific Inquiry			21			24		27
Science, Technology, and Human Activity			7			6		6

Table 3. 2: Content Coverage: MAP 2008 Item Maps

Communication Arts Grade 3						
Content Standard		Session 1		Session 2	Session 3	Session 4
		CR Item #	SR Item #	CR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12		10, 11, 18, 19, 20, 21, 27, 28, 38, 39	
2	Reading—Fiction/Poetry/Drama				12, 13, 14, 15, 16, 17, 22, 23, 24, 25, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37	1, 2, 3A
3	Reading—Nonfiction	3, 4, 5, 6A	1, 2		1, 2, 3, 4, 5, 6, 7, 8, 9	
4	Writing Formally & Informally	6B, 6C		1WP		3B
5	Combined Reading from Standards 2 & 3	3, 4, 5, 6A	1, 2		1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 13, 14, 15, 16, 17, 22, 23, 24, 25, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37	1, 2, 3A

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Communication Arts Grade 4					
Content Standard		Session 1		Session 2	Session 3
		CR Item #	SR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12	13, 14, 29, 30	
2	Reading—Fiction/Poetry/Drama			1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 31, 32, 33, 34, 35, 36, 37, 38, 39	1, 2, 3
3	Reading—Nonfiction	3, 4, 5A, 6	1, 2		
4	Writing Formally & Informally	5B, 5C			
5	Combined Reading from Standards 2 & 3	3, 4, 5A, 6	1, 2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 31, 32, 33, 34, 35, 36, 37, 38, 39	1, 2, 3

Communication Arts Grade 5					
Content Standard		Session 1		Session 2	Session 3
		CR Item #	SR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12	17, 18, 19, 20, 21, 28, 29	
2	Reading—Fiction/Poetry/Drama			1, 2, 3, 4, 5, 6, 7, 22, 23, 24, 25, 26, 27, 34, 35, 36, 37, 38, 39	
3	Reading—Nonfiction	3, 4, 5A, 6A	1, 2	8, 9, 10, 11, 12, 13, 14, 15, 16, 30, 31, 32, 33	1, 2, 3
4	Writing Formally & Informally	5B, 6B			
5	Combined Reading from Standards 2 & 3	3, 4, 5A, 6A	1, 2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 22, 23, 24, 25, 26, 27, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39	1, 2, 3

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Communication Arts Grade 6					
Content Standard		Session 1		Session 2	Session 3
		CR Item #	SR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12	13, 14, 19, 20, 31, 32, 33	
2	Reading—Fiction/Poetry/Drama	3, 4, 5, 6A	1, 2	1, 2, 3, 4, 5, 6, 7, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26	
3	Reading—Nonfiction			8, 9, 10, 11, 12, 27, 28, 29, 30, 34, 35, 36, 37, 38, 39	1, 2, 3
4	Writing Formally & Informally	6B			
5	Combined Reading from Standards 2 & 3	3, 4, 5, 6A	1, 2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 34, 35, 36, 37, 38, 39	1, 2, 3

Communication Arts Grade 7						
Content Standard		Session 1		Session 2	Session 3	Session 4
		CR Item #	SR Item #	CR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11, 12, 13, 14, 31, 32	
2	Reading—Fiction/Poetry/Drama	3, 4, 5, 6A	1, 2		5, 6, 7, 8, 9, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 33, 36, 37, 38, 39	
3	Reading—Nonfiction				1, 2, 3, 4, 33, 34, 35	1, 2, 3A
4	Writing Formally & Informally	6B		1WP		3B, 3C
5	Combined Reading from Standards 2 & 3	3, 4, 5, 6A	1, 2		1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 33, 34, 35, 36, 37, 38, 39	1, 2, 3A

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Communication Arts Grade 6					
Content Standard		Session 1		Session 2	Session 3
		CR Item #	SR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12, 13, 14, 15, 16	18, 19, 20, 25, 26	
2	Reading—Fiction/Poetry/Drama	3, 4, 5, 6A	1, 2	12, 13, 14, 15, 16, 17, 21, 22, 23, 24	
3	Reading—Nonfiction			1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39	1, 2, 3A
4	Writing Formally, & Informally	6B, 6C			3B
5	Combined Reading from Standards 2 & 3	3, 4, 5, 6A	1, 2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39	1, 2, 3A

Communication Arts Grade 11						
Content Standard		Session 1		Session 2	Session 3	Session 4
		CR Item #	SR Item #	CR Item #	SR Item # (including TN Reading NRT)	CR Item #
1	Speaking/Writing Standard English		7, 8, 9, 10, 11, 12, 13, 14, 15, 16		13, 15, 27, 39	
2	Reading—Fiction/Poetry/Drama	3, 4, 5, 6A	1, 2		16, 17, 18, 19, 20	
3	Reading—Nonfiction				1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38	1, 2, 3A
4	Writing Formally & Informally	6B, 6C		1 WP	32	3B, 3C
5	Combined Reading from Standards 2 & 3	3, 4, 5, 6A				1, 2, 3A

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Mathematics Grade 3				
GLE Strand		Session 1	Session 2	Session 3
1	Number and Operations	1, 5, 9, 14, 17, 21	1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 15, 18, 23, 24	3, 6
2	Algebraic Relationships	3, 6, 7, 10, 11, 18, 19, 22	11, 28, 30	2
3	Geometric and Spatial Relationships	2, 4, 12, 15, 23	17, 19, 26, 27	4, 7
4	Measurement	8, 13, 16, 20	10, 13, 16, 29	1
5	Data and Probability	14, 20, 21, 22, 25		5

Mathematics Grade 4					
GLE Strand		Session 1		Session 2 TN NRT SR Item #	Session 3 CR Item #
		SR Item #	PE Item #		
1	Number and Operations			1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 17, 20, 21, 22, 23, 27, 29	1, 6
2	Algebraic Relationships	3, 7, 14	31	9, 10	2, 7
3	Geometric and Spatial Relationships	2, 6, 13, 16, 18, 19, 23		16, 24, 30, 31	3
4	Measurement	1, 4, 8, 10, 11, 17, 21		18, 19, 26, 32	4, 8
5	Data and Probability	5, 9, 12, 15, 20, 22		13, 14, 15	5, 9

Mathematics Grade 5				
GLE Strand		Session 1 SR Item #	Session 2 TN NRT SR Item #	Session 3 CR Item #
1	Number and Operations		1, 2, 3, 4, 5, 7, 8, 9, 10, 12, 13, 15, 22, 26, 29, 30	
2	Algebraic Relationships	1, 6, 9, 13, 17, 21	6, 11, 28	3, 6
3	Geometric and Spatial Relationships	4, 7, 10, 14, 18, 22	25, 31, 32	2, 7
4	Measurement	3, 5, 11, 15, 19, 23	14, 16, 20, 21, 23, 24	4
5	Data and Probability	2, 8, 12, 16, 20	17, 18, 19, 27	1, 5

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Mathematics Grade 6				
GLE Strand		Session 1 SR Item #	Session 2 TN NRT SR Item #	Session 3 CR Item #
1	Number and Operations		1, 2, 3, 4, 5, 6, 7, 8, 12, 18, 19, 20, 21, 22, 25, 27, 31	
2	Algebraic Relationships	1, 5, 10, 16, 20	13, 14, 15, 28	1, 6
3	Geometric and Spatial Relationships	4, 6, 9, 12, 14, 19, 22, 23	17, 30	3
4	Measurement	3, 7, 13, 17, 18	23, 24, 26, 29	2, 5
5	Data and Probability	2, 8, 11, 15, 21	9, 10, 11, 16	4, 7

Mathematics Grade 7				
GLE Strand		Session 1 SR Item #	Session 2 TN NRT SR Item #	Session 3 CR Item #
1	Number and Operations		1, 2, 3, 4, 5, 6, 7, 8, 9, 13, 17, 18, 20, 21, 24, 32	
2	Algebraic Relationships	1, 3, 6, 9, 12, 15, 20	16, 25	1, 4
3	Geometric and Spatial Relationships	4, 7, 11, 16, 19, 23	14, 22, 26, 30	5, 7
4	Measurement	2, 8, 13, 17, 21	15, 27, 28, 29	3, 6
5	Data and Probability	5, 10, 14, 18, 22	10, 11, 12, 19, 23, 31	2

Mathematics Grade 8					
GLE Strand		Session 1		Session 2 TN NRT SR Item #	Session 3 CR Item #
		SR Item #	PE Item #		
1	Number and Operations			1, 2, 3, 4, 5, 6, 7, 10, 11, 15, 23, 24, 26, 30	
2	Algebraic Relationships	1, 5, 8, 12, 16, 19, 21, 23	31	16, 18, 19, 27, 31	3, 6
3	Geometric and Spatial Relationships	2, 6, 9, 13, 15, 17, 22		14, 17, 20, 25	2, 8
4	Measurement	3, 7, 10, 14, 18		12, 13, 28	5, 9
5	Data and Probability	4, 11, 20		8, 9, 21, 22, 29	1, 4, 7

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Mathematics Grade 10				
GLE Strand		Session 1 SR Item #	Session 2 TN NRT SR Item #	Session 3 CR Item #
1	Number and Operations		1, 2, 3, 4, 5, 10, 16, 17, 19, 20, 21	
2	Algebraic Relationships	1, 5, 7, 9, 11, 14, 17, 19, 23	6, 8, 23	1, 6, 10
3	Geometric and Spatial Relationships	2, 6, 12, 15, 18, 20, 22	22, 24, 25	4, 8
4	Measurement	3, 8, 16	11, 13, 14, 15	3, 7
5	Data and Probability	4, 10, 13, 21	7, 9, 12, 18	2, 5, 9

Science Grade 5				
GLE Strand	Session 1 CR Item #	Session 2		Session 3 PE Item #
		TN NRT SR Item #	CR Item #	
1	Matter and Energy		13, 20	32, 33, 35, 36
2	Force and Motion	5	11	26, 28
3	Living Organisms	8	2, 16, 17, 25	29
4	Ecology	1, 4	6, 12, 21	30, 31
5	Earth Systems	2, 11	3, 24	30, 31
6	Universe	3, 6, 7	4, 22	
7	Scientific Inquiry		1, 8, 14, 15, 23	1, 2, 3, 4, 5, 6, 7, 8, 9
8	Science, Technology, and Human Activity	9, 10	7, 9, 10	27

Table 3.2: Content Coverage: MAP 2008 Item Maps (cont'd)

Science Grade 8					
GLE Strand	Session 1 CR Item #	Session 2		Session 3 PE Item #	
		TN NRT SR Item #	CR Item #		
1	Matter and Energy	4, 5, 9, 11	2, 4	35	
2	Force and Motion	1, 10	7, 9, 25		
3	Living Organisms	2	1, 15, 17, 19	26, 28, 32	
4	Ecology	7	13, 18	30, 36	
5	Earth Systems	3, 11, 12, 21, 22		27, 33, 37	
6	Universe	3, 6, 12		34	
7	Scientific Inquiry		5, 6, 8, 10, 14, 16, 23		1, 2, 3, 4, 5, 6, 7, 8, 9, 10
8	Science, Technology, and Human Activity	8	20	29, 31	

Science Grade 11					
GLE Strand	Session 1 CR Item #	Session 2 PE Item #	Session 3		
			TN NRT SR Item #	CR Item #	
1	Matter and Energy	1, 2, 18		14, 18, 19, 25	28
2	Force and Motion	6, 9, 12		15, 20	26
3	Living Organisms	4, 5, 17		6, 16, 22, 23, 24	
4	Ecology	7, 8, 11		30	
5	Earth Systems	14, 15		3, 9, 10	29
6	Universe	13, 16			27, 31
7	Scientific Inquiry		1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	1, 4, 7, 8, 11, 12, 13	
8	Science, Technology, and Human Activity	3, 10		2, 17	

3.2 Item Development

Planning and preparation for the development of item content for the 2008 and 2009 MAP Operational Test forms began in 2004. The plan specified an item development and selection cycle that included an initial item writing/passage selection workshop, a local pilot study, a content and bias review, item refinements and form construction, a subsequent round of formal field testing, the selection of operational forms based on statistical data from the field test, and ultimately, operational testing at grade levels 3 through 8 and high school. Each of these steps is described in greater detail below.

3.3 Item Writing

Communication Arts and Mathematics: In February 2005, a group comprised of Missouri educators, Regional Instructional Facilitators, DESE staff, and CTB personnel participated in an Item Writing Workshop (IWW) for Communication Arts and Mathematics at the Resort at Port Arrowhead, located at Lake Ozark, Missouri. The workshops were conducted with more than 30 teacher participants per content area. Teacher participants were selected by DESE to represent educational sites throughout Missouri. During the first day of the workshop, Communication Arts participants selected reading passages. During the next three days, Communication Arts participants used selected passages as a basis for writing constructed-response (CR) items and writing prompts for the 2008 and 2009 Operational forms for grades 3–8 and 11. The Mathematics participants wrote CR items and performance-event (PE) items along with scoring guides to create a pool of items for the 2008 and 2009 Operational forms for grades 3–8 and 10. The content developed at the workshop was based specifically upon the Missouri Show-Me Standards and Grade Level Expectations (GLEs). Some selected-response (SR) items were developed by CTB after the workshop to help supplement the item pool and reviewed by DESE. Items were refined after the initial item writing workshop which led to the production of local pilot test forms.

Science: In November 2004, a group comprised of Missouri educators, Regional Instructional Facilitators, DESE staff, and CTB personnel participated in a four-day Science Item Development Workshop (IDW) in Columbia, Missouri. The IDW was conducted with 37 teacher participants selected by DESE on the basis of their prior experience and expertise in item development for MAP Science and to represent educational sites throughout Missouri. The purpose of the IDW was to revise existing items and write new items to ensure a well-balanced item pool for the 2008 and 2009 MAP Science operational tests. The existing items came from the MAP Science item pool previously developed for operational testing at grades 3, 7, and 10. During the first two days of the IDW, the existing items were revised to target the new MAP Science GLEs. These new GLEs were developed for the 2008 assessment to be administered at grades 5, 8, and 11. During the third and fourth days of the IDW, Science participants wrote new CR items and performance events. A new MAP Science Performance event development template was introduced at the IDW. This template specified the types of tasks and numbers of items that comprise a Performance event.

Overall, the item writing workshops in November 2004 and February 2005 provided a basis upon which items written for the Communication Arts, Mathematics, and Science assessments could be selected for use on small-scale local pilot tests administered throughout Missouri.

3.4 Local Pilot Test

Small-scale pilot tests were administered in March 2005 (Science) and November 2005 (Communication Arts and Mathematics) in a limited number of classrooms throughout Missouri. Teachers who administered the pilot tests were generally selected by DESE from the pool of Item Writing Workshop participants.

Six Communication Arts forms per grade were piloted, consisting of approximately two SR items and six CR items for each of grades 4, 5, 6, and 8. The six Communication Arts pilot forms for grades 3, 7, and 11 each contained two selected-response items, four constructed-response items, and one writing prompt. Six Mathematics forms per grade were piloted, consisting of approximately twelve SR items and two CR items for each of grades 3, 5, 6, and 7. The six Mathematics pilot forms for grades 4, 8, and 10 each contained twelve SR items, four CR items, and one performance event. Ten Science forms per grade, consisting of approximately 15 CR items, were piloted for each of grades 5, 8, and 11. In addition to these ten pilot forms, eight performance events were piloted at each grade level.

3.5 Score, Revise, Rewrite Workshop

In April 2005 (Science) and February 2006 (Communication Arts and Mathematics), the results of the pilot studies underwent further evaluation during Score, Revise, and Rewrite (SRR) Workshops.

The purpose of the SRR Workshop was for the participants to score the items piloted in Missouri classrooms and to revise the items and rubrics/scoring guides based on the scoring process, student results, and subsequent discussion. DESE invited approximately 5 to 7 participants per grade/content area, resulting in the direct participation of approximately 100 Missouri educators in this step of the development process. CTB and DESE personnel were present to facilitate the SRR Workshop. The participants individually scored the students' pilot forms, tallied the results, and then reviewed the items as a group. Regional Instructional Facilitators (RIFs) were also present and participated in the process. Overall, the goal of the workshop was to improve the item quality prior to the next step in the process, Content and Bias Review, and to ensure that quality items were developed for future use in the Missouri Assessment Program. Most participants commented that this workshop was successful in this regard.

3.6 Content and Bias Review Workshop

Content and Bias Review (CBR) workshops were conducted in May 2005 (Science) and May 2006 (Communication Arts and Mathematics) with DESE, Missouri educators, Regional Facilitators, and CTB staff involved. Both of the CBR workshops were

conducted in Columbia, Missouri. For the Content Review, DESE invited participants from educational sites throughout Missouri to review items, writing prompts, performance events and scoring guides for content accuracy and grade level appropriateness. In Communication Arts, participants also reviewed passages. In addition, participants in all three content areas verified each item's alignment to the Missouri curriculum by reviewing the Content Standard, Process Standard, and GLE assignment for each item at the review. The Content Review was accomplished over the course of one or two days, and was followed a one- or two-day Bias Review. The Bias review committee was comprised of representatives from various backgrounds whose purpose was to screen the items for any racial, socioeconomic, gender, or other sensitivity issues. This committee could revise or reject items because of issues related to possible bias. Only four Communication Arts items and no Mathematics items were rejected from their respective pools. The remaining items were either accepted or accepted with revisions.

For each content area, over 30 Missouri educators participated in the process to help ensure content validity. Greater than 90% of reviewed items were accepted by the CBR committees. The general consensus was that the items as a group were well written and edited, and that the changes made during and after the conclusion of the Score, Revise, Rewrite Workshop had contributed to a smooth CBR workshop. The accepted items became candidates for the next step in the process, the MAP field test.

3.7 Field Test Selection and Administration

The items approved by CBR committees became the basis for the formation of stand-alone Field Test forms administered in 2006 and 2007. The custom-written material was arranged into test forms using *TerraNova* Survey as a common anchor across forms. (The same anchor would become the NRT portion of the 2008 operational test and is described in more detail in the following section). Field test items were selected and placed into forms so that the combined coverage of the NRT and customized portions of the test met the established blueprint requirements for content coverage; each field test form was constructed using the same design.

The MAP Spring 2006 Science Field Test consisted of four parallel forms per grade level, which were successfully administered at grades 5, 8, and 11 in May 2006. The MAP Spring 2007 Communication Arts and Mathematics Field Tests consisted of six parallel forms per grade/content area which were successfully administered at grades 3–8 and high school in May 2007. All field test forms were reviewed and approved by DESE prior to administration. The field tests generated item statistics that were used to help select two years of parallel operational forms, to be administered in 2008 and 2009.

3.8 Operational Test Selection

The use of an embedded *TerraNova* Survey provides both a horizontal anchor in the MAP tests and a norm-referenced (NRT) subtest, which is a requirement of the MAP. For most grade/content areas, the intact *TerraNova* Survey Form D was embedded in the 2006 and 2007 Field Tests and again in the 2008 operational tests. Grade 11,

Communication Arts and Science used alternate forms of *TerraNova* Survey (Form A and Form C, respectively) because no grade 11 tests were produced for *TerraNova* Form D. For grade 8, Communication Arts, one passage and item set (Session 2, items 12-17) was also selected from an alternate form of *TerraNova* Survey (Form A) due to an author's denial of permissions. A small number of items from the Language Arts section of *TerraNova* Survey were identified by DESE as being aligned to Missouri's "Writing Standard English" content standard. To supplement the custom items and fulfill the blueprint, a selection of these *TerraNova* Language Arts items, plus the intact *TerraNova* Reading section of Form D, were embedded in the 2007 Communication Arts Field Test and the 2008 Communication Arts operational test. For 2008, NRT scores were generated and reported from the *TerraNova* Survey component of the test using only the Reading items for Communication Arts and the intact survey for Mathematics and Science. *TerraNova* Survey Form D will be used again in 2009 but will be replaced by *TerraNova* Form E for 2010 and 2011. Form D is scheduled to be used again in 2012 followed by Form E in 2013.

The use of the *TerraNova* Survey and its match/alignment to the Missouri standards plays an important role in planning for the entire development process leading up to the time of item selection. This is because the test blueprint is applied to the entire test, which includes both the norm-referenced (NRT) and custom portions. As an NRT product, *TerraNova* items are pre-classified to an existing set of *TerraNova* Reading, Language, Mathematics, or Science standards². In many cases, the alignment of *TerraNova* items to Missouri standards could be considered equivalent; nevertheless, the item development process provided for a DESE review of how the items in the *TerraNova* Survey were matched to the Missouri standards. The alignment of *TerraNova* items to Missouri standards was initially assessed by DESE in 2004 and then verified by DESE in October of 2007.

Operational item selections for 2008 and 2009 were performed in September–October of 2007 by CTB. The selection process followed strict statistical criteria specified by CTB's Research department and approved by DESE. The selection criteria were based on both content requirements and statistical criteria, including the following:

1. *TerraNova* Survey Form D is the anchor for all grades and content areas, with exceptions, as noted above
2. Test length and item types match the DESE-approved test design.
3. Content coverage matches DESE-approved test blueprint.
4. The following items are to be avoided, whenever possible:
 - a. For CR items: 3+ point items, where more than 50% were able to attain the top score points.
 - b. P-value ≤ 0.20 or ≥ 0.90
 - c. Omitted rates $\geq 5\%$
 - d. Poor Fit statistics (Q1)

² It's important to note that the Communication Arts MAP is comprised of both Language and Reading items that are scaled together. In the *TerraNova* family of tests, Language and Reading are administered in a single booklet but are scaled separately.

- e. Significant DIF statistics.
 - i. If an item with DIF must be included for blueprint coverage, examine the item to determine if any content reason exists for the DIF flag (sometimes items will demonstrate statistical bias but no content reason can be determined for the bias).
 - ii. Obtain DESE permission to use the DIF item (meaning someone from DESE should examine the item and agree that no content reason can be determined for the statistical bias).
- 5. Statistical properties of the test:
 - a. ITEMWIN software must be used to select forms.
 - b. Standard Error of Measurement (SEM) and Test Characteristic Curve (TCC) of 2008 operational test must match within 5% of 2007 MAP (not applicable to Science, which tested a new GLE framework at new grade levels in 2008)
 - c. Percent difference between 2008 and 2009 selections should be less than 5% (final summaries). NOTE: A larger percent difference may be allowed for Communication Arts in order to adjust the TCCS so that they are ordinal.

Upon receipt of DESE approval of item selections, production of the 2008 operational test forms commenced. Items were ordered and placed into test books in preparation for operational testing, and the standard process of page reviews between CTB and DESE ensued until final approvals were in place in January 2008. Then, test books and ancillary materials were printed and distributed in support of the Spring 2008 Operational Test, which was administered March 31–April 25, 2008.

3.9 Accommodations and Universal Design

Assessments that are universally designed allow participation of the widest possible range of students, resulting in more valid inferences about students' performance. Universally designed assessments may reduce the need for accommodations by reducing or eliminating access barriers associated with the tests themselves. Table 2.5 presents the elements of universal design (Thompson & Thurlow, 2002). The elements of Universal Design are relevant to both item development and form construction. This section addresses how the elements of Universal Design were addressed in the construction of the spring 2008 test forms.

Table 3.3: Elements of Universal Design

Element	Explanation
Inclusive Assessment Population	Tests designed for state, district, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct irrelevant cognitive, sensory, emotional, and physical barriers can be removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	The test design facilitates the use of needed accommodations (e.g., all items can be Brailled).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and presented in understandable language.
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words are kept to a minimum) to produce readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, to tables, figures, and illustrations, and to response formats.

Universal design requires that assessments need to measure the performance of students with a wide range of abilities and skill repertoires, ensuring that students with diverse learning needs receive opportunities to demonstrate competence on the same content. Because field test items are embedded on operational forms and multiple forms are spiraled within classrooms, field test items are administered to students with a wide range of disabilities, students with limited English proficiency, and students across racial, ethnic, and socioeconomic lines. Students with disabilities or who are English Language Learners may be provided test administration accommodation based on their Individualized Education Plan (IEP). Accommodation code definitions can be found on the DESE website at:

<http://www.dese.mo.gov/divimprove/assess/special.html>.

To accommodate the greatest number of students within MAP, the regular print assessment includes simple, clear, and intuitive instructions and procedures, maximum readability and comprehensibility, and maximum legibility. All of these design components are addressed primarily through the physical layout and formatting of the test books. The page specifications and template for test book pages define how directions and test items are placed on the pages, the location and appearance of headers and footers, spacing between an item stem and answer choices, and other page elements to ensure a consistent, legible appearance of printed test books. Written instructions in the test books at the beginning of each test session are clearly and simply stated, and the wording of such instructions is standardized as much as possible across content areas and grade levels to ensure clarity and consistency.

The MAP test books are designed to minimize distractions and to support navigation through the test book. In Grade 3 Communication Arts, the test items are read aloud to the students. In all grade-level and content areas, a “full-page stop” at the end of each testing session indicates that the students cannot turn the page until instructed by the test examiner. Right-facing pages within a session have a “go on” arrow at the bottom right-hand corner to indicate that the test session continues on the next page. Any pages that

are intentionally left blank are labeled “Do Not Mark on this Page” to indicate that there are no test materials on that page.

Braille and large print versions were constructed for each grade/content area to enable visually impaired students to participate in MAP testing. Two meetings were conducted in 2008 with DESE and a committee of teachers of visually-challenged students. During the first meeting, in September 2007, the entire pool of items available for 2008 and 2009 operational test selections were reviewed to determine which could not be Brailled and to make recommendations for how to transcribe those that were appropriate for the Braille version. Specific recommendations were provided to the transcribers and an Independent Braille expert, who collaborated to produce the Braille proof and the teacher’s notes that accompany the Braille form. During the second review meeting (January 2008), DESE and a teacher committee reviewed the 2008 Braille version of Form 1 of each grade level and made recommendations, as needed, for how to modify the transcription to best serve their students’ needs.

While the goal is to maximize the number of items on the Braille form, it was not possible to transcribe all items into Braille, as some items represent concepts that are simply not appropriate for students who take the Braille form. At some grade levels/content areas, it was necessary to omit items from the Braille version due to bias issues or excessive difficulty associated with the Braille transcription. Table 3.4 lists the items that were omitted from the 2008 Braille versions. The concerns noted by the committee for items that were dropped from the Braille form will be brought to the attention of assessment editors and item writers to guide future item development. Note that the use of item response theory (IRT) models to construct MAP assessments means that it is possible to drop items from the assessment and still provide scores of comparable quality to the full MAP form.

Table 3. 4: Items Omitted from the MAP Spring 2008 Braille Version

Grade	Content Area	Session	Item
3	Mathematics	1	15
		2	26
4	Mathematics	2	24
5	Mathematics	3	7
	Science	2	9
		3	2
		3	4
6	Mathematics	1	12
		1	22
		3	3
7	Mathematics	1	23
		3	7
8	Mathematics	1	17
		2	17
	Science	1	4
		2	4
		3	6
10	Mathematics	3	4
		3	8
11	Science	2	3

Table 3. 5: MAP 2008 Content Standard Item/Point Distributions, Communication Arts

Grade	Content Standard	TN NRT Items	CR/PE Items	SR Items	Total Items	TN Points	CR/PE Points	SR Points	Total Points	% of Total Points
03	Speaking/Writing Standard English	1	0	15	16	1	0	15	16	24%
	Reading Fiction/Poetry/Drama	20	3	0	23	20	6	0	26	38%
	Reading Nonfiction	9	4	2	15	9	8	2	19	28%
	Writing Formally & Informally	0	4	0	4	0	7	0	7	10%
	Combined Reading from Standards 2 & 3	29	7	2	38	29	14	2	45	66%
	Total	30	11	17	58	30	21	17	68	100%
04	Speaking/Writing Standard English	0	0	10	10	0	0	10	10	16%
	Reading Fiction/Poetry/Drama	35	3	0	38	35	6	0	41	65%
	Reading Nonfiction	0	4	2	6	0	8	2	10	16%
	Writing Formally & Informally	0	2	0	2	0	2	0	2	3%
	Combined Reading from Standards 2 & 3	35	7	2	44	35	14	2	51	81%
	Total	35	9	12	56	35	16	12	63	100%
05	Speaking/Writing Standard English	0	0	13	13	0	0	13	13	21%
	Reading Fiction/Poetry/Drama	19	0	0	19	19	0	0	19	30%
	Reading Nonfiction	13	7	2	22	13	14	2	29	46%
	Writing Formally & Informally	0	2	0	2	0	2	0	2	3%
	Combined Reading from Standards 2 & 3	32	16	27	41	67	30	27	48	76%
	Total	32	9	15	56	32	16	15	63	100%
06	Speaking/Writing Standard English	0	0	13	13	0	0	13	13	21%
	Reading Fiction/Poetry/Drama	17	4	2	23	17	8	2	27	44%
	Reading Nonfiction	14	3	1	18	14	6	1	21	34%
	Writing Formally & Informally	0	1	0	1	0	1	0	1	2%
	Combined Reading from Standards 2 & 3	31	7	3	41	31	14	3	48	77%
	Total	31	8	16	55	31	15	16	62	100%
07	Speaking/Writing Standard English	0	0	16	16	0	0	16	16	22%
	Reading Fiction/Poetry/Drama	27	4	2	33	27	8	2	37	50%
	Reading Nonfiction	7	3	0	10	7	6	0	13	18%
	Writing Formally & Informally	0	4	0	4	0	8	0	8	11%
	Combined Reading from Standards 2 & 3	34	7	2	43	34	14	2	50	68%
	Total	34	11	18	63	34	22	18	74	100%
08	Speaking/Writing Standard English	0	0	15	15	0	0	15	15	22%
	Reading Fiction/Poetry/Drama	10	4	2	16	10	8	2	20	29%
	Reading Nonfiction	24	3	0	27	24	6	0	30	43%
	Writing Formally & Informally	0	3	0	3	0	4	0	4	6%
	Combined Reading from Standards 2 & 3	34	7	2	43	34	14	2	50	72%
	Total	34	10	17	61	34	18	17	69	100%
11	Speaking/Writing Standard English	0	0	14	14	0	0	14	14	19%
	Reading Fiction/Poetry/Drama	5	4	2	11	5	8	2	15	21%
	Reading Nonfiction	29	3	0	32	29	6	0	35	48%
	Writing Formally & Informally	0	5	1	6	0	8	1	9	12%
	Combined Reading from Standards 2 & 3	34	7	2	43	34	14	2	50	68%
	Total	34	12	17	63	34	22	17	73	100%

Table 3. 6: MAP 2008 GLE Strand Item/Point Distributions, Mathematics

Grade	Content Standard	TN NRT Items	CR/PE Items	SR Items	Total Items	TN Points	CR/PE Points	SR Points	Total Points	% of Total Points
03	Number and Operations	14	2	6	22	14	4	6	24	36%
	Algebraic Relationships	3	1	8	12	3	2	8	13	19%
	Geometric and Spatial Relationships	4	2	5	11	4	4	5	13	19%
	Measurement	4	1	4	9	4	2	4	10	15%
	Data and Probability	5	1	0	6	5	2	0	7	10%
	Total	30	7	23	60	30	14	23	67	100%
04	Number and Operations	17	2	0	19	17	4	0	21	27%
	Algebraic Relationships	4	3	3	10	4	8	3	15	19%
	Geometric and Spatial Relationships	4	1	7	12	4	2	7	13	17%
	Measurement	4	2	7	13	4	4	7	15	19%
	Data and Probability	3	2	6	11	3	4	6	13	17%
	Total	32	10	23	65	32	22	23	77	100%
05	Number and Operations	16	0	0	16	16	0	0	16	23%
	Algebraic Relationships	3	2	6	11	3	4	6	13	19%
	Geometric and Spatial Relationships	3	2	6	11	3	4	6	13	19%
	Measurement	6	1	6	13	6	2	6	14	20%
	Data and Probability	4	2	5	11	4	4	5	13	19%
	Total	32	7	23	62	32	14	23	69	100%
06	Number and Operations	17	0	0	17	17	0	0	17	25%
	Algebraic Relationships	4	2	5	11	4	4	5	13	19%
	Geometric and Spatial Relationships	2	1	8	11	2	2	8	12	18%
	Measurement	4	2	5	11	4	4	5	13	19%
	Data and Probability	4	2	5	11	4	4	5	13	19%
	Total	31	7	23	61	31	14	23	68	100%
07	Number and Operations	16	0	0	16	16	0	0	16	23%
	Algebraic Relationships	2	2	7	11	2	4	7	13	19%
	Geometric and Spatial Relationships	4	2	6	12	4	4	6	14	20%
	Measurement	4	2	5	11	4	4	5	13	19%
	Data and Probability	6	1	5	12	6	2	5	13	19%
	Total	32	7	23	62	32	14	23	69	100%
08	Number and Operations	14	0	0	14	14	0	0	14	18%
	Algebraic Relationships	5	3	8	16	5	8	8	21	28%
	Geometric and Spatial Relationships	4	2	7	13	4	4	7	15	20%
	Measurement	3	2	5	10	3	4	5	12	16%
	Data and Probability	5	3	3	11	5	6	3	14	18%
	Total	31	10	23	64	31	22	23	76	100%
10	Number and Operations	11	0	0	11	11	0	0	11	16%
	Algebraic Relationships	3	3	9	15	3	8	9	20	29%
	Geometric and Spatial Relationships	3	2	7	12	3	4	7	14	20%
	Measurement	4	2	3	9	4	4	3	11	16%
	Data and Probability	4	3	4	11	4	6	4	14	20%
	Total	25	10	23	58	25	22	23	70	100%

Table 3. 7: MAP 2008 GLE Strand Item/Point Distributions, Science

Grade	Content Standard	TN NRT Items	CR/PE Items	SR Items	Total Items	TN Points	CR/PE Points	SR Points	Total Points	% of Total Points
05	Matter and Energy	2	4	2	6	2	8	2	10	13%
	Force and Motion	1	3	1	4	1	6	1	7	9%
	Living Organisms	4	2	4	6	4	4	4	8	10%
	Ecology	3	3	3	6	3	6	3	9	11%
	Earth Systems	2	4	2	6	2	8	2	10	13%
	Universe	2	3	2	5	2	6	2	8	10%
	Scientific Inquiry	5	9	5	14	5	14	5	19	24%
	Science, Technology, & Human Activity	3	3	3	6	3	6	3	9	11%
Total	22	31	22	53	22	58	22	80	100%	
08	Matter and Energy	2	5	2	7	2	10	2	12	14%
	Force and Motion	3	2	3	5	3	4	3	7	8%
	Living Organisms	4	4	4	8	4	8	4	12	14%
	Ecology	2	3	2	5	2	6	2	8	9%
	Earth Systems	5	3	5	8	5	6	5	11	13%
	Universe	0	4	0	4	0	8	0	8	9%
	Scientific Inquiry	7	10	7	17	7	15	7	22	25%
	Science, Technology, & Human Activity	1	3	1	4	1	6	1	7	8%
Total	24	34	24	58	24	63	24	87	100%	
11	Matter and Energy	4	4	4	8	4	8	4	12	13%
	Force and Motion	2	4	2	6	2	8	2	10	11%
	Living Organisms	5	3	0	8	5	6	0	11	12%
	Ecology	0	4	5	4	0	8	5	8	9%
	Earth Systems	3	3	3	6	3	7	3	10	11%
	Universe	0	4	0	4	0	8	0	8	9%
	Scientific Inquiry	7	11	7	18	7	20	7	27	29%
	Science, Technology, & Human Activity	2	2	2	4	2	4	2	6	7%
Total	23	35	23	58	23	69	23	92	100%	

Test content evidence of validity is provided for the MAP with the specification of each of the Process Standards that are influential in acquiring the skills tested in the items/tasks used in each of the MAP tests. If teachers teach by the Process Standards as intended, then student performance should improve on those items that were identified as implicitly tapping these habits of mind. The following charts provide the distribution of items and points by Process Standards deemed addressable using MAP paper-and-pencil items.

Table 3. 8: MAP 2008 Number of Items/Points Measuring Process Standards, Communication Arts

Grade Level	Process Standard	NRT Items (SR)	Custom/CR/, Other Items	Total Items	NRT Points (SR)	Custom/CR/, Other Pts	Total Points
03	1.5	13	0	13	13	0	13
	1.6	8	9	17	8	16	24
	2.1	0	4	4	0	7	7
	2.2	0	15	15	0	15	15
	3.5	9	0	9	9	0	9
04	1.5	3	0	3	3	0	3
	1.6	20	6	26	20	11	31
	2.1	0	2	2	0	2	2
	2.2	0	10	10	0	10	10
	3.1	0	1	1	0	2	2
	3.5	12	2	14	12	3	15
05	1.4	0	1	1	0	1	1
	1.5	1	0	1	1	0	1
	1.6	18	4	22	18	8	26
	1.8	0	1	1	0	1	1
	2.1	0	5	5	0	5	5
	2.2	0	9	9	0	9	9
	3.1	0	1	1	0	2	2
	3.5	13	3	16	13	5	18
06	1.4	0	1	1	0	1	1
	1.6	24	5	29	24	9	33
	1.8	0	1	1	0	1	1
	2.2	0	13	13	0	13	13
	2.4	1	0	1	1	0	1
	3.1	0	1	1	0	2	2
	3.5	6	3	9	6	5	11
07	1.5	5	0	5	5	0	5
	1.6	12	3	15	12	5	17
	1.7	0	1	1	0	1	1
	1.8	0	1	1	0	2	2
	2.1	0	3	3	0	6	6
	2.2	0	16	16	0	16	16
	2.4	3	2	5	3	4	7
	3.1	0	1	1	0	2	2
	3.5	13	2	15	13	4	17

Table 3.8: MAP 2008 Number of Items/Points Measuring Process Standards, Communication Arts (cont'd)

Grade Level	Process Standard	NRT Items (SR)	Custom/CR/, Other Items	Total Items	NRT Points (SR)	Custom/CR/, Other Pts	Total Points
08	1.5	10	0	10	10	0	10
	1.6	22	3	25	22	5	27
	1.8	0	1	1	0	2	2
	2.1	0	2	2	0	2	2
	2.2	0	15	15	0	15	15
	2.4	2	0	2	2	0	2
	3.1	0	1	1	0	2	2
	3.5	0	5	5	0	9	9
11	1.5	1	0	1	1	0	1
	1.6	25	4	29	25	6	31
	2.1	0	6	6	0	9	9
	2.2	0	14	14	0	14	14
	2.4	0	1	1	0	2	2
	3.5	8	4	12	8	8	16

Table 3. 9: MAP 2008 Number of Items/Points Measuring Process Standards, Mathematics

Grade Level	Process Standard	NRT Items (SR)	Custom/CR/, Other Items	Total Items	NRT Points (SR)	Custom/CR/, Other Pts	Total Points
03	1.10	0	3	3	0	4	4
	1.2	0	1	1	0	2	2
	1.5	6	0	6	6	0	6
	1.6	8	12	20	8	13	21
	3.1	1	2	3	1	2	3
	3.2	0	1	1	0	2	2
	3.3	13	7	20	13	9	22
	3.4	0	1	1	0	2	2
	3.5	2	0	2	2	0	2
	3.6	0	3	3	0	3	3
04	1.10	2	7	9	2	8	10
	1.5	0	2	2	0	2	2
	1.6	5	10	15	5	14	19
	1.7	0	1	1	0	2	2
	1.8	0	2	2	0	3	3
	3.1	11	1	12	11	1	12
	3.2	0	1	1	0	1	1
	3.3	14	6	20	14	10	24
	3.6	0	3	3	0	4	4
05	1.1	0	1	1	0	1	1
	1.10	0	1	1	0	2	2
	1.5	4	0	4	4	0	4
	1.6	0	9	9	0	10	10
	1.8	0	1	1	0	2	2
	3.1	2	6	8	2	6	8
	3.2	0	2	2	0	3	3
	3.3	24	5	29	24	7	31
	3.5	1	1	2	1	2	3
	3.6	0	4	4	0	4	4
3.7	1	0	1	1	0	1	
06	1.10	0	5	5	0	6	6
	1.5	4	0	4	4	0	4
	1.6	3	6	9	3	9	12
	1.7	3	0	3	3	0	3
	1.8	0	3	3	0	4	4
	3.1	6	8	14	6	8	14
	3.2	0	2	2	0	2	2
	3.3	12	4	16	12	5	17
	3.5	2	1	3	2	2	4
	3.6	0	1	1	0	1	1
	3.7	1	0	1	1	0	1

Table 3.9: MAP 2008 Number of Items/Points Measuring Process Standards, Mathematics (cont'd)

Grade Level	Process Standard	NRT Items (SR)	Custom/CR/, Other Items	Total Items	NRT Points (SR)	Custom/CR/, Other Pts	Total Points
07	1.10	0	5	5	0	8	8
	1.4	0	1	1	0	2	2
	1.5	3	0	3	3	0	3
	1.6	2	9	11	2	11	13
	3.1	8	4	12	8	4	12
	3.2	0	2	2	0	2	2
	3.3	18	3	21	18	4	22
	3.5	0	4	4	0	4	4
	3.6	0	2	2	0	2	2
	3.7	1	0	1	1	0	1
08	1.10	0	1	1	0	2	2
	1.2	0	1	1	0	1	1
	1.4	0	2	2	0	2	2
	1.5	6	1	7	6	2	8
	1.6	4	12	16	4	18	22
	1.8	1	0	1	1	0	1
	3.1	2	4	6	2	5	7
	3.2	0	1	1	0	1	1
	3.3	18	2	20	18	3	21
	3.4	0	2	2	0	3	3
	3.5	0	2	2	0	2	2
	3.6	0	4	4	0	4	4
	3.8	0	1	1	0	2	2
10	1.1	0	1	1	0	2	2
	1.10	2	3	5	2	5	7
	1.2	0	1	1	0	1	1
	1.5	3	2	5	3	2	5
	1.6	6	12	18	6	16	22
	3.1	5	4	9	5	5	10
	3.2	0	3	3	0	4	4
	3.3	8	1	9	8	2	10
	3.4	0	1	1	0	1	1
	3.5	1	3	4	1	5	6
	3.6	0	3	3	0	3	3
	3.8	0	1	1	0	1	1

Table 3. 10: MAP 2008 Number of Items/Points Measuring Process Standards, Science

Grade Level	Process Standard	NRT Items (SR)	Custom/CR/, Other Items	Total Items	NRT Points (SR)	Custom/CR/, Other Pts	Total Points
05	1.1	0	3	3	0	4	4
	1.10	9	8	17	9	16	25
	1.3	2	3	5	2	5	7
	1.5	4	2	6	4	2	6
	1.6	3	11	14	3	21	24
	1.8	0	1	1	0	4	4
	3.1	0	1	1	0	2	2
	3.2	0	1	1	0	2	2
	3.5	2	1	3	2	2	4
08	1.1	0	2	2	0	2	2
	1.10	16	11	27	16	21	37
	1.3	1	4	5	1	7	8
	1.5	3	2	5	3	3	6
	1.6	4	9	13	4	17	21
	1.7	0	1	1	0	1	1
	1.8	0	1	1	0	4	4
	3.5	0	2	2	0	4	4
	3.6	0	1	1	0	2	2
	3.8	0	1	1	0	2	2
11	1.1	0	1	1	0	1	1
	1.10	17	12	29	17	22	39
	1.3	0	5	5	0	10	10
	1.5	2	1	3	2	1	3
	1.6	4	10	14	4	21	25
	1.7	0	1	1	0	2	2
	1.8	0	2	2	0	6	6
	3.5	0	2	2	0	4	4
	3.8	0	1	1	0	2	2

CHAPTER 4: TEST ADMINISTRATION

Chapter 4 of the Technical Report describes the processes and activities implemented and information disseminated to help ensure standardized test administration procedures and, thus, uniform test administration conditions for students. According to the *Standards for Educational and Psychological Testing*, the “usefulness and interpretability of test scores require that a test be administered and scored according to the developer’s instructions” (61). Chapter 4 examines how test administration procedures implemented for the MAP strengthen and support the intended score interpretations and reduce construct-irrelevant variance that could threaten the validity of score interpretations. Chapter 4 demonstrates adherence to AERA/APA/NCME standards 3.19, 3.20, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, and 5.10 in the MAP program.

4.1 Training of Districts

To ensure that the MAP tests are administered and scored in accordance with the department’s mandates, DESE takes a primary role in communicating with and training district personnel. The development of the MAP tests is a collaborative effort between DESE and CTB/McGraw-Hill. DESE conveys to districts the purpose of the MAP program and that test administration must be consistent with test industry standards, meeting State Board of Education policies and the mandates of both state and federal legislation. To accomplish these goals, DESE provides train-the-trainer opportunities for the Regional Instructional Facilitators (RIFs). The RIFs convey test administration training to districts. The RIFs also conduct Quality Assurance visits during testing to ensure district adherence to the standardized administration of the tests.

The RIFs are responsible to districts within their region. The RIFs disseminate information to each district, offer assistance with test administration and serve as the liaisons between DESE and the districts. DESE departmental staff also communicate directly with districts answering questions particular to the MAP program as well as general assessment questions. DESE staff also provide assistance with MAP data and interpretation for MAP test results.

The Director of Assessment and the Assistant Director of Assessment trained the RIFs in the following components of MAP test administration: the *Test Coordinator’s Manual*, *Test Examiner’s Manual*, dates for testing, appropriate protocols for test administration and security, guidance on the timing and administration of tests and changes made to the test since 2007. Appendix A contains DESE’s presentations on the *Test Coordinator’s Manual* and the *Examiner’s Manual*. The RIFs, in turn, used this information to train district-level staff. Appendix A also contains one of these presentations that was compiled by RIFs in the St. Louis region. It is representative of the information that all RIFs would use in their presentations.

4.2 Ancillary Materials

Test administration ancillary materials for the MAP contribute to the body of evidence of the validity of score interpretation. This section examines how the test materials address the AERA/APA/NCME standards related to test administration procedures.

For the spring 2008 test administration, CTB/McGraw-Hill produced two types of administration manuals: the *Test Coordinator's Manual* and the *Examiner's Manual*. DESE Curriculum and Assessment staff review, provide feedback, and give final approval for each manual.

The *Test Coordinator's Manual* is common to all grades and content areas. It provides an overview of MAP and any changes made to MAP in 2008. It gives guidelines for testing, such as the inclusion of special populations, the use of translators, and the invalidation procedures. It also details the Test Coordinator's role in the testing process by outlining nine steps the Test Coordinator should follow. These steps are:

- Step 1: Review Testing Materials
- Step 2: Distribute Testing Materials
- Step 3: Collect Testing Materials
- Step 4: Check the Organization of Materials Collected
- Step 5: Check the Student Information Sheet (SIS)
- Step 6: Check the Group Information Sheet (GIS)
- Step 7: Complete the School/Group List
- Step 8: Organize Materials for the District Test Coordinator
- Step 9: Package and Ship Testing Materials

The *Examiner's Manuals* are specific to each grade. The MAP *Examiner's Manuals* also outline steps that should be followed when administering MAP. These steps include:

- Step 1: Preparing for Testing
- Step 2: Organize Your Classroom
- Step 3: Check your Testing Materials
- Step 4: Before Testing
- Step 5: Administer the Test
- Step 6: Invalidations and Makeups
- Step 7: After Testing: Student Status Coding
- Step 8: Assemble Materials for Return

These steps provide instructions on pre-test and post-test procedures, such as:

- Test security
- Standardized testing protocols for norm-referenced information
- Using student barcode labels
- Completing the student information sheet, including recording test accommodations

This section presents the AERA/APA/NCME standards relevant to test administration and how information in the MAP *Examiner's Manuals* and *Test Coordinator Manuals* address these standards.

Standard 3.19 *Present directions with sufficient clarity so others can replicate administration conditions.*

The MAP *Examiner's Manuals* provide instructions for before-, during-, and after-testing activities with sufficient detail and clarity to support reliable test administrations by qualified test administrators. To ensure uniform administration conditions throughout the state, instructions in the *Examiner's Manuals* describe the following: the materials that the examiner and students need for testing; how to verify that pre-coded student information on student barcode labels is correct; how to fill out the Student Information Sheet if the student barcode label is incorrect; how to prepare the testing environment; how to administer the tests; and the test schedule, including testing times.

Standard 3.20 *Instructions to test takers should contain sufficient detail so they can respond to tasks in the manner intended by test developer.*

To ensure clarity of instructions to students, the manuals include scripts that the examiner is instructed to read verbatim to students. Examiners are instructed to follow the script and to repeat any part of the directions as many times as needed, but to not modify the words used. Examiners may use professional judgment to respond to student questions, but they may not reword test items, suggest answers, or evaluate student work during the testing session. A sample of a script is presented in Figure 4.1.

Sample test items are provided in each content area to familiarize students with how to fill in answers. Sample items are also provided in the *Examiner's Manual*.

Standard 5.1 *Test administrators should follow standardized procedures for administration.*

To ensure the usefulness and interpretability of test scores and to minimize sources of construct-irrelevant variance, it is essential that the MAP is administered according to the prescribed test schedule. The *Test Coordinator's Manual* includes instructions for scheduling the test within the state testing window of March 31 through May 2, 2008. The *Examiner's Manuals* contain the schedule for timing each test session and whether timing is to be strictly enforced. The test timing schedule is presented in Table 4.1.

Standard 5.2 *Document modifications or disruptions of standardized test administration procedures or scoring.*

DESE staff administer reports on Testing Concerns which have a wide range of types of improper activities that may occur during testing including the following: copying and reviewing MAP test questions with students; cueing students during testing either verbally or with written materials on the classroom walls; cueing students nonverbally, such as tapping or nodding the head; using a calculator on parts of the test where it is not allowed; allowing too much time on *TerraNova* sections of the test; allowing students to correct or complete answers after tests have been returned to the teacher; splitting sessions into two parts; ignoring the standardized directions in the test books; reading the Communication Arts test to students; paraphrasing parts of the test to students; changing or completing (or allowing other school personnel to change or complete) answers of students; allowing accommodations that are not written in the IEP; allowing non-IEP students accommodations; allowing students to use dictionaries on parts of MAP other than writing prompt; defining terms on test.

Testing concerns are gathered from school officials, students, parents, and other interested parties who call DESE to state their allegation. A narrative of the conversation is written and read back to them. The superintendent of the district in which the allegation was made is then contacted and read the narrative. A letter is sent to confirm the conversation and to ask the superintendent to investigate the claim. A District Report Form is sent for the superintendent to use for replying to the allegation. The District Report Form is shown in Figure 4.2.

All of these narratives, letters, and reports are given to the Data, Accountability, and Accreditation section in order to make accountability decisions.

Standard 5.4 *Testing environment should furnish reasonable comfort with minimal distractions.*

Step 2 in the *Examiner's Manual* overviews the steps that teachers should take to prepare their classroom for administering the MAP test. These include:

- Plan for the distribution and collection of materials.
- Plan seating arrangements. Allow enough space between students to prevent the sharing of answers.
- Eliminate distractions such as bells or telephones.
- Use a Do Not Disturb sign on the door of the testing room.
- Make sure classroom maps, charts, and any other materials that relate to the content and processes of the test are out of the students' view.
- When administering the timed portion of the test, write on the board the starting and stopping times for the test.

Standard 5.6 *Reasonable efforts should be made to ensure the integrity of test scores.*

The *Examiner's Manual* and *Test Coordinator's Manual* present instructions for post-test activities to ensure that test materials are handled properly to ensure the integrity of student information and test scores. Detailed instructions guide test examiners in completing required information on students' scannable test books. For students who were administered a large print or Braille version of the MAP, examiners are instructed to transcribe students' responses from the large print test or Braille test book to a regular-edition test book exactly as they responded in the large print or Braille test book.

Standard 5.7 *Test users are responsible for the security of test materials at all times.*

Throughout the manuals, test coordinators and examiners are reminded of test security requirements and procedures to maintain test security. Specific actions that are direct violations of test security are so noted. Detailed information about test security procedures are presented in Section 4.3.

4.2.1 Return Material Forms and Guidelines

The *Test Coordinator's Manual* instructs test coordinators in procedures for organizing and packing materials and returning them to CTB/McGraw-Hill for scanning and scoring. DESE curriculum and assessment staff have opportunities to review, provide feedback and have final approval. The purpose of the instructions is to ensure that used and unused test materials are properly accounted for and student answer documents are organized properly for return shipment. Proper organization of materials and accurate completion of the school/group list document contributes to accurate score reports and helps in delivery of such reports in a timely manner.

4.2.2 Security Forms

As soon as books are received by a district, the district test coordinator assures that the first and last security barcode on the tests match the packing list they received. The district test coordinator then packages the tests to be sent to schools. Upon returning tests to CTB/McGraw-Hill, school and district test coordinators are required to complete and submit a "Test Book Accountability Form" that details the number of scorable and nonscorable books returned. This form also requires that districts/schools document nonstandard situations, including lost, damaged, destroyed, extra, or missing books). The Test Book Accountability Form is shown in Figure 4.3.

4.2.3 Interpretive Guides

Essential to making valid interpretations of test scores is an understanding of what the test scores mean and how to interpret score reports. The *Guide to Interpreting Results* is written for Missouri teachers and administrators who receive MAP score reports from the 2008 administration. More detail about the guide can be found in Chapter 7.6.

4.3 Test security measures

Maintaining the security of all test materials is crucial to preventing the possibility of random or systematic errors, such as unauthorized exposure of test items, that would affect the valid interpretation of test scores. Several test security measures are implemented for the MAP. Test security procedures are discussed throughout the *Test Examiner* and *Test Coordinator's Manuals*.

Test coordinators and examiners are instructed to keep all test materials in locked storage, except during actual test administration, and access to secure materials must be restricted to authorized individuals only (e.g., test examiners and the school test coordinator). During the testing sessions, test examiners are directly responsible for the security of the MAP and must account for all test materials at all times. The test examiner must supervise the test administration at all times.

4.4 Test Administration

The 2008 test was administered to students within the state testing window of March 31 to May 2, 2008. Systems chose when and how to administer the MAP within this window. Each session within each content area of the MAP was required to be administered in one block of time.

4.4.1 Time

Each section of each content area test was timed to provide sufficient time for students to attempt all items. The *Examiner's Manuals* provided examiners with timing guidelines for the custom portions of MAP. Strict timing guidelines were given for the norm-referenced portions of the test. For MAP's custom sessions, examiners are instructed to allow students to complete the assessment if s/he is making adequate progress. For the norm-referenced portion of the test, students may receive an accommodation for additional time if so needed and documented on their individualized education plan. The timing schedule of the MAP is presented in Table 4.2.

4.4.2 Accommodations

Regular education students do not receive testing accommodations. Test accommodations may be used with students who qualify under IDEA and have an IEP, Section 504 of the Americans' with Disabilities Act and have a 504 plan, or who are identified as an English Language Learner. Accommodations must be specified in the qualifying student's individual plan and must be consistent with accommodations used during daily classroom instruction and testing. The use of any accommodation must be indicated on the student information sheet at the time of test administration.

The grade-specific MAP *Examiner's Manual* contains the list of accommodations permissible for the MAP assessments. The table of accommodations presented in the *Examiner's Manual* is shown in Table 4.5. If a specific accommodation is not on the list of accommodations in the *Examiner's Manual*, the accommodation may still be permitted. However, for accountability purposes, there are some accommodations that

will invalidate a student's test results, such as an oral administration of the Communications Test or paraphrasing any of the tests. Detailed information regarding testing accommodations can be found at the DESE website:

<http://dese.mo.gov/divimprove/assess/ancillaries.html>

Braille and large print forms are provided to students with vision disabilities.

Tables 4.4 through 4.6 summarize the numbers of reportable students receiving accommodations by accommodation type for the 2008 MAP, the Braille edition of the 2007 MAP, and the large print edition of the 2008 MAP. The analyses in Tables 4.4 through 4.6 are based on census data and include only those students receiving accommodations and who received a scale score on the Communication Arts, Mathematics, or Science MAP.

In 2008, setting and timing accommodations appear to be the most frequently used for the Communication Arts, Mathematics, and Science MAP. For the Science and Mathematics MAP, having the test read aloud was also among the more frequently used accommodations. For the Mathematics MAP, using calculators was also among the more frequently used accommodations.

On the Braille and large print editions of MAP, the setting and timing accommodations are again among the most frequently used accommodations. Common accommodations for both the Braille and large print editions include using a scribe for the Communication Arts, Mathematics, and Science MAPs, having the test read aloud for the Mathematics and Science MAPs, and using a calculator for the Mathematics MAP.

4.5 Summary

In summary, the overall purpose of each of the test administration workshops and the ancillary materials is to keep districts informed about policies and procedures related to testing in general and the MAP program in particular. The information imparted is clearly related to standardizing the administration of the MAP, maintaining the security of the assessment, allowing access to the assessments for special populations by clearly delineating appropriate accommodations, and by providing guidance on appropriate interpretations of the test results. These communication and training efforts by DESE and the ancillary information developed by CTB/McGraw-Hill address multiple best practices of the testing industry but in particular are related to the following *Standards for Educational and Psychological Testing*:

- Standard 5.1—Test administrators should follow carefully the standardized procedures for administration and scoring specified by the test developer, unless the situation or a test taker's disability dictates that an exception should be made.
- Standard 5.2—Modifications or disruptions of standardized test administration procedures or scoring should be documented.

- Standard 5.3—When formal procedures have been established for requesting and receiving accommodation, test takers should be informed of these procedures in advance of testing.
- Standard 5.4—The testing environment should furnish reasonable comfort with minimal distractions.
- Standard 5.6—Reasonable efforts should be made to assure the integrity of test scores by eliminating opportunities for test takers to attain scores by fraudulent means.
- Standard 5.7—Test users have the responsibility of protecting the security of test materials at all times.
- Standard 5.10—When test score information is released to students, parents, legal representatives, teachers, clients, or the media, those responsible for testing programs should provide appropriate interpretations. The interpretations should describe in simple language what the test covers, what test scores mean, the precision of scores, common misinterpretations of scores, and how scores will be used.

Table 4. 1: MAP Administration Schedule

Mathematics Timed Sessions	
Grades 3, 5 and 7	
Session 1	40-55 (timing guideline)
Session 2-Part 1	10 (strictly timed)
Session 2-Part 2	30 (strictly timed)
Session 3	35-45 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Mathematics Timed Sessions	
Grade 6	
Session 1	40-45 (timing guideline)
Session 2-Part 1	10 (strictly timed)
Session 2-Part 2	30 (strictly timed)
Session 3	35-45 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Mathematics Timed Sessions	
Grade 10*	
Session 1	35-45 (timing guideline)
Session 2-Part 1	10 (strictly timed)
Session 2-Part 2	30 (strictly timed)
Session 3	45-60 (timing guideline)

*Timings and item counts are approximate and may vary by grade level

Communications Arts Timed Sessions	
Grade 3*	
Session 1	45-55 (timing guideline)
Session 2	60-90 (timing guideline)
Session 3-Part 1	28 (strictly timed)
Session 3-Part 2	24 (strictly timed)
Session 4	50-65 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Communications Arts Timed Sessions	
Grade 7*	
Session 1	45-55 (timing guideline)
Session 2	60-90 (timing guideline)
Session 3-Part 1	27 (strictly timed)
Session 3-Part 2	26 (strictly timed)
Session 4	50-65 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Communications Arts Timed Sessions	
Grades 4 and 5*	
Session 1	45-55 (timing guideline)
Session 2-Part 1	28 (strictly timed)
Session 2-Part 2	24 (strictly timed)
Session 3	50-65 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Science Timed Sessions	
Grades 5 and 8*	
Session 1	65-85 (timing guideline)
Session 2-Part 1	25 (strictly timed)
Session 2-Part 2	40-55 (timing guideline)
Session 3	90-105 (timing guideline)

*Timings and item counts are approximate and may vary by grade level

Science Timed Sessions	
Grade 11*	
Session 1	60-80 (timing guideline)
Session 2	55-75 (timing guideline)
Session 3-Part 1	25 (strictly timed)
Session 3-Part 2	20-30 (timing guideline)

*Timings and item counts are approximate and may vary by grade level

Communications Arts Timed Sessions	
Grades 6 and 8*	
Session 1	45-55 (timing guideline)
Session 2-Part 1	27 (strictly timed)
Session 2-Part 2	26 (strictly timed)
Session 3	50-65 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Communications Arts Timed Sessions	
Grade 11*	
Session 1	45-55 (timing guideline)
Session 2	60-90 (timing guideline)
Session 3-Part 1	27 (strictly timed)
Session 3-Part 2	26 (strictly timed)
Session 4	30 (timing guideline)

*Timings and item counts are approximate and may vary by grade level.

Mathematics Timed Sessions	
Grade 4 and 8*	
Session 1-Part 1	40-55 (timing guideline)
Session 1-Part 2	15-20 (timing guideline)
Session 2-Part 1	10 (strictly timed)
Session 2-Part 2	30 (strictly timed)
Session 3	50-70 (timing guideline)

*Timings and item counts are approximate and may vary by grade level

Table 4. 2: MAP Accommodations for students who are English Language Learners

<u>Accommodations List for Students Who Are English Language Learners (ELL)</u>			
The following are the only accommodations allowed for ELL students:			
Code	Invalidates	Administration Accommodations	Description
04		Oral reading of assessment (<i>Not permissible for Communication Arts</i>)	The test examiner reads items verbatim to the student in an isolated setting so that other students will not benefit or be disturbed.
11		Oral reading in native language (<i>Not permissible for Communication Arts</i>)	
		Timing Accommodations	Description
20		Extend time allotted to complete TerraNova Survey. See Note 3.	
21		Administer test using more than allotted periods	Dates for taking the MAP must occur within the MAP testing window.
22		Other: Specify	Other timing accommodations.
		Response Accommodations	Description
35		Use of scribe to record student response in test booklet	The student conveys verbal responses to a scribe in an isolated, individual setting so that other students cannot benefit or be disturbed. The scribe cannot suggest ideas, words, or concepts. The scribe records the student's answers verbatim. The student should indicate capitalization and punctuation if language mechanics are being assessed.
		Oral response	The student provides an oral response to the examiner.
43		Use of bilingual dictionary (<i>Not permissible for Communication Arts</i>)	
		Setting Accommodations	Description
50		Testing individually	The room should be free of noises, conversation, and distractions from adjoining rooms. Individual testing is appropriate when, for example, responses are given orally or questions are paraphrased.
51		Testing with small groups	The location should be free of noises, conversation, and distractions from adjoining rooms. Students may not interact with one another about questions or answers. The examiner must be present at all times. Testing in small groups is not appropriate for students who give responses orally or require paraphrasing of questions.
53		Other: Specify	Other setting accommodations.

Table 4. 3: MAP Accommodations for Students with Disabilities

Accommodations List for Students with Disabilities			
Code	Invalidates	Administration Accommodations	Description
01		Braille edition of assessment	Braille editions of the assessment require special processing. Consult your Braille edition test materials for specific instructions.
02		Large Print edition of assessment	Large Print editions of the assessment require special processing. Consult your Large Print test materials for specific instructions.
04	✓	Oral reading of assessment. See Note 1.	The test examiner reads items verbatim to the student in an isolated setting so that other students will not benefit or be disturbed.
04		Oral reading of assessment to Blind/Partial Sight students. See Note 1.	The test examiner reads items verbatim to the student who cannot read Braille in an isolated setting so that other students will not benefit or be disturbed.
05		Signing	A certified sign language interpreter or deaf education instructor signs the Mathematics and/or Science test (directions and test items are allowed) and/or the directions only of the Communication Arts test to the student.
06	✓	Paraphrasing See Note 2.	The test examiner paraphrases questions to help student understanding in an isolated setting. Terms may be defined as long as they: 1) are not the actual concept or content being assessed, 2) would not give clues, or 3) would not disclose the answer.
10		Other administration accommodations	
		Use of assistive devices	An assistive device, which permits a student to read and/or respond to the assessment, is used. Examples of assistive devices include computers that assist students with fine-motor problems, text enlargers that enable students to independently read and answer test questions, or augmentative communication devices.
		Use of visual aids: Specify	Visual aids include any type of optical or non-optical devices used to enhance visual capability. Examples of visual aids include bold-line felt-tip markers, lamps, filters, bold-lined paper, writing guides, or other adaptations that alter the visual environment by adjusting the space, illumination, color, contrast, or other physical features of the environment.
		Timing Accommodations	Description
20		Extend time allotted to complete <i>TerraNova</i> Survey. See Note 3.	Extended time to complete the <i>TerraNova</i> Survey is allowed for a student whose disability may cause him/her to be unable to meet time constraints.
21		Administer assessment using more than allotted periods	Students with disabilities may need to complete the assessments over more than one test period as a result of fatigue and/or loss of concentration. Some students may require additional breaks. Dates for taking the MAP must occur within the MAP testing window.
22		Other: Specify	Other timing accommodations
		Response Accommodations	Description
35		Use of scribe to record student response in test booklet	The student conveys verbally or signs responses to a scribe in an isolated, individual setting so that other students cannot benefit or be disturbed. The scribe cannot suggest ideas, words, or concepts. The scribe records the student's answers verbatim. The student should indicate capitalization and punctuation if language mechanics are being assessed.
		Student taped response	The student speaks responses into a tape recorder in an isolated setting so that other students cannot benefit or be disturbed. The test examiner must be present at all times.
		Signed response	The student uses sign language to convey responses. A certified sign language interpreter or deaf education instructor records responses.
		Pointing to respond	The student points to correct responses and the administrator records responses in the MAP test booklet.
		Oral response	The student provides oral responses to the test examiner.
		Use of a Braille	A student records responses using a Braille. Examples of a Braille include a Braillewriter, a slate and stylus, or an electronic Braille note taker.
		Use of a communication device	The student uses a communication device to provide responses to the test examiner.
		Use of a computer/word processor/typewriter to respond	The student uses a computer/word processor to write the responses. (Provide a non-networked computer to avoid inappropriate use of the computer to access answers.) The student uses a typewriter to write the responses.
39		Use of a calculator/math table/ abacus	In sessions of the MAP where calculators are allowed, the accommodation code should not be marked. The use of a calculator represents an accommodation when it is used on a section of the assessment for which calculator use is not allowed. Students may use talking calculators, but only in an isolated setting. Students use tables to assist in simple addition, subtraction, multiplication, and division facts using whole numbers. Students use an abacus to perform mathematical computations by sliding beads along rods.
44		Other: Specify. See Note 4.	Other response accommodations
		Setting Accommodations	Description
50		Testing individually	The location should be free of noises, conversation, and distractions from adjoining rooms. Individual testing is appropriate when, for example, responses are given orally or questions are paraphrased.
51		Testing in small groups	The location should be free of noises, conversation, and distractions from adjoining rooms. Students may not interact with one another about questions or answers. The test examiner must be present at all times. Testing in small groups is not appropriate for students who give responses orally or require paraphrasing of questions.
53		Other: Specify	Other setting accommodations

Table 4. 4: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Regular Edition

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
3	Regular Edition	66,141	100.00%	66,220	100.00%		
	Oral reading	35	0.05%	4,304	6.50%		
	Oral reading blind	3	0.00%				
	Signing of assessment	4	0.01%	22	0.03%		
	Paraphrasing	2	0.00%	4	0.01%		
	Other administration	130	0.20%	91	0.14%		
	Oral reading in native language	1	0.00%	145	0.22%		
	Extend time— <i>TerraNova</i> session	2,548	3.85%	2,621	3.96%		
	Administer using > allotted periods	2,511	3.80%	2,473	3.73%		
	Other timing	488	0.74%	498	0.75%		
	Use of scribe	1,871	2.83%	1,655	2.50%		
	Use of calculator, math table, etc.	80	0.12%	1,340	2.02%		
	Use of bilingual dictionary	10	0.02%	32	0.05%		
	Other response	70	0.11%	71	0.11%		
	Testing individually	1,917	2.90%	1,820	2.75%		
	Testing in small group	4,321	6.53%	4,535	6.85%		
	Other setting	252	0.38%	248	0.37%		
4	Regular Edition	66,825	100.00%	66,895	100.00%		
	Oral reading	52	0.08%	5,054	7.56%		
	Oral reading blind	8	0.01%				
	Signing of assessment	10	0.01%	20	0.03%		
	Paraphrasing	3	0.00%	5	0.01%		
	Other administration	163	0.24%	101	0.15%		
	Oral reading in native language	1	0.00%	165	0.25%		
	Extend time— <i>TerraNova</i> session	2,938	4.40%	3,024	4.52%		
	Administer using > allotted periods	2,947	4.41%	2,928	4.38%		
	Other timing	620	0.93%	641	0.96%		
	Use of scribe	2,063	3.09%	1,915	2.86%		
	Use of calculator, math table, etc.	80	0.12%	1,881	2.81%		
	Use of bilingual dictionary	5	0.01%	47	0.07%		
	Other response	126	0.19%	116	0.17%		
	Testing individually	2,212	3.31%	2,124	3.18%		
	Testing in small group	5,118	7.66%	5,359	8.01%		
	Other setting	257	0.38%	245	0.37%		
5	Regular Edition	65,489	100.00%	65,583	100.00%	65,534	100.00%
	Oral reading	28	0.04%	5,217	7.95%	5,049	7.70%
	Oral reading blind	3	0.00%				
	Signing of assessment	2	0.00%	12	0.02%	13	0.02%
	Paraphrasing	5	0.01%	7	0.01%	7	0.01%
	Other administration	169	0.26%	115	0.18%	90	0.14%

Table 4.4: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Regular Edition (Cont'd)

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
5	Oral reading in native language	4	0.01%	143	0.22%	149	0.23%
	Extend time— <i>TerraNova</i> session	3,070	4.69%	3,158	4.82%	2,934	4.48%
	Administer using > allotted periods	3,090	4.72%	3,109	4.74%	2,885	4.40%
	Other timing	603	0.92%	605	0.92%	583	0.89%
	Use of scribe	1,839	2.81%	1,790	2.73%	1,836	2.80%
	Use of calculator, math table, etc.	100	0.15%	2,148	3.28%	689	1.05%
	Use of bilingual dictionary	6	0.01%	33	0.05%	42	0.06%
	Other response	97	0.15%	98	0.15%	94	0.14%
	Testing individually	1,933	2.95%	1,864	2.84%	1,814	2.77%
	Testing in small group	5,670	8.66%	5,980	9.12%	5,664	8.64%
	Other setting	260	0.40%	257	0.39%	245	0.37%
	6	Regular Edition	65,630	100.00%	65,673	100.00%	
Oral reading		36	0.05%	4,197	6.39%		
Oral reading blind		5	0.01%				
Signing of assessment		7	0.01%	26	0.04%		
Paraphrasing		10	0.02%	11	0.02%		
Other administration		141	0.21%	106	0.16%		
Oral reading in native language		2	0.00%	110	0.17%		
Extend time— <i>TerraNova</i> session		2,529	3.85%	2,599	3.96%		
Administer using > allotted periods		2,619	3.99%	2,577	3.92%		
Other timing		706	1.08%	723	1.10%		
Use of scribe		1,245	1.90%	1,076	1.64%		
Use of calculator, math table, etc.		120	0.18%	2,912	4.43%		
Use of bilingual dictionary		17	0.03%	76	0.12%		
Other response		62	0.09%	51	0.08%		
Testing individually		1,411	2.15%	1,270	1.93%		
Testing in small group	5,657	8.62%	5,893	8.97%			
Other setting	202	0.31%	197	0.30%			
7	Regular Edition	66,648	100.00%	66,677	100.00%		
	Oral reading	41	0.06%	3,796	5.69%		
	Oral reading blind	3	0.00%				
	Signing of assessment	4	0.01%	14	0.02%		
	Paraphrasing	11	0.02%	10	0.01%		
	Other administration	81	0.12%	77	0.12%		
	Oral reading in native language	1	0.00%	117	0.18%		
	Extend time— <i>TerraNova</i> session	2,299	3.45%	2,338	3.51%		
	Administer using > allotted periods	2,425	3.64%	2,364	3.55%		
	Other timing	496	0.74%	506	0.76%		
	Use of scribe	854	1.28%	619	0.93%		
	Use of calculator, math table, etc.	168	0.25%	3,524	5.29%		
	Use of bilingual dictionary	23	0.03%	97	0.15%		

Table 4.4: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Regular Edition (Cont'd)

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
7	Other response	41	0.06%	33	0.05%		
	Testing individually	951	1.43%	824	1.24%		
	Testing in small group	5,831	8.75%	6,072	9.11%		
	Other setting	155	0.23%	157	0.24%		
8	Regular Edition	6,7247	100.00%	67,282	100.00%	67,179	100.00%
	Oral reading	34	0.05%	3,475	5.16%	3,387	5.04%
	Oral reading blind	7	0.01%				
	Signing of assessment	15	0.02%	22	0.03%	22	0.03%
	Paraphrasing	4	0.01%	2	0.00%	4	0.01%
	Other administration	61	0.09%	53	0.08%	56	0.08%
	Oral reading in native language	0	0.00%	123	0.18%	106	0.16%
	Extend time— <i>TerraNova</i> session	2,028	3.02%	2,045	3.04%	1,924	2.86%
	Administer using > allotted periods	2,202	3.27%	2,222	3.30%	2,055	3.06%
	Other timing	442	0.66%	446	0.66%	411	0.61%
	Use of scribe	634	0.94%	520	0.77%	579	0.86%
	Use of calculator, math table, etc.	204	0.30%	3,430	5.10%	1,983	2.95%
	Use of bilingual dictionary	8	0.01%	95	0.14%	103	0.15%
	Other response	42	0.06%	32	0.05%	43	0.06%
	Testing individually	768	1.14%	694	1.03%	691	1.03%
	Testing in small group	5,491	8.17%	5,722	8.50%	5,487	8.17%
Other setting	158	0.23%	160	0.24%	155	0.23%	
11	Regular Edition	61,010	100.00%	68,743	100.00%	62,101	100.00%
	Oral reading	61	0.10%	2,211	3.22%	1,735	2.79%
	Oral reading blind	10	0.02%				
	Signing of assessment	14	0.02%	19	0.03%	18	0.03%
	Paraphrasing	16	0.03%	10	0.01%	11	0.02%
	Other administration	34	0.06%	54	0.08%	36	0.06%
	Oral reading in native language	3	0.00%	67	0.10%	61	0.10%
	Extend time— <i>TerraNova</i> session	1,863	3.05%	1,968	2.86%	1,718	2.77%
	Administer using > allotted periods	1,048	1.72%	1,192	1.73%	1,030	1.66%
	Other timing	260	0.43%	274	0.40%	233	0.38%
	Use of scribe	247	0.40%	141	0.21%	215	0.35%
	Use of calculator, math table, etc.	53	0.09%	3,001	4.37%	786	1.27%
	Use of bilingual dictionary	0	0.00%	27	0.04%	52	0.08%
	Other response	20	0.03%	14	0.02%	19	0.03%
	Testing individually	328	0.54%	302	0.44%	309	0.50%
	Testing in small group	3,983	6.53%	5,075	7.38%	3,963	6.38%
Other setting	104	0.17%	84	0.12%	105	0.17%	

Table 4. 5: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Braille Edition

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
3	Braille Edition	7	100.00%	7	100.00%		
	Oral reading	0	0.00%	4	57.14%		
	Oral reading blind	2	28.57%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	0	0.00%	0	0.00%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	1	14.29%	2	28.57%		
	Administer using > allotted periods	4	57.14%	4	57.14%		
	Other timing	0	0.00%	0	0.00%		
	Use of scribe	5	71.43%	5	71.43%		
	Use of calculator, math table, etc.	2	28.57%	3	42.86%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		
	Other response	2	28.57%	1	14.29%		
	Testing individually	5	71.43%	5	71.43%		
	Testing in small group	0	0.00%	0	0.00%		
	Other setting	0	0.00%	0	0.00%		
4	Braille Edition	6	100.00%	6	100.00%		
	Oral reading	0	0.00%	4	66.67%		
	Oral reading blind	1	16.67%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	1	16.67%		
	Other administration	0	0.00%	0	0.00%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	2	33.33%	2	33.33%		
	Administer using > allotted periods	1	16.67%	2	33.33%		
	Other timing	0	0.00%	0	0.00%		
	Use of scribe	4	66.67%	5	83.33%		
	Use of calculator, math table, etc.	0	0.00%	0	0.00%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		
	Other response	0	0.00%	0	0.00%		
	Testing individually	4	66.67%	5	83.33%		
	Testing in small group	0	0.00%	0	0.00%		
	Other setting	0	0.00%	0	0.00%		
5	Braille Edition	4	100.00%	4	100.00%	4	100.00%
	Oral reading	0	0.00%	2	50.00%	2	50.00%
	Oral reading blind	0	0.00%				
	Signing of assessment	0	0.00%	0	0.00%	0	0.00%
	Paraphrasing	0	0.00%	0	0.00%	0	0.00%
	Other administration	0	0.00%	0	0.00%	0	0.00%

Table 4.5: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Braille Edition (Cont'd)

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
5	Oral reading in native language	0	0.00%	0	0.00%	0	0.00%
	Extend time— <i>TerraNova</i> session	3	75.00%	2	50.00%	2	50.00%
	Administer using > allotted periods	2	50.00%	2	50.00%	2	50.00%
	Other timing	2	50.00%	2	50.00%	2	50.00%
	Use of scribe	3	75.00%	3	75.00%	3	75.00%
	Use of calculator, math table, etc.	0	0.00%	1	25.00%	0	0.00%
	Use of bilingual dictionary	0	0.00%	0	0.00%	0	0.00%
	Other response	2	50.00%	2	50.00%	2	50.00%
	Testing individually	3	75.00%	3	75.00%	3	75.00%
	Testing in small group	1	25.00%	1	25.00%	1	25.00%
	Other setting	0	0.00%	0	0.00%	0	0.00%
6	Braille Edition	7	100.00%	8	100.00%		
	Oral reading	1	14.29%	2	25.00%		
	Oral reading blind	0	0.00%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	0	0.00%	0	0.00%		
	Oral reading in native language	0	0.00%	1	12.50%		
	Extend time— <i>TerraNova</i> session	1	14.29%	1	12.50%		
	Administer using > allotted periods	4	57.14%	5	62.50%		
	Other timing	0	0.00%	0	0.00%		
	Use of scribe	4	57.14%	5	62.50%		
	Use of calculator, math table, etc.	0	0.00%	0	0.00%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		
	Other response	0	0.00%	0	0.00%		
Testing individually	4	57.14%	5	62.50%			
Testing in small group	3	42.86%	3	37.50%			
Other setting	0	0.00%	0	0.00%			
7	Braille Edition	6	100.00%	7	100.00%		
	Oral reading	0	0.00%	2	28.57%		
	Oral reading blind	2	33.33%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	0	0.00%	0	0.00%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	3	50.00%	2	28.57%		
	Administer using > allotted periods	2	33.33%	3	42.86%		
	Other timing	0	0.00%	0	0.00%		
	Use of scribe	3	50.00%	3	42.86%		
	Use of calculator, math table, etc.	0	0.00%	3	42.86%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		

Table 4.5: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Braille Edition (Cont'd)

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
7	Other response	0	0.00%	0	0.00%		
	Testing individually	5	83.33%	6	85.71%		
	Testing in small group	1	16.67%	1	14.29%		
	Other setting	0	0.00%	0	0.00%		
8	Braille Edition	6	100.00%	7	100.00%	6	100.00%
	Oral reading	0	0.00%	3	42.86%	3	50.00%
	Oral reading blind	3	50.00%				
	Signing of assessment	0	0.00%	0	0.00%	0	0.00%
	Paraphrasing	0	0.00%	0	0.00%	0	0.00%
	Other administration	1	16.67%	1	14.29%	1	16.67%
	Oral reading in native language	0	0.00%	0	0.00%	0	0.00%
	Extend time— <i>TerraNova</i> session	0	0.00%	0	0.00%	0	0.00%
	Administer using > allotted periods	2	33.33%	2	28.57%	2	33.33%
	Other timing	0	0.00%	0	0.00%	0	0.00%
	Use of scribe	4	66.67%	4	57.14%	4	66.67%
	Use of calculator, math table, etc.	0	0.00%	3	42.86%	1	16.67%
	Use of bilingual dictionary	0	0.00%	0	0.00%	0	0.00%
	Other response	0	0.00%	0	0.00%	0	0.00%
	Testing individually	4	66.67%	4	57.14%	4	66.67%
Testing in small group	2	33.33%	2	28.57%	2	33.33%	
Other setting	0	0.00%	0	0.00%	0	0.00%	
11	Braille Edition	10	100.00%	7	100.00%	11	100.00%
	Oral reading	0	0.00%	0	0.00%	3	27.27%
	Oral reading blind	3	30.00%				
	Signing of assessment	0	0.00%	0	0.00%	0	0.00%
	Paraphrasing	0	0.00%	0	0.00%	0	0.00%
	Other administration	1	10.00%	0	0.00%	1	9.09%
	Oral reading in native language	0	0.00%	0	0.00%	0	0.00%
	Extend time— <i>TerraNova</i> session	3	30.00%	1	14.29%	3	27.27%
	Administer using > allotted periods	5	50.00%	3	42.86%	5	45.45%
	Other timing	0	0.00%	0	0.00%	0	0.00%
	Use of scribe	5	50.00%	3	42.86%	5	45.45%
	Use of calculator, math table, etc.	0	0.00%	4	57.14%	1	9.09%
	Use of bilingual dictionary	0	0.00%	0	0.00%	0	0.00%
	Other response	0	0.00%	0	0.00%	0	0.00%
	Testing individually	5	50.00%	3	42.86%	5	45.45%
Testing in small group	3	30.00%	2	28.57%	5	45.45%	
Other setting	0	0.00%	1	14.29%	0	0.00%	

Table 4. 6: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Large Print Edition

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
3	Large Print Edition	31	100.00%	31	100.00%		
	Oral reading	0	0.00%	13	41.94%		
	Oral reading blind	3	9.68%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	1	3.23%	1	3.23%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	7	22.58%	6	19.35%		
	Administer using > allotted periods	9	29.03%	10	32.26%		
	Other timing	1	3.23%	0	0.00%		
	Use of scribe	14	45.16%	14	45.16%		
	Use of calculator, math table, etc.	0	0.00%	3	9.68%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		
	Other response	1	3.23%	0	0.00%		
	Testing individually	12	38.71%	12	38.71%		
	Testing in small group	9	29.03%	9	29.03%		
	Other setting	1	3.23%	1	3.23%		
4	Large Print Edition	42	100.00%	43	100.00%		
	Oral reading	1	2.38%	16	37.21%		
	Oral reading blind	2	4.76%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	1	2.38%	1	2.33%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	9	21.43%	10	23.26%		
	Administer using > allotted periods	16	38.10%	15	34.88%		
	Other timing	1	2.38%	1	2.33%		
	Use of scribe	16	38.10%	16	37.21%		
	Use of calculator, math table, etc.	1	2.38%	5	11.63%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		
	Other response	0	0.00%	0	0.00%		
	Testing individually	17	40.48%	18	41.86%		
	Testing in small group	18	42.86%	17	39.53%		
	Other setting	0	0.00%	0	0.00%		
5	Large Print Edition	51	100.00%	49	100.00%	48	100.00%
	Oral reading	0	0.00%	20	40.82%	20	41.67%
	Oral reading blind	2	3.92%				
	Signing of assessment	0	0.00%	0	0.00%	0	0.00%
	Paraphrasing	0	0.00%	0	0.00%	0	0.00%
	Other administration	4	7.84%	3	6.12%	2	4.17%

Table 4.6: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Large Print Edition (Cont'd)

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
5	Oral reading in native language	0	0.00%	0	0.00%	0	0.00%
	Extend time— <i>TerraNova</i> session	14	27.45%	14	28.57%	13	27.08%
	Administer using > allotted periods	27	52.94%	27	55.10%	23	47.92%
	Other timing	3	5.88%	4	8.16%	4	8.33%
	Use of scribe	22	43.14%	21	42.86%	20	41.67%
	Use of calculator, math table, etc.	1	1.96%	12	24.49%	3	6.25%
	Use of bilingual dictionary	0	0.00%	0	0.00%	0	0.00%
	Other response	2	3.92%	2	4.08%	1	2.08%
	Testing individually	22	43.14%	21	42.86%	21	43.75%
	Testing in small group	19	37.25%	19	38.78%	18	37.50%
	Other setting	1	1.96%	2	4.08%	2	4.17%
6	Large Print Edition	35	100.00%	35	100.00%		
	Oral reading	0	0.00%	12	34.29%		
	Oral reading blind	1	2.86%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	1	2.86%	0	0.00%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	15	42.86%	16	45.71%		
	Administer using > allotted periods	17	48.57%	17	48.57%		
	Other timing	1	2.86%	0	0.00%		
	Use of scribe	11	31.43%	11	31.43%		
	Use of calculator, math table, etc.	0	0.00%	10	28.57%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		
	Other response	1	2.86%	1	2.86%		
Testing individually	14	40.00%	13	37.14%			
Testing in small group	10	28.57%	11	31.43%			
Other setting	0	0.00%	0	0.00%			
7	Large Print Edition	47	100.00%	43	100.00%		
	Oral reading	1	2.13%	13	30.23%		
	Oral reading blind	7	14.89%				
	Signing of assessment	0	0.00%	0	0.00%		
	Paraphrasing	0	0.00%	0	0.00%		
	Other administration	0	0.00%	0	0.00%		
	Oral reading in native language	0	0.00%	0	0.00%		
	Extend time— <i>TerraNova</i> session	16	34.04%	15	34.88%		
	Administer using > allotted periods	17	36.17%	16	37.21%		
	Other timing	1	2.13%	1	2.33%		
	Use of scribe	24	51.06%	20	46.51%		
	Use of calculator, math table, etc.	4	8.51%	16	37.21%		
	Use of bilingual dictionary	0	0.00%	0	0.00%		

Table 4.6: Number and Percent of Students Receiving Accommodations by Accommodation Type, MAP 2008 Large Print Edition (Cont'd)

Grade	Accommodation	Communication Arts		Mathematics		Science	
		Frequency	Percent	Frequency	Percent	Frequency	Percent
7	Other response	3	6.38%	3	6.98%		
	Testing individually	27	57.45%	22	51.16%		
	Testing in small group	14	29.79%	15	34.88%		
	Other setting	0	0.00%	0	0.00%		
8	Large Print Edition	25	100.00%	23	100.00%	24	100.00%
	Oral reading	0	0.00%	3	13.04%	4	16.67%
	Oral reading blind	1	4.00%				
	Signing of assessment	0	0.00%	0	0.00%	0	0.00%
	Paraphrasing	0	0.00%	0	0.00%	0	0.00%
	Other administration	0	0.00%	0	0.00%	0	0.00%
	Oral reading in native language	0	0.00%	0	0.00%	0	0.00%
	Extend time— <i>TerraNova</i> session	5	20.00%	3	13.04%	2	8.33%
	Administer using > allotted periods	3	12.00%	1	4.35%	1	4.17%
	Other timing	0	0.00%	0	0.00%	0	0.00%
	Use of scribe	10	40.00%	8	34.78%	8	33.33%
	Use of calculator, math table, etc.	0	0.00%	4	17.39%	0	0.00%
	Use of bilingual dictionary	0	0.00%	0	0.00%	0	0.00%
	Other response	0	0.00%	0	0.00%	0	0.00%
	Testing individually	8	32.00%	8	34.78%	8	33.33%
	Testing in small group	10	40.00%	9	39.13%	8	33.33%
Other setting	0	0.00%	0	0.00%	0	0.00%	
11	Large Print Edition	21	100.00%	26	100.00%	21	100.00%
	Oral reading	0	0.00%	2	7.69%	3	14.29%
	Oral reading blind	0	0.00%				
	Signing of assessment	0	0.00%	0	0.00%	0	0.00%
	Paraphrasing	0	0.00%	1	3.85%	0	0.00%
	Other administration	0	0.00%	0	0.00%	0	0.00%
	Oral reading in native language	0	0.00%	0	0.00%	0	0.00%
	Extend time— <i>TerraNova</i> session	5	23.81%	9	34.62%	5	23.81%
	Administer using > allotted periods	3	14.29%	6	23.08%	4	19.05%
	Other timing	1	4.76%	0	0.00%	0	0.00%
	Use of scribe	9	42.86%	6	23.08%	8	38.10%
	Use of calculator, math table, etc.	0	0.00%	6	23.08%	1	4.76%
	Use of bilingual dictionary	0	0.00%	0	0.00%	0	0.00%
	Other response	0	0.00%	0	0.00%	0	0.00%
	Testing individually	8	38.10%	6	23.08%	8	38.10%
	Testing in small group	9	42.86%	14	53.85%	10	47.62%
Other setting	0	0.00%	0	0.00%	3	14.29%	

Figure 4. 1: Sample Script of Examiner’s Manual

Directions for Administering Communication Arts

SESSION 1

Distribute the test books. If this is the first day of testing, check to see that students write their names and district/school on their test books. Ensure that all students use a non-mechanical No. 2 pencil.

Before administering the test, take a moment to have your students look through the test book. Point out the STOP pages. Tell the students that whenever they see one of the STOP pages, they should not continue.

SAY

In Session 1 you will find two types of questions. For some questions, you will choose from a list of answers. Fill in **only** the circle that goes with the answer you choose. Fill in the circle completely and make your mark solid and dark. If you want to change the answer, completely erase the mark you made and fill in a different circle. For the remaining questions, you will write your own answers in the test book. **For scoring purposes, write or print your answers clearly and stay within the area provided.**

You should read each question carefully and do your best to answer clearly and completely. Your score on these questions will depend on how well you show your understanding of what you read. You may choose to look over the questions **before** reading the passage. Read the passage, and then answer the questions.

Some of the questions in this session are worth more than one point, so you can earn partial credit.

Open your test book to Session 1 of Communication Arts. Be sure to stay on the pages that are marked “Session 1” at the bottom of the page.

CHAPTER 5: CONSTRUCTED-RESPONSE SCORING

In this section, we first describe the scoring process used for MAP. In particular, we focus on the MAP handscoring process. At the end of this section, we describe and report the results of the inter-rater reliability study conducted on the handscoring of MAP constructed-response items.

Chapter 5 adheres to AERA/APA/NCME standards 3.22, 3.23, 3.24, 5.9, and 6.7.

5.1 MAP Scoring Process

Multiple-choice items were scored by CTB using electronic scanning equipment. Constructed-response items were scored by human raters who were trained by CTB.

5.1.1 Handscoring Process Used for MAP

Evidence of validity is provided by the procedures described below for handscoring.

5.1.2 Selection of Scoring Evaluators

CTB/McGraw-Hill and Kelly Services strive to develop a highly qualified, experienced core of evaluators so that the integrity of all projects is appropriately maintained.

Recruitment

The MAP 2008 project was staffed with a large number of returning evaluators and team leaders who had previous experience with MAP and other handscoring projects. Kelly Services also recruited new team leaders and evaluators for employment. Recruitment sources included advertisements in newspapers in Indianapolis, Indiana; Mather, California, and nearby areas and Internet sources.

CTB requires that all evaluators and team leaders possess a bachelor's degree or higher. Kelly Services carefully screened all new applicants and required them to produce either a transcript or a copy of the degree. Kelly Services also required a one- to two-hour interview/screening process. Individuals who did not present proper documentation or had less than desirable work records were eliminated during this process. Kelly Services verified that 100% of all potential evaluators met the degree requirement. All experienced evaluators and team leaders had already successfully completed the screening process.

The Interview Process

All potential evaluators completed a pre-interview activity. For some parts of the pre-interview activity, applicants were shown examples of test responses and were supplied with a scoring guide. In a brief introduction, they became acquainted with the application of a rubric. After the introduction, applicants applied the scoring guide to score the sample responses. The applicant's scores were used for discussion during the interview process to determine the applicant's trainability as well as his/her ability to understand and implement the standards set forth in the sample scoring guide.

Kelly Services interviewed each applicant and determined the applicant's suitability for a specific content area and grade level. Applicants with strong leadership skills were questioned further to determine whether they were qualified to be team leaders.

When Kelly Services determined applicants were qualified, the applicants were recommended for employment. All assignments were made according to availability and suitability. Before being hired, all employees were required to read, agree to, and sign a nondisclosure agreement outlining the CTB/McGraw-Hill business ethics and security procedures.

5.1.3 Handscoring Process

Training Material Development

All materials necessary for scoring were developed by CTB. These materials include the scoring guides and training papers used to complete the handscoring of constructed-response and extended-response items (writing essays and performance events).

Missouri operational items have been previously field tested. Prior to actual scoring, handscoring supervisors assembled materials based on the rubrics. Student answer documents were randomly sampled to ensure that a representative sample of possible responses was used. Supervisors selected anchor papers and training papers and recommended clarifications to rubrics. All materials were presented during the Training Material Review Meeting (TMRM) and scores and annotations were approved by DESE participants.

From this point, training and qualifying materials were developed based on the rubric and scoring philosophies discussed during the TMRM.

Training Material Review Meeting

CTB prepared all anchors, scoring guides, and student response samples for DESE and Missouri participant review. Each response, score, and annotation was reviewed and updated as needed within the outlined limitations.

Training and Qualifying Procedures

Handscoring involves training and qualifying team leaders and evaluators, monitoring scoring accuracy and production, and ensuring security of both the test materials and the scoring facilities. An explanation of the training and qualification procedures follows.

All readers were trained and qualified in specific RIBs (Rater Item Block) consisting of one item to be scored. Evaluators and team leaders were trained using the following steps:

- Reviewing the student response booklet
- Reviewing rubrics
- Reviewing anchor papers
- Explaining scoring strategies, followed by a question-and-answer period

- Scoring a training set, followed by sharing established scores, discussing responses, and answering questions arising from scores
- Scoring and discussing additional training sets
- Qualifying Round 1
- Qualifying Round 2 (if necessary)
- Explaining condition codes and sensitive paper procedures
- Explaining nonstandard response or computer-generated response (nsr/cgr) procedures
- Explaining un-scannable image procedures

All evaluators were trained and qualified using the same procedures and criteria used for the team leaders. Qualification standards for every item were predetermined by DESE. In order to score an item, readers must have met the specific standards for that item. The qualification standards were:

- 4-point item: 80% qualification
- 2-point item: 90% qualification
- 1-point item: 100% qualification

Team Leader Training (TLT)

DESE and participants joined CTB team leaders during training. During this time, regional facilitators and team leaders were trained and qualified. These participants served as trainers and team leaders at each of the distributed handscoring sites (DHS). Once qualification was completed, checksets were presented for approval.

5.1.4 Monitoring the Scoring Process

Daily Accuracy Checks

Throughout the course of handscoring, calibration sets of pre-scored papers (checksets/validity sets) were administered daily to each scorer to monitor scoring accuracy and to maintain a consistent focus on the established rubrics and guidelines. Checksets were executed via imaging software that provided images in such a way that the reader did not know when a checkset was administered. All checkset scores had been approved by DESE participants during and immediately following the TMRM.

In addition to the checkset process, CTB’s handscoring protocol included the use of read-behinds. The read-behind was another valuable rater-reliability monitoring technique that allowed a team leader to review a reader’s scored documents, providing feedback and counseling as appropriate.

Approximately 5% of Communication Arts, Mathematics, and Science papers were scored by a second reader to establish inter-rater reliability statistics for all constructed-response items. This procedure is called a “double-blind read,” because the second reader does not know the first reader’s score.

Recalibration of Raters

Recalibration in handscoring refers to the process in which scorers/raters who begin to drift away from scoring accuracy are realigned to correct scoring.

5.1.5 Security

Security guards were on site whenever employees were present in the building. All employees were issued photo identification badges and were required to wear them in plain view at all times. Visitors and employees who forgot their badges were issued visitors' badges and were required to wear them in plain view. All employees and visitors were subject to inspection of their personal effects.

5.2 Inter-Rater Reliability

Approximately 5% of papers in Communication Arts, Mathematics, and Science were scored independently by a second reader. To determine the reliability of scoring, the percent of perfect agreement and adjacent agreement between the two readers was calculated. A weighted kappa was calculated to reflect the level of improvement beyond the chance level in the consistency of scoring. The statistics for the inter-rater reliability were calculated by form. For Grades 3 through 8 where six different test forms were administered, a mean and standard deviation of the weighted kappa averaged across the six forms will be presented. Item-level rater agreement rates, average weighted Kappas, and the standard deviation of the weighted kappas are presented in Tables 5.1, 5.2, and 5.3. High inter-rater reliability is expected for items that have few points.

All Communication Arts, Mathematics, and Science items show good inter-rater agreement. As shown in Table 5.1, raters demonstrated at least 95% perfect and adjacent agreement for all items. The Kappa statistic for the Communication Arts items ranged from an average weighted kappa of .58, with a standard deviation of .05 (Grade 5, Session 3, Item 2) to an average weighted kappa of .96, with a standard deviation of .02 (Grade 7, Session 4, Item 3B).

As shown in Table 5.2, raters demonstrated above 97% adjacent agreement for all Mathematics items. The Kappa statistic for the Mathematics items ranged from an average weighted kappa of .81, with a standard deviation of .02 (Grade 8, Session 3, Item 3) to an average weighted kappa of .99, with a standard deviation of 0.0 (Grade 3, Session 3, Item 4; Grade 5, Session 3, Item 2; Grade 10, Session 3, Item 4).

As shown in Table 5.3, raters demonstrated above 92% adjacent agreement for all Science items. The Kappa statistic for the Science items ranged from an average weighted kappa of .38, with a standard deviation of .04 (Grade 8, Session 3, Item 3) to an average weighted kappa of .99, with a standard deviation of 0.01 (Grade 8, Session 3, Item 4). The Grade 8 item with an averaged weighted kappa of .38 was a difficult item (p -value = .18). The percent of perfect agreement on this item was 83%.

Table 5. 1: Inter-rater Reliability, Communication Arts

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD
3	3	1	2	83	17	100	0.82	0.02
	4	1	2	78	21	99	0.77	0.03
	5	1	2	87	12	99	0.88	0.02
	6A	1	2	78	21	99	0.77	0.04
	6B	1	1	99	1	100	0.86	0.06
	6C	1	1	99	1	100	0.88	0.03
	1	2	4	68	31	99	0.70	0.03
	1	4	2	82	17	99	0.79	0.01
	2	4	2	90	9	99	0.90	0.01
	3A	4	2	90	9	99	0.94	0.01
	3B	4	1	95	4	99	0.90	0.03
4	3	1	2	86	13	99	0.81	0.03
	4	1	2	86	14	100	0.82	0.01
	5A	1	2	80	20	100	0.76	0.04
	5B	1	1	99	1	100	0.90	0.03
	5C	1	1	96	4	100	0.68	0.06
	6	1	2	78	21	99	0.80	0.01
	1	3	2	94	6	100	0.93	0.01
	2	3	2	79	20	99	0.78	0.01
3	3	2	85	14	99	0.87	0.02	
5	3	1	2	87	13	100	0.75	0.04
	4	1	2	88	11	99	0.86	0.02
	5A	1	2	78	20	98	0.74	0.02
	5B	1	1	93	7	100	0.72	0.03
	6A	1	2	84	15	99	0.86	0.03
	6B	1	1	97	3	100	0.87	0.02
	1	3	2	77	21	98	0.74	0.02
	2	3	2	63	36	99	0.58	0.05
	3	3	2	81	17	98	0.84	0.02
6	3	1	2	87	9	96	0.81	0.02
	4	1	2	78	17	95	0.74	0.03
	5	1	2	70	26	96	0.66	0.02
	6A	1	2	79	19	98	0.76	0.01
	6B	1	1	98	2	100	0.95	0.02
	1	3	2	77	22	99	0.79	0.01
	2	3	2	91	8	99	0.94	0.01
	3	3	2	85	15	100	0.80	0.01
7	3	1	2	75	24	99	0.75	0.02
	4	1	2	74	24	98	0.74	0.03
	5	1	2	80	15	95	0.78	0.03

Table 5. 1: Inter-rater Reliability, Communication Arts (continued)

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD
7	6A	1	2	77	21	98	0.75	0.03
	6B	1	2	90	9	99	0.88	0.01
	1	2	4	75	24	99	0.68	0.04
	1	4	2	83	15	98	0.85	0.01
	2	4	2	93	5	98	0.86	0.02
	3A	4	2	85	14	99	0.86	0.03
	3B	4	1	98	2	100	0.96	0.02
	3C	4	1	95	5	100	0.79	0.04
8	3	1	2	75	23	98	0.74	0.03
	4	1	2	71	26	97	0.70	0.02
	5	1	2	82	16	98	0.78	0.03
	6A	1	2	80	17	97	0.78	0.02
	6B	1	1	98	2	100	0.93	0.01
	6C	1	1	97	3	100	0.77	0.05
	1	3	2	74	24	98	0.72	0.03
	2	3	2	62	36	98	0.59	0.03
	3A	3	2	77	20	97	0.80	0.03
	3B	3	2	83	16	99	0.85	0.01
11	3	1	2	77	18	95	0.77	-
	4	1	2	85	13	98	0.82	-
	5	1	2	75	20	95	0.75	-
	6A	1	2	76	22	98	0.75	-
	6B	1	1	95	5	100	0.87	-
	6C	1	1	99	1	100	0.93	-
	1	2	4	66	33	99	0.71	-
	1	4	2	72	27	99	0.71	-
	2	4	2	81	17	98	0.82	-
	3A	4	2	71	27	98	0.73	-
	3B	4	1	97	3	100	0.94	-
	3C	4	1	99	1	100	0.95	-

* The percent perfect & adjacent may not add up to 100 for 1-point items due to the percent discrepant. The percent discrepant includes the cases where one rater assigned a score and the other rater assigned a condition code. With 2- or more point items, it also refers to the cases where the assigned score varied by more than 1 point.

Table 5. 2: Inter-rater Reliability, Mathematics

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD
3	1	3	2	94	6	100	0.96	0.01
	2	3	2	92	8	100	0.92	0.01
	3	3	2	98	2	100	0.98	0.00
	4	3	2	99	0	99	0.99	0.00
	5	3	2	93	7	100	0.94	0.01
	6	3	2	93	7	100	0.95	0.01
	7	3	2	98	1	99	0.98	0.01
4	31	1	4	85	12	97	0.93	0.01
	1	3	2	96	4	100	0.94	0.02
	2	3	2	93	7	100	0.94	0.02
	3	3	2	98	2	100	0.97	0.01
	4	3	2	97	2	99	0.98	0.01
	5	3	2	98	2	100	0.96	0.01
	6	3	2	96	3	99	0.94	0.01
	7	3	2	95	5	100	0.97	0.00
	8	3	2	98	2	100	0.98	0.00
9	3	2	93	7	100	0.94	0.01	
5	1	3	2	89	8	97	0.90	0.02
	2	3	2	99	1	100	0.99	0.00
	3	3	2	98	2	100	0.98	0.01
	4	3	2	96	4	100	0.97	0.01
	5	3	2	93	7	100	0.91	0.01
	6	3	2	97	2	99	0.96	0.01
	7	3	2	97	2	99	0.98	0.01
6	1	3	2	85	14	99	0.90	0.01
	2	3	2	91	8	99	0.95	0.01
	3	3	2	96	3	99	0.97	0.01
	4	3	2	91	8	99	0.89	0.01
	5	3	2	96	4	100	0.95	0.01
	6	3	2	94	6	100	0.91	0.02
	7	3	2	86	14	100	0.85	0.03
7	1	3	2	89	11	100	0.91	0.01
	2	3	2	95	5	100	0.95	0.01
	3	3	2	91	8	99	0.94	0.01
	4	3	2	94	6	100	0.95	0.01
	5	3	2	98	2	100	0.98	0.01
	6	3	2	98	2	100	0.98	0.01
	7	3	2	95	5	100	0.96	0.01
8	31	1	4	88	11	99	0.95	0.01
	1	3	2	94	6	100	0.96	0.01

Table 5. 2: Inter-rater Reliability, Mathematics (continued)

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD
8	2	3	2	97	3	100	0.97	0.01
	3	3	2	85	15	100	0.81	0.02
	4	3	2	98	2	100	0.98	0.01
	5	3	2	93	7	100	0.96	0.00
	6	3	2	96	4	100	0.96	0.01
	7	3	2	92	7	99	0.95	0.01
	8	3	2	98	2	100	0.98	0.01
	9	3	2	94	5	99	0.96	0.01
10	1	3	2	94	5	99	0.96	-
	2	3	2	93	7	100	0.96	-
	3	3	2	95	4	99	0.96	-
	4	3	2	98	1	99	0.99	-
	5	3	2	95	4	99	0.97	-
	6	3	2	92	7	99	0.92	-
	7	3	2	96	4	100	0.96	-
	8	3	2	90	9	99	0.94	-
	9	3	2	89	10	99	0.92	-
	10	3	4	86	13	99	0.87	-

* The percent perfect & adjacent may not add up to 100 for 1-point items due to the percent discrepant. The percent discrepant includes the cases where one rater assigned a score and the other rater assigned a condition code in addition to the cases where the score assigned varied by more than 1 point.

Table 5. 3: Inter-rater Reliability, Science

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD	
5	1	1	2	88	12	100	0.86	0.02	
	2	1	2	97	2	99	0.96	0.01	
	3	1	2	84	15	99	0.86	0.01	
	4	1	2	82	17	99	0.81	0.03	
	5	1	2	74	23	97	0.72	0.03	
	6	1	2	72	26	98	0.73	0.02	
	7	1	2	90	9	99	0.89	0.01	
	8	1	2	97	3	100	0.97	0.01	
	9	1	2	83	16	99	0.82	0.01	
	10	1	2	81	17	98	0.72	0.03	
	11	1	2	74	22	96	0.71	0.04	
	26	2	2	85	15	100	0.72	0.04	
	27	2	2	78	21	99	0.81	0.02	
	28	2	2	88	12	100	0.90	0.01	
	29	2	2	76	23	99	0.79	0.02	
	30	2	2	77	21	98	0.72	0.02	
	31	2	2	81	17	98	0.85	0.02	
	32	2	2	83	16	99	0.85	0.01	
	33	2	2	88	12	100	0.86	0.01	
	34	2	2	86	13	99	0.83	0.03	
	35	2	2	82	17	99	0.81	0.03	
	36	2	2	92	8	100	0.90	0.02	
	1	3	2	96	4	100	0.96	0.01	
	2	3	4	79	13	92	0.89	0.01	
	3	3	1	91	9	100	0.81	0.03	
	4	3	1	99	1	100	0.98	0.01	
	5	3	2	70	27	97	0.69	0.02	
	6	3	1	100	0	100	0.97	0.02	
	7	3	1	90	10	100	0.72	0.03	
	8	3	1	82	17	99	0.59	0.04	
	9	3	1	97	3	100	0.94	0.02	
	8	1	1	2	98	2	100	0.98	0.01
		2	1	2	90	10	100	0.70	0.05
		3	1	2	85	14	99	0.82	0.02
		4	1	2	94	6	100	0.85	0.01
		5	1	2	95	5	100	0.94	0.01
6		1	2	91	9	100	0.93	0.02	
7		1	2	84	16	100	0.81	0.02	
8		1	2	74	25	99	0.67	0.03	
9		1	2	84	15	99	0.80	0.01	

Table 5.3: Inter-rater Reliability, Science (continued)

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD
8	10	1	2	98	1	99	0.97	0.01
	11	1	2	81	17	98	0.76	0.02
	12	1	2	80	19	99	0.78	0.02
	26	2	2	97	3	100	0.96	0.01
	27	2	2	83	16	99	0.84	0.02
	28	2	2	83	16	99	0.83	0.02
	29	2	2	84	14	98	0.86	0.01
	30	2	2	89	10	99	0.91	0.01
	31	2	2	89	9	98	0.94	0.01
	32	2	2	88	11	99	0.90	0.02
	33	2	2	91	9	100	0.81	0.04
	34	2	2	94	5	99	0.90	0.02
	35	2	2	80	18	98	0.83	0.02
	36	2	2	82	17	99	0.82	0.02
	37	2	2	87	13	100	0.85	0.03
	1	3	1	95	5	100	0.91	0.02
	2	3	3	76	21	97	0.77	0.02
	3	3	1	83	17	100	0.38	0.04
	4	3	1	99	1	100	0.99	0.01
	5	3	1	88	12	100	0.75	0.04
6	3	4	84	13	97	0.90	0.01	
7	3	1	95	4	99	0.86	0.03	
8	3	1	94	5	99	0.89	0.02	
9	3	1	84	16	100	0.66	0.03	
10	3	1	86	13	99	0.58	0.06	
11	1	1	2	97	3	100	0.96	-
	2	1	2	95	5	100	0.95	-
	3	1	2	88	11	99	0.87	-
	4	1	2	82	16	98	0.71	-
	5	1	2	86	13	99	0.86	-
	6	1	2	90	8	98	0.93	-
	7	1	2	84	15	99	0.81	-
	8	1	2	83	16	99	0.80	-
	9	1	2	91	9	100	0.85	-
	10	1	2	78	20	98	0.74	-
	11	1	2	82	17	99	0.82	-
	12	1	2	88	11	99	0.82	-
	13	1	2	86	12	98	0.90	-
	14	1	2	83	16	99	0.83	-
	15	1	2	89	10	99	0.87	-

Table 5. 3: Inter-rater Reliability, Science (continued)

Grade	Item #	Session	# Points	% Perfect	% Adjacent	% Perfect & Adjacent*	Kappa Mean	Kappa SD
11	16	1	2	85	14	99	0.84	-
	17	1	2	90	8	98	0.88	-
	18	1	2	94	5	99	0.95	-
	1	2	2	88	11	99	0.90	-
	2	2	2	69	26	95	0.71	-
	3	2	4	68	28	96	0.87	-
	4	2	1	99	1	100	0.98	-
	5	2	1	98	1	99	0.96	-
	6	2	2	75	22	97	0.78	-
	7	2	1	99	1	100	0.98	-
	8	2	1	86	13	99	0.72	-
	9	2	1	90	9	99	0.79	-
	10	2	3	71	25	96	0.80	-
	11	2	2	80	17	97	0.84	-
	26	3	2	90	9	99	0.85	-
	27	3	2	80	19	99	0.78	-
	28	3	2	81	19	100	0.72	-
29	3	3	87	12	99	0.90	-	
30	3	2	71	26	97	0.61	-	
31	3	2	88	11	99	0.87	-	

* The percent perfect & adjacent may not add up to 100 for 1-point items due to the percent discrepant. The percent discrepant includes the cases where one rater assigned a score and the other rater assigned a condition code in addition to the cases where the score assigned varied by more than 1 point.

CHAPTER 6: OPERATIONAL DATA ANALYSES

This chapter of the MAP Technical Report describes the analyses that occurred on the operational data. These analyses include a classical item analysis and examination of the raw scores and an IRT analyses involving calibrating, scaling, and linking. All of these analyses were conducted using the calibration sample.

In this section, we first discuss the calibration sample. Next, we present the classical item statistics, including aggregate raw score statistics and individual item-level statistics. Then, we discuss the item response theory (IRT) models used for calibrating the data and address how well these models fit the Missouri data. If the IRT model fits the empirical item response distributions for the population for which we want to make generalizations, i.e., Missouri students, then the claim that the scores are valid indicators of an underlying ability is strengthened. The lowest obtainable scale score (LOSS) and highest obtainable scale score (HOSS) for MAP are presented. Finally, we provide a general overview of scaling and discuss the methods used to link the MAP results to the *TerraNova* scale.

Chapter 6 demonstrates adherence to AERA/APA/NCME standards 1.5, 1.13, 2.1, 2.2, 2.4, 2.8, 2.14, 3.18, 4.1, 4.2, 4.11, 6.4, 6.5, and 7.1 in the MAP program.

6.1 Calibration Sample

In 2008, the grade-level calibration samples were comprised of, at least, 80% of the total student population for that grade. Several large school districts were identified for inclusion in the 80% sample. These districts are listed in Table 6.1. Data from these districts had to be included in the calibration sample before data analyses procedures could begin. These large districts were identified because past data processing has demonstrated that large districts often return data at the end of the data-return window while small districts often return data early in the data-return window. Since the calibration sample was going to be based on the first 80% of data to be returned, it was important to identify large districts to ensure the calibration data were representative of the state.

Tables 6.2 through 6.4 examine the representativeness of the calibration sample compared to the census data. These tables demonstrate that the calibration sample was representative of the state.

6.2 Classical Item Statistics

In this section, we present summary test statistics for each grade/content area MAP. This is followed by item-level statistics for each grade/content area MAP.

6.2.1. Test-Level Statistics

Tables 6.5 to 6.7 present for each grade level of Communication Arts, Mathematics, and Science, respectively, the number of items and score points on each test, as well as the mean and standard deviation of the raw scores, *p-values* and item-total correlations (also known as item discrimination values). For Grade 11, the means of the *p-values* and item-total correlations were based on items from both the Grade 11 Communication Arts general and breach form. The data from the two forms were combined into a single data set for the purposes of operational analyses.

The mean *p-value* is the average of all item *p-values* of a specific grade/content area. The mean item-total correlation (R_{it}) is the average of all item biserial correlations of a specific grade/content area. The *p-value* and item-total correlation are explained in the next section, *Item-Level Statistics*.

6.2.2. Item-Level Statistics

Tables 6.8 to 6.15 present the item statistics for each item by grade/content area. The tables include test session, item booklet number and part (if applicable), number of score points, *p-values*, and item-total correlations (R_{it}) for each item by grade/content area. The constructed-response (CR) items appear in the tables first, followed by the multiple-choice (MC) items. Table 6.15 shows the item-level statistics for the Session 3 items from the breach form. Session 3 items from the general form are included in Table 6.14. Grade 11 Communication Arts items from Sessions 1, 2, and 4 are based on student responses from both the breach and general form.

P-Value: The *p-value* is a measure of item difficulty. For a multiple-choice item, the *p-value* is calculated from the number of students who correctly respond to an item divided by the total number of students attempting the item. The value is reported as a proportion. For a constructed-response item, the *p-value* is calculated from the average score for the item divided by the maximum points possible and is also reported as a proportion.

In terms of *p-values*, test scores tend to be more precise when their average *p-values* are in the mid 0.50s to low 0.70s. However, in building a criterion-referenced test, it is important to select items on the basis of content rather than on purely statistical criteria. As demonstrated in Tables 7.1 and 7.2, the average *p-values* associated with the Communication Arts MAP range from .68 (Grade 8) to .74 (Grade 4). The average *p-values* associated with the Mathematics MAP range from .57 (Grade 8) to .75 (Grade 3). The average *p-values* associated with the Science MAP range from .47 (Grade 11) to .57 (Grade 5). The average *p-values* for Grade 8 and Grade 11 Science tests indicate that these tests are difficult for the Grade 8 and Grade 11 students.

It is important that one examines the range of *p-values* and not just the average *p-value* to determine whether a test measures well. It is desirable for the test to measure well throughout the range of skills present at a given grade. That is, it is important that the items measure the performance of both low-scoring and high-scoring students, as well as students in the center of the distribution. Having a range of *p-values* also helps to prevent floor and/or ceiling effects so that the test does not have large numbers of students at the

minimum or maximum possible scores. The Communication Arts MAP has items with p -values ranging from the mid 0.10 to the 0.90s (see Tables 6.8 through 6.15). The p -values on the Mathematics MAP tend to range from the 0.20s and 0.30s to the 0.90s (see Tables 6.8 through 6.14). The Science MAP has items with p -values ranging from 0.10s to the 0.90s. (see Tables 6.10, 6.13, and 6.14). Such a broad range of p -values indicates that the items measure well throughout the range of skills at a given grade, and hence supports the accuracy of the MAP test scores.

Item-Total Correlations: An item-total correlation is the correlation between an item and the total test score where the item score is included in the total score. It indicates how well an item differentiates between low- and high-achieving students. In general, items with correlations below .20 are said to be poorly discriminating. The majority of the items in the MAP had item-test correlations above this threshold. Any item with an item-total correlation below the .20 threshold was further analyzed to assure that the item was correctly keyed.

Omit Rates: The omit rate for each item indicate the percentage of students who did not answer the item. Omit rates can be used to examine possible speededness issues on tests. A test may be speeded if students do not have adequate time to answer all questions on the test. As a rule of thumb, an item is said to have a high-omit rate if more than 5% of students failed to respond to the item.

In general, very few students omitted items across all of the assessments; however, there were a few items with high omit rates. Items 22 and 23 from Session 2 of the Grade 3 Mathematics had omit rates above 5%. These items occurred at the end of an administration section within Session 2, indicating that students may not have enough time to complete the section. Similarly items 11 and 12 from Session 2 of the Grade 4 Mathematics had omit rates above 5%, and these items occurred at the end of an administration section. Item 9 from Session 2 of the Grade 7 Mathematics had an omit rate above 5% and occurred at the end of an administration section.

Item 18 in Session 2 of the Grade 8 Science test had an omit rate above 5%. This may have been due to the layout of the item on the page. The items were displayed in two columns, and Item 18 was in the left-hand column. Students may not have seen the item.

Items 34 and 35 in Session 2 and item 4 in Session 3 of the Grade 8 Science test had omit rates above 5%. These were CR items, and it appears the students skipped the items. Items 13 and 17 in Session 1 and items 9, 10, and 11 in Session 2 of the Grade 11 Science MAP had omit rates above 5%. These were CR items. Items 9, 10, and 11 in Session 2 also appeared at the end of an administration section, so timing of the section may have been an issue for some students.

6.3 Item Response Theory

A marginal maximum-likelihood procedure was used to simultaneously estimate the item parameters using the 3PL/2PPC IRT models (Bock & Aitkin, 1981; Thissen, 1982).

Under the 3PL model, the probability that a student with trait or scale score θ will respond correctly to multiple-choice item j is

$$P_j(\theta) = c_j + (1 - c_j) / [1 + \exp(-1.7a_j(\theta - b_j))]. \quad (1)$$

In equation (1), a_j is the item discrimination, b_j is the item difficulty, and c_j is the probability of a correct response by a very low-scoring student. Under the 2PPC model, the probability that a student with trait or scale score θ will respond in category k to partial-credit item j is

$$P_{jk}(\theta) = \exp(z_{jk}) / \sum_{i=1}^{m_j} \exp(z_{ji}), \quad (2)$$

$$\text{where } z_{jk} = (k - 1)f_j - \sum_{i=0}^{k-1} g_{ji}, \text{ and } g_{j0} = 0 \text{ for all } j.$$

The summary output of the 3PL and 2PPC models is in two different metrics. The location and discrimination parameters for the MC items are in the traditional 3PL metric, and are labeled b and a , respectively. In the 2PPC model, f (alpha) and g (gamma) are analogous to b and a , where alpha is the discrimination parameter and gamma over alpha (g/f) is the location where adjacent trace lines cross on the ability scale. Because of the different metrics used, the 3PL parameters b and a are not directly comparable to the 2PPC parameters f and g ; however, they can be converted to a common metric. The two metrics are related by $b = g/f$ and $a = f / 1.7$ (Burket, 1995). As a result of this procedure, the MC and CR items are placed on the same scale. Note that for the 2PPC model, there are $m_j - 1$ (where m_j is a score level j) independent g 's and one f , for a total of m_j independent parameters estimated for each item, while there is one a and one b per item in the 3PL model.

6.3.1. Model Fit

A procedure developed by Yen (1981) was used to assess model-to-data fit for all test items. In this procedure, students are rank ordered on the basis of their $\hat{\theta}$ values and sorted into ten cells with ten percent of the sample in each cell. Each item j in each decile i has a response from N_{ij} examinees. The fitted IRT models are used to calculate an expected proportion E_{ijk} of examinees who respond to item j in category k . The observed proportion O_{ijk} is also tabulated for each decile, and the approximate chi-square statistic

$$Q_{1j} = \sum_{i=1}^{10} \sum_{k=1}^{m_j} \frac{N_{ij} (O_{ijk} - E_{ijk})^2}{E_{ijk}},$$

Q_{1j} should be approximately chi-square distributed with degrees of freedom (DF) equal to the number of "independent" cells, $10(m_j - 1)$, minus the number of estimated parameters. For the 3PL model $m_j = 2$, so $DF = 10(2 - 1) - 3 = 7$. For the 2PPC model,

$DF = 10(m_j - 1) - m_j = 9m_j - 10$. Since DF differs between MC and CR items and between CR items with different score levels m_j , Q_{1j} is transformed, yielding the test statistic

$$Z_j = \frac{Q_{1j} - DF}{\sqrt{2DF}}.$$

This statistic is useful for flagging items that fit relatively poorly. Z_j is sensitive to sample size, and cutoff values for flagging an item based on Z_j have been developed and were used to identify items for the item review. The cutoff value is $(N/1500 \times 4)$ for a given test, where N is the sample size.

Twenty-two MAP operational items were flagged for poor fit. In Communication Arts, one item was flagged for poor fit in Grades 3, 4, and 6, two items were flagged for poor fit in Grade 5, and three items were flagged for poor fit in Grade 8. One item was flagged for poor fit on the Grade 11 regular test. One item was flagged for poor fit on the Grade 11 breach form. In Mathematics, one item was flagged for poor fit in Grades 4, 6, and 7; two items were flagged for poor fit in Grades 8 and 10, and three items were flagged for poor fit in Grade 5. In Science, one item was flagged for poor fit in Grades 5, 8, and 11. Table 6.12 shows the chi-square statistic and the Z -statistic for each flagged item. The average percent across ten cells of observed percentage correct and predicted percentage correct is also provided. The difference between the observed and predicted percentages provides an indication of how well the modeled response curves reflect the empirical curves.

Each of the flagged items was examined more closely by studying its item characteristic curve (ICC) at each non-zero score point. The ICC models the relationship between the examinees' performance on an item and the examinees' underlying ability. In almost all cases for which model misfit occurs, relatively few students occupy these scale score ranges which are at the lower and upper tails of the distribution. Poor fit may occur in one region of the underlying ability distribution when there are relatively few students at that particular point in the distribution. The model tend to show good model-data fit for the flagged items in the middle of the theta distribution where the majority of students perform.

Figures 6.1 to 6.23 show the item characteristic curves for each of the misfitting MAP items. The smooth line in each of these figures represents predicted relationship between examinee performance on the item and examinee ability, and the jagged line represents the observed relationship³. Large differences between the two lines indicate poor fit. Each figure also shows the distribution of theta scores, so that the fit between observed and predicted performance at different ability levels can be interpreted in light of the overall distribution of examinees.

³ For constructed-response items, there will be one graph for each score level. For example, a 2-point item will have three graphs for 0, 1, and 2 score points.

With large numbers of observations such as there are for the Missouri calibration samples, items may be flagged for statistically significant differences; however, these differences may not be of practical importance. In the case of the 22 MAP items flagged for misfit, the differences do not seem to be of practical importance. Misfitting items that have content validity are often retained for use in an assessment and monitored over a period of usage. A large number of misfitting items in an assessment would indicate that caution should be exercised in the interpretation of the overall score. No MAP test had more than three items flagged for misfit.

Figure 6.1 presents the ICC for Session 4, Item 1 (2-point constructed-response item) on the Grade 3 Communication Arts test. As shown, there is poor fit at the lower end of level 1 (students who scored 0 out of 2). Levels 2 and 3 show spikes at the lower and higher ends of the ability distribution for students who scored 1 out of 2 and 2 out of 2, respectively.

Figure 6.2 presents the ICC for Session 2, Item 17 (selected-response item) on the Grade 4 Communication Arts test. This figure shows this is an easy item. There may be somewhat poor fit in the middle of the ability distribution.

Figure 6.3 presents the ICC for Session 1, Item 3 (2- point constructed-response item) on the Grade 5 Communication Arts test. There is poor fit at the lower end of the ability distribution of level 1 and throughout the ability distribution of level 2 and 3.

Figure 6.4 presents the ICC for Session 1, Item 4 (2- point constructed-response item) on the Grade 5 Communication Arts test. There is poor fit at the lower end of the ability distribution of level 1 and throughout the ability distribution of level 2, and at the upper end of the ability distribution for level 3.

Figure 6.5 presents the ICC for Session 1, Item 6A (2- point constructed-response item) on the Grade 6 Communication Arts test. There is poor fit at the lower end of the ability distribution of level 1 and throughout the ability distribution of level 2, and at the upper end of the ability distribution for level 3.

Figure 6.6 presents the ICC for Session 1, Item 4 (2- point constructed-response item) on the Grade 8 Communication Arts test. There is poor fit at the lower end of the ability distribution of level 1 and throughout the ability distribution of level 2, and at the upper end of the ability distribution for level 3.

Figure 6.7 presents the ICC for Session 2, Item 11 (selected-response item) on the Grade 8 Communication Arts test. As shown, there is poor fit throughout the ability distribution for this item.

Figure 6.8 presents the ICC for Session 2, Item 38 (selected-response item) on the Grade 8 Communication Arts test. As shown, there is poor fit through the lower and upper end of the ability distribution for this item.

Figure 6.9 presents the ICC for Session 3, Item 29 (selected-response item) on the Grade 11 Communication Arts test. As shown, there is poor fit throughout the lower end of the ability distribution for this item.

Figure 6.10 presents the ICC for Session 3, Item 33 (selected-response item) on the Grade 11 Communication Arts breach test. As shown, there is poor fit through the lower and upper end of the ability distribution for this item.

Figure 6.11 presents the ICC for Session 3, Item 8 (2- point constructed-response item) on the Grade 4 Mathematics test. There is poor fit at the lower end of the ability distribution for level 1. There is poor fit throughout the ability distribution for level 2. Level 3 shows poor fit at the upper end of the ability distribution.

Figure 6.12 presents the ICC for Session 3, Item 3 (2- point constructed-response item) on the Grade 5 Mathematics test. As shown, there is reasonably good fit throughout the ability distribution for levels 1 and 3. There is poor fit just below the center of the ability distribution for level 2.

Figure 6.13 presents the ICC for Session 3, Item 4 (2- point constructed-response item) on the Grade 5 Mathematics test. There is somewhat poor fit throughout the ability distribution for all levels.

Figure 6.14 presents the ICC for Session 3, Item 5 (2- point constructed-response item) on the Grade 5 Mathematics test. There is poor fit throughout the ability distribution for all levels.

Figure 6.15 presents the ICC for Session 3, Item 7 (2- point constructed-response item) on the Grade 6 Mathematics test. As shown, there is poor fit throughout the ability distribution for levels 1 and 2. There is poor fit at the upper end of the ability distribution for level 3.

Figure 6.16 presents the ICC for Session 3, Item 3 (2- point constructed-response item) on the Grade 7 Mathematics test. As shown, there is poor fit at the lower end of the ability distribution of level 1, throughout level 2, and at the upper end of the distribution of level 3.

Figure 6.17 presents the ICC for Session 2, Item 16 (selected-response item) on the Grade 8 Mathematics test. There is poor fit throughout the ability distribution.

Figure 6.18 presents the ICC for Session 3, Item 5 (2- point constructed-response item) on the Grade 8 Mathematics test. There is poor fit at the throughout the ability distribution for level 1 and 2 and at the upper end of the ability distribution for level 3.

Figure 6.19 presents the ICC for Session 2, Item 9 (selected-response item) on the Grade 10 Mathematics test. There is poor fit throughout the ability distribution.

Figure 6.20 presents the ICC for Session 3, Item 5 (2-point constructed-response item) on the Grade 10 Mathematics test. As shown, there is reasonably good fit throughout the ability distribution for level 1. Levels 2 and 3 shows a poor fit at the upper end of the ability distribution.

Figure 6.21 presents the ICC for Session 3, Item 2 (4-point constructed-response item) on the Grade 5 Science test. As shown, there is reasonably good fit throughout the ability distribution for level 1. Levels 2 through 4 demonstrate poor fit throughout the ability distribution. Level 5 shows poor fit at the upper end of the ability distribution.

Figure 6.22 presents the ICC for Session 2, Item 32 (2-point constructed-response item) on the Grade 8 Science test. All three levels show poor fit throughout the ability distribution.

Figure 6.23 presents the ICC for Session 3, Item 15 (selected-response item) on the Grade 11 Science test. There is poor fit throughout the ability distribution for this item.

6.4 Scaling

The purpose of scaling a test is to enhance its validity by increasing the comparability of test takers' scores. The MAP scores are produced using a three-parameter logistic, two-parameter partial credit (3PL/2PPC) IRT model that assumes that each of the items and tasks is an independent indicator of the underlying ability governing the propensity for students to answer an item correctly (or with greater correctness in the case of the multilevel constructed-response items).

Scaling and linking of complex assessment data were performed using PARDUX (Burket, 1995), which is proprietary software developed by CTB/McGraw-Hill. PARDUX is designed to produce a single scale by jointly analyzing data resulting from students' responses to both multiple-choice (MC) items and constructed-response (CR) items. In PARDUX, items are calibrated based on IRT, using the 3PL model (Lord and Novick, 1968) for MC items and the 2PPC model (Yen, 1993) for CR items. PARDUX is also used to link the scales developed by two calibrations through the common-item procedure developed by Stocking and Lord (1983).

6.4.1. Linking Method and Results

After the initial IRT item calibration, items parameters were linked to the *TerraNova* scale using the Stocking & Lord (1983) equating procedure. This approach takes advantage of the vertical properties of the *TerraNova* scale and allows for comparison of the results from the 2007 administration to results from future administrations.

For the linking, the intact *TerraNova* Survey items served as anchors. Linking was performed using a test characteristic curve (TCC) method proposed by Stocking & Lord (1983), and implemented using PARDUX (Burket, 1995).

Appendix B provides further details on the linking methodology that was used to put the 2008 MAP onto the MAP scale. Appendix B also provides details on the methodology used to evaluate the anchor items. Finally, Appendix B discusses the breach form.

6.4.2. Vertical Scale

The scale on which the MAP scale scores are reported is based in part on a standardized achievement test, which makes it possible to report national percentile scores in addition to the criterion-referenced scale scores of MAP. Although the MAP scale is unique to Missouri, the characteristic growth seen on the scale from grade to grade for the standardized test has been utilized and built upon to give MAP its vertical scale characteristics. The vertical scale is sometimes referred to as a growth scale.

Evidence of the validity of the MAP growth scale is provided by the increase of the scale score at selected percentiles as grade level increases. Figures 6.24, 6.25, and 6.26 display the scale scores for several points on the score distributions for each grade of the Communication Arts, Mathematics, and Science MAP, respectively. These scale scores indicate the growth, or change, in score by grade at the 1st, 10th, 20th, . . . , 90th, and 99th percentiles. In the ideal, the scale score associated with each percentile will increase from grade to grade. Figure 6.24 shows the selected percentiles for the Communication Arts MAP. Considering all but the 1st and 99th percentile, the scale scores progress upward from Grades 3 through 5 and then flatten from Grade 5 through 7 before continuing to progress upward again from Grade 7 through 11.

Figure 6.25 shows the selected percentiles for the Mathematics MAP. Except for the 1st percentile, there is an upward progression of scale scores across all grades with only a slight flattening out between Grades 6 and 7. At the 1st percentile, there is a decrease in scale score between Grades 4 and 5 and a flattening out between Grades 8 and 10.

Figure 6.26 shows the selected percentiles for the Science MAP. Except for the 1st percentile, there is an upward progression of scale scores across the three Science grades. At the 1st percentile, there is a decrease in scale score between Grades 8 and 11.

Figure 6.27 shows the TCCs by grade for the MAP Communication Arts, Figure 6.28 shows the TCCs by grade for the MAP Mathematics, and Figure 6.29 shows the TCCs by grade for the MAP Science. Because these tests were linked to the *TerraNova* scale, they have an underlying vertical scale. By plotting the TCCs together, we can demonstrate that the tests increase in difficulty as the grade levels increase. Figure 6.27 shows that the TCCs for Grades 5, 6, and 7 overlap. The Grades 5 and 6 TCC are very close to each other, separating only in the middle of the TCCs. The Grade 7 TCC crosses the Grade 5 and 6 TCC at the lower end of those TCCs. During the selection of the forms, the pre-equated TCCs were examined and efforts were made to further separate the Grades 5 through 7 TCCs while, at the same time, protecting against scale drift. The available item pool was insufficient to create tests that resulted in the optimal increases in test difficulty. For Grade 7, the mean scale score is higher than Grade 5 and 6. The Grades 5 and 6 mean scale scores were nearly identical. DESE continues to work on differentiating skills in these grades, which may help pull apart the Grades 5 and 6 TCCs.

For both Mathematics (Figure 6.28) and Science (Figure 6.29), the TCCs indicate that test difficulty increases with grade level.

6.4.3. Lowest and Highest Obtainable Scale Scores

A maximum likelihood procedure cannot produce scale score estimates for students with perfect scores or scores below the level expected by guessing. Also, although maximum likelihood estimates are available for students with extreme scores other than zero or perfect, occasionally these estimates have standard errors of measurement that are very large, and differences between these extreme values have little meaning. Therefore, scores are established for these students based on a rational but necessarily non-maximum likelihood procedure. These values, which are set separately by grade, are called the lowest obtainable scale score (LOSS) and the highest obtainable scale score (HOSS). Table 6.17 shows the LOSS and HOSS values used for each grade of the Communication Arts, Mathematics, and Science MAPs.

6.5 Item-Pattern Scoring

Simply put, a scale score can be interpreted as a highly probable estimate of a student's ability in a given content area. These scores are based on the student's responses to all items on a given test, and the scores account for the characteristics of the items that are in the test (such as item difficulty).

The scores on the MAP are derived using item-pattern scoring. There are two IRT-based scoring methods generally used for large scale assessments: number-correct scoring and item-pattern scoring. Item-pattern scoring tends to be preferred over number-correct scoring, especially on shorter tests, because it tends to produce more accurate and reliable scores.

Item-pattern scoring produces more accurate scores for individual students. Specifically, it produces a smaller standard error of measurement (SEM) across the scale score range for a given test compared to number-correct scoring. The smaller the SEM, the more confident one can be in the accuracy of the test results. The increase in accuracy provided by item-pattern scoring is equivalent, on average, to approximately a 15% to 20% increase in test length (Yen, 1984; Yen & Candell, 1991).

Second, reliability tends to be higher using item-pattern scoring, which means a) fewer items are needed to achieve a given level of reliability, and that b) a given test with a given number of items will have higher reliability than when using number-correct scoring. Yen (1984) has demonstrated that an equivalent level of reliability for a 20-item test scored by the number-correct scoring method could be obtained with a 16- or 17-item test scored by the item-pattern scoring method.

The procedures applied here are similar to those followed in the development of the *TerraNova* test (CTB/McGraw-Hill, 1997), *TerraNova 2nd* Edition (CTB/McGraw-Hill,

2000). For additional information on the technical details of the item-pattern scoring, readers can also refer to Yen & Candell (1991) and to *TerraNova 2nd Edition* (CTB/McGraw-Hill, 2000).

Table 6. 1: Large Districts that were Included in the 80% Calibration Sample

District Name
Columbia
St Joseph
North Kansas
Springfield
Blue Springs
Lee's Summit
Kansas City
Fort Zumwalt
Francis Howell
Hazelwood
Ferguson Florissant
Rockwood
Mehlville
Parkway
St. Louis City

Table 6. 2: Summary of Calibration and Census Data: Communication Arts

	Communication Arts, Grade 3				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	54794		66179		
Gender					
Male	28069	51.23	33879	51.19	0.04
Female	26621	48.58	32180	48.63	-0.05
Unknown	104	0.19	120	0.18	0.01
Race/Ethnicity					
White	40791	74.44	49942	75.47	-1.03
Black	10318	18.83	11935	18.03	0.80
Hispanic	2284	4.17	2690	4.06	0.11
Asian/Pacific Islander Native	1096	2.00	1231	1.86	0.14
American/Alaskan	200	0.37	258	0.39	-0.02
Unknown	105	0.19	123	0.19	0.00
	Communication Arts, Grade 4				
All Students	56866		66873		
Gender					
Male	29172	51.30	34297	51.29	0.01
Female	27616	48.56	32476	48.56	0.00
Unknown	78	0.14	100	0.15	-0.01
Race/Ethnicity					
White	42450	74.65	50497	75.51	-0.86
Black	10677	18.78	12164	18.19	0.59
Hispanic	2377	4.18	2659	3.98	0.20
Asian/Pacific Islander Native	1072	1.89	1187	1.78	0.11
American/Alaskan	208	0.37	262	0.39	-0.02
Unknown	82	0.14	104	0.16	-0.02
	Communication Arts, Grade 5				
All Students	52300		65544		
Gender					
Male	26590	50.84	33344	50.87	-0.03
Female	25627	49.00	32099	48.97	0.03
Unknown	83	0.16	101	0.15	0.01
Race/Ethnicity					
White	39043	74.65	49831	76.03	-1.38
Black	9849	18.83	11699	17.85	0.98
Hispanic	2069	3.96	2444	3.73	0.23
Asian/Pacific Islander Native	1034	1.98	1175	1.79	0.19
American/Alaskan	209	0.40	280	0.43	-0.03
Unknown	96	0.18	115	0.18	0.00

Table 6.2: Summary of Calibration and Census Data: Communication Arts (Cont'd)

	Communication Arts, Grade 6				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	53590		65672		
Gender					
Male	27178	50.71	33419	50.89	-0.18
Female	26343	49.16	32176	49.00	0.16
Unknown	69	0.13	77	0.12	0.01
Race/Ethnicity					
White	40463	75.50	50403	76.75	-1.25
Black	9871	18.42	11415	17.38	1.04
Hispanic	1903	3.55	2235	3.40	0.15
Asian/Pacific Islander Native	1041	1.94	1248	1.90	0.04
American/Alaskan	238	0.44	289	0.44	0.00
Unknown	74	0.14	82	0.12	0.02
	Communication Arts, Grade 7				
All Students	58764		66701		
Gender					
Male	30252	51.48	34334	51.47	0.01
Female	28425	48.37	32266	48.37	0.00
Unknown	87	0.15	101	0.15	0.00
Race/Ethnicity					
White	44644	75.97	50853	76.24	-0.27
Black	10553	17.96	11873	17.80	0.16
Hispanic	2092	3.56	2303	3.45	0.11
Asian/Pacific Islander Native	1141	1.94	1298	1.95	-0.01
American/Alaskan	262	0.45	291	0.44	0.01
Unknown	72	0.12	83	0.12	0.00
	Communication Arts, Grade 8				
All Students	56312		67278		
Gender					
Male	28767	51.09	34384	51.11	-0.02
Female	27449	48.74	32766	48.70	0.04
Unknown	96	0.17	128	0.19	-0.02
Race/Ethnicity					
White	42612	75.67	51293	76.24	-0.57
Black	10598	18.82	12200	18.13	0.69
Hispanic	1801	3.20	2149	3.19	0.01
Asian/Pacific Islander Native	919	1.63	1175	1.75	-0.12
American/Alaskan	283	0.50	330	0.49	0.01
Unknown	99	0.18	131	0.19	-0.01

Table 6. 2: Summary of Calibration and Census Data: Communication Arts (Cont'd)

	Communication Arts, Grade 11				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	60838		61041		
Gender					
Male	30257	49.73	30370	49.75	-0.02
Female	30449	50.05	30536	50.03	0.02
Unknown	132	0.22	135	0.22	0.00
Race/Ethnicity					
White	49119	80.74	49255	80.69	0.05
Black	8752	14.39	8810	14.43	-0.04
Hispanic	1416	2.33	1420	2.33	0.00
Asian/Pacific Islander Native	1115	1.83	1117	1.83	0.00
American/Alaskan	274	0.45	274	0.45	0.00
Unknown	162	0.27	165	0.27	0.00
	Communication Arts, Grade 11 Breach				
All Students	1267		1299		
Gender					
Male	572	45.15	586	45.11	0.04
Female	686	54.14	702	54.04	0.10
Unknown	9	0.71	11	0.85	-0.14
Race/Ethnicity					
White	270	21.31	273	21.02	0.29
Black	831	65.59	853	65.67	-0.08
Hispanic	128	10.10	134	10.32	-0.22
Asian/Pacific Islander Native	30	2.37	31	2.39	-0.02
American/Alaskan	1	0.08	1	0.08	0.00
Unknown	7	0.55	7	0.54	0.01

Table 6. 3: Summary of Calibration and Census Data: Mathematics

	Mathematics, Grade 3				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	54907		66258		
Gender					
Male	28123	51.22	33911	51.18	0.04
Female	26683	48.60	32230	48.64	-0.04
Unknown	101	0.18	117	0.18	0.00
Race/Ethnicity					
White	40817	74.34	49946	75.38	-1.04
Black	10333	18.82	11938	18.02	0.80
Hispanic	2326	4.24	2731	4.12	0.12
Asian/Pacific Islander Native	1128	2.05	1264	1.91	0.14
American/Alaskan	200	0.36	258	0.39	-0.03
Unknown	103	0.19	121	0.18	0.01
	Mathematics, Grade 4				
All Students	56976		66944		
Gender					
Male	29233	51.31	34333	51.29	0.02
Female	27664	48.55	32510	48.56	-0.01
Unknown	79	0.14	101	0.15	-0.01
Race/Ethnicity					
White	42475	74.55	50498	75.43	-0.88
Black	10698	18.78	12172	18.18	0.60
Hispanic	2414	4.24	2697	4.03	0.21
Asian/Pacific Islander Native	1097	1.93	1211	1.81	0.12
American/Alaskan	209	0.37	261	0.39	-0.02
Unknown	83	0.15	105	0.16	-0.01
	Mathematics, Grade 5				
All Students	52429		65636		
Gender					
Male	26670	50.87	33396	50.88	-0.01
Female	25676	48.97	32140	48.97	0.00
Unknown	83	0.16	100	0.15	0.01
Race/Ethnicity					
White	39071	74.52	49841	75.94	-1.42
Black	9884	18.85	11707	17.84	1.01
Hispanic	2096	4.00	2474	3.77	0.23
Asian/Pacific Islander Native	1071	2.04	1218	1.86	0.18
American/Alaskan	210	0.40	281	0.43	-0.03
Unknown	114	0.17	115	0.18	-0.01

Table 6.3: Summary of Calibration and Census Data: Mathematics (Cont'd)

	Mathematics, Grade 6				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	53666		65716		
Gender					
Male	27221	50.72	33445	50.89	-0.17
Female	26375	49.15	32193	48.99	0.16
Unknown	70	0.13	78	0.12	0.01
Race/Ethnicity					
White	40477	75.42	50392	76.68	-1.26
Black	9882	18.41	11417	17.37	1.04
Hispanic	1927	3.59	2260	3.44	0.15
Asian/Pacific Islander Native	1070	1.99	1278	1.94	0.05
American/Alaskan	237	0.44	288	0.44	0.00
Unknown	73	0.14	81	0.12	0.02
	Mathematics, Grade 7				
All Students	58830		66727		
Gender					
Male	30289	51.49	34340	51.46	0.03
Female	28455	48.37	32287	48.39	-0.02
Unknown	86	0.15	100	0.15	0.00
Race/Ethnicity					
White	44651	75.90	50834	76.18	-0.28
Black	10555	17.94	11859	17.77	0.17
Hispanic	2125	3.61	2336	3.50	0.11
Asian/Pacific Islander Native	1167	1.98	1325	1.99	-0.01
American/Alaskan	261	0.44	291	0.44	0.00
Unknown	71	0.12	82	0.12	0.00
	Mathematics, Grade 8				
All Students	56387		67312		
Gender					
Male	28807	51.09	34398	51.10	-0.01
Female	27484	48.74	32789	48.71	0.03
Unknown	96	0.17	125	0.19	-0.02
Race/Ethnicity					
White	42635	75.61	51289	76.20	-0.59
Black	10605	18.81	12198	18.12	0.69
Hispanic	1829	3.24	2175	3.23	0.01
Asian/Pacific Islander Native	936	1.66	1192	1.77	-0.11
American/Alaskan	284	0.50	331	0.49	0.01
Unknown	98	0.17	127	0.19	-0.02

Table 6. 3: Summary of Calibration and Census Data: Mathematics (Cont'd)

	Mathematics, Grade 10				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	66374		68776		
Gender					
Male	33481	50.44	34672	50.41	0.03
Female	32691	49.25	33900	49.29	-0.04
Unknown	202	0.30	204	0.30	0.00
Race/Ethnicity					
White	51729	77.94	53731	78.12	-0.18
Black	11115	16.75	11329	16.47	0.28
Hispanic	1914	2.88	1950	2.84	0.04
Asian/Pacific Islander	1088	1.64	1223	1.78	-0.14
Native					
American/Alaskan	318	0.48	330	0.48	0.00
Unknown	210	0.32	213	0.31	0.01

Table 6. 4: Summary of Calibration and Census Data: Science

	Science, Grade 5				
	Calibration Sample		Census Data		Diff (Calib % - Census %)
	N	%	N	%	
All Students	52385		65586		
Gender					
Male	26649	50.87	33371	50.88	-0.01
Female	25655	48.97	32116	48.97	0.00
Unknown	81	0.15	99	0.15	0.00
Race/Ethnicity					
White	39046	74.54	49809	75.94	-1.40
Black	9870	18.84	11692	17.83	1.01
Hispanic	2094	4.00	2473	3.77	0.23
Asian/Pacific Islander Native	1070	2.04	1217	1.86	0.18
American/Alaskan	210	0.40	281	0.43	-0.03
Unknown	95	0.18	114	0.17	0.01
	Science, Grade 8				
All Students	56291		67209		
Gender					
Male	28754	51.08	34335	51.09	-0.01
Female	27443	48.75	32750	48.73	0.02
Unknown	94	0.17	124	0.18	-0.01
Race/Ethnicity					
White	42591	75.66	51238	76.24	-0.58
Black	10565	18.77	12155	18.09	0.68
Hispanic	1820	3.23	2167	3.22	0.01
Asian/Pacific Islander Native	938	1.67	1194	1.78	-0.11
American/Alaskan	282	0.50	329	0.49	0.01
Unknown	95	0.17	126	0.19	-0.02
	Science, Grade 11				
All Students	50645		62133		
Gender					
Male	25085	49.53	30834	49.63	-0.10
Female	25424	50.20	31143	50.12	0.08
Unknown	136	0.27	156	0.25	0.02
Race/Ethnicity					
White	39992	78.97	49362	79.45	-0.48
Black	8154	16.10	9634	15.51	0.59
Hispanic	1280	2.53	1553	2.50	0.03
Asian/Pacific Islander Native	874	1.73	1155	1.86	-0.13
American/Alaskan	203	0.40	268	0.43	-0.03
Unknown	142	0.28	161	0.26	0.02

Table 6. 5: MAP Means, Standard Deviations for Raw Scores, P-Values, Item-Total Correlation (R_{it}): Communication Arts 2008

Grade	Total Items	Total Points	Mean Raw Score (SD)	Mean P-Value (SD)	Mean R_{it} (SD)
3	58	68	48.09 (11.49)	.71 (.15)	.42 (.10)
4	56	63	44.92 (10.36)	.74 (.17)	.43 (.08)
5	56	63	43.43 (10.94)	.70 (.15)	.42 (.09)
6	55	62	41.65 (10.80)	.69 (.16)	.41 (.08)
7	62	73	48.77 (12.31)	.69 (.17)	.42 (.09)
8	61	69	45.98 (11.62)	.68 (.20)	.41 (.10)
11	63	73	47.45 (12.77)	.62 (.18)	.39 (.11)
11b	63	73	38.67 (13.68)		

Table 6. 6: MAP Means, Standard Deviations for Raw Scores, P-Values, Item-Total Correlation (R_{it}): Mathematics 2008

Grade	Total Items	Total Points	Mean Raw Score (SD)	Mean P-Value (SD)	Mean R_{it} (SD)
3	60	67	49.10 (11.75)	.75 (.14)	.42 (.10)
4	65	77	53.91 (13.26)	.72 (.16)	.41 (.10)
5	62	69	47.51 (11.93)	.70 (.15)	.40 (.10)
6	61	68	44.71 (12.22)	.67 (.13)	.41 (.09)
7	62	69	40.83 (13.05)	.61 (.17)	.42 (.10)
8	64	76	41.72 (14.65)	.57 (.19)	.42 (.12)
10	58	70	39.53 (14.73)	.60 (.17)	.47 (.10)

Table 6. 7: MAP Means, Standard Deviations for Raw Scores, P-Values, Item-Total Correlation (R_{it}): Science 2008

Grade	Total Items	Total Points	Mean Raw Score (SD)	Mean P-Value (SD)	Mean R_{it} (SD)
5	53	79	42.33 (13.37)	.57 (.23)	.42 (.10)
8	58	82	38.79 (14.94)	.51 (.25)	.45 (.12)
11	58	88	40.02 (16.65)	.47 (.17)	.45 (.10)

Table 6. 8: Item Statistics: Grade 3

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
1	1	0.70	0.30	0.08%	1	1	0.80	0.45	0.68%
1	2	0.84	0.52	0.11%	1	2	0.79	0.41	0.18%
1	3	0.73	0.60	0.41%	1	3	0.77	0.54	0.48%
1	4	0.64	0.34	0.43%	1	4	0.83	0.32	0.24%
1	5	0.69	0.56	0.36%	1	5	0.48	0.27	0.36%
1	6A	0.70	0.54	0.53%	1	6	0.83	0.51	0.22%
1	6B	0.97	0.30	0.56%	1	7	0.61	0.55	0.36%
1	6C	0.98	0.25	0.56%	1	8	0.88	0.39	0.24%
1	7	0.58	0.39	0.34%	1	9	0.79	0.44	0.39%
1	8	0.73	0.38	0.37%	1	10	0.79	0.47	0.46%
1	9	0.54	0.31	0.45%	1	11	0.68	0.60	0.38%
1	10	0.58	0.37	0.37%	1	12	0.65	0.36	0.35%
1	11	0.73	0.32	0.37%	1	13	0.86	0.43	0.16%
1	12	0.42	0.15	0.48%	1	14	0.78	0.50	0.47%
2	1	0.72	0.56	0.20%	1	15	0.55	0.40	0.20%
3	1	0.94	0.41	0.10%	1	16	0.78	0.26	0.20%
3	2	0.96	0.37	0.17%	1	17	0.77	0.45	0.53%
3	3	0.76	0.43	0.25%	1	18	0.79	0.55	0.34%
3	4	0.87	0.33	0.32%	1	19	0.69	0.40	1.17%
3	5	0.96	0.34	0.17%	1	20	0.45	0.34	0.25%
3	6	0.97	0.33	0.28%	1	21	0.83	0.48	0.21%
3	7	0.87	0.44	0.28%	1	22	0.73	0.35	0.27%
3	8	0.79	0.43	1.07%	1	23	0.87	0.37	0.31%
3	9	0.62	0.31	1.94%	2	1	0.92	0.29	0.16%
3	10	0.90	0.47	0.31%	2	2	0.91	0.31	0.26%
3	11	0.71	0.33	0.38%	2	3	0.76	0.36	0.42%
3	12	0.69	0.32	0.58%	2	4	0.61	0.47	1.96%
3	13	0.50	0.38	0.78%	2	5	0.74	0.50	1.95%
3	14	0.39	0.37	2.03%	2	6	0.93	0.33	0.69%
3	15	0.79	0.53	0.75%	2	7	0.74	0.37	1.70%
3	16	0.59	0.34	1.13%	2	8	0.51	0.39	2.68%
3	17	0.76	0.35	1.49%	2	9	0.91	0.38	0.42%
3	18	0.79	0.51	2.48%	2	10	0.87	0.33	0.13%

Table 6.8: Item Statistics: Grade 3 (Cont'd)

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
3	19	0.78	0.47	2.82%	2	11	0.99	0.19	0.22%
3	20	0.83	0.53	3.56%	2	12	0.93	0.35	0.71%
3	21	0.89	0.38	4.19%	2	13	0.91	0.31	0.40%
3	22	0.84	0.48	0.25%	2	14	0.93	0.32	0.26%
3	23	0.69	0.52	0.63%	2	15	0.71	0.51	0.40%
3	24	0.81	0.30	0.68%	2	16	0.93	0.40	0.80%
3	25	0.54	0.37	0.67%	2	17	0.95	0.29	0.37%
3	26	0.81	0.49	1.05%	2	18	0.87	0.39	0.42%
3	27	0.38	0.28	1.62%	2	19	0.74	0.35	1.23%
3	28	0.53	0.44	2.25%	2	20	0.96	0.32	0.23%
3	29	0.65	0.34	0.19%	2	21	0.61	0.46	1.01%
3	30	0.53	0.51	0.21%	2	22	0.74	0.30	5.37%
3	31	0.61	0.42	0.34%	2	23	0.86	0.40	5.27%
3	32	0.79	0.49	0.59%	2	24	0.88	0.45	0.82%
3	33	0.65	0.48	0.93%	2	25	0.80	0.42	0.62%
3	34	0.78	0.54	0.93%	2	26	0.77	0.27	0.30%
3	35	0.82	0.37	0.87%	2	27	0.78	0.35	0.28%
3	36	0.82	0.37	0.37%	2	28	0.53	0.52	0.46%
3	37	0.70	0.44	0.54%	2	29	0.71	0.50	0.50%
3	38	0.65	0.42	0.67%	2	30	0.60	0.48	0.87%
3	39	0.49	0.35	1.04%	3	1	0.48	0.63	0.24%
4	1	0.58	0.39	0.46%	3	2	0.68	0.53	0.19%
4	2	0.75	0.64	0.58%	3	3	0.54	0.60	0.37%
4	3A	0.62	0.61	0.67%	3	4	0.65	0.49	0.25%
4	3B	0.71	0.37	0.95%	3	5	0.42	0.46	0.60%
					3	6	0.61	0.64	0.40%
					3	7	0.78	0.51	0.32%

Table 6. 9: Item Statistics: Grade 4

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
1	1	0.78	0.49	0.08%	1	1	0.74	0.20	0.20%
1	2	0.54	0.42	0.10%	1	2	0.74	0.30	0.18%
1	3	0.42	0.47	0.43%	1	3	0.74	0.46	0.16%
1	4	0.28	0.42	0.26%	1	4	0.81	0.28	0.22%
1	5A	0.75	0.46	0.43%	1	5	0.69	0.33	0.33%
1	5B	0.94	0.38	0.43%	1	6	0.58	0.34	0.21%
1	5C	0.93	0.38	0.43%	1	7	0.82	0.52	0.28%
1	6	0.50	0.51	0.66%	1	8	0.66	0.50	0.24%
1	7	0.73	0.28	0.22%	1	9	0.32	0.28	0.29%
1	8	0.46	0.24	0.27%	1	10	0.79	0.44	0.41%
1	9	0.58	0.35	0.47%	1	11	0.58	0.47	0.23%
1	10	0.89	0.38	0.45%	1	12	0.71	0.49	0.24%
1	11	0.82	0.29	0.50%	1	13	0.79	0.35	0.22%
1	12	0.89	0.38	0.65%	1	14	0.76	0.44	0.31%
2	1	0.83	0.40	0.10%	1	15	0.39	0.38	0.35%
2	2	0.98	0.24	0.10%	1	16	0.50	0.28	0.27%
2	3	0.87	0.42	0.18%	1	17	0.71	0.33	0.28%
2	4	0.89	0.44	0.52%	1	18	0.77	0.38	0.70%
2	5	0.85	0.47	0.22%	1	19	0.86	0.41	0.25%
2	6	0.80	0.38	0.52%	1	20	0.95	0.35	0.27%
2	7	0.96	0.36	0.43%	1	21	0.68	0.38	0.30%
2	8	0.73	0.47	0.30%	1	22	0.41	0.35	0.30%
2	9	0.49	0.47	0.44%	1	23	0.79	0.27	0.27%
2	10	0.76	0.43	0.35%	1	31	0.35	0.60	0.38%
2	11	0.51	0.26	0.33%	2	1	0.90	0.24	0.17%
2	12	0.81	0.47	0.53%	2	2	0.76	0.47	0.32%
2	13	0.42	0.31	0.44%	2	3	0.71	0.40	0.97%
2	14	0.82	0.41	0.75%	2	4	0.67	0.40	1.43%
2	15	0.83	0.48	0.96%	2	5	0.86	0.48	0.49%
2	16	0.86	0.42	0.75%	2	6	0.61	0.57	1.08%
2	17	0.93	0.41	0.95%	2	7	0.49	0.44	2.04%
2	18	0.88	0.53	1.69%	2	8	0.87	0.43	2.66%
2	19	0.94	0.44	1.13%	2	9	0.91	0.39	3.50%

Table 6.9: Item Statistics: Grade 4 (Cont'd)

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
2	20	0.63	0.44	1.51%	2	10	0.71	0.52	4.63%
2	21	0.87	0.40	1.49%	2	11	0.70	0.34	5.99%
2	22	0.82	0.52	0.21%	2	12	0.72	0.35	7.36%
2	23	0.86	0.42	0.23%	2	13	0.93	0.33	0.15%
2	24	0.87	0.47	0.44%	2	14	0.86	0.25	0.26%
2	25	0.84	0.46	0.32%	2	15	0.69	0.43	2.20%
2	26	0.64	0.44	0.56%	2	16	0.92	0.17	0.31%
2	27	0.70	0.47	0.84%	2	17	0.89	0.40	0.47%
2	28	0.84	0.39	0.43%	2	18	0.50	0.50	0.43%
2	29	0.86	0.44	0.40%	2	19	0.67	0.50	0.34%
2	30	0.89	0.42	0.47%	2	20	0.67	0.54	0.43%
2	31	0.78	0.60	1.64%	2	21	0.78	0.37	0.58%
2	32	0.84	0.33	3.24%	2	22	0.84	0.41	0.36%
2	33	0.54	0.25	0.93%	2	23	0.72	0.51	1.00%
2	34	0.82	0.53	1.05%	2	24	0.93	0.23	0.49%
2	35	0.46	0.36	1.34%	2	25	0.83	0.45	0.54%
2	36	0.82	0.40	1.51%	2	26	0.95	0.25	0.35%
2	37	0.72	0.56	1.71%	2	27	0.84	0.47	0.33%
2	38	0.71	0.51	1.94%	2	28	0.82	0.46	0.46%
2	39	0.72	0.51	2.05%	2	29	0.86	0.54	0.37%
3	1	0.62	0.57	0.78%	2	30	0.98	0.23	0.31%
3	2	0.66	0.39	0.57%	2	31	0.95	0.35	0.32%
3	3	0.43	0.39	0.48%	2	32	0.63	0.36	0.34%
					3	1	0.78	0.47	0.28%
					3	2	0.57	0.57	0.35%
					3	3	0.81	0.51	0.25%
					3	4	0.37	0.50	0.80%
					3	5	0.84	0.53	0.22%
					3	6	0.84	0.55	0.34%
					3	7	0.51	0.54	0.40%
					3	8	0.70	0.52	0.36%
					3	9	0.56	0.48	0.62%

Table 6. 10: Item Statistics: Grade 5

Communication Arts					Mathematics					Science				
Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate
1	1	0.46	0.29	0.12%	1	1	0.66	0.24	0.19%	1	1	0.70	0.44	0.24%
1	2	0.40	0.21	0.13%	1	2	0.59	0.40	0.13%	1	2	0.74	0.36	0.81%
1	3	0.52	0.30	0.37%	1	3	0.59	0.27	0.22%	1	3	0.52	0.50	0.58%
1	4	0.49	0.49	0.39%	1	4	0.80	0.32	0.10%	1	4	0.50	0.49	0.34%
1	5A	0.77	0.58	0.49%	1	5	0.51	0.23	0.30%	1	5	0.39	0.45	1.10%
1	5B	0.87	0.45	0.49%	1	6	0.71	0.32	0.23%	1	6	0.49	0.52	0.68%
1	6A	0.65	0.59	0.55%	1	7	0.82	0.40	0.15%	1	7	0.25	0.50	1.08%
1	6B	0.89	0.27	0.85%	1	8	0.86	0.27	0.28%	1	8	0.34	0.54	0.31%
1	7	0.66	0.35	0.23%	1	9	0.53	0.35	0.22%	1	9	0.51	0.50	0.80%
1	8	0.67	0.32	0.20%	1	10	0.67	0.49	0.27%	1	10	0.25	0.35	1.59%
1	9	0.61	0.27	0.30%	1	11	0.40	0.33	0.31%	1	11	0.32	0.43	1.08%
1	10	0.72	0.39	0.46%	1	12	0.47	0.27	0.29%	2	1	0.91	0.21	0.13%
1	11	0.83	0.37	0.50%	1	13	0.40	0.27	0.33%	2	2	0.93	0.29	0.17%
1	12	0.50	0.34	0.54%	1	14	0.85	0.37	0.23%	2	3	0.92	0.26	0.15%
2	1	0.91	0.32	0.22%	1	15	0.61	0.40	0.29%	2	4	0.89	0.24	0.18%
2	2	0.83	0.45	0.30%	1	16	0.72	0.27	0.41%	2	6	0.88	0.35	0.41%
2	3	0.62	0.44	0.27%	1	17	0.71	0.38	0.15%	2	7	0.91	0.30	0.45%
2	4	0.87	0.36	0.23%	1	18	0.82	0.36	0.31%	2	8	0.94	0.30	0.30%
2	5	0.64	0.40	0.45%	1	19	0.57	0.36	0.41%	2	9	0.96	0.24	0.38%
2	6	0.83	0.44	0.73%	1	20	0.68	0.48	0.27%	2	10	0.88	0.39	1.75%
2	7	0.78	0.40	1.17%	1	21	0.59	0.46	0.24%	2	11	0.70	0.20	4.17%
2	8	0.77	0.46	0.46%	1	22	0.77	0.39	0.19%	2	12	0.67	0.19	0.29%
2	9	0.70	0.49	0.59%	1	23	0.47	0.34	0.24%	2	13	0.55	0.27	0.27%
2	10	0.88	0.41	1.58%	2	1	0.75	0.36	0.27%	2	14	0.84	0.44	0.31%
2	11	0.79	0.29	2.66%	2	2	0.76	0.32	0.37%	2	15	0.72	0.47	0.44%
2	12	0.89	0.47	0.61%	2	3	0.77	0.39	2.99%	2	16	0.57	0.53	0.80%
2	13	0.60	0.24	0.85%	2	4	0.86	0.50	0.28%	2	17	0.67	0.44	1.27%
2	14	0.74	0.35	1.96%	2	5	0.73	0.45	0.54%	2	20	0.49	0.40	0.27%
2	15	0.63	0.39	1.17%	2	6	0.72	0.54	0.92%	2	21	0.63	0.42	0.48%
2	16	0.84	0.51	2.71%	2	7	0.85	0.39	2.57%	2	22	0.52	0.40	1.47%
2	17	0.93	0.38	1.41%	2	8	0.81	0.38	2.38%	2	23	0.50	0.51	0.37%
2	18	0.75	0.49	1.87%	2	9	0.57	0.35	3.41%	2	24	0.33	0.33	0.62%
2	19	0.69	0.51	4.78%	2	10	0.91	0.38	0.16%	2	25	0.42	0.50	0.34%

Table 6.10: Item Statistics: Grade 5 (Cont'd)

Communication Arts					Mathematics					Science				
Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate
2	20	0.77	0.30	2.33%	2	11	0.84	0.55	0.34%	2	26	0.86	0.36	0.18%
2	21	0.80	0.44	2.64%	2	12	0.75	0.54	0.66%	2	27	0.53	0.47	2.15%
2	22	0.95	0.42	0.17%	2	13	0.88	0.45	1.13%	2	28	0.56	0.52	0.35%
2	23	0.80	0.55	0.21%	2	14	0.71	0.46	1.46%	2	29	0.52	0.58	0.55%
2	24	0.70	0.25	0.34%	2	15	0.53	0.45	2.02%	2	30	0.53	0.42	0.77%
2	25	0.70	0.49	0.28%	2	16	0.69	0.49	0.17%	2	31	0.46	0.51	1.37%
2	26	0.52	0.47	0.60%	2	17	0.95	0.30	0.17%	2	32	0.40	0.53	0.45%
2	27	0.46	0.34	0.94%	2	18	0.84	0.41	0.24%	2	33	0.43	0.51	1.38%
2	28	0.63	0.51	0.36%	2	19	0.82	0.38	0.35%	2	34	0.49	0.51	0.62%
2	29	0.66	0.49	0.39%	2	20	0.83	0.42	0.62%	2	35	0.32	0.49	1.67%
2	30	0.85	0.47	0.31%	2	21	0.94	0.20	0.70%	2	36	0.17	0.36	0.97%
2	31	0.77	0.50	0.33%	2	22	0.81	0.45	0.47%	3	1	0.23	0.51	2.41%
2	32	0.77	0.45	0.53%	2	23	0.82	0.43	0.44%	3	2	0.48	0.46	1.65%
2	33	0.93	0.44	0.37%	2	24	0.60	0.48	0.48%	3	3	0.38	0.49	1.94%
2	34	0.77	0.59	0.63%	2	25	0.83	0.30	0.57%	3	4	0.73	0.55	0.80%
2	35	0.61	0.37	0.66%	2	26	0.75	0.34	0.39%	3	5	0.50	0.45	0.86%
2	36	0.44	0.34	0.56%	2	27	0.90	0.35	0.48%	3	6	0.92	0.24	0.44%
2	37	0.80	0.55	0.93%	2	28	0.62	0.47	1.50%	3	7	0.24	0.40	1.15%
2	38	0.73	0.46	1.08%	2	29	0.57	0.49	0.50%	3	8	0.30	0.42	2.38%
2	39	0.37	0.27	1.18%	2	30	0.59	0.48	0.70%	3	9	0.34	0.44	0.74%
3	1	0.73	0.45	0.30%	2	31	0.55	0.34	0.73%					
3	2	0.53	0.46	0.59%	2	32	0.79	0.38	0.49%					
3	3	0.56	0.47	1.07%	3	1	0.61	0.54	0.45%					
					3	2	0.75	0.50	0.31%					
					3	3	0.74	0.65	0.50%					
					3	4	0.71	0.62	0.47%					
					3	5	0.71	0.31	0.26%					
					3	6	0.27	0.51	0.40%					
					3	7	0.53	0.48	0.61%					

Table 6. 11: Item Statistics: Grade 6

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
1	1	0.81	0.41	0.10%	1	1	0.72	0.45	0.19%
1	2	0.74	0.46	0.14%	1	2	0.78	0.38	0.16%
1	3	0.69	0.44	0.58%	1	3	0.78	0.39	0.23%
1	4	0.66	0.53	0.41%	1	4	0.79	0.28	0.30%
1	5	0.36	0.52	0.85%	1	5	0.68	0.43	0.29%
1	6A	0.59	0.52	0.70%	1	6	0.34	0.35	0.27%
1	6B	0.75	0.34	0.73%	1	7	0.75	0.22	0.21%
1	7	0.76	0.39	0.21%	1	8	0.57	0.42	0.37%
1	8	0.68	0.37	0.92%	1	9	0.40	0.19	0.41%
1	9	0.77	0.26	0.25%	1	10	0.61	0.22	0.22%
1	10	0.71	0.11	0.20%	1	11	0.62	0.28	0.31%
1	11	0.67	0.41	0.25%	1	12	0.75	0.41	0.44%
1	12	0.74	0.39	0.29%	1	13	0.65	0.44	0.32%
2	1	0.75	0.44	0.21%	1	14	0.46	0.38	0.50%
2	2	0.83	0.36	0.25%	1	15	0.56	0.42	0.34%
2	3	0.93	0.46	0.58%	1	16	0.63	0.48	0.80%
2	4	0.92	0.37	0.92%	1	17	0.88	0.44	0.38%
2	5	0.88	0.41	1.76%	1	18	0.65	0.25	0.44%
2	6	0.70	0.41	0.26%	1	19	0.66	0.39	0.39%
2	7	0.78	0.39	0.29%	1	20	0.51	0.31	0.40%
2	8	0.87	0.34	0.27%	1	21	0.76	0.47	0.35%
2	9	0.64	0.41	0.45%	1	22	0.78	0.40	0.31%
2	10	0.42	0.37	0.69%	1	23	0.67	0.39	0.47%
2	11	0.92	0.45	0.61%	2	1	0.77	0.26	0.18%
2	12	0.93	0.35	1.02%	2	2	0.77	0.33	0.33%
2	13	0.68	0.47	0.57%	2	3	0.75	0.28	1.90%
2	14	0.91	0.46	0.63%	2	4	0.76	0.45	0.39%
2	15	0.86	0.48	0.83%	2	5	0.84	0.42	0.69%
2	16	0.68	0.44	1.13%	2	6	0.72	0.44	1.15%
2	17	0.67	0.29	1.39%	2	7	0.60	0.42	1.65%
2	18	0.81	0.37	1.66%	2	8	0.84	0.42	2.21%
2	19	0.63	0.36	1.68%	2	9	0.94	0.29	0.17%
2	20	0.52	0.34	1.87%	2	10	0.85	0.35	0.35%

Table 6.11: Item Statistics: Grade 6 (Cont'd)

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
2	21	0.70	0.48	0.38%	2	11	0.59	0.35	3.46%
2	22	0.63	0.37	0.69%	2	12	0.69	0.45	0.24%
2	23	0.71	0.47	0.40%	2	13	0.75	0.47	0.41%
2	24	0.56	0.31	0.29%	2	14	0.82	0.34	0.47%
2	25	0.63	0.47	0.35%	2	15	0.79	0.52	0.41%
2	26	0.66	0.49	0.59%	2	16	0.77	0.53	0.49%
2	27	0.75	0.32	0.64%	2	17	0.72	0.37	0.43%
2	28	0.75	0.44	0.52%	2	18	0.83	0.49	0.35%
2	29	0.80	0.26	2.03%	2	19	0.76	0.45	0.65%
2	30	0.70	0.43	0.51%	2	20	0.78	0.53	0.76%
2	31	0.43	0.34	0.48%	2	21	0.51	0.42	0.79%
2	32	0.92	0.44	0.40%	2	22	0.70	0.45	1.14%
2	33	0.72	0.53	0.46%	2	23	0.62	0.42	0.95%
2	34	0.85	0.44	0.66%	2	24	0.77	0.59	0.46%
2	35	0.82	0.48	0.80%	2	25	0.66	0.45	0.61%
2	36	0.46	0.29	1.13%	2	26	0.66	0.49	0.80%
2	37	0.50	0.32	1.36%	2	27	0.76	0.50	0.69%
2	38	0.77	0.54	1.43%	2	28	0.63	0.39	0.81%
2	39	0.45	0.33	1.82%	2	29	0.57	0.48	1.05%
3	1	0.62	0.53	0.31%	2	30	0.65	0.30	1.02%
3	2	0.42	0.54	0.82%	2	31	0.62	0.38	1.17%
3	3	0.23	0.34	0.67%	3	1	0.47	0.39	0.41%
					3	2	0.60	0.59	0.91%
					3	3	0.54	0.46	1.69%
					3	4	0.31	0.53	1.89%
					3	5	0.81	0.50	0.52%
					3	6	0.44	0.56	0.48%
					3	7	0.59	0.45	0.57%

Table 6. 12: Item Statistics: Grade 7

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
1	1	0.66	0.22	0.13%	1	1	0.52	0.35	0.48%
1	2	0.81	0.29	0.16%	1	2	0.57	0.49	0.25%
1	3	0.53	0.45	0.46%	1	3	0.50	0.40	0.33%
1	4	0.49	0.48	3.71%	1	4	0.80	0.34	0.37%
1	5	0.62	0.50	1.08%	1	5	0.58	0.41	0.33%
1	6A	0.76	0.54	3.99%	1	6	0.58	0.43	0.22%
1	6B	0.81	0.44	3.98%	1	7	0.69	0.40	0.19%
1	7	0.30	0.31	0.18%	1	8	0.47	0.39	0.45%
1	8	0.69	0.26	0.32%	1	9	0.74	0.45	0.31%
1	9	0.64	0.31	0.25%	1	10	0.30	0.40	0.29%
1	10	0.85	0.42	0.99%	1	11	0.85	0.27	0.18%
1	11	0.62	0.40	0.34%	1	12	0.52	0.42	0.60%
1	12	0.86	0.38	0.48%	1	13	0.64	0.53	0.61%
1	13	0.86	0.33	1.36%	1	14	0.43	0.19	0.33%
1	14	0.91	0.40	1.14%	1	15	0.88	0.47	0.34%
1	15	0.59	0.30	0.62%	1	16	0.68	0.46	0.34%
1	16	0.72	0.39	0.61%	1	17	0.73	0.54	0.48%
2	1	0.68	0.55	0.42%	1	18	0.32	0.22	0.51%
3	1	0.94	0.36	0.21%	1	19	0.42	0.24	0.41%
3	2	0.63	0.35	0.26%	1	20	0.53	0.16	0.36%
3	3	0.87	0.44	0.41%	1	21	0.66	0.51	0.54%
3	4	0.78	0.51	0.68%	1	22	0.47	0.50	0.63%
3	5	0.88	0.21	0.35%	1	23	0.42	0.23	0.43%
3	6	0.88	0.44	0.37%	2	1	0.72	0.32	0.26%
3	7	0.49	0.30	0.62%	2	2	0.73	0.42	0.26%
3	8	0.89	0.34	0.44%	2	3	0.53	0.37	0.92%
3	9	0.91	0.35	0.68%	2	4	0.64	0.53	1.01%
3	10	0.92	0.38	0.86%	2	5	0.57	0.39	0.84%
3	11	0.85	0.42	0.35%	2	6	0.59	0.37	1.32%
3	12	0.76	0.44	1.01%	2	7	0.51	0.47	2.08%
3	13	0.80	0.41	0.48%	2	8	0.65	0.34	4.21%
3	14	0.71	0.46	0.63%	2	9	0.62	0.43	6.04%
3	15	0.82	0.49	1.09%	2	10	0.99	0.14	0.22%

Table 6.12: Item Statistics: Grade 7 (Cont'd)

Communication Arts					Mathematics				
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
3	16	0.65	0.34	3.61%	2	11	0.86	0.47	0.30%
3	17	0.73	0.51	1.36%	2	12	0.90	0.34	1.05%
3	18	0.73	0.51	1.52%	2	13	0.91	0.37	2.37%
3	19	0.77	0.37	1.68%	2	14	0.84	0.37	0.30%
3	20	0.75	0.39	1.85%	2	15	0.66	0.42	0.42%
3	21	0.80	0.49	0.31%	2	16	0.89	0.43	0.60%
3	22	0.60	0.43	0.46%	2	17	0.92	0.35	0.33%
3	23	0.70	0.47	0.44%	2	18	0.62	0.43	0.42%
3	24	0.43	0.35	1.12%	2	19	0.82	0.41	1.68%
3	25	0.74	0.44	0.54%	2	20	0.66	0.46	0.80%
3	26	0.75	0.44	0.52%	2	21	0.45	0.41	0.53%
3	27	0.85	0.30	0.74%	2	22	0.54	0.36	1.59%
3	28	0.52	0.37	0.98%	2	23	0.77	0.49	0.51%
3	29	0.73	0.52	0.67%	2	24	0.61	0.51	1.29%
3	30	0.43	0.28	1.21%	2	25	0.76	0.41	0.44%
3	31	0.56	0.35	0.49%	2	26	0.63	0.50	0.80%
3	32	0.60	0.38	0.57%	2	27	0.66	0.41	0.52%
3	33	0.80	0.50	1.02%	2	28	0.73	0.52	0.59%
3	34	0.79	0.43	1.08%	2	29	0.50	0.49	0.86%
3	35	0.69	0.49	1.51%	2	30	0.66	0.43	0.95%
3	36	0.65	0.47	1.15%	2	31	0.62	0.60	0.75%
3	37	0.82	0.55	1.51%	2	32	0.52	0.48	1.18%
3	38	0.64	0.41	1.33%	3	1	0.66	0.36	0.79%
3	39	0.53	0.42	1.51%	3	2	0.34	0.53	1.71%
4	1	0.34	0.49	1.17%	3	3	0.43	0.58	0.60%
4	2	0.14	0.30	1.20%	3	4	0.50	0.60	1.40%
4	3A	0.35	0.52	1.59%	3	5	0.25	0.59	1.84%
4	3B	0.75	0.46	1.59%	3	6	0.29	0.46	0.81%
4	3C	0.87	0.39	1.59%	3	7	0.42	0.57	0.74%

Table 6. 13: Item Statistics: Grade 8

Communication Arts					Mathematics					Science				
Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate
1	1	0.89	0.27	0.29%	1	1	0.70	0.36	0.21%	1	1	0.44	0.61	0.32%
1	2	0.36	0.20	0.34%	1	2	0.61	0.59	0.26%	1	2	0.10	0.27	2.23%
1	3	0.65	0.56	1.14%	1	3	0.58	0.47	0.27%	1	3	0.21	0.47	1.98%
1	4	0.53	0.52	1.77%	1	4	0.72	0.21	0.26%	1	4	0.11	0.35	1.45%
1	5	0.57	0.58	1.27%	1	5	0.49	0.28	0.40%	1	5	0.26	0.55	1.87%
1	6A	0.73	0.61	1.96%	1	6	0.64	0.17	0.36%	1	6	0.41	0.65	1.52%
1	6B	0.87	0.40	1.96%	1	7	0.27	0.33	0.42%	1	7	0.35	0.54	2.15%
1	6C	0.96	0.28	1.96%	1	8	0.49	0.25	0.54%	1	8	0.33	0.46	2.79%
1	7	0.90	0.22	0.27%	1	9	0.39	0.49	0.45%	1	9	0.28	0.36	2.93%
1	8	0.80	0.29	0.49%	1	10	0.33	0.40	0.29%	1	10	0.37	0.52	1.45%
1	9	0.79	0.40	0.48%	1	11	0.52	0.18	0.34%	1	11	0.26	0.47	1.12%
1	10	0.36	0.27	1.12%	1	12	0.76	0.41	0.41%	1	12	0.50	0.63	2.48%
1	11	0.33	0.25	0.40%	1	13	0.68	0.51	0.46%	2	1	0.93	0.20	0.35%
1	12	0.82	0.36	0.72%	1	14	0.48	0.51	0.42%	2	2	0.91	0.33	0.37%
1	13	0.70	0.39	0.86%	1	15	0.54	0.36	0.43%	2	3	0.90	0.42	0.40%
1	14	0.50	0.32	1.05%	1	16	0.40	0.29	0.48%	2	4	0.88	0.30	0.39%
1	15	0.44	0.33	1.71%	1	17	0.48	0.37	0.31%	2	5	0.97	0.26	0.39%
1	16	0.83	0.37	0.96%	1	18	0.44	0.52	0.47%	2	6	0.81	0.38	0.45%
2	1	0.96	0.33	0.21%	1	19	0.63	0.43	0.34%	2	7	0.72	0.36	0.47%
2	2	0.94	0.33	0.26%	1	20	0.90	0.36	0.42%	2	8	0.73	0.38	0.42%
2	3	0.89	0.41	0.32%	1	21	0.39	0.33	0.51%	2	9	0.82	0.37	0.60%
2	4	0.93	0.35	0.63%	1	22	0.31	0.21	0.57%	2	10	0.87	0.38	0.68%
2	5	0.81	0.49	0.27%	1	23	0.73	0.31	0.45%	2	11	0.73	0.36	0.60%
2	6	0.89	0.26	0.39%	1	31	0.52	0.72	1.16%	2	12	0.76	0.38	0.60%
2	7	0.51	0.25	0.50%	2	1	0.80	0.35	0.28%	2	13	0.65	0.50	0.61%
2	8	0.97	0.38	0.25%	2	2	0.78	0.38	0.41%	2	14	0.76	0.39	0.64%
2	9	0.83	0.54	0.30%	2	3	0.76	0.41	0.65%	2	15	0.78	0.50	0.71%
2	10	0.80	0.48	0.39%	2	4	0.95	0.26	0.32%	2	16	0.63	0.54	0.64%
2	11	0.65	0.25	0.49%	2	5	0.40	0.33	0.50%	2	17	0.72	0.26	0.56%
2	12	0.76	0.46	1.05%	2	6	0.42	0.39	0.49%	2	18	0.56	0.40	5.37%
2	13	0.76	0.47	1.16%	2	7	0.75	0.38	0.58%	2	19	0.71	0.27	0.52%
2	14	0.44	0.36	0.67%	2	8	0.77	0.32	0.40%	2	20	0.71	0.46	0.85%
2	15	0.82	0.50	2.43%	2	9	0.37	0.39	0.35%	2	21	0.60	0.29	0.74%

Table 6.13: Item Statistics: Grade 8 (Cont'd)

Communication Arts					Mathematics					Science				
Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate
2	16	0.92	0.46	0.90%	2	10	0.83	0.43	0.36%	2	22	0.43	0.44	1.35%
2	17	0.83	0.45	1.20%	2	11	0.68	0.32	0.42%	2	23	0.92	0.27	0.77%
2	18	0.59	0.38	0.72%	2	12	0.83	0.38	0.98%	2	25	0.48	0.26	0.60%
2	19	0.33	0.27	0.85%	2	13	0.74	0.41	0.56%	2	26	0.67	0.60	0.71%
2	20	0.35	0.28	1.12%	2	14	0.60	0.43	0.53%	2	27	0.38	0.54	2.47%
2	21	0.75	0.42	0.28%	2	15	0.89	0.43	0.54%	2	28	0.24	0.46	3.67%
2	22	0.94	0.41	0.32%	2	16	0.80	0.28	0.49%	2	29	0.60	0.58	3.17%
2	23	0.79	0.51	0.44%	2	17	0.86	0.32	0.46%	2	30	0.43	0.61	1.04%
2	24	0.77	0.41	0.41%	2	18	0.61	0.51	0.60%	2	31	0.58	0.64	7.29%
2	25	0.74	0.49	0.30%	2	19	0.64	0.58	0.61%	2	32	0.52	0.42	3.63%
2	26	0.81	0.45	0.36%	2	20	0.55	0.33	0.71%	2	33	0.11	0.43	2.51%
2	27	0.55	0.43	0.53%	2	21	0.87	0.40	0.53%	2	34	0.14	0.40	3.29%
2	28	0.88	0.42	0.48%	2	22	0.82	0.36	0.61%	2	35	0.34	0.61	7.93%
2	29	0.67	0.41	0.56%	2	23	0.71	0.41	0.56%	2	36	0.47	0.59	2.51%
2	30	0.87	0.50	0.55%	2	24	0.79	0.32	0.57%	2	37	0.30	0.55	3.34%
2	31	0.42	0.32	1.10%	2	25	0.35	0.33	0.65%	3	1	0.60	0.35	1.40%
2	32	0.47	0.28	1.83%	2	26	0.67	0.57	0.65%	3	2	0.24	0.52	3.24%
2	33	0.88	0.47	2.24%	2	27	0.43	0.35	1.15%	3	3	0.18	0.25	2.98%
2	34	0.60	0.46	1.05%	2	28	0.61	0.56	1.81%	3	4	0.69	0.48	7.22%
2	35	0.81	0.47	0.62%	2	29	0.39	0.36	1.78%	3	5	0.38	0.52	4.04%
2	36	0.65	0.46	0.98%	2	30	0.41	0.37	1.05%	3	6	0.21	0.56	2.16%
2	37	0.38	0.38	0.70%	2	31	0.36	0.48	1.15%	3	7	0.20	0.34	1.92%
2	38	0.41	0.28	0.93%	3	1	0.50	0.53	2.72%	3	8	0.58	0.61	4.32%
2	39	0.26	0.27	0.93%	3	2	0.28	0.57	1.12%	3	9	0.39	0.39	2.26%
3	1	0.42	0.49	0.84%	3	3	0.31	0.52	2.11%	3	10	0.21	0.43	4.82%
3	2	0.46	0.48	1.11%	3	4	0.62	0.50	0.87%					
3	3A	0.62	0.64	4.34%	3	5	0.43	0.68	2.70%					
3	3B	0.68	0.47	4.33%	3	6	0.26	0.61	1.53%					
					3	7	0.44	0.65	2.53%					
					3	8	0.36	0.52	1.16%					
					3	9	0.34	0.64	4.25%					

Table 6. 14: Item Statistics: Grade 11 (Communication Arts, Science), Grade 10 (Mathematics)

Communication Arts					Mathematics					Science				
Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate
1	1	0.69	0.41	0.33%	1	1	0.43	0.40	0.66%	1	1	0.27	0.59	0.54%
1	2	0.32	0.32	0.35%	1	2	0.66	0.37	0.22%	1	2	0.30	0.51	3.56%
1	3	0.51	0.45	1.80%	1	3	0.46	0.56	0.52%	1	3	0.41	0.31	2.12%
1	4	0.36	0.42	1.79%	1	4	0.42	0.40	0.25%	1	4	0.18	0.45	3.62%
1	5	0.47	0.54	4.76%	1	5	0.40	0.32	0.38%	1	5	0.39	0.49	2.60%
1	6A	0.74	0.51	2.15%	1	6	0.31	0.49	0.31%	1	6	0.47	0.61	3.60%
1	6B	0.80	0.41	2.15%	1	7	0.42	0.43	0.31%	1	7	0.44	0.54	1.80%
1	6C	0.94	0.28	2.15%	1	8	0.90	0.29	0.27%	1	8	0.55	0.46	1.53%
1	7	0.86	0.28	0.40%	1	9	0.56	0.38	0.54%	1	9	0.26	0.50	2.18%
1	8	0.36	0.25	0.69%	1	10	0.83	0.33	0.34%	1	10	0.56	0.45	2.87%
1	9	0.47	0.51	0.59%	1	11	0.77	0.28	0.28%	1	11	0.44	0.56	0.80%
1	10	0.47	0.44	0.97%	1	12	0.67	0.41	0.28%	1	12	0.34	0.48	3.14%
1	11	0.58	0.18	1.28%	1	13	0.92	0.39	0.25%	1	13	0.67	0.42	8.29%
1	12	0.54	0.30	0.58%	1	14	0.50	0.32	0.43%	1	14	0.31	0.46	4.21%
1	13	0.62	0.25	0.64%	1	15	0.71	0.50	0.32%	1	15	0.27	0.50	2.71%
1	14	0.89	0.35	0.54%	1	16	0.56	0.42	0.34%	1	16	0.42	0.49	2.09%
1	15	0.89	0.32	0.78%	1	17	0.57	0.61	0.39%	1	17	0.26	0.54	10.22%
1	16	0.67	0.30	0.52%	1	18	0.62	0.55	0.36%	1	18	0.30	0.62	4.95%
2	1	0.73	0.59	0.93%	1	19	0.29	0.40	0.64%	2	1	0.45	0.44	2.23%
3	1	0.54	0.55	0.66%	1	20	0.65	0.31	0.36%	2	2	0.53	0.58	3.32%
3	2	0.95	0.38	0.63%	1	21	0.75	0.51	0.35%	2	3	0.55	0.66	1.59%
3	3	0.78	0.43	0.66%	1	22	0.55	0.58	0.44%	2	4	0.50	0.54	1.57%
3	4	0.76	0.39	0.62%	1	23	0.37	0.47	0.49%	2	5	0.16	0.32	3.14%
3	5	0.28	0.30	0.74%	2	1	0.74	0.48	0.37%	2	6	0.46	0.43	2.98%
3	6	0.82	0.37	0.72%	2	2	0.69	0.39	0.40%	2	7	0.59	0.50	2.52%
3	7	0.81	0.46	0.66%	2	3	0.56	0.45	0.67%	2	8	0.64	0.47	3.38%
3	8	0.75	0.43	0.69%	2	4	0.84	0.38	0.48%	2	9	0.34	0.38	5.21%
3	9	0.72	0.50	0.70%	2	5	0.38	0.49	0.62%	2	10	0.32	0.57	6.09%
3	10	0.84	0.39	0.91%	2	6	0.63	0.53	1.15%	2	11	0.48	0.51	8.60%
3	11	0.77	0.37	0.64%	2	7	0.67	0.52	0.49%	3	1	0.81	0.43	0.59%
3	12	0.98	0.29	0.66%	2	8	0.80	0.48	0.49%	3	2	0.70	0.35	0.58%
3	13	0.69	0.40	0.74%	2	9	0.91	0.23	0.32%	3	3	0.54	0.34	2.36%
3	14	0.76	0.43	0.91%	2	10	0.76	0.56	0.45%	3	4	0.69	0.44	0.58%

Table 6.14: Item Statistics: Grade 11 (Communication Arts, Science), Grade 10 (Mathematics)
(Cont'd)

Communication Arts					Mathematics					Science				
Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate	Ses-sion	Item	P-Value	R _{it}	Omit Rate
3	15	0.65	0.38	0.82%	2	11	0.89	0.43	0.82%	3	6	0.61	0.33	1.34%
3	16	0.81	0.57	1.13%	2	12	0.90	0.39	1.64%	3	7	0.62	0.49	0.94%
3	17	0.80	0.50	1.23%	2	13	0.67	0.56	1.79%	3	8	0.87	0.34	0.96%
3	18	0.76	0.52	1.33%	2	14	0.65	0.49	0.78%	3	9	0.73	0.32	0.96%
3	19	0.72	0.51	1.71%	2	15	0.56	0.38	0.71%	3	10	0.60	0.33	0.91%
3	20	0.75	0.37	1.88%	2	16	0.65	0.41	0.73%	3	11	0.75	0.37	0.97%
3	21	0.87	0.40	0.72%	2	17	0.61	0.44	1.09%	3	12	0.78	0.45	0.85%
3	22	0.87	0.53	0.67%	2	18	0.64	0.46	0.58%	3	13	0.65	0.35	0.76%
3	23	0.59	0.39	0.74%	2	19	0.70	0.53	0.73%	3	14	0.57	0.33	0.85%
3	24	0.89	0.50	0.70%	2	20	0.78	0.50	0.72%	3	15	0.66	0.14	0.67%
3	25	0.55	0.45	0.76%	2	21	0.53	0.40	0.97%	3	16	0.54	0.29	0.56%
3	26	0.55	0.32	1.01%	2	22	0.75	0.48	1.23%	3	17	0.53	0.42	0.73%
3	27	0.59	0.43	0.72%	2	23	0.59	0.36	1.36%	3	18	0.34	0.44	3.83%
3	28	0.56	0.38	0.85%	2	24	0.62	0.34	1.31%	3	19	0.44	0.42	0.65%
3	29	0.40	0.35	0.82%	2	25	0.62	0.53	1.52%	3	20	0.37	0.40	0.70%
3	30	0.72	0.46	1.42%	3	1	0.56	0.69	3.09%	3	22	0.38	0.21	0.92%
3	31	0.27	0.10	1.60%	3	2	0.60	0.59	1.69%	3	23	0.46	0.45	0.83%
3	32	0.71	0.36	0.82%	3	3	0.24	0.56	3.42%	3	24	0.42	0.33	0.78%
3	33	0.39	0.32	1.15%	3	4	0.28	0.55	1.42%	3	25	0.42	0.40	0.88%
3	34	0.55	0.41	1.14%	3	5	0.37	0.66	3.03%	3	26	0.35	0.45	1.67%
3	35	0.74	0.49	1.41%	3	6	0.53	0.60	2.79%	3	27	0.65	0.40	1.58%
3	36	0.81	0.48	1.99%	3	7	0.57	0.55	1.69%	3	28	0.37	0.43	1.71%
3	37	0.57	0.31	2.64%	3	8	0.40	0.62	3.39%	3	29	0.23	0.57	4.47%
3	38	0.64	0.27	2.02%	3	9	0.65	0.58	2.68%	3	30	0.33	0.45	2.22%
3	39	0.30	0.21	1.85%	3	10	0.28	0.66	1.48%	3	31	0.33	0.44	2.97%
4	1	0.70	0.58	1.30%										
4	2	0.60	0.53	2.48%										
4	3A	0.60	0.47	3.00%										
4	3B	0.70	0.42	2.99%										
4	3C	0.87	0.41	2.99%										

Table 6. 15: Item Statistics: Grade 11 Communication Arts Breach Form

Communication Arts									
Session	Item	P-Value	R _{it}	Omit Rate	Session	Item	P-Value	R _{it}	Omit Rate
3	1	0.36	0.15	0.96%	3	36	0.27	0.20	2.24%
3	2	0.87	0.39	1.20%	3	37	0.51	0.28	2.24%
3	3	0.33	0.49	1.20%	3	38	0.58	0.51	2.32%
3	4	0.36	0.37	1.20%	3	39	0.55	0.47	2.48%
3	5	0.49	0.47	1.12%	3	40	0.51	0.50	2.48%
3	6	0.67	0.42	0.88%	3	41	0.50	0.37	2.24%
3	7	0.52	0.35	0.80%					
3	8	0.75	0.43	1.20%					
3	9	0.60	0.44	1.04%					
3	10	0.64	0.33	0.88%					
3	11	0.67	0.38	0.80%					
3	12	0.23	0.16	1.44%					
3	13	0.38	0.42	1.36%					
3	14	0.62	0.50	1.36%					
3	15	0.65	0.39	1.36%					
3	16	0.55	0.52	1.76%					
3	17	0.50	0.37	2.24%					
3	18	0.16	0.11	3.04%					
3	19	0.50	0.20	2.08%					
3	20	0.53	0.37	2.00%					
3	23	0.77	0.42	1.68%					
3	24	0.38	0.22	1.84%					
3	25	0.58	0.46	2.00%					
3	26	0.46	0.46	2.56%					
3	27	0.40	0.35	2.24%					
3	28	0.66	0.49	1.68%					
3	29	0.69	0.25	1.68%					
3	30	0.50	0.39	1.68%					
3	31	0.76	0.45	2.24%					
3	32	0.83	0.45	1.92%					
3	33	0.64	0.52	2.24%					
3	34	0.53	0.54	2.00%					
3	35	0.46	0.46	2.24%					

Table 6. 16: Item Fit Statistics for Misfitting Items

Content	Grade	Session	Item	Chi Square	DF	Total N	Z	Observed	Predicted	Obs-Pred
CA	3	4	1	2016.58	17	54523	342.93	0.58	0.58	0.00
CA	4	2	17	630.04	7	56315	166.51	0.93	0.92	0.00
CA	5	1	3	2353.57	17	52100	400.72	0.52	0.52	0.00
CA	5	1	4	1352.43	17	52087	229.02	0.49	0.49	0.00
CA	6	1	6A	1038.62	17	53204	175.21	0.59	0.59	0.00
CA	8	1	4	1218.33	17	55308	206.03	0.53	0.54	0.00
CA	8	2	11	776.22	7	56030	205.58	0.65	0.66	0.00
CA	8	2	38	869.90	7	55783	230.62	0.41	0.40	0.01
CA	11	3	29	623.77	7	60361	164.84	0.40	0.42	-0.02
CA	11b	3	33	20.48	7	1220	3.60	0.64	0.64	0.00
MA	4	3	8	2525.10	17	56717	430.14	0.70	0.70	0.00
MA	5	3	3	1346.92	17	52156	228.08	0.74	0.73	0.00
MA	5	3	4	897.84	17	52173	151.06	0.71	0.71	0.00
MA	5	3	5	2257.06	17	52278	384.17	0.71	0.71	0.00
MA	6	3	7	1736.63	17	53342	294.91	0.59	0.59	0.00
MA	7	3	3	1030.34	17	58459	173.79	0.43	0.43	0.00
MA	8	2	16	977.58	7	56097	259.40	0.80	0.80	0.00
MA	8	3	5	1494.45	17	54849	253.38	0.43	0.43	0.00
MA	10	2	9	1175.79	7	66119	312.37	0.91	0.90	0.00
MA	10	3	5	1695.49	17	64320	287.86	0.37	0.37	0.00
SC	5	3	2	1993.79	35	51514	234.12	0.48	0.48	0.00
SC	8	2	32	1091.72	17	54250	184.31	0.52	0.52	0.00
SC	11	3	15	554.55	7	50304	146.34	0.66	0.66	0.00

Table 6. 17: LOSS and HOSS Values by Grade and Content Area

Grade	Communication Arts		Mathematics		Science	
	LOSS	HOSS	LOSS	HOSS	LOSS	HOSS
3	455	790	450	780		
4	470	820	465	805		
5	485	840	480	830	470	855
6	505	855	495	845		
7	515	865	510	860		
8	530	875	525	885	540	895
10			555	910		
11	545	885			550	970

Figure 6. 1: Item characteristic curve for Grade 3 Communication Arts, Session 4, Item 1 (Z = 342.93)

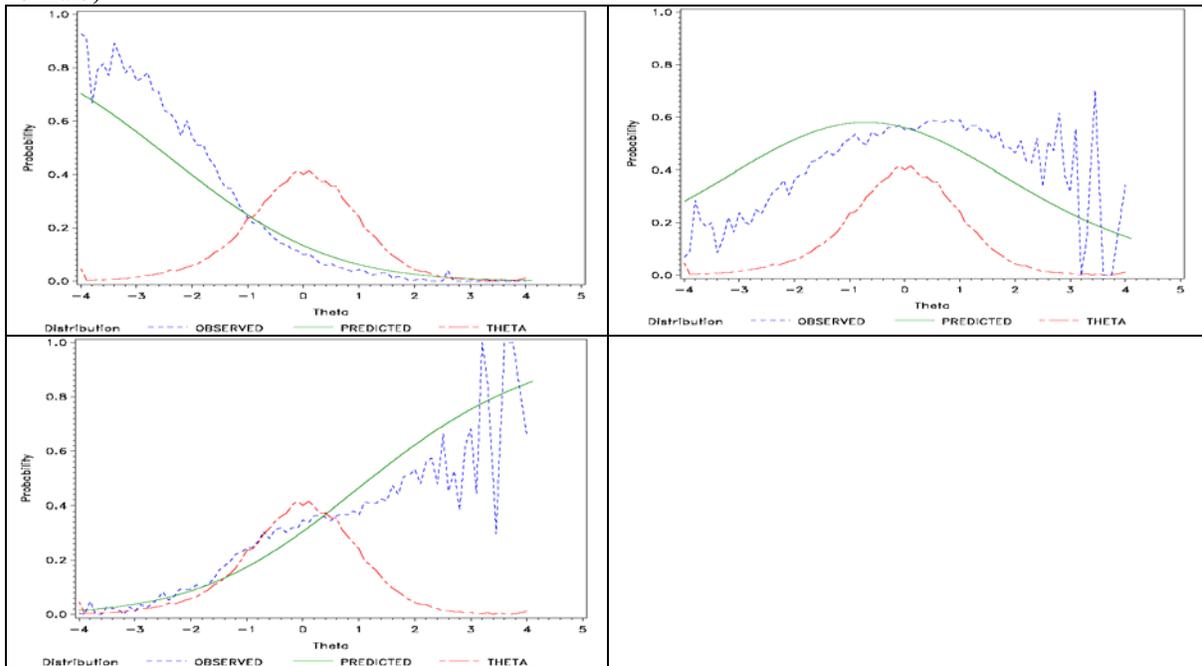


Figure 6. 2: Item characteristic curve for Grade 4 Communication Arts, Session 2, Item 17 (Z = 166.51)

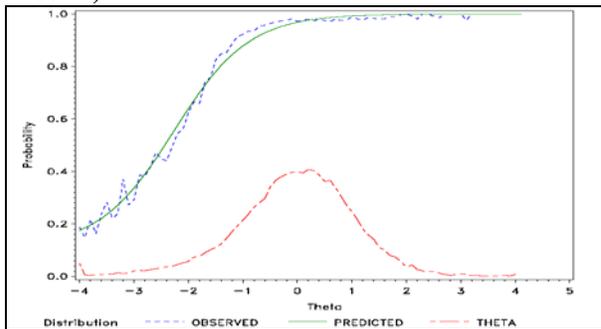


Figure 6. 3: Item characteristic curve for Grade 5 Communication Arts, Session 1, Item 3 (Z = 400.72)

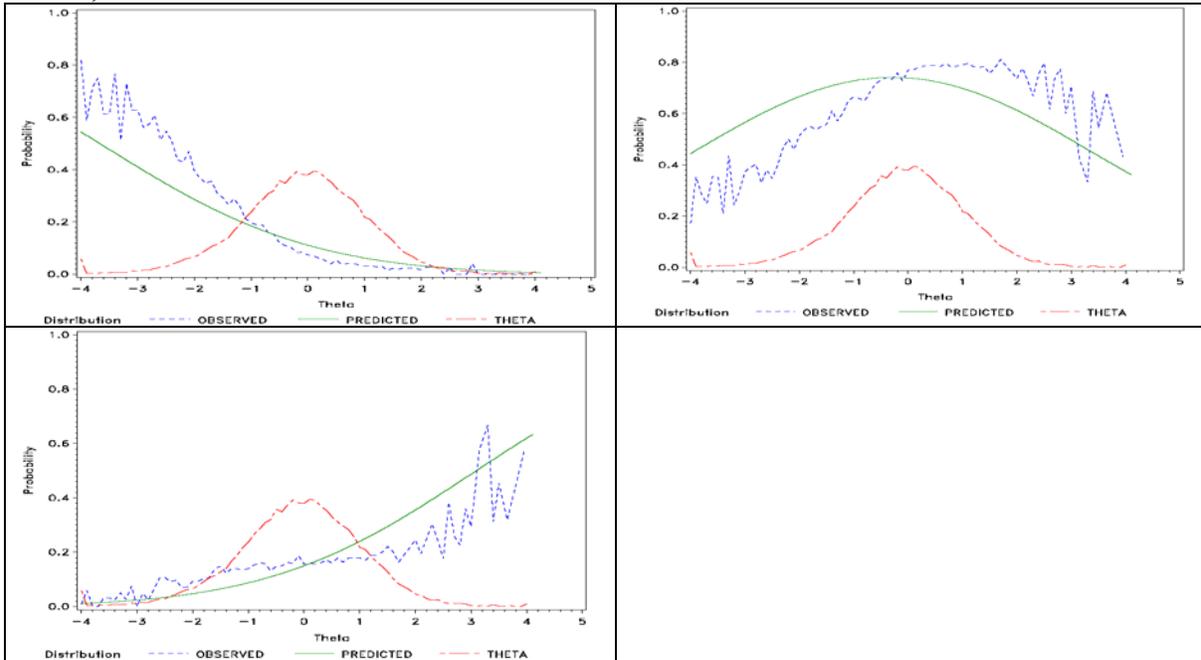


Figure 6. 4: Item characteristic curve for Grade 5 Communication Arts, Session 1, Item 4 (Z = 229.02)

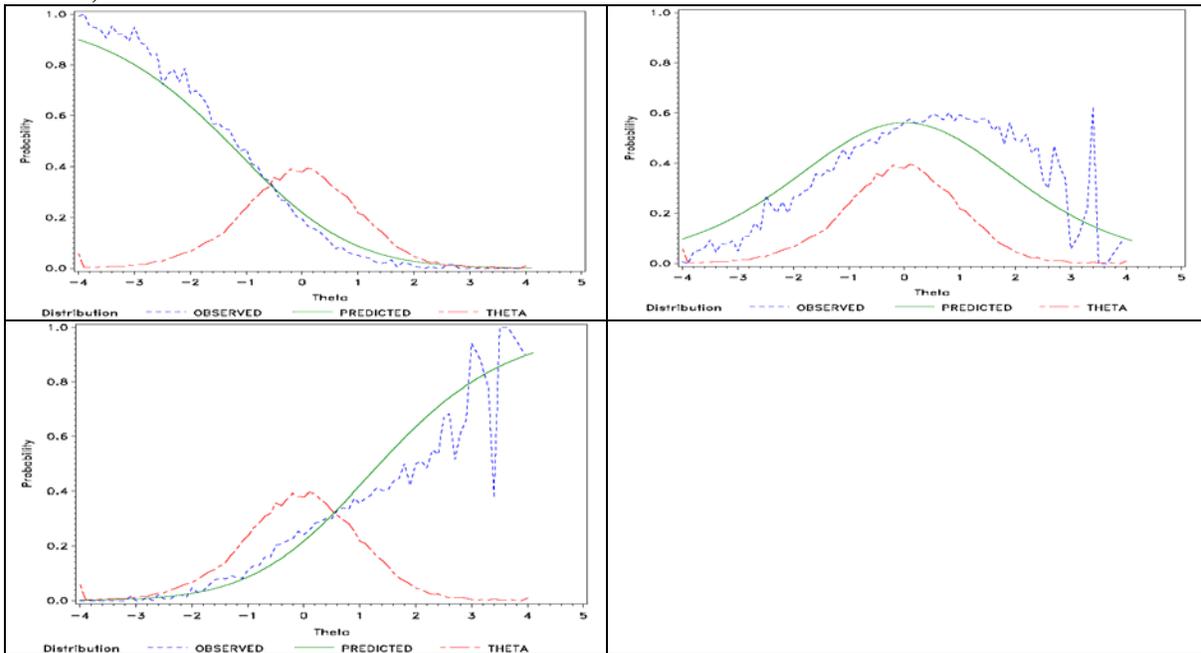


Figure 6. 5: Item characteristic curve for Grade 6 Communication Arts, Session 1, Item 6A (Z = 175.21)

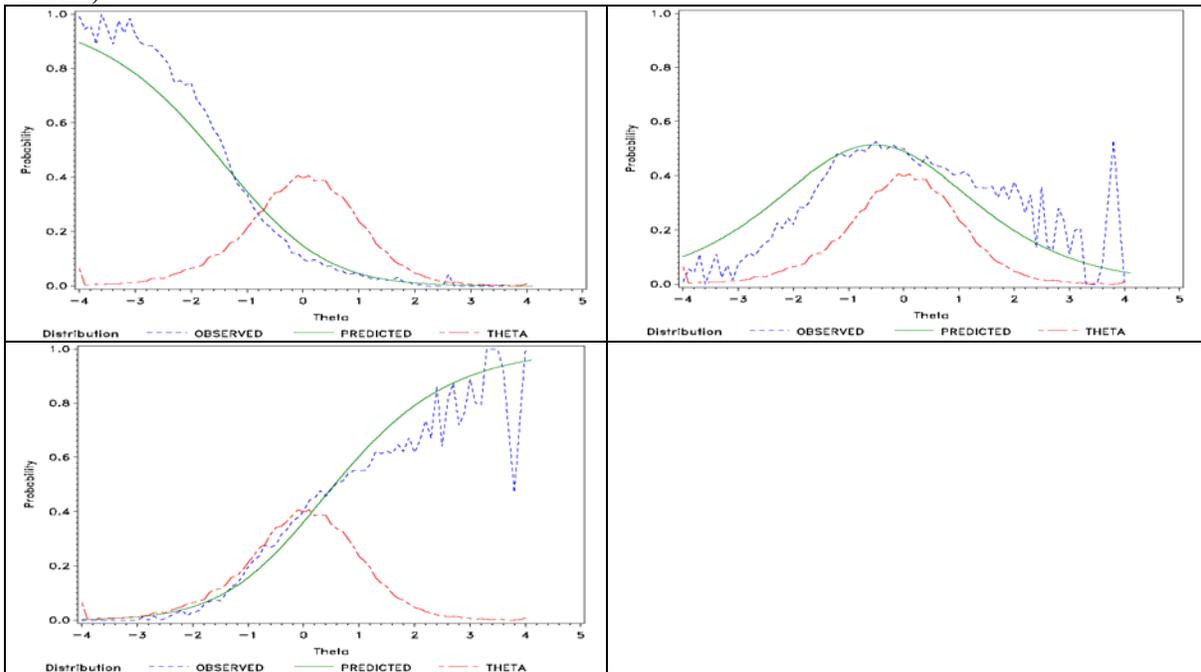


Figure 6. 6: Item characteristic curve for Grade 8 Communication Arts, Session 1, Item 4 (Z = 206.03)

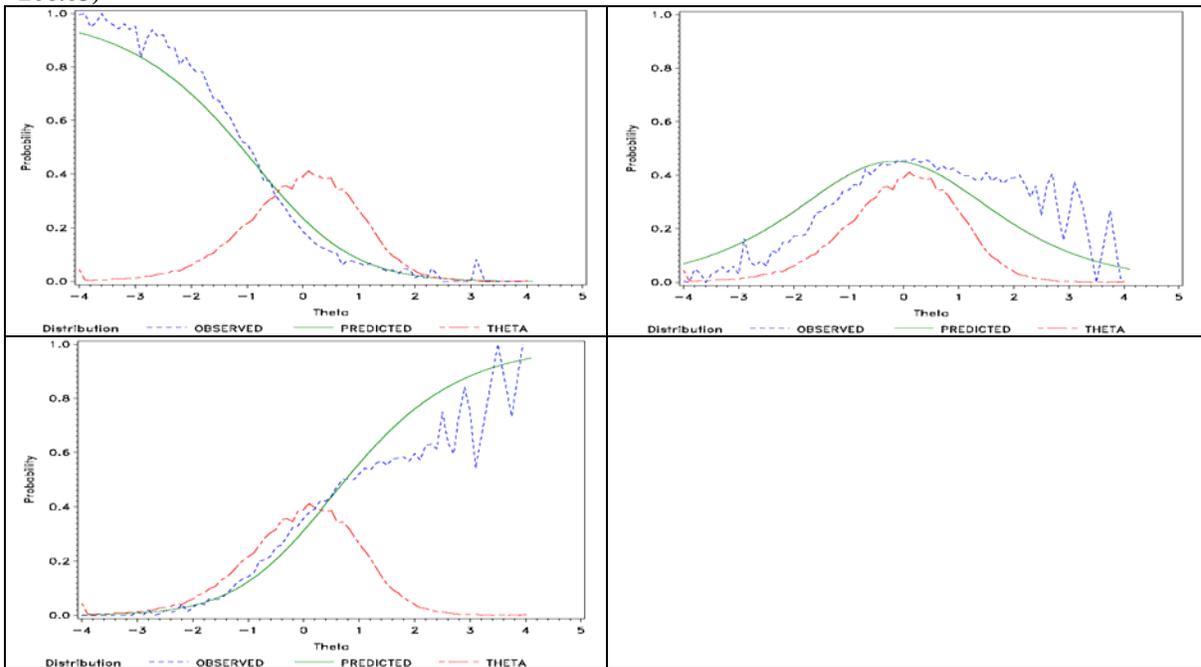


Figure 6. 7: Item characteristic curve for Grade 8 Communication Arts, Session 2, Item 11 (Z = 205.58)

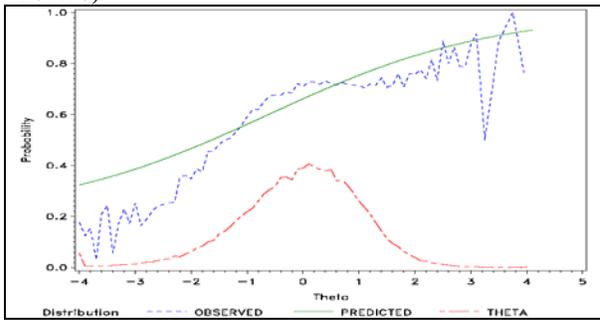


Figure 6. 8: Item characteristic curve for Grade 8 Communication Arts, Session 2, Item 38 (Z = 230.62)

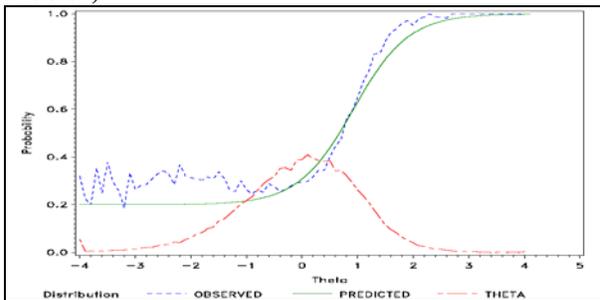


Figure 6. 9: Item characteristic curve for Grade 11 Communication Arts, Session 3, Item 29 (Z = 164.84)

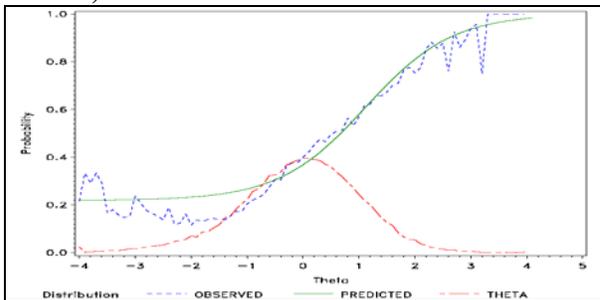


Figure 6. 10: Item characteristic curve for Grade 11 Communication Arts breach form, Session 3, Item 33 (Z = 3.60)

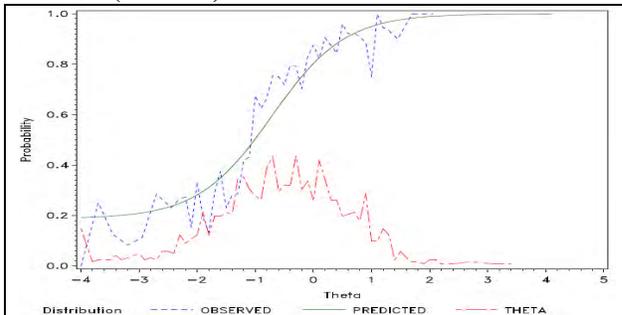


Figure 6. 11: Item characteristic curve for Grade 4 Mathematics, Session 3, Item 8 (Z = 430.14)

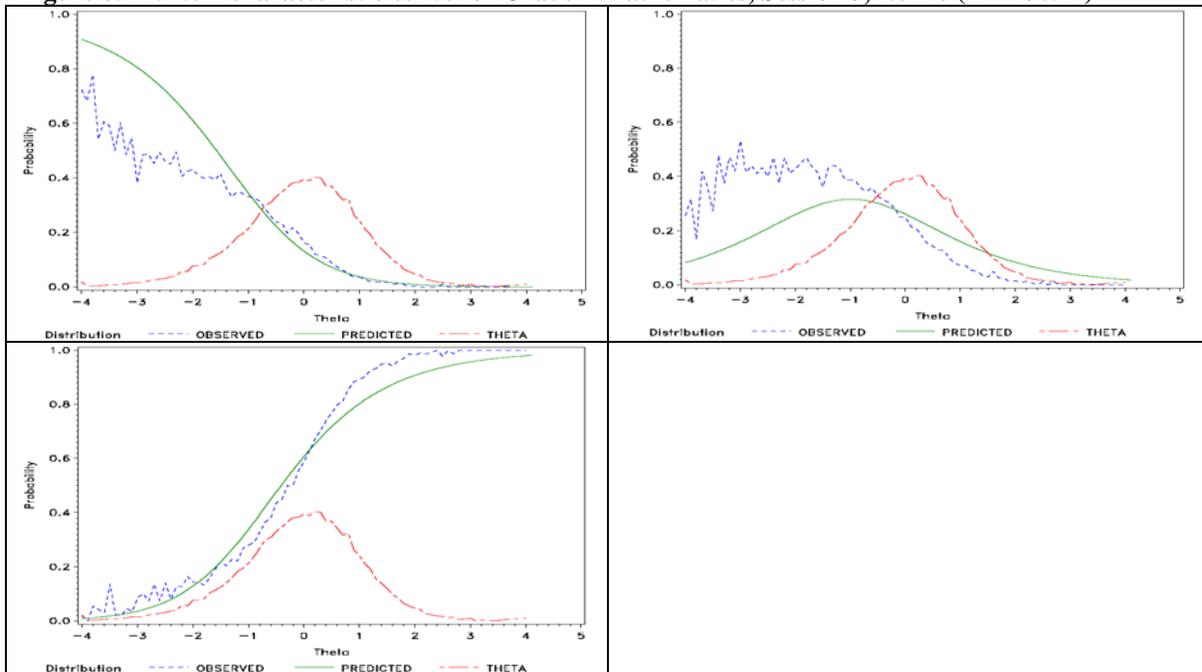


Figure 6. 12: Item characteristic curve for Grade 5 Mathematics, Session 3, Item 3 (Z = 228.08)

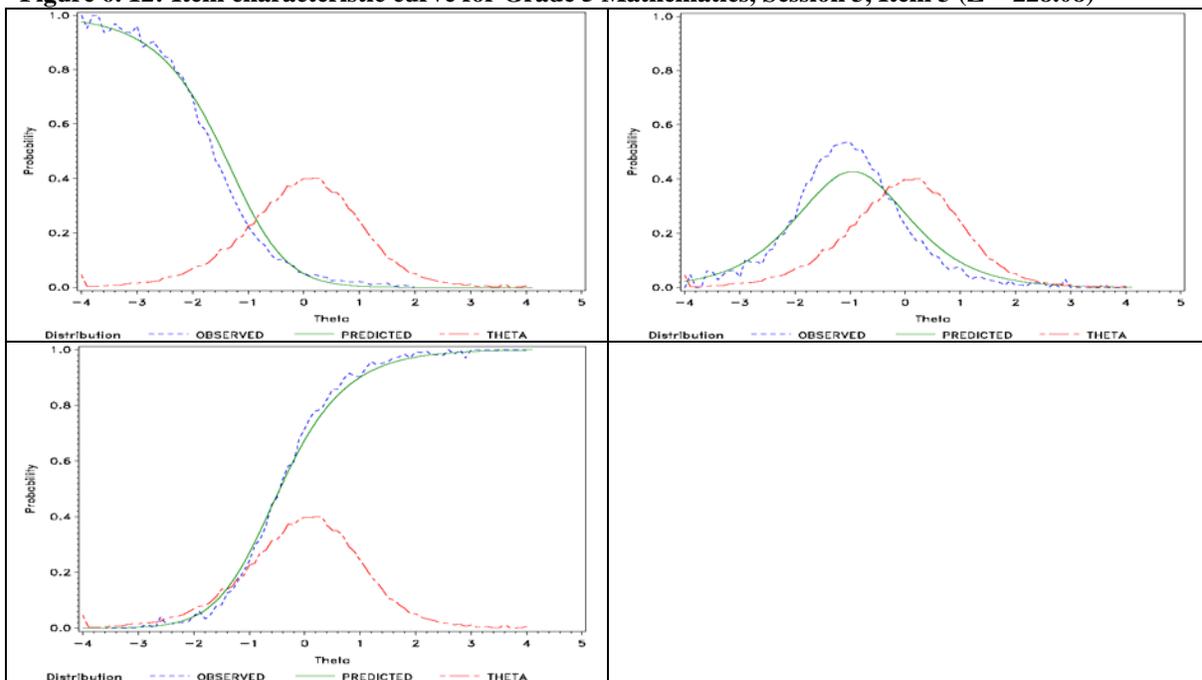


Figure 6. 13: Item characteristic curve for Grade 5 Mathematics, Session 3, Item 4 (Z = 151.06)

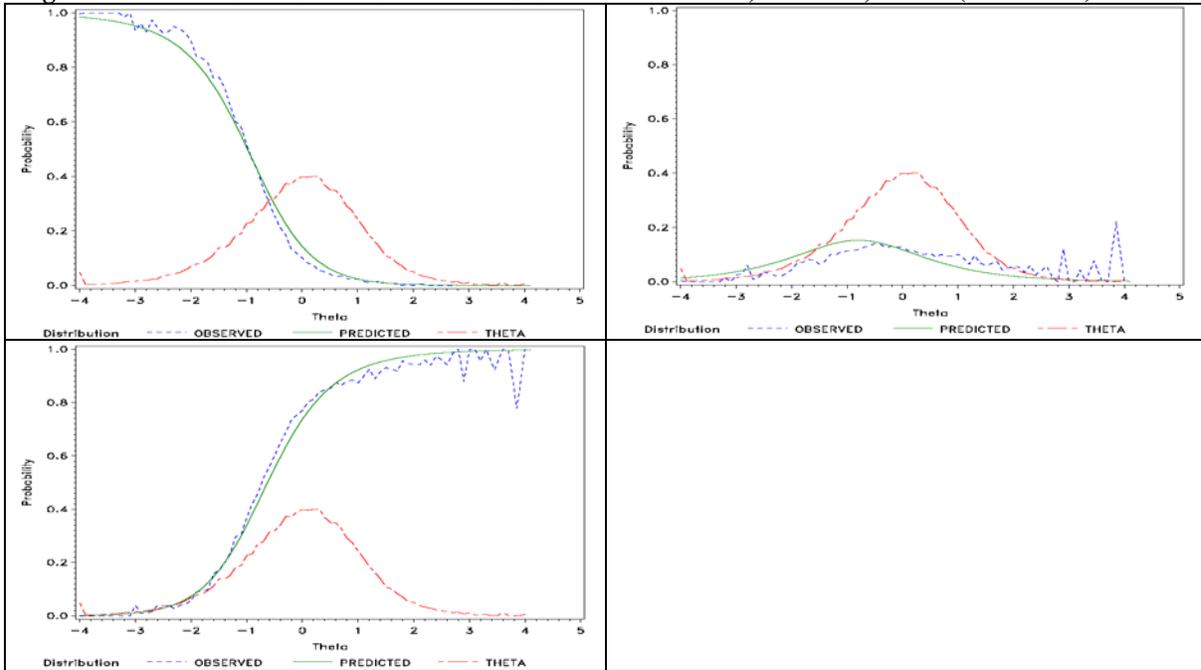


Figure 6. 14: Item characteristic curve for Grade 5 Mathematics, Session 3, Item 5 (Z = 384.17)

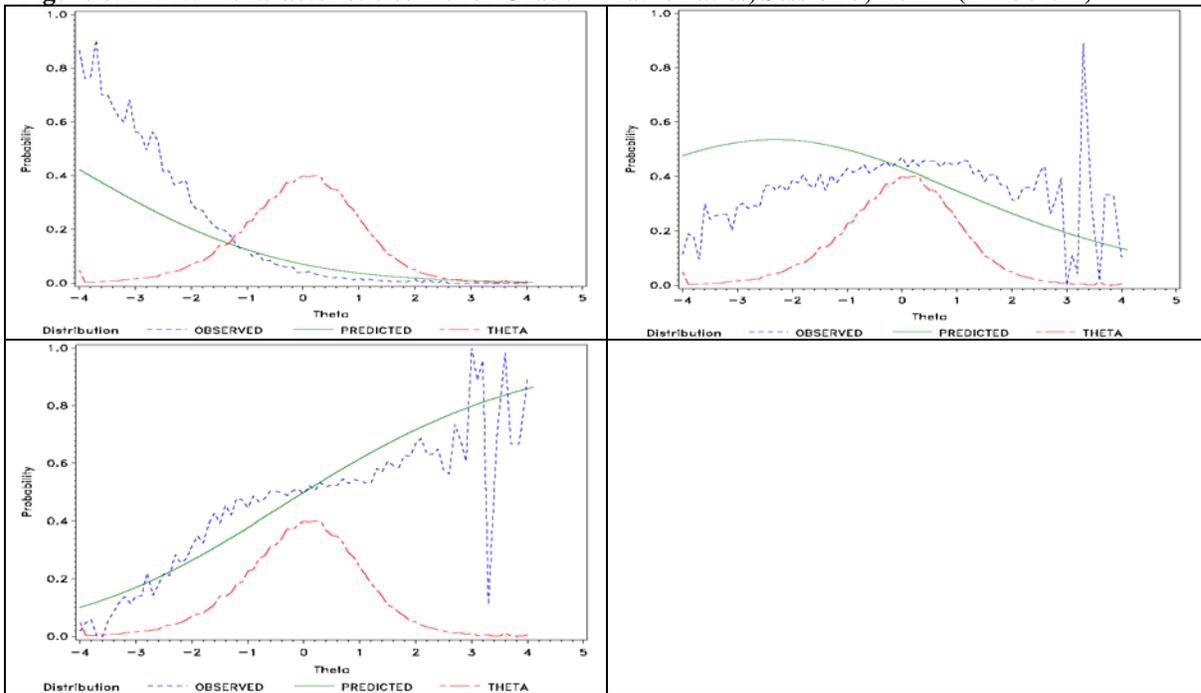


Figure 6.15: Item characteristic curve for Grade 6 Mathematics, Session 3, Item 7 (Z = 294.91)

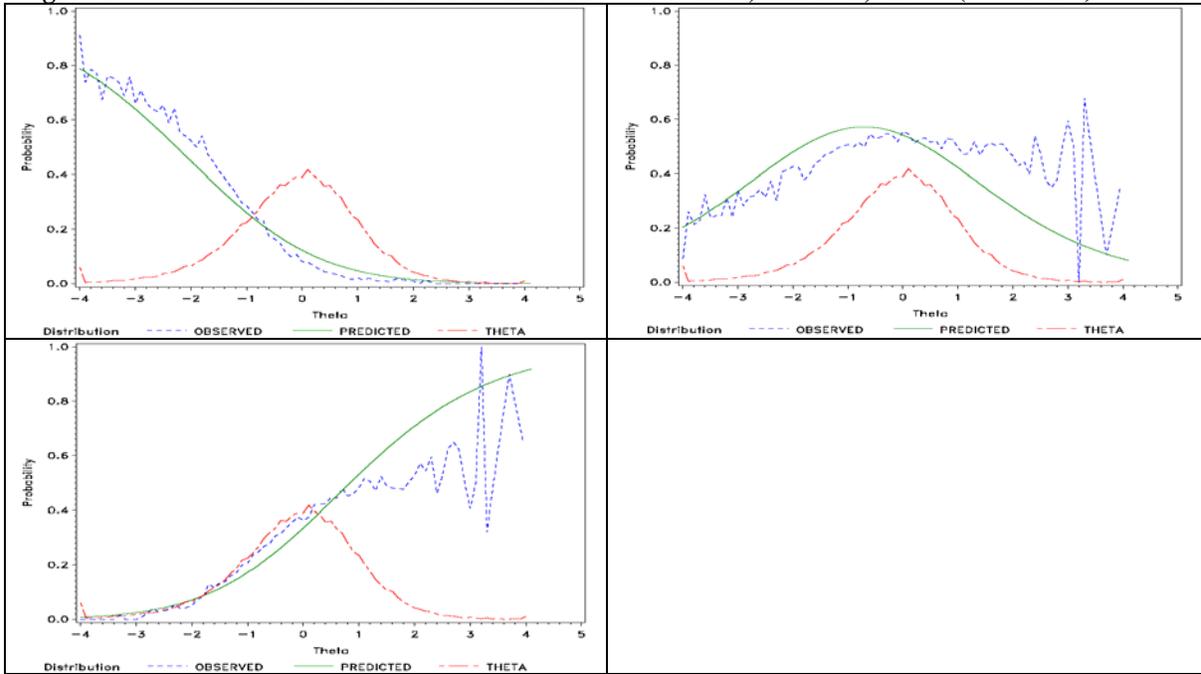


Figure 6.16: Item characteristic curve for Grade 7 Mathematics, Session 3, Item 3 (Z = 173.79)

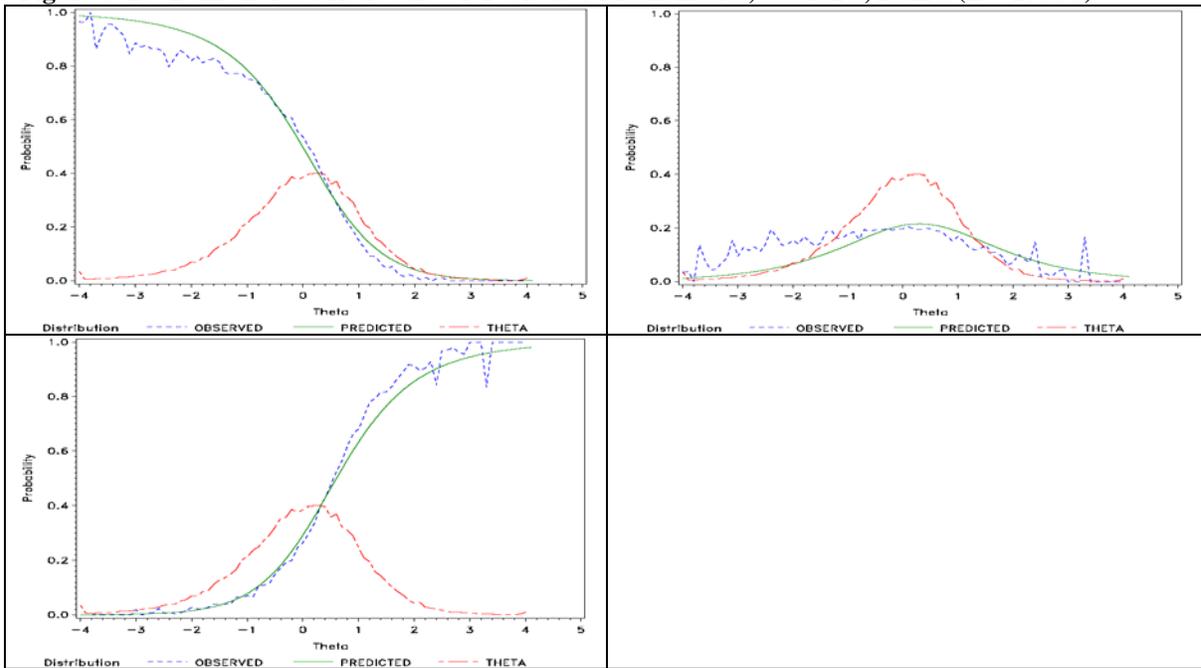


Figure 6. 17: Item characteristic curve for Grade 8 Mathematics, Session 2, Item 16 (Z = 259.40)

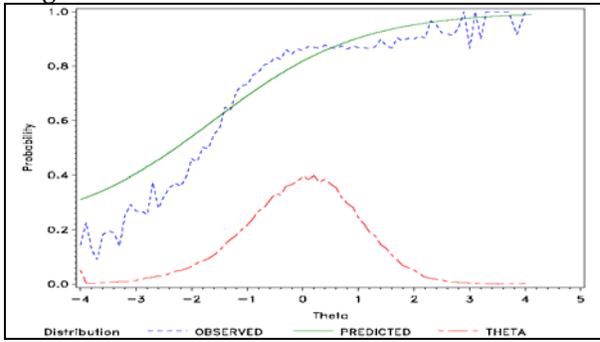


Figure 6. 18: Item characteristic curve for Grade 8 Mathematics, Session 3, Item 5 (Z = 253.38)

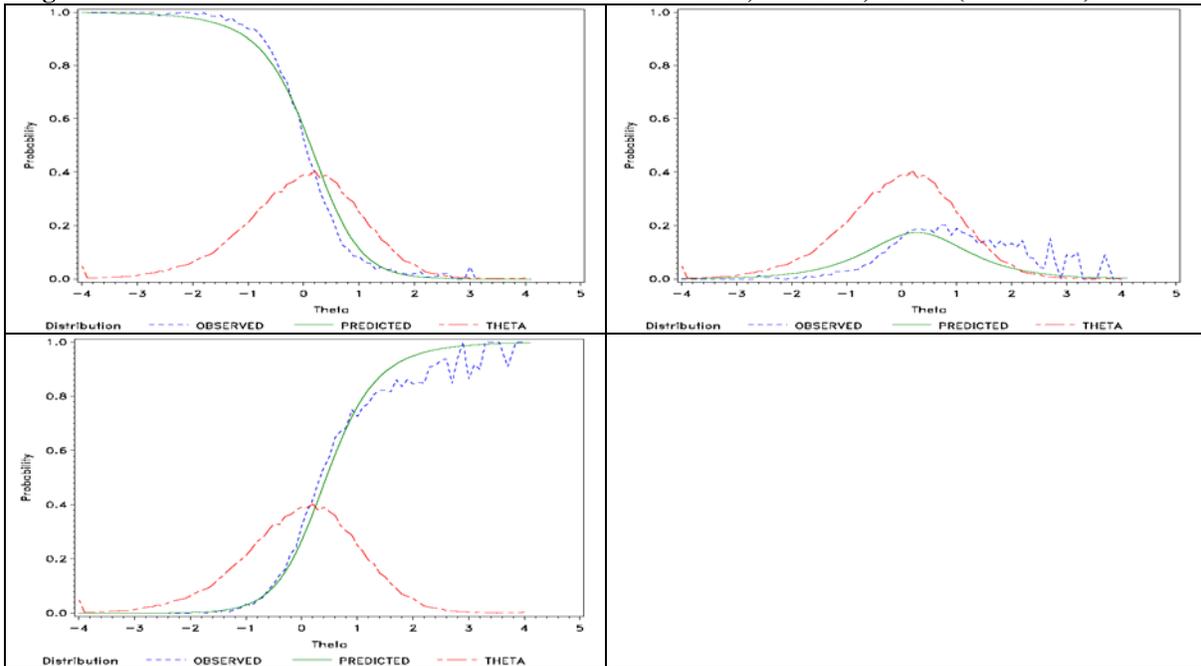


Figure 6. 19: Item characteristic curve for Grade 10 Mathematics, Session 2, Item 9 (Z = 312.37)

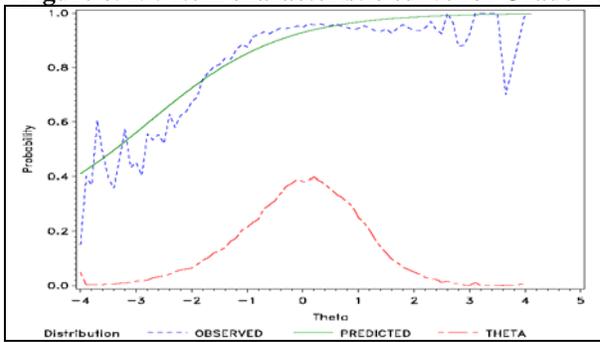


Figure 6. 20: Item characteristic curve for Grade 10 Mathematics, Session 3, Item 5 (Z = 287.86)

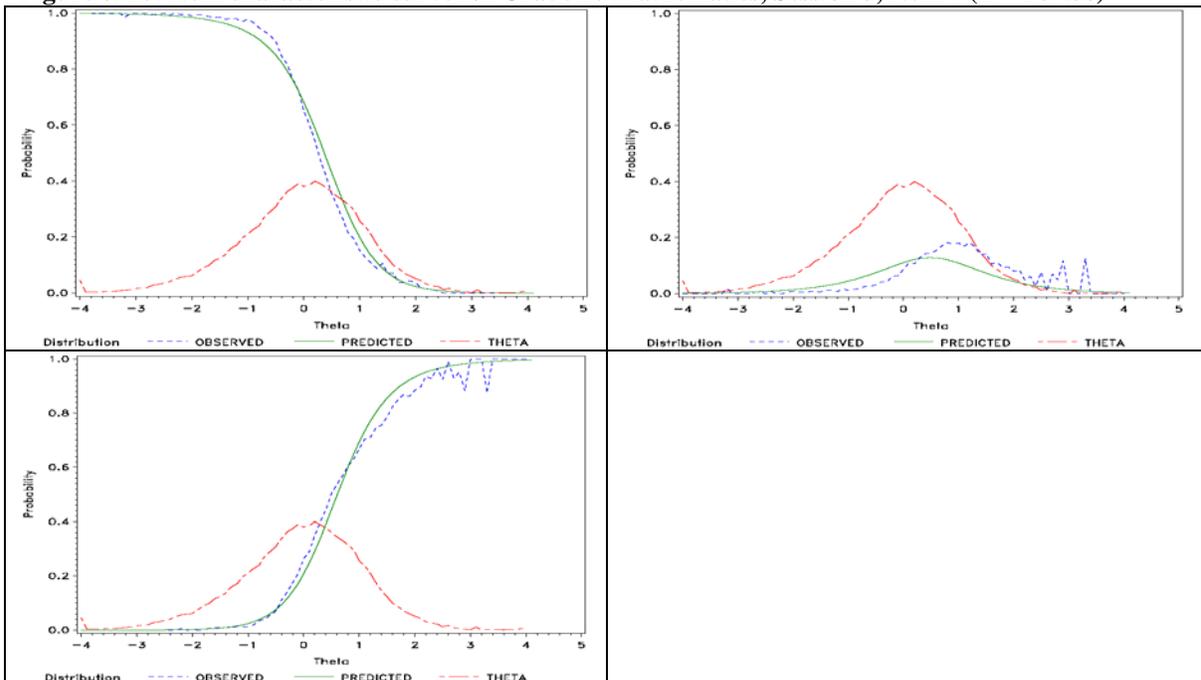


Figure 6. 21: Item characteristic curve for Grade 5 Science, Session 3, Item 2 (Z = 234.12)

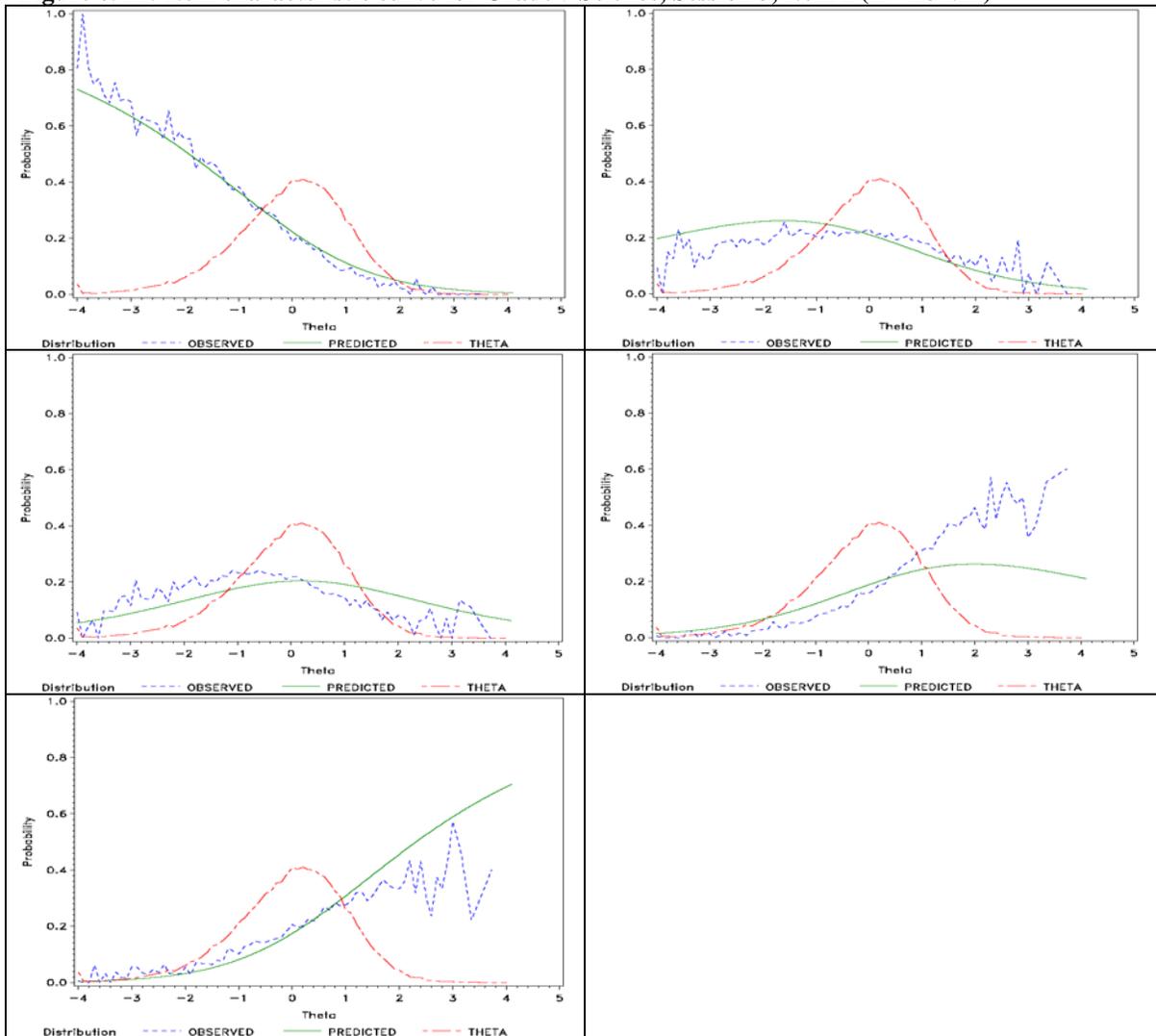


Figure 6.22: Item characteristic curve for Grade 8 Science, Session 2, Item 32 (Z = 184.31)

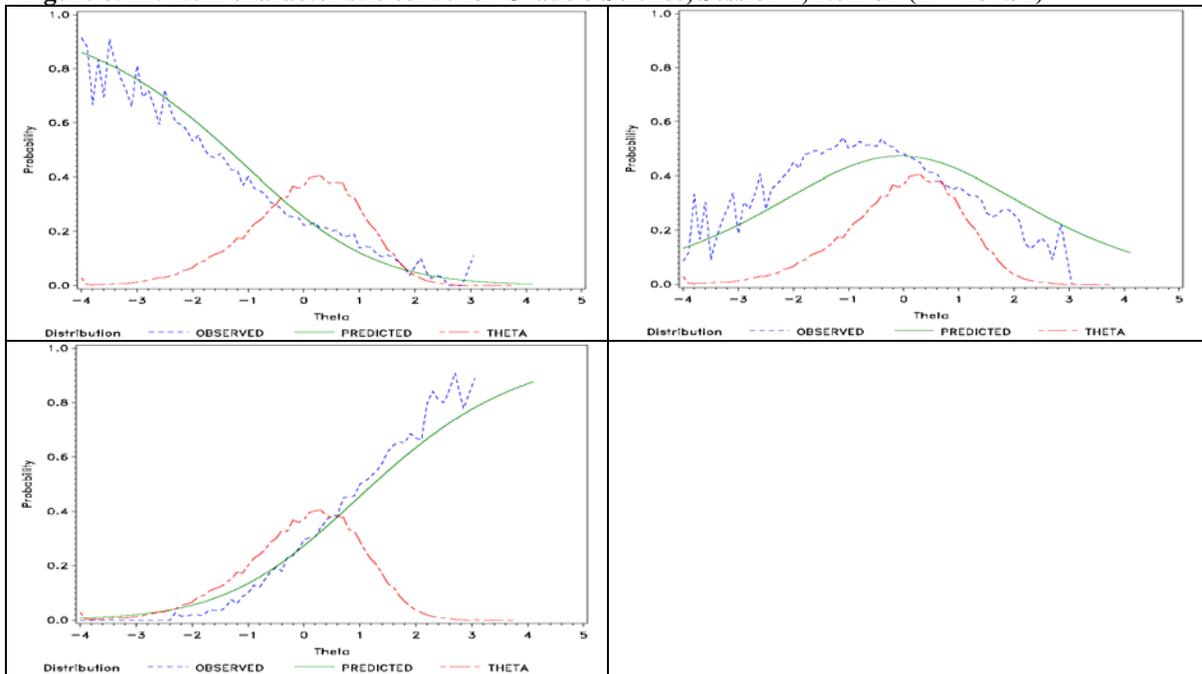


Figure 6.23: Item characteristic curve for Grade 11 Science, Session 3, Item 15 (Z = 146.34)

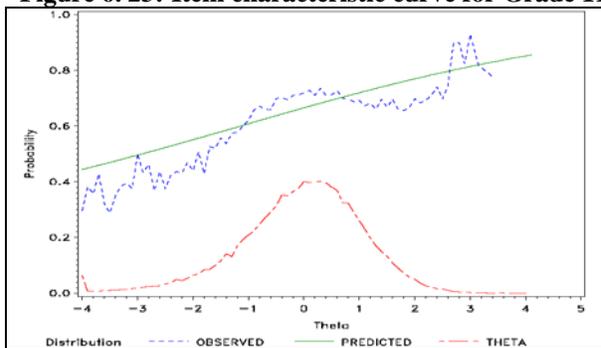


Figure 6. 24: Cross-Grade Articulation of Scale Scores at Selected Percentiles, Communication Arts MAP

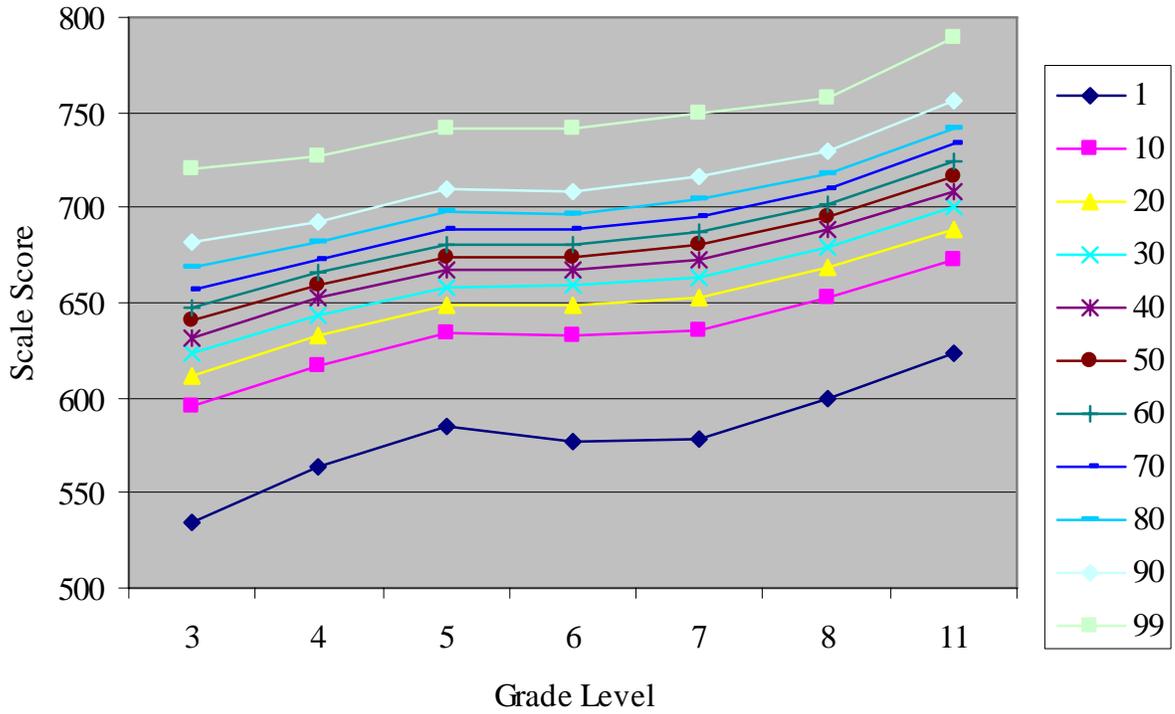


Figure 6. 25: Cross-Grade Articulation of Scale Scores at Selected Percentiles, Mathematics MAP

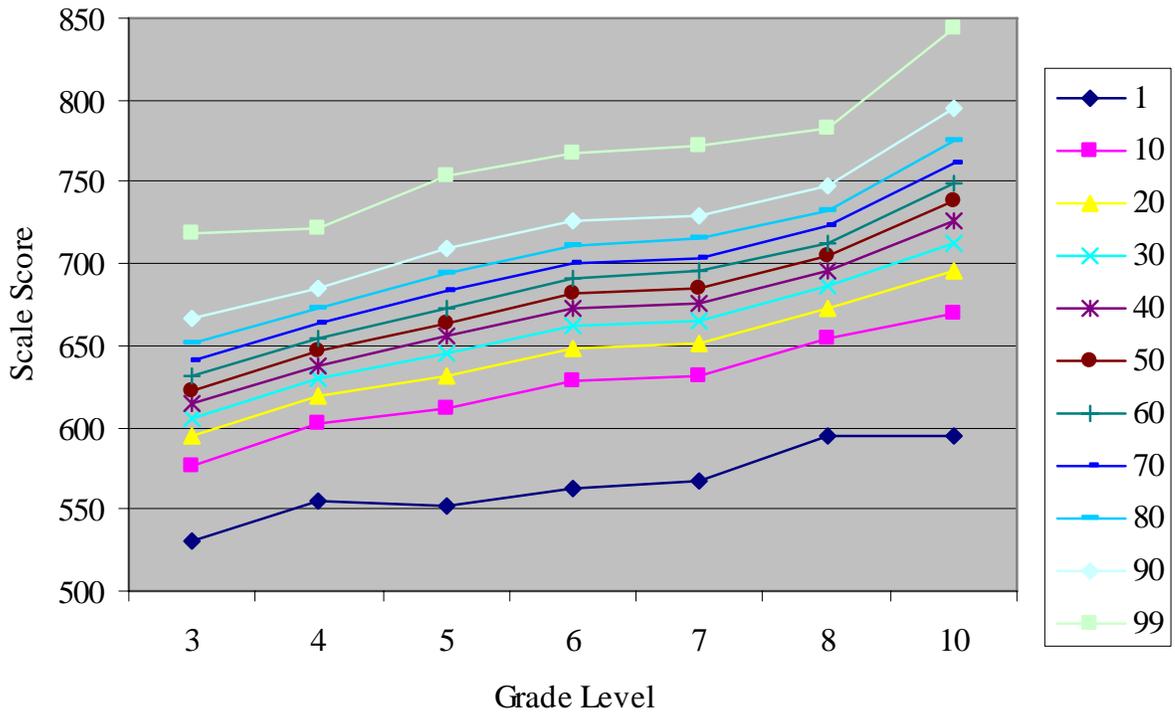


Figure 6. 26: Cross-Grade Articulation of Scale Scores at Selected Percentiles, Science MAP

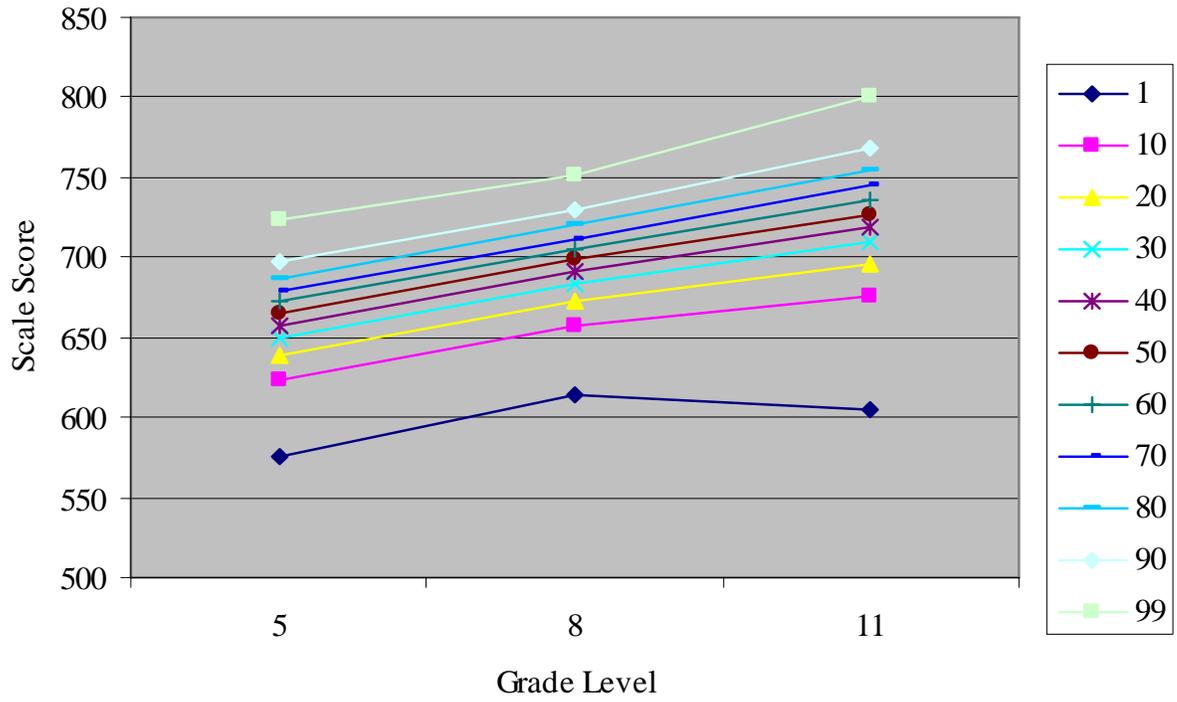


Figure 6. 27: Communication Arts Test Characteristic Curves (TCC) by grade, 2008

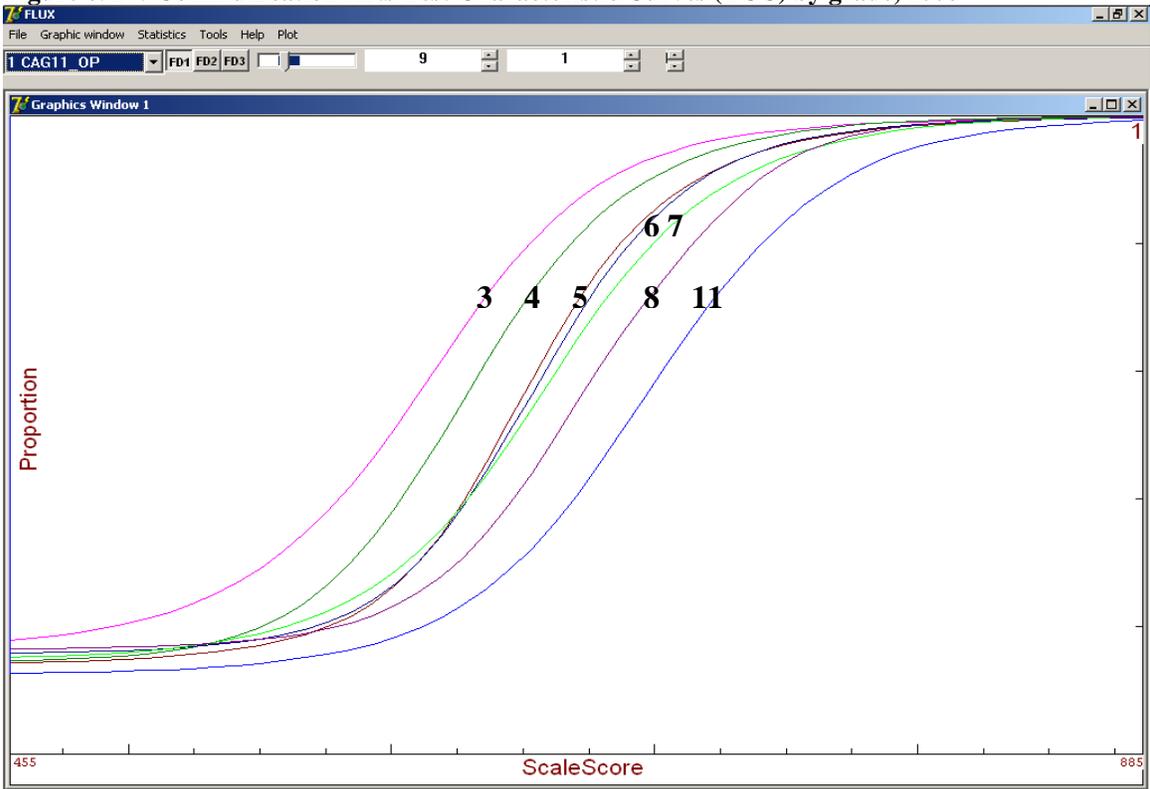


Figure 6. 28: Mathematics Test Characteristic Curves (TCC) by grade, 2008

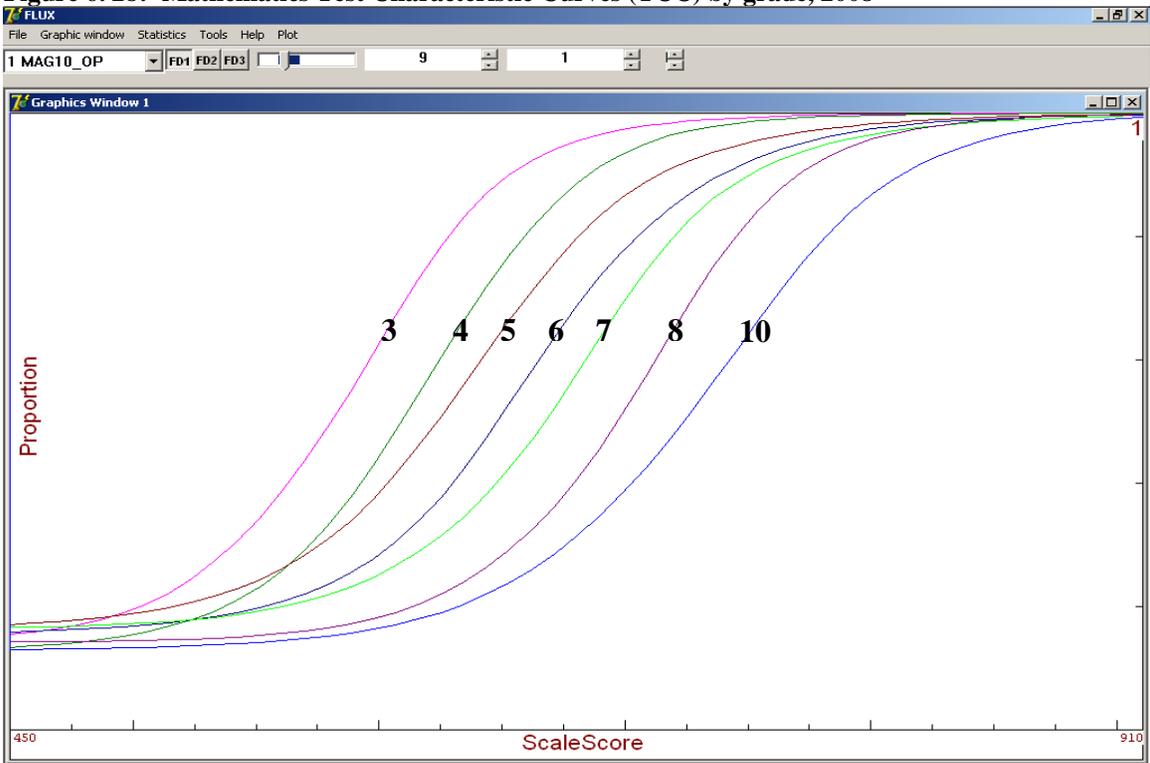
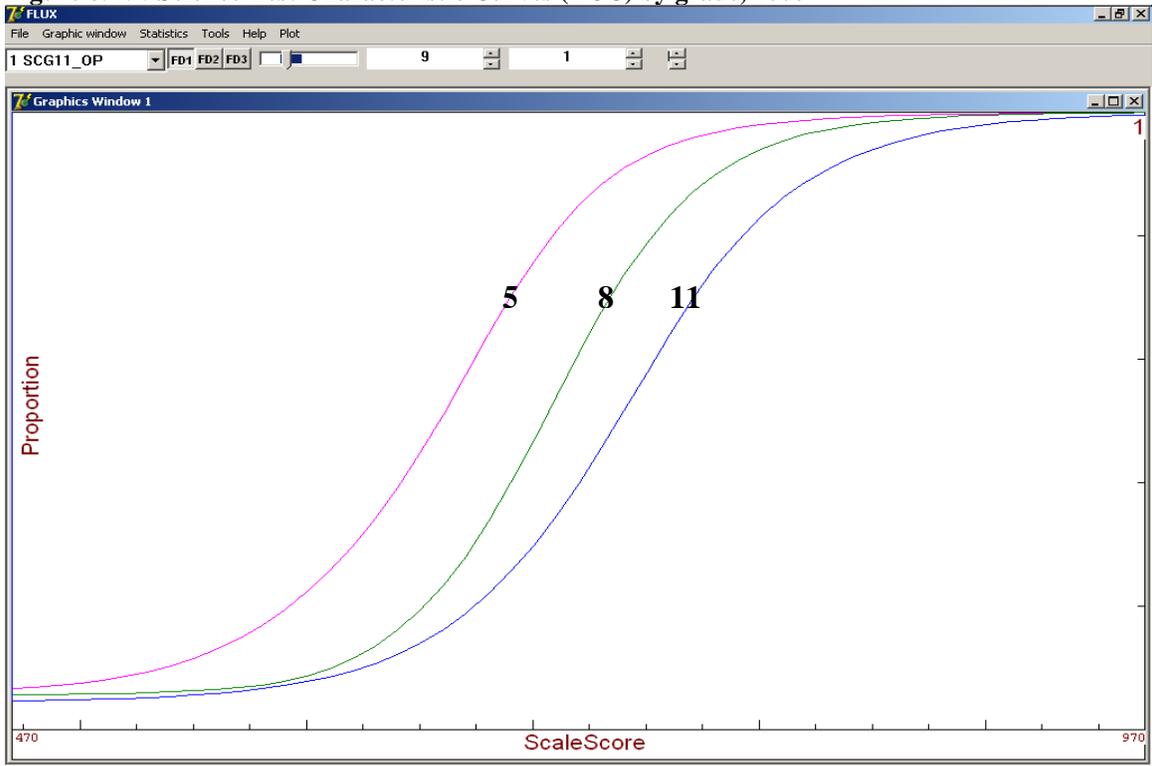


Figure 6. 29: Science Test Characteristic Curves (TCC) by grade, 2008



CHAPTER 7. TEST RESULTS

This chapter of the Technical Report contains information on the results of the Spring 2008 administration of the MAP. The scale score results are presented here. Performance level information is also provided. Presenting the results by performance level translates the quantitative scale provided through scale scores into a qualitative description of student performance: *Below Basic*, *Basic*, *Proficient*, and *Advanced*.

While the scale score provides an essential quantitative reference to student performance, the performance level information speaks directly to requirements of the NCLB Act, as well as to parents, students, and educators. Combined, scale scores, performance levels, and Lexile scores provide a comprehensive set of tools to assess Missouri student performance by content and grade level.

This chapter also provides the description on the score reports, data structure, and interpretive guide. The AERA/APA/NCME standards addressed in Chapter 6 include: 5.10, 6.1, 6.2, 6.3, 6.4, 13.15, and 13.19.

Results presented below are based on census data. The results presented here may differ slightly from the official state summary report of all student population due to ongoing resolution of test materials and student information. The results in the following tables are presented as evidence of reliability and validity of the scores from the MAP assessments, and should not be used for state accountability purposes.

7.1 Student Participation

For the MAP, the following are subgroups reported during the administration of MAP (other demographic information is collected separately and merged to MAP data after CTB sends DESE the General Research File):

- Gender: Female and Male
- Race and Ethnicity: White, Black, Hispanic, Asian/Pacific Islander, and Native American/Alaskan
- Accommodations: Students receiving testing accommodations

For the purposes of this report, participation rate is defined as the percent of students who received a valid scale score given the total number of students who received a test book. These participation rates are summarized in Tables 7.1 through 7.9. Tables 7.1 through 7.9 show both the number of students classified as reportable and those classified as accountable. Reportable students include all students with a valid scale score. Accountable students include all students for whom a test book was submitted. These are students who should have received a MAP scale score, but did not take the test and could not be assigned a scale score.

7.2 Current Administration Data

The Communication Arts assessment was administered to students in Grades 3 through 8 and 11. Mathematics was administered to students in Grades 3 through 8 and 10. MAP assessments in Science were administered to students in Grades 5, 8, and 11.

Tables 7.10 through 7.12 provide a summary of the scale scores based on the state population for the 2008 administration of the MAP. In compliance with Standard 13.19 of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999), these tables present the number of students, mean, standard deviation, and scale scores at specific percentile points.

7.3 Cross-year, Cross-sectional Comparisons

It is often desirable to examine the scores of students across time. The data in this section compare student performance on the MAP, using census data from 2006, 2007, and 2008. It should be noted that beginning in 2008, invalidated students were assigned to the lowest obtainable scale score (LOSS) and to the *Below Basic* achievement level. Prior to 2008, invalidated students did not receive a scale score.

Table 7.13 compares the state-level aggregate means. Table 7.13 shows that in most grades there was little change in the mean scale score between 2006, 2007, and 2008. The largest change in mean scale score from 2007 to 2008 occurred on the Grade 10 Mathematics test, where the mean scale score increased by 5.6 scale score points from 2007. A decrease in the mean scale score was observed in Grades 3, 4, and 11 Communication Arts and Grade 3, 4, and 6 Mathematics when 2008 scale scores were compared to 2007 scale scores.

Table 7.14 shows the percent of students in each achievement level in 2006, 2007, and 2008 on the Communication Arts test. In Grades 3, 4, 5, and 11, the percentage of students at or above Proficient decreased from 2007. It should be noted that, in Grades 4 and 5, the 2008 percentage represents an increase in the percentage of students at or above Proficient from the 2006 test. The percentages of students at or above Proficient increased from 2007 in Grads 6, 7, and 8.

Table 7.15 shows the percent of students in each achievement level in 2006, 2007, and 2008 on the Mathematics test. As compared to 2007, decreases in the percentage of students at or above Proficient were observed in Grades 3, 4, and 5; however, it should again be noted that the 2008 percentages represent an increase in the percentage of students at or above Proficient from the 2006 test. The percentages of students at or above Proficient increased from 2007 in Grades 6, 7, 8, and 10.

Table 7.16 shows the percent of students in each achievement level in 2008 on the Science test. This is the first year of the Science test; and thus no cross-year comparisons can be made.

7.4 Reports

Score reports are the primary means of communicating test scores to relevant district personnel (e.g. testing coordinators or superintendents), teachers, and parents. Standard 5.10 from the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) states:

When test score information is released to students, parents, legal representatives, teachers, clients, or the media, those responsible for testing programs should provide appropriate interpretations. The interpretations should describe in a simple language what the test covers, what scores mean, the precision of the scores, common misinterpretations of test scores, and how scores will be used.

Interpretations related to the test score are disseminated in two ways: (1) the individual score report, and (2) the *Guide to Interpreting Results* (CTB/McGraw-Hill, 2008).

The individual student report is the primary means for sharing student test results with parents. As such, it should be a stand-alone document from which parents can glean relevant information so they understand their child's test score. In 2008, the individual MAP student reports were redesigned so that they were more parent-friendly. These changes include improved interpretations of the MAP scale score, *TerraNova* scale score, and Lexile score. In addition, the state mean score is now provided as are activities that parents may engage in to help their children improve their skills within the content area in accordance with the Missouri Curriculum Framework. The new score reports also simplify the way in which the scale score and performance level are presented and interpreted. Finally, in 2008, parents no longer receive scores for content/knowledge standards or for process/performance standards.

The *Guide to Interpreting Results* is intended for use by school and district personnel so that they can interpret their score reports. It provides a context for the score reports in that it outlines the history and purpose of MAP. It also overviews the Missouri Show-Me Standards/GLE Strands. It provides greater details on the types of scores reported on the individual student report and it provides all of the abbreviated achievements level descriptors (ALD) as well as the web location of the detailed ALDs. Finally, it outlines each piece of the individual student report and overviews the student label. The *Guide to Interpreting Results* is located on DESE's website:
<http://dese.mo.gov/divimprove/assess/map/>

7.4.1. Description of Each Type of Report

In this section, descriptions for the following reports are provided: Individual Student Report, Student Score Label, online Crystal Reports, and School/District Summary Reports. Table 7.17 shows each report type and for whom the report is intended.

Individual Student Report

One copy of the Individual Student Report (ISR) is provided to schools to be sent home to the parents. On the left side of the page, results for a given content area are shown, including the student's MAP scale score, the state mean MAP scale score, the National Percentile score from the *TerraNova* section of MAP, and a brief definition of the National Percentile. On the Communication Arts ISR, the student's Lexile measure is also reported, along with a brief explanation of the Lexile measure and a website where more information may be obtained.

In the middle of the page, the student's scale score is again shown as is the achievement level associated with that scale score. This is followed by a brief explanation of what the achievement level means. When a student does not receive a scale score, then he or she will receive either "Level Not Determined" (LND) or "Invalidated" in place of the MAP scale score. For the LND students, no achievement level is assigned. Invalidated students are assigned to the LOSS and to the *Below Basic* achievement level. A brief explanation also accompanies the meaning of LND or invalidated.

On the right side of the pages are recommended activities based on the child's achievement level. These are generic activities that are targeted to all students within an achievement level, not specific activities targeted at the individual student. A sample report is provided in Appendix C, Figure C.1.

Student Score Label

The Student Score Label is designed so that each student's test results can be placed in the student's permanent record. A label is provided for every student who participated in the spring 2008 administration of the MAP. Each label has a self-adhesive backing so that it can be peeled from the sheet and placed in the student's cumulative school record. The label presents a snapshot of the student's results on the MAP. Separate labels are generated for each grade and content area; thus, a child will have multiple labels for each of the content areas administered within a grade. The label lists the student's scale score and National Percentile for each content area. For the Communication Arts MAP, the label also lists the student's Lexile measure.

CTB/McGraw-Hill provided one label per student submitted for scoring. The labels are provided in print only. A sample report is provided in Appendix C, Figure C.2

Online Crystal Reports

Schools and districts are able to access summary level reports through the online Crystal Reports tool. This tool allows district and school administrators to create on-the-fly reports containing information relevant to their data needs. There are several reporting options available through the Crystal Reports tool, including administrative reports, AYP reports, achievement level reports, content standard reports, and item analysis reports. Table 7.18 lists each of the major report headings and the sub-reports found under each reporting type.

For each sub-report, a user selects various filters such as year, grade/content area, and level of reporting (state, district, school) in order to create the desired report. For the Content Standard Reports, the user may also disaggregate results by various subgroups (e.g. race, disability).

A detailed discussion of all reports is beyond the scope of this document. Only those reports that are first-level analyses of MAP data will be discussed. The Achievement Level-5 reports will not be discussed as these are summaries of the pre-NCLB testing program. The AYP reports also will not be discussed nor will some of the Administrative Reports, including the High School Career Education Student Summary, Level Not Determined, and Map Alternate reports. Examples of all reports discussed are provided in Appendix C.

The Crystal Reports tool is accessed through DESE's website. Each school and/or district is assigned a user name and password so that it can access the site.

Administrative Reports

These reports provide student-level test data. Based on only the MAP test results, four reports are generated: MAP Scale Score Summary, MAP Student Demographic, Student Achievement Level, and Student Report.

MAP Scale Score Summary: This report lists each student in the school or district along with his/her MOSIS ID, testing year, content area, grade level, MAP scale score, achievement level, and *TerraNova* National Percentile. An example is included in Appendix C, Figure C.3.

MAP Student Demographic: This report lists all of the students in the school or district along with their date of birth (DOB), content area, CTB number, MOSIS ID, district ID, and relevant demographic information, including if the student has been in the district for less than a year, if the student has been in the building for less than a year, if the student is Limited English Proficiency (LEP), the student's race, if the student qualifies for free and reduced lunch (SES), if the student has an individualized education plan (IEP), if the student is an English-language learner (ELL)/LEP who has been in the school for less than one year and in the country for less than three years, if the student is an LEP/ELL Title 3, the number of months the LEP/ELL student has been in U.S., the student's disability diagnosis, if the student took MAP-A, and if the student is Title I. An example is included in Appendix C, Figure C.4.

Student Achievement Level: This report lists all of the students in a school or district along with the year of testing, content area, grade-level, achievement level, and MOSIS ID. An example is included in Appendix C, Figure C.5.

Student Report: For each school or district, this report contains the following information: student name, DOB, district student number, CTB student number, MOSIS ID, content area tested, grade level, and achievement level, scale score, and *TerraNova* National Percentile for each content area tested. An example is included in Appendix C, Figure C.6.

Achievement Level-4 Levels

These reports contain summary information on school or district performance in terms of the four MAP achievement levels. There are two types of achievement level reports: Achievement Level 4 Charts and Achievement Level 4 Report.

Achievement Level 4 Chart: This report charts the percentage of students classified as Proficient or Advanced across all grade levels tested in a particular content area. State-level, district-level, and/or school-level performance may be displayed on the chart. An example is included in Appendix C, Figure C.7.

Achievement Level 4 Report: This report summarizes the number and percentage of students in each achievement level. This report is comprised of 19 columns: Total, content area, grade, year, number of accountable (ACC) students, number of reportable (REP) students, number and percentage of students whose achievement level was not determined (LND), number and percentage of students classified in the Below Basic (BB) achievement level, number and percentage of students classified in the Basic (B) achievement level, number and percentage of students classified in the Proficient (P) achievement level, number and percentage of students classified in the Advanced (A) achievement level, MAP index score, mean MAP scale score, and the median *TerraNova* national percentile. The first column, Total, shows if aggregate or disaggregated information is being shown. A key to the abbreviations is found in the bottom left corner as is the computation details for the MAP Index score. An example is included in Appendix C, Figure C.8.

Content Standard

The content standard reports summarize information about the content standards.

Content Standards Report: This report has 14 columns: content area, grade level, category/type, year, percentage of points earned on content standard 1 (CS-1), points possible (PP) on CS-1, percentage of points earned on CS-2, PP on CS-2, percentage of points earned on CS-3, PP on CS-3, percentage of points earned on CS-4, PP on CS-4, percentage of points earned on CS-5, and PP on CS-5. The category/type column indicates if the data is aggregated or disaggregated data. An example is included in Appendix C, Figure C.9.

Content Standards Detail: This report shows the percentage of points each student achieved on each content standard (CS) within a particular content area. An example is included in Appendix C, Figure C.10.

Item Analysis Expanded

This set of reports provides detailed item-level results for the school or district aggregated either by the content or process standard.

Content Standard IBD EX: The Content Standard Item Benchmark Descriptions (IBD) Extended (EX) report contains item-level detail aggregated by content standard. The

report is comprised of 11 columns: school code (SC), grade level (GR), standard number and description (desc.), code for the grade-level expectation (GLE), description of the GLE, depth of knowledge (DOK) of the item, session/item number where the item was in the operational test, question type (QT), points possible for the item, average points (avg pts) earned by students in the district on that item, and percentage of points earned by the students in the district on that item. An example is included in Appendix C, Figure C.11.

Goal Process Standard IBD EX: The Goal Process Standard Item Benchmark Descriptions (IBD) Extended (EX) report contains item-level detail aggregated by the goal process standard. The report is comprised of 12 columns: school code (SC), grade level (GR), goal, standard description (desc.), code for the grade-level expectation (GLE), description of the GLE, depth of knowledge (DOK) of the item, session/item number where the item was in the operational test, question type (QT), points possible for the item, the average points (avg pts) earned by students in the district on that item, and percentage of points earned by the students in the district on that item. An example is included in Appendix C, Figure C.12.

School/District Summary Report

CTB provides DESE with school and district summary reports for each school and district in the state. These reports are intended for the sole use of DESE and are not distributed to schools and districts. These reports provide performance information for all students within a school or district who took the MAP.

The school (or district) is listed in the left-most column along with the purpose of the report along with the report purpose. The main section of the Summary Report consists of a table that divides students from the school (or district) into achievement levels. The *Reportable* column shows the number of students with valid MAP scale scores. The *Accountable* column should equal the grade-level enrollment at the time the MAP was administered.

Within both the *Reportable* and *Accountable* columns, students are categorized as *Advanced*, *Proficient*, *Basic*, or *Below Basic*. The number and percent of students falling into each achievement level is reported. A short description of the knowledge skills and abilities associated with each achievement level is also reported. Students who are not assigned to an achievement level will be classified as *Level Not Determined*. A short descriptor is also associated with this categorization.

Below this table, the norm-referenced summary statistics are reported for each school (or district). The norm-referenced information includes the National Percentile (NP) associated with the Mean Normal Curve Equivalent, the median NP, and the number of students with *TerraNova* scores.

On the back of these reports, the terms *Reportable* and *Accountable* are defined. A sample of the School/District Summary Report is provided in Appendix E, Figure E.1.7.

7.5 Data Structures

The data file (referred to as General Research File (GRF)), which contains demographic information for each student as well as item responses, raw score, content and process standard raw scores, and scale score data for each content area, was provided to DESE by CTB/McGraw-Hill. It contains one record for every test book submitted.

7.5.1. General Research File

The layout for a state level GRF is included in Appendix C.

7.6 Interpreting Test Results

Individual Student Reports and Student Labels

The *Guide to Interpreting Results* was written for Missouri teachers and administrators who receive score reports from the 2008 administration of the MAP. The *Guide to Interpreting Results* was developed collaboratively by CTB/McGraw-Hill and DESE staff. DESE staff have opportunities to review, provide feedback and have final approval.

This guide has six sections. The first section presents an overview of key terms and test-related concepts. The second section presents the Show-Me Content Standards/GLE Strands. The third section presents the Show-Me Performance Standards. The fourth section discusses assessment terms and the types of scores that will be presented on the score reports. The fifth section presents the achievement-level descriptors for all grade/content areas. Finally, the sixth section presents sample score reports.

The 2008 edition is available on the DESE website at:

<http://dese.mo.gov/divimprove/assess/map/mapgenresources.html>

Crystal Reports

Training for the Crystal Report tool is provided through DESE's Regional Instructional Facilitators as well as through online help tools. Appendix C contains an example of a training session provided by the St. Louis RIFs.

Table 7.1: Participation Rates: All Students

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	66,357	99.73%	66,357	99.85%	-	-
4	67,049	99.74%	67,049	99.84%	-	-
5	65,734	99.71%	65,734	99.85%	65,734	99.77%
6	65,830	99.76%	65,830	99.83%	-	-
7	66,923	99.67%	66,923	99.71%	-	-
8	67,574	99.56%	67,574	99.61%	67,574	99.46%
HS	61,512	99.23%	69,220	99.36%	62,700	99.10%
11b	1,309	99.24%	-	-	-	-

Table 7.2: Participation Rates: Males

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	33,965	99.75%	33,965	99.84%	-	-
4	34,395	99.72%	34,395	99.82%	-	-
5	33,457	99.66%	33,457	99.82%	33,457	99.74%
6	33,510	99.73%	33,510	99.81%	-	-
7	34,454	99.65%	34,454	99.67%	-	-
8	34,571	99.46%	34,571	99.50%	34,571	99.32%
HS	30,603	99.24%	34,941	99.23%	31,120	99.08%
11b	592	98.99%	-	-	-	-

Table 7.3: Participation Rates: Females

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	32,271	99.72%	32,271	99.87%	-	-
4	32,553	99.76%	32,553	99.87%	-	-
5	32,175	99.76%	32,175	99.89%	32,175	99.82%
6	32,239	99.80%	32,239	99.86%	-	-
7	32,362	99.70%	32,362	99.77%	-	-
8	32,871	99.68%	32,871	99.75%	32,871	99.63%
HS	30,757	99.28%	34,065	99.52%	31,407	99.16%
11b	706	99.43%	-	-	-	-

Table 7.4: Participation Rates: White

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	50,010	99.86%	50,010	99.87%	-	-
4	50,571	99.85%	50,571	99.86%	-	-
5	49,907	99.85%	49,907	99.87%	49,907	99.80%
6	50,470	99.87%	50,470	99.85%	-	-
7	50,929	99.85%	50,929	99.81%	-	-
8	51,436	99.72%	51,436	99.71%	51,436	99.62%
HS	49,534	99.44%	54,001	99.50%	49,713	99.29%
11b	275	99.27%	-	-	-	-

Table 7.5: Participation Rates: Black

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	11,961	99.78%	11,961	99.81%	-	-
4	12,197	99.73%	12,197	99.80%	-	-
5	11,732	99.72%	11,732	99.79%	11,732	99.66%
6	11,445	99.74%	11,445	99.76%	-	-
7	11,939	99.45%	11,939	99.33%	-	-
8	12,296	99.22%	12,296	99.20%	12,296	98.85%
HS	8,965	98.27%	11,469	98.78%	9,812	98.19%
11b	859	99.30%	-	-	-	-

Table 7.6: Participation Rates: Hispanic

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	2,737	98.28%	2,737	99.78%	-	-
4	2,700	98.48%	2,700	99.89%	-	-
5	2,478	98.63%	2,478	99.84%	2,478	99.80%
6	2,262	98.81%	2,262	99.91%	-	-
7	2,349	98.04%	2,349	99.45%	-	-
8	2,181	98.53%	2,181	99.72%	2,181	99.36%
HS	1,429	99.37%	1,967	99.14%	1,566	99.17%
11b	136	98.53%	-	-	-	-

Table 7.7: Participation Rates: Asian/Pacific Islander

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	1,265	97.31%	1,265	99.92%	-	-
4	1,214	97.78%	1,214	99.75%	-	-
5	1,218	96.47%	1,218	100.00%	1,218	99.92%
6	1,279	97.58%	1,279	99.92%	-	-
7	1,325	97.96%	1,325	100.00%	-	-
8	1,195	98.33%	1,195	99.75%	1,195	99.92%
HS	1,126	99.20%	1,228	99.59%	1,163	99.31%
11b	31	100.00%	-	-	-	-

Table 7.8: Participation Rates: Native American/Alaskan

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	259	99.61%	259	99.61%	-	-
4	262	100.00%	262	99.62%	-	-
5	281	99.64%	281	100.00%	281	100.00%
6	289	100.00%	289	99.65%	-	-
7	293	99.32%	293	99.32%	-	-
8	331	99.70%	331	100.00%	331	99.40%
HS	276	99.28%	332	99.40%	270	99.26%
11b	1	100.00%	-	-	-	-

Table 7.9: Participation Rates: Students Receiving Accommodations

Grade	Accountable in Comm. Arts	Percent Reportable in Comm. Arts	Accountable in Mathematics	Percent Reportable in Mathematics	Accountable in Science	Percent Reportable in Science
3	6,185	99.85%	6,398	99.80%	-	-
4	7,160	99.89%	7,382	99.95%	-	-
5	7,496	99.91%	7,759	99.96%	7,440	99.83%
6	7,079	99.89%	7,269	99.78%	-	-
7	7,048	99.76%	7,264	99.77%	-	-
8	6,598	99.35%	6,887	99.62%	6,604	99.33%
HS	4,800	99.46%	5,934	99.66%	4,738	100.00%
11b	55	100.00%	-	-	-	-

Table 7.10: Summary Statistics for Communication Arts

Grade	N	Mean	Std. Dev.	Scale Scores by Percentiles				
				10	25	50	75	90
3	66,179	637.60	37.54	593	617	640	661	681
4	66,873	655.61	33.63	617	638	658	677	693
5	65,544	671.48	33.71	633	654	673	692	710
6	65,672	671.27	33.50	633	654	674	692	708
7	66,701	675.87	35.08	633	657	679	699	716
8	67,278	691.05	33.57	651	673	694	713	729
11	61,041	713.56	35.89	670	694	716	737	755
11b	1,299	693.14	43.28	642	671	697	722	743

Table 7.11: Summary Statistics for Mathematics

Grade	N	Mean	Std. Dev.	Scale Scores by Percentiles				
				10	25	50	75	90
3	66,258	621.65	36.92	577	600	623	644	664
4	66,944	644.18	34.19	602	624	646	666	684
5	65,636	661.43	40.73	612	638	664	687	708
6	65,716	678.46	41.13	629	656	681	705	726
7	66,727	681.15	41.38	630	658	684	708	729
8	67,312	701.30	39.40	653	679	704	727	747
10	68,776	729.29	49.55	666	700	733	763	787

Table 7.12: Summary Statistics for Science

Grade	N	Mean	Std. Dev.	Scale Scores by Percentiles				
				10	25	50	75	90
5	65,586	661.64	31.52	623	644	665	683	697
8	67,209	694.36	30.67	655	677	698	716	730
11	62,133	722.19	40.11	673	701	726	749	768

Table 7.13: Comparison of State-Level Means, 2006, 2007, and 2008 Census Data

Grade	Year	Communication Arts			Mathematics			Science		
		N	Mean SS	S.D. SS	N	Mean SS	S.D. SS	N	Mean SS	S.D. SS
3	2006	64,486	639.86	36.84	64,763	621.59	39.11			
	2007	66,347	639.58	38.04	66,640	622.40	38.72			
	2008	66,179	637.60	37.54	66,258	621.65	36.92			
4	2006	65,179	654.55	38.56	65,306	643.88	37.07			
	2007	65,274	656.11	39.51	65,363	644.47	36.56			
	2008	66,873	655.61	33.63	66,944	644.18	34.19			
5	2006	66,007	668.18	37.09	66,123	660.06	39.99			
	2007	65,461	671.01	37.14	65,498	663.21	41.50			
	2008	65,544	671.48	33.71	65,636	661.43	40.73	65,586	661.64	31.52
6	2006	66,948	666.85	33.70	67,017	673.30	39.80			
	2007	66,247	667.99	34.63	66,332	676.31	41.75			
	2008	65,672	671.27	33.50	65,716	678.46	41.13			
7	2006	70,290	671.63	37.06	70,698	675.38	41.27			
	2007	67,167	672.11	36.26	67,554	677.41	42.62			
	2008	66,701	675.87	35.08	66,727	681.15	41.38			
8	2006	72,483	686.85	37.87	72,542	697.73	40.37			
	2007	70,187	686.90	37.54	70,204	698.33	41.98			
	2008	67,278	691.05	33.57	67,312	701.30	39.40	67,209	694.36	30.67
High School	2006	60,004	716.69	31.42	68,083	724.46	51.18			
	2007	61,482	715.86	31.32	68,330	723.69	47.83			
	2008	61,041	713.56	35.89	68,776	729.29	49.55	62,133	722.19	40.11

**Table 7.14: Comparison of Percent of Students in each Achievement Level, Communication Arts
2006, 2007, and 2008 Census Data**

Grade	Year	N	No Level	Below Basic	Basic	Proficient	Advanced	Prof & Adv
3	2006	65,344	1.3	8.8	47.5	25.7	16.7	42.4
	2007	67,259	1.4	9.4	46.6	25.8	16.8	42.6
	2008	66,357	0.3	9.3	50.2	25.2	15.1	40.3
4	2006	65,849	1.0	10.6	44.5	28.8	15.0	43.8
	2007	65,982	1.1	10.5	43.4	28.2	16.8	45.1
	2008	67,049	0.3	8.0	46.7	33.4	11.7	45.1
5	2006	66,704	1.0	9.1	44.8	29.6	15.4	45.0
	2007	66,098	1.0	8.3	42.9	29.8	18.0	47.8
	2008	65,734	0.3	6.4	45.1	32.2	15.9	48.1
6	2006	67,709	1.1	11.9	44.8	31.6	10.6	42.2
	2007	67,045	1.2	11.2	44.0	31.8	11.7	43.6
	2008	65,830	0.2	9.0	43.5	34.0	13.4	47.4
7	2006	71,632	1.9	13.7	41.8	30.5	12.2	42.7
	2007	68,404	1.8	13.1	40.7	32.8	11.6	44.4
	2008	66,923	0.3	10.0	40.7	36.1	12.9	49.0
8	2006	73,516	1.4	9.1	48.0	26.6	15.0	41.5
	2007	71,200	1.4	8.7	48.3	26.9	14.6	41.6
	2008	67,574	0.4	5.7	45.8	33.1	15.0	48.1
11	2006	61,593	2.6	10.1	46.1	30.7	10.6	41.3
	2007	62,946	2.3	10.2	46.9	31.5	9.0	40.6
	2008	61,512	0.8	13.8	46.2	27.5	11.7	39.2

Table 7.15: Comparison of Percent of Students in each Achievement Level, Mathematics 2006, 2007, and 2008 Census Data

Grade	Year	N	No Level	Below Basic	Basic	Proficient	Advanced	Prof & Adv
3	2006	65,325	0.9	7.2	48.7	33.3	10.0	43.3
	2007	67,257	0.9	7.2	46.9	35.0	10.0	45.0
	2008	66,357	0.1	6.5	49.6	35.0	8.8	43.8
4	2006	65,845	0.8	8.3	47.5	34.4	9.0	43.4
	2007	65,975	0.9	8.1	46.5	35.2	9.3	44.5
	2008	67,049	0.2	7.6	48.0	36.0	8.2	44.2
5	2006	66,703	0.9	8.1	47.8	32.7	10.6	43.3
	2007	66,075	0.9	7.6	44.9	33.1	13.4	46.6
	2008	65,734	0.1	7.5	46.5	34.4	11.4	45.8
6	2006	67,706	1.0	11.1	44.1	34.4	9.5	43.9
	2007	67,039	1.1	11.1	40.0	35.5	12.3	47.8
	2008	65,830	0.2	9.5	39.6	37.8	12.9	50.7
7	2006	71,575	1.2	17.4	38.5	32.7	10.2	42.9
	2007	68,405	1.2	16.7	37.1	33.2	11.7	44.9
	2008	66,923	0.3	13.9	36.3	36.7	12.8	49.5
8	2006	73,523	1.3	21.1	37.8	27.6	12.2	39.8
	2007	71,190	1.4	21.4	36.6	26.6	14.0	40.6
	2008	67,574	0.4	18	37.7	29.9	13.9	43.8
10	2006	69,229	1.7	24.1	32.7	31.4	10.1	41.5
	2007	69,488	1.7	23.8	34.5	32.1	8.0	40.1
	2008	69,220	0.6	21.7	32.1	34.3	11.3	45.6

Table 7.16: Percent of Students in each Achievement Level, Science 2008 Census Data

Grade	Year	N	No Level	Below Basic	Basic	Proficient	Advanced	Prof & Adv
5	2008	65,734	0.2	11.2	44.0	29.6	14.9	44.5
8	2008	67,574	0.5	19.3	37.0	36.7	6.5	43.2
11	2008	62,700	0.9	24.9	27.0	38.3	8.9	47.2

Table 7.17: Summary of Score Reports for Spring 2008

Score Report		Paper Report		Electronic Report		
		Parent	Teacher	Principal	System	DESE
Student Score Labels			X			
Individual Student Report		X				
Performance Summary Report	School Performance Summary Report					X
	District Performance Summary Report					X
	Crystal Reports			X	X	

Table 7.18: Types of Reports Available to Districts through Crystal Reports

Crystal Report	Sub Reports
Administrative Report	HS Career Ed Student Summary Level Not Determined MAP Alternate MAP Scale Score Summary MAP Student Demographic Student Achievement Level Student Report
AYP	AYP AYP Additional Indicator AYP Growth Report AYP Summary
Achievement Level-4 Levels	Achievement Level 4 Chart Achievement Level 4 Report
Achievement Level-5 Levels	Achievement Level 5 Chart Top2 Bottom2 Achievement Level 5 Chart Top2 Bottom2 State District Achievement Level 5 Report MAP Index Chart 5 Levels MAP Index Chart 5 Levels District State
Content Standard	Content Standards Content Standards Detail
Item Analysis Expanded	Content Standard IBD EX Goal Process Standard IBD EX

CHAPTER 8: ACHIEVEMENT-LEVEL SETTING

A Bookmark standard setting was held in 2005 to establish cut scores for the Missouri Assessment Program (MAP) for the Communication Arts and Mathematics MAP. Another Bookmark standard setting was held in 2008 to establish cut scores for the Science MAP. In this chapter, we briefly describe the MAP achievement-level setting and we present the cut scores established and the achievement-level descriptors derived from the achievement-level setting. A detailed discussion of the Communication Arts and Mathematics achievement-level setting may be found in the *Missouri Assessment Program Final Bookmark Standard Setting Technical Report* (2005). A detailed discussion of the Science achievement-level setting may be found in the *Missouri Assessment Program Bookmark Standard Setting Technical Report 2008 for Missouri Achievement-Level Setting Grades 5, 8, and 11 Science* (2008).

In describing the standard setting processes for MAP, chapter 8 functions to address Standard 4.19 of the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999): “When proposed score interpretation involves one or more cut scores, the rationale and procedures used for establishing cut scores should be clearly documented.” In terms of the validity of the MAP, it is essential to understand that descriptors and cut scores are established in a collaborative, participatory process, largely driven by the input of Missouri teachers and educators. In addition, as cited in the *Standards*, validity extends to the interpretation of test scores. The descriptors clearly establish, in plain language, the proper frame of reference for understanding how to interpret test scores, and cut scores in particular.

8.1. Legislation Affecting MAP Standard Setting

A modified Bookmark Standard Setting Procedure (BSSP) was used to establish cut scores for the Communication Arts and Mathematics MAPs for Grades 3 through 8 and high school and Science for Grades 5, 8, and 11. A modification of Bookmark was used to meet the requirements of Senate Bill 1080, which requires that NAEP-like cut scores be established for the MAPs.

Senate Bill 1080 was interpreted such that the *Proficient* achievement level met, but did not exceed, the NAEP performance standards. In other words, the percentage of students who attain *Proficient* on the MAP should be similar to or slightly higher than the percentage attaining *Proficient* on NAEP. The percentage of students in the other three achievement levels would be allowed to vary between NAEP and the MAP.

For the purposes of the MAP standard setting, participants were allowed to recommend *Proficient* cut scores within a pre-specified range. This range was based on the percentage of students who could be classified as either *Proficient* or *Advanced*. For Communication Arts and Mathematics, no fewer than 26% of students could be classified as either *Proficient* or *Advanced*, and no more than 44% of students could be classified as *Proficient* or *Advanced*. For Science, no fewer than 27% of students could be classified as

either *Proficient* or *Advanced*, and no more than 48% of students could be classified as *Proficient* or *Advanced*.

The pre-specified range was determined using the results from NAEP and MAP. For all three subject areas, the high end of the range (in terms of scale score points) was based on NAEP results. This was the lowest percentage of students classified as *Proficient* or *Advanced* on the NAEP test for Grades 4 and 8 Reading, Mathematics, and Science using both national and state data.

The low end of the range (in terms of scale score points) was based on the 2005 MAP results for the Communication Arts and Mathematics standard setting and was based on the 2007 MAP results for Science. This was the highest percentage of students classified as *Proficient* or *Advanced* on the previous years' tests.

8.2. Bookmark Standard Setting Procedure

A modified BSSP was used to establish cut scores on the Communication Arts, Mathematics, and Science MAP. At both workshops, the BSSP involved three rounds of discussion and voting. Here, we discuss the major activities of the three rounds. The interested reader is referred to the specific Technical Reports available for each standard setting for a detailed discussion of the procedure.

Round 1: In this round, panelists discussed target students (the student for whom they were placing cut scores), took the test, studied and discussed the test items in order of difficulty, and made initial recommendations of cut scores.

Round 2: In this round, panelists were shown their Round 1 recommendations and the percentage of students in each achievement level as a result of their Round 1 recommendations. They discussed their Round 1 recommendations for cut scores and made another recommendation based on their Round 2 discussions.

Round 3: In this round, panelists were shown their Round 2 recommendations and the percentage of students in each achievement level as a result of their Round 2 recommendations. They discussed their Round 2 recommendations for cut scores and made another recommendation based on their Round 3 discussions.

Following Round 3, panelists wrote draft achievement- level descriptors which were later edited by CTB and DESE staff.

The Missouri State Board of Education approved the cut scores as recommended by the standard setting panelists.

8.3. Cut Scores

In this section, we present the cut scores for each grade/content area of MAP. Table 8.1 shows the cut scores for Communication Arts. Table 8.2 shows the cut scores for Mathematics. Table 8.3 shows the cut scores for Science.

Table 8. 1: Communication Arts Cut Scores

Grade	Cut Scores		
	Basic	Proficient	Advanced
3	592	648	673
4	612	662	691
5	625	675	702
6	631	676	704
7	634	680	712
8	639	696	723
11	679	725	753

Table 8. 2: Mathematics Cut Scores

Grade	Cut Scores		
	Basic	Proficient	Advanced
3	568	628	667
4	596	651	688
5	605	668	706
6	628	681	721
7	640	685	724
8	670	710	741
11	695	738	785

Table 8. 3: Science Cut Scores

Grade	Cut Scores		
	Basic	Proficient	Advanced
5	626	669	692
8	671	703	735
11	702	729	770

8.4. Achievement-Level Descriptors

In this section, we present the short achievement-level descriptors that were drafted during the standard setting and finalized between CTB and DESE staff after the standard setting.

Communication Arts

GRADE 3—SHORT DESCRIPTORS

Below Basic

Reading—Students locate information in text; identify an obvious main idea; define simple words and phrases. Writing—Students show minimal awareness of beginning, middle, end, audience, purpose and controlling idea; attempt to create friendly letters; use graphic organizers.

Basic

Reading—Students make simple comparisons; recall simple sequence of events; make obvious inferences and predictions; use context clues to determine word meaning.

Writing—Students use basic parts of speech correctly in simple sentences; show minimal awareness of beginning, middle, end, audience, purpose and controlling idea.

Proficient

Reading—Students locate/identify supporting details, obvious cause and effect; make inferences; use context clues to determine word meaning; make comparisons; recall detailed sequence of events; identify solutions and fact vs. fiction; recognize figurative language; draw obvious conclusions. Writing—Students generally use rules of Standard English; show awareness of audience, purpose, controlling idea, relevant details, beginning, middle and end.

Advanced

Reading—Students identify relevant/supporting information to make predictions and draw conclusions; infer word meaning; infer main idea; make complex comparisons; make complex inferences; categorize information; identify correct sequence of events.

Writing—Students consistently apply rules of Standard English; construct complex sentences; use details effectively; have a clear controlling idea, awareness of audience and purpose, beginning, middle and end.

GRADE 4—SHORT DESCRIPTORS

Below Basic

Reading—Students locate information in text; recall stated information; draw obvious conclusions; make simple comparisons and descriptions. Writing—Students write simple letters, minimally use the rules of Standard English; attempt to organize information.

Basic

Reading—Students identify appropriate details; use context clues; make obvious inferences; select vocabulary using context clues. Writing—Students write simple letters with an awareness of an intended audience and purpose; generally use the rules of Standard English.

Proficient

Reading—Students make simple inferences; recall, identify, and use relevant information; draw conclusions; explain figurative language and main idea; use context clues to select vocabulary; identify character traits, sensory details, and simple cause and effect. Writing—Students show organization and awareness of an intended audience and purpose; use the rules of Standard English; use a writing process to revise, edit, and proofread.

Advanced

Reading—Students make complex inferences and comparisons; evaluate simple information; infer cause/effect and word meaning; interpret figurative language; identify author's purpose; identify complex problems/solutions; explain complex main ideas. Writing—Students consistently use the rules of Standard English.

GRADE 5—SHORT DESCRIPTORS

Below Basic

Reading—Students locate/identify information in text; draw simple conclusions; make obvious inferences and predictions; identify character traits. Writing—Students use correct letter writing format; partially organize information.

Basic

Reading—Students identify supporting details, problems/solutions; use context clues; make obvious inferences; give partial summary of action. Writing—Students edit for Standard English.

Proficient

Reading—Students interpret figurative language; infer main idea; identify author's purpose, point of view, the sequence of information, cause/effect, the meaning of vocabulary; summarize; distinguish between fact and opinion; draw conclusions; make inferences and comparisons; support a position. Writing—Students use the rules of Standard English; construct complex sentences; edit for appropriate support, organize information.

Advanced

Reading—Students interpret and draw conclusions from complex information; analyze complex characters; infer author's purpose and word meaning; categorize information; make simple evaluations and judgments; determine the appropriateness of a source and the accuracy of information. Writing—Students consistently use the rules of Standard English; use a writing process to organize information.

GRADE 6—SHORT DESCRIPTORS

Below Basic

Reading—Students locate/identify information in text; make simple inferences; identify main idea, sensory information, figurative language, simple problems or solutions.

Writing—Students show awareness of audience and letter format; use simple organizational techniques and graphic organizers; use simple rules of Standard English.

Basic

Reading—Students identify supporting information, simple cause/effect relationships, conflicts, point of view and problem-solving processes. Writing—Students use correct letter writing format; generally use the rules of Standard English including spelling; revise; have a controlling idea.

Proficient

Reading—Students identify author's purpose, supporting details, point of view; describe character traits, plot; identify problems/solutions; support a position with text-based details; draw conclusions; interpret figurative language; make inferences and predictions; locate resources. Writing—Students use the rules of Standard English; construct complex sentences; write for an audience and purpose; organize information.

Advanced

Reading—Students make complex connections; analyze complex characters; evaluate the accuracy and importance of information; draw conclusions and make inferences from complex information, analyze complex characters; determine cause and effect; paraphrase. Writing—Students demonstrate consistent use of a controlling idea and Standard English.

GRADE 7—SHORT DESCRIPTORS

Below Basic

Reading—Students locate and apply information in text; identify figurative language, text elements, and problems/solutions, character traits; make obvious predictions. Writing—Students organize information; use some components of letter writing format; generally stay on topic; show awareness of audience and purpose; minimally use rules and conventions of Standard English.

Basic

Reading—Students identify text-based details; identify main idea; make simple summaries; identify the meaning of figurative language; draw simple conclusions; make simple inferences. Writing—Students use a writing process; edit for appropriate support; revise for a controlling idea; generally use the rules of Standard English.

Proficient

Reading—Students make inferences; summarize; make comparisons and predictions using complex text; analyze characters; determine word meaning, point of view, supporting information; locate resources. Writing—Students stay on topic; write for a specific audience and purpose; demonstrate consistent use of a controlling idea; use rules and conventions of Standard English; use complex sentences, cohesive devices, clear and varied sentences.

Advanced

Reading—Students interpret complex figurative language and vocabulary; support a position; make predictions; summarize, analyze, and synthesize information and techniques; paraphrase ideas. Writing—Students consistently use the rules and conventions of Standard English; use logical order, cohesive devices, clear and varied sentences, writing techniques; target specific audience and purpose.

GRADE 8—SHORT DESCRIPTORS

Below Basic

Reading—Students identify author's purpose, figurative language, plot, and setting; use context clues to choose vocabulary. Writing—Students create a graphic organizer; write a basic paragraph; show some awareness of audience.

Basic

Reading—Students define simple vocabulary; identify main idea; draw simple conclusions; make simple inferences; recall details from text; determine reliability of resources. Writing—Students write a paragraph to a specific audience.

Proficient

Reading—Students summarize; infer vocabulary meaning and cause/effect; interpret figurative language; analyze text features; follow multi-step directions; identify author's technique; analyze text; make inferences, interpretations, predictions, comparisons, using complex material; evaluate evidence, reliability of resources. Writing—Students edit for relevant details and purpose; organize and edit text; consistently use rules/conventions of Standard English.

Advanced

Reading—Students analyze complex information, author's purpose, characters; synthesize information; summarize complex ideas; make complex inferences. Writing—Students edit text correctly applying the rules/conventions of Standard English.

GRADE 11—SHORT DESCRIPTORS

Below Basic

Reading—Students identify figurative language, plot, setting, and main idea; explain main idea; determine pertinent details. Writing—Students use correct letter writing format; address a topic with some detail; defend a position; create and label an appropriate graphic organizer.

Basic

Reading—Students determine pertinent details; categorize information; determine reliability of information; explain cause/effect; make simple comparisons; infer meaning; identify main idea; analyze figurative language. Writing—Students attempt to use the rules/conventions of Standard English; create and label an appropriate graphic organizer; compose an essay with a beginning, middle, and end.

Proficient

Reading—Students analyze significant details, figurative language, use of information, characters, reasoning, author's purpose, cause/effect; infer meaning; summarize complex ideas/information; determine tone; identify supporting information; make sophisticated comparisons; evaluate style; distinguish fact/opinion. Writing—Students use the rules/conventions of Standard English; write with a controlling idea for a specific audience; use cohesive devices and complex sentences.

Advanced

Reading—Students evaluate reliability of sources; analyze complex figurative language; infer connections; evaluate adequacy of support; make inferences and comparisons from complex text. Writing—Students consistently apply the rules/conventions of Standard English; follow a writing process to compose a well-developed and organized essay with a controlling idea, relevant details; use precise language.

Mathematics

GRADE 3—SHORT DESCRIPTORS

Below Basic

Students use multiplication to model situations; recognize that addition and subtraction are inverse operations; add 2-digit numbers; apply subtraction skills; extend shapes or numbers in a pattern; use number sentences to model situations; use transformations to check congruency of shapes; recognize a line of symmetry; use an appropriate unit on a ruler to measure length; estimate length; interpret information from graphs.

Basic

Students estimate with less-than and greater-than; sort items by size; apply regrouping for adding and subtracting 3-digit numbers; order 3-digit whole numbers; count using numbers and pictures; identify and explain a pattern; use an appropriate unit of measurement; read thermometers; read analog clocks to nearest 5 minutes; use a ruler to measure to the nearest centimeter; compare data; transfer data to graphs.

Proficient

Students identify odd/even numbers; locate landmark numbers; describe change using increase/decrease; perform basic division of 2-digit whole numbers; identify and locate fractional parts; set up/solve simple word problems; recognize 2-D and 3-D shapes; combine 3-D solids; identify 2-D faces of 3-D objects; determine perimeter of polygons; identify appropriate units of measure; add monetary values up to \$5.00; use calendars to determine dates; estimate length with fractions.

Advanced

Students estimate and justify results of addition/subtraction of numbers; represent a mathematical situation as a number sentence or an expression; identify multiple lines of symmetry; determine change from \$5.00 including different combinations of coins; predict events as likely or unlikely.

GRADE 4—SHORT DESCRIPTORS

Below Basic

Students write and compare decimals to the hundredths place; identify fraction as a part of a whole; describe the results of combining shapes; identify parallel lines; estimate linear measurements; read and compare data on a bar graph; complete tables; create tables or graphs to represent data.

Basic

Students use multiplication to solve problems; analyze patterns using words, tables, and graphs; identify the missing value in a number sentence; identify 2-D and 3-D shapes and attributes; identify the results of transformations; tell time to the nearest minute; use benchmarks to estimate linear measurements; transfer numerical data to a graph; propose and justify conclusions that are based on data.

Proficient

Students compare parts of a whole as fractions; identify place value up to 6-digit whole numbers; decompose/compose whole numbers; represent multiplication using sets/arrays; divide 3-digit by 1-digit numbers; write a number sentence; describe movement on grid using geometric vocabulary; identify lines of symmetry; use standard/metric units to measure; add/subtract money values to \$10.00; determine area on grid; read/interpret data on a line plot; analyze and explain data.

Advanced

Students describe constant rates of change; identify strategies to solve problems; describe numeric and geometric patterns; solve problems using graphs, tables, or number sentences; construct a figure with one line of symmetry; determine differences in measures; estimate measurement of angles; determine change from \$10.00; identify equivalent linear measures within a system; count combinations of items.

GRADE 5—SHORT DESCRIPTORS

Below Basic

Students recognize equivalent representations of numbers by composing and decomposing numbers up to 5 digits; order decimals to thousandths place; interpret place value to hundred-thousands; determine operations used in numeric patterns; use symmetry to complete figures; make generalizations about geometric patterns; describe attributes of 2-D shapes; identify data on a line graph; make and justify predictions using data; describe, compare, and organize data in a bar graph.

Basic

Students identify place value to the millions place; read, write, and compare unit fractions and decimals to the thousandths place; identify lines of symmetry; identify appropriate units of area; identify appropriate units of measure; use data to create a bar graph and perform calculations using numbers between given intervals.

Proficient

Students multiply decimals to the hundredths place; use estimation in computations; divide 3-digit by 2-digit numbers; add fractions with like denominators; solve problems involving rates of change; extend numeric patterns; complete number sentences; identify faces of 3-D and similar figures; interpret direction on a coordinate grid; calculate area using a grid; compute elapsed time in hours; analyze data in line graphs and tables; explain the probability of a simple event.

Advanced

Students use addition/subtraction of money in a real-world situation; explain and justify the results of calculations; justify and model the results of calculations involving constant rates; use number sentences to model a mathematical situation; analyze characteristics of and identify 3-D figures, quadrilaterals, and angle measures; use a coordinate grid to describe paths and determine distances between points; convert between standard units of measurement.

GRADE 6—SHORT DESCRIPTORS

Below Basic

Students compare and order integers, positive rational numbers, and percents; describe patterns in tables and pictures; identify properties of 2-D and 3-D shapes; identify acute, obtuse, or right angles; identify transformations of 2-D shapes; identify equivalent algebraic expressions using the associative property; read and interpret line and circle graphs.

Basic

Students generate equivalent forms of percents, fractions and decimals; determine a rule for a geometric or numeric pattern; use coordinate geometry to construct and identify 2-D shapes using ordered pairs; use models to compare and explain probabilities; estimate and interpret data in graphs.

Proficient

Students add/subtract positive rational numbers; identify least common multiple and greatest common factor; estimate quotients; determine rate of increase; analyze rates of change; use variables; compare spatial views of 3-D objects; construct polygons; describe transformations; determine area of rectangles; measure angles; convert within a system of measure; interpret and complete a table based on probability; compare/explain data; calculate measures of center.

Advanced

Students estimate and convert measurements; describe solutions to algebraic equations; recognize similarities between 2-D shapes; use properties of basic figures to draw conclusions about angle size; determine area of triangles; solve elapsed time problems; apply formula for perimeter; estimate area of a figure using a coordinate grid; interpret stem-and-leaf plots; determine appropriate data collection methods and questions; interpret data to solve problems.

GRADE 7—SHORT DESCRIPTORS

Below Basic

Students place integers on a number line; identify shapes from a group of 2-D shapes based on a common property; transform 2-D shapes; analyze precision and accuracy using measurement tools; identify unit of measure for volume; interpret bar graphs; use representations of data from bar graphs, circle graphs, stem-and-leaf plots, and box-and-whisker plots; predict outcomes using probability.

Basic

Students multiply and divide positive rational numbers; identify bases and exponents of numbers in exponential form; recognize equivalent numerical representations; solve 2-step problems; use variables to solve inequalities and equations; analyze patterns represented numerically or graphically; read and interpret graphs.

Proficient

Students read/write numbers up to hundred-millions place; compare integers, rational numbers, percents; perform operations with mixed numbers; use circle graphs to recognize relationship of parts to whole; solve fraction/decimal/percent problems; solve proportion/scale problems; use models to solve problems; model with equations; describe and classify 2-D/3-D shapes; apply spatial reasoning to estimate area; solve time problems; solve area problems; calculate measures of center.

Advanced

Students calculate totals involving percents in multi-step problems; extend non-linear patterns; model with inequalities; apply the relationship of corresponding and similar angles; use scale factors on a grid to dilate shapes; describe corresponding angles and sides of similar polygons; solve problems using time conversions; find circumference and area of circles; make conversions using proportions.

GRADE 8—SHORT DESCRIPTORS

Below Basic

Students generalize numeric patterns; generalize relationships between attributes of 2-D shapes; identify the results of subdividing 3-D shapes; identify 3-D figures using a 2-D representation; solve problems involving area; use scales to estimate distance; interpret graphs; find the mean value of a data set; select graphical representations of data; interpret data; make conjectures based on theoretical probability.

Basic

Students perform operations with rational numbers; solve and interpret one-step linear equations; extend geometric patterns; generalize patterns to find a specific term; identify relationships in 3-D objects; calculate the theoretical probability of an event; interpret a scatter plot to determine the relationship between two variables.

Proficient

Students identify equivalent representations of a number; identify mental strategies to solve problems; solve multi-step equations; use symbolic algebra; identify transformations; classify angles; create similar polygons; use coordinate geometry; solve problems involving area; identify appropriate units of measure; convert standard units within a system of measurement; interpret graphic organizers; calculate measures of center.

Advanced

Students estimate the value of square roots; write numbers using scientific notation; solve two-step inequalities; analyze slope and intercept in linear equations; apply the Pythagorean Theorem using coordinate geometry; identify polygons based on their attributes; identify coordinates of vertices of a transformed polygon; use a protractor to measure angles; solve problems involving surface area; select, create, and use appropriate graphical representation of data.

GRADE 10—SHORT DESCRIPTORS

Below Basic

Students use rules of exponents to create equivalent numbers; extend and generalize numerical patterns; use visual models to represent 3-D figures; use constructions to represent reflections; determine volume of geometric figures; use diagrams to estimate measurements; predict outcomes using probability.

Basic

Students compute numbers mentally; determine an equation that represents a quantitative relationship; use algebraic relationships to solve problems; solve for angles of parallel lines cut by a transversal; identify characteristics of surveys and sample groups; calculate central tendencies; use graphical representation of data.

Proficient

Students justify estimates; solve problems using coordinate planes; solve quadratics; evaluate algebraic expressions; identify an appropriate mathematical model based on data; identify equivalent algebraic expressions; solve area problems; analyze precision of computations; use unit analysis to solve problems; use proportions to solve conversion and rate problems; apply statistical concepts; describe shape of distribution of data.

Advanced

Students apply formulas to solve problems on a coordinate plane; use rules of exponents to simplify expressions; determine the equation of a straight line on the coordinate plane; compare linear properties of functions; compare area of similar polygons; determine compound probability; evaluate accuracy of a graph.

Science

GRADE 5—SHORT DESCRIPTORS

Below Basic

Students identify the relationship between mass and force; classify bodies of water; identify weather instruments and their uses; identify characteristics of the solar system; compare amounts/measurements given in a simple format; identify appropriate tools for simple scientific measurements; identify how technological advances may be helpful to humans.

Basic

Students explain the relationship between mass and force; describe how specialized body structures help animals survive; match environments to the plants and animals they support; identify environmental problems and find solutions; construct part of a graph; determine the appropriate scientific tool and its function in an investigation; determine how technological advances address problems and enhance life.

Proficient

Students describe changes in properties of matter; identify uses of simple machines; explain how work is done; identify forces of magnetism; describe the motion of objects; identify plant parts and their functions; classify vertebrates and invertebrates; classify producers, consumers, or decomposers; predict changes in food chains; identify the effects of human activities on other organisms; describe the Sun as a source of light and heat, or the moon as a reflector of light; explain the day/night cycle; identify characteristics and variables of a fair test; interpret data and make predictions; draw conclusions based on evidence; distinguish between man-made and natural objects; apply problem solving skills to a situation.

Advanced

Students identify energy transformations; predict the effect of heat energy on water; diagram a complete electrical circuit; predict how simple machines affect the force needed to do work; describe the effects of weathering and erosion on Earth's surface; describe relationships in weather data; explain how the Sun's position and the length and position of shadows relate to the time of day; interpret and apply knowledge from a data table; identify appropriate steps, tools and metric units in an investigation; construct a graph and plot data; formulate a question for an investigation.

GRADE 8—SHORT DESCRIPTORS

Below Basic

Students identify simple terms related to matter and energy; demonstrate beginning understanding of properties of light and how it travels; identify structures of plants and animals needed for survival; identify levels of organization in multicellular organisms; read simple graphs and make simple data comparisons.

Basic

Students identify an example of a force; demonstrate simple understanding of how traits are passed from one generation to the next; have a basic understanding of climate; identify a simple hypothesis; recognize a trend in a data table; demonstrate some awareness of how various factors influence and are influenced by science and technology.

Proficient

Students classify types of motion; calculate the speed of an object; demonstrate simple understanding of life processes; classify and/or show relationships between organisms; explain how adaptations help organisms survive; explain how species are affected by environmental change; understand and describe a food web; explain rock and fossil evidence of changes in the Earth; explain how Earth's systems interact; draw conclusions from tables or graphs; demonstrate basic understanding of the solar system; recognize the need for, and calculate, averages; understand the importance of constants in investigations; use appropriate tools and methods to collect data; describe tools and discoveries that advance scientific knowledge.

Advanced

Students explain the physical and chemical properties of matter; apply knowledge of energy and energy transfer; demonstrate understanding of physical and chemical processes of organisms; evaluate the effects of balanced and unbalanced forces; predict the impact of environmental change in ecosystems; justify how adaptations help organisms survive; demonstrate understanding of the water cycle; compare and contrast weather and climate; explain the cause of seasons on Earth; demonstrate understanding of the solar system; apply the concept of light years; construct a complete graph; evaluate experimental design; create testable questions and hypotheses; apply awareness of the influence of science and technology in society.

GRADE 11—SHORT DESCRIPTORS

Below Basic

Students describe that matter is made up of particles too small to be seen; describe how the mass of an object affects its motion; identify that organisms are made up of cells; define a species in terms of ability to reproduce; identify major gasses that make up the atmosphere; identify an advantage or disadvantage of tools used to study the universe; identify a valid justification of a conclusion; communicate minimal information about the data collected in a experiment; identify how one environmental factor can be impacted by technology.

Basic

Students identify some properties that can be used to classify substances; perform basic calculations related to force and motion of an object when given a formula; identify and describe cell structures and their functions; explain community interaction to maintain a balanced ecosystem; identify the processes involved in the rock cycle; identify a property of the electromagnetic spectrum as it relates to the universe; select appropriate investigation methods; use data to formulate an explanation; describe a reason theories change.

Proficient

Students describe properties and changes in matter; explain gravitational force between objects; apply Newton's Laws of Motion; describe the relationship between force and work; describe photosynthesis and cellular respiration; describe energy flow in a food web; explain natural selection; identify processes of the water cycle; explain the motion of the Earth, moon, and Sun; describe the role of the electromagnetic spectrum in the universe; design scientific investigations; analyze data, form conclusions and communicate results; identify the constants and variables in an investigation required for reliable results; create testable hypotheses; identify technology used to increase scientific knowledge.

Advanced

Students explain energy and energy transfer; apply an understanding of the Periodic Table; apply the Law of Conservation of Matter and Energy; evaluate the Law of Gravity; compare the efficiency of machines; apply the principles of Mendelian genetics; explain cell division; analyze genetic diversity; predict changes in a food web; explain factors of climate; justify land use based on natural resources; explain the processes and energy sources of plate tectonics; provide evidence for the revolution and rotation of Earth; construct data tables and graphs; justify constants and variables; design a valid scientific investigation; gather and interpret qualitative and quantitative data; determine the effects of technology.

CHAPTER 9: EVIDENCE OF CONSTRUCT-RELATED VALIDITY

Evidence for construct-related validity—the meaning of test scores and the inferences they support—is the central concept underlying the MAP validation process. In this section, CTB presents evidence of construct-related validity through studies of test reliability, convergent validity, and divergent validity. All analyses in this section are based on census data.

Chapter 9 of this report demonstrates the adherence to AERA/APA/NCME standards 1.2, 1.6, 1.7, 2.1, 2.2, 2.10, 2.11, 2.14, 2.15, 3.1, 3.5, 3.6, 3.7, 3.9, 3.11, 6.4, and 6.15.

9.1 Minimization of Construct-Irrelevant Variance and Under-Representation

Minimization of construct-irrelevant variance and construct under-representation is addressed in the following steps of the test development process: 1) specification, 2) item writing, 3) review, 4) field testing, 5) test construction, and 6) calibration (see Chapter 3 for more information on 1 through 5 and Chapter 6 for more information on calibration).

Construct-irrelevant variance refers to error variance that is caused by factors unrelated to the constructs measured by the test. For example, when tests are not administered under standardized conditions (e.g., one administration may be timed, but another administration may be untimed), differences in student performance related to different administration conditions may result. Careful specification of content and review of the items representing that content are first steps in minimizing construct-irrelevant variance. Then, empirical evidence, especially item-level data, is used to infer construct irrelevance.

Construct under-representation occurs when the content of the assessment does not reflect the full range of content that the assessment is expected to cover. MAP is designed to represent the Show-Me Standards/GLE strands. Specification and review, in which test blueprints are developed and reviewed, are primary steps in the development process designed to ensure that content is appropriately represented.

9.2 Reliability

Reliability refers to the consistency of the students' test scores on parallel forms of a test. A reliable test is one that produces scores that are expected to be relatively stable if the test is administered repeatedly under similar conditions. Often, however, it is impractical to administer multiple forms of the test, and reliability is estimated on a single administration of the test. This type of reliability, known as internal consistency, provides an estimate of how consistently examinees perform across items within a test during a single test administration (Crocker & Algina, 1986). Reliability is a necessary but not sufficient condition of validity.

The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) indicate:

. . . reliability evidence may be reported in terms of variances or standard deviations of measurement errors, in terms of one or more coefficients, or in terms of IRT-based test information functions. (27)

In accordance with the AERA/APA/NCME *Standards* and developing and maintaining tests of the highest quality, CTB has calculated the reliability of each MAP test in a variety of ways: reliability of raw scores, overall standard error of measurement, IRT-based conditional standard error of measurement, and decision consistency of achievement-level classifications.

9.2.1. Test Reliability

The reliability of raw scores on the MAP tests was evaluated using Cronbach's (1951) coefficient alpha, which is a lower-bound estimate of test reliability. The reliability coefficient is a ratio of the variance of true test scores to those of the observed scores, with the values ranging from 0 to 1. The closer the value of the reliability coefficient is to 1, the more consistent the scores, where 1 refers to a perfectly consistent test. As a rule of thumb, reliability coefficients that are equal to or greater than 0.8 are considered acceptable for tests of moderate lengths.

Cronbach's coefficient alpha was computed using the formula

$$\alpha = \frac{n}{n-1} \left[1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_x^2} \right], \quad (9.1)$$

where n is the number of items on the test, σ_i^2 is the variance of item i and σ_x^2 is the variance of the total test score.

Total test reliability measures such as Cronbach's coefficient alpha and standard error of measurement consider the consistency (reliability) of performance over all test questions in a given form, the results of which imply how well the questions measure the content domain and could continue to do so over repeated administrations. The number of items in the test influences these statistics; a longer test can be expected to be more reliable than a shorter test.

The reliability coefficients for the MAP are reported in Tables 9.1, 9.2, and 9.3 for Communication Arts, Mathematics, and Science respectively. These reliability coefficients were computed using the census data. All reliability statistics are over .90 for

all tests indicating acceptable reliability. The reliability statistics by subgroup are reported and discussed in Chapter 10.

9.2.2. Standard Error of Measurement

The reliability of reported test scores can be characterized by the standard errors associated with the scores. The standard error of measurement (SEM) may be used to determine the range within which a student's true score is likely to fall. An observed score should be regarded not as a student's true score, but as an estimate of a student's true score. It is expected that 68% of the time a student's score obtained from a single test administration would fall within one SEM of the student's true score and that 95% of the time the obtained score would fall within approximately two standard errors of the true score. The standard error of measurement (SEM) is an index of the random variability in test scores and is defined as follows:

$$SEM = SD\sqrt{1 - R_{xx'}}, \quad (9.2)$$

where SD represents standard deviation of the raw score distribution, and $R_{xx'}$ is estimated by $\hat{\alpha}$ as expressed in Equation 9.1.

The overall SEM is expressed in scale score units and is a test level statistic. The SEM is summarized in Table 9.6 with respect to all students and each subgroup. There were some observable trends in SEM by the subgroups in all grades and contents. The SEM for females tended to be smaller than for males. The SEM was smaller for nonaccommodated students than for accommodated students.

9.2.3. Conditional Standard Error of Measurement

In contrast to SEM, the conditional standard errors of measurement (CSEM) express the degree of measurement error in scale score units and are conditioned on the ability of the student. The CSEMs are defined as the reciprocal of the square root of the test information function and can be estimated across all points of the ability continuum (Hambleton & Swaminathan, 1985):

$$CSEM(\theta_i) = \frac{1}{\sqrt{I(\theta_i)}}, \quad (9.3)$$

where $I(\theta_i)$ is the test information function, as a sum of item information function², obtained as

$$I(\theta_i) = \sum_j \frac{p'_{ij}(\theta_i)^2}{p_{ij}(\theta_i)q_{ij}(\theta_i)}, \quad (9.4)$$

where $p'_{ij}(\theta_i)$ is the derivative of $p_{ij}(\theta_i)$ and $q_{ij}(\theta_i) = 1 - p_{ij}(\theta_i)$.

Note that the CSEMs vary in magnitude across the entire range of student ability estimates (i.e., scale scores) and are smaller in the middle of the score distribution and higher at the tails. This pattern is seen for all MAP CSEMs and is to be expected when IRT methods are used. The CSEMs at the three cut scores that define the performance levels are presented in Table 9.7 and range from 7 to 12 scale score points.

Figures 9.1 through 9.18 display the CSEM curves for each grade/content area. The estimates of measurement error tend to be higher at the low and high ends of the scale score range. The measurement error increases when there are few observations at a particular ability level. Generally, there are few students with extreme scores, and these score levels can not be estimated as accurately as levels toward the middle of the ability range. Figures 9.1 through 9.18 demonstrate that the tests are designed so that measurement error is minimized in the middle of the scale range where the majority of students are located.

9.2.4. Decision Accuracy and Consistency

The *Standards* also make reference to an additional measurement concern that bears on evidence for validity:

Some authorities have proposed that a semantic distinction be made between “reliability of scores” and “degree of agreement in classification.” The former term would be reserved for analysis of score variation under repeated measurement. The term *classification consistency* . . . , rather than reliability, would be used in discussions of consistency of classification. Adoption of such usage would make it clear that the importance of an error of any given size depends on the proximity of the examinee’s score to the cut score. (30)

Decision Consistency: Classification consistency or decision consistency is defined as the extent to which the classifications of students agree on the basis of two independent administrations of the test, or one administration of two parallel test forms. However, it is difficult to obtain data from repeated administrations of the same form because of cost, time, and students’ recall of the first administration. Also, it is difficult to construct two parallel forms. Therefore, a common practice is to estimate decision consistency from one administration of a test.

Decision Accuracy: Decision accuracy is defined as the extent to which the actual classifications of test takers agree with classifications that would be made on the basis of their true scores (Livingston & Lewis, 1995). That is, decision consistency refers to the agreement between two observed scores, while decision accuracy refers to the agreement between the observed score and the true score. It is common to estimate decision accuracy by assuming the psychometric model to find true scores corresponding to observed scores.

CTB used the Livingston-Lewis (1995) methodology to calculate these statistics on the 2007 MAP results. The Livingston-Lewis procedure utilizes a beta-binomial model that

requires two steps: 1) fitting proportion-correct true scores to a four-parameter beta distribution; and 2) using the binomial distribution to estimate classification accuracy and consistency. All calculations for decision accuracy and consistency are based on census data.

Table 9.4 reports the decision accuracy and consistency classifications conditioned on each level of achievement for each grade/content area. In Table 9.4, the accuracy conditioned on level of achievement indicates the percentage of students correctly classified into a level of achievement given their true score status. For example, 87% of the Grade 8 Mathematics students who were estimated to have a true status of *Below Basic* were correctly classified on the Grade 8 Mathematics test by their observed score.

In Table 9.4, the consistency conditioned on level of achievement indicates the percentage of students whose classification would be in the same level of achievement based on a hypothetical alternate form of the test. For example, 66% of Grade 7 Communication Arts students whose performance was classified as *Proficient* would be classified in the same level based on the hypothetical alternate form, if they had taken it.

Perhaps the most important indices for accountability systems are those for the accuracy and consistency of classification decisions made at specific cut points. These results are reported in Table 9.5. To evaluate decisions at specific cut points, the joint distribution of all the performance levels is collapsed into a dichotomized distribution around that specific cut point. As an example, the dichotomization at the cut point between the *Basic* and *Proficient* classifications was formed. The proportion of correct classifications below this particular cut point is equal to the sum of all the cells at the levels *Below Basic* and *Basic*, and the proportion of correct classifications above that particular cut point is equal to the sum of all the cells at the levels *Proficient* and *Advanced*. As shown in Table 9.5, all accuracy statistics conditioned on cut point are above 90% and all decision statistics conditioned on cut point are above 86%.

The magnitude of classification consistency and accuracy measures is influenced by key features of the test design including the number of items, number of cut scores, as well as the reliability and associated standard error of measurement. The classification consistency and accuracy results established for 2008 suggest that consistent and accurate performance level classifications are being made for students in Missouri based on the MAP assessments.

9.2.5. Convergent Validity

Convergent validity is a subtype of construct validity that can be estimated by the extent to which measures of constructs that theoretically should be related to each other are, in fact, observed as related to each other. Analyses of the internal structure of a test can indicate the extent to which the relationships among test items conform to the construct the test purports to measure. For example, the MAP Mathematics test is designed to measure a single overall construct—Mathematics achievement; therefore, the items comprising the Mathematics MAP should only measure Mathematics, not Science, Language, or Reading.

This Technical Report summarizes additional statistics that contribute to construct validity (Cronbach's coefficient alpha reported previously in this section and item fit reported in Chapter 6). The internal consistency coefficient (Cronbach's alpha) reported above is a measure of item homogeneity. In order for a group of items to be homogeneous, they must measure the same construct (construct validity) or represent the same content domain (content validity). Because IRT models were used to calibrate test items and to report student scores, item fit is also relevant to construct validity. The extent to which test items function as the IRT model prescribes is relevant to the validation of test scores. As shown in Chapter 6, only 23 items total were flagged for poor model/data fit across all 17 grade/content area MAPs.

9.3 Principal Components Analysis

As another measure of construct validity, CTB examined the unidimensionality of each grade-level MAP test. One of the underlying assumptions of the IRT models used to scale MAP is that the tests being calibrated are unidimensional, that is, items comprising MAP in each grade/content area measure a single content domain. For example, Mathematics items should measure Mathematics ability and not measure Reading skills. The unidimensionality assumption is in practice a testable hypothesis that is commonly evaluated through Principal Components Analysis (PCA). This analysis seeks evidence that there exists a single primary factor, the first principal component, that accounts for much of the relationship between items. The presence of a single or dominant factor suggests that a test is sufficiently unidimensional (i.e., measures one underlying construct).

A principal components factor analysis was conducted on each grade/content area MAP. A large first principal component is evident in each analysis. In Figures 9.19 to 9.36, scree plots (Cattell, 1966) of eigenvalues are presented to illustrate the relative dominance of the first principal component in each of MAP test. It is common to have additional eigenvalues greater than 1.0, which may suggest the presence of other factors.

For all grade/content area MAPs, the ratio of the variance accounted for by the first factor to the second and third is sufficiently large to support the claim that these tests are essentially unidimensional. All of the MAP subject area tests exhibit first principal components accounting for more than 15% of the test variance (see Tables 9.8 to 9.10). To further investigate the unidimensionality of the Communication Arts, Mathematics, and Science tests, the ratio of first eigenvalue to the second eigenvalue was explored (see Tables 9.8 to 9.10). These ratios show that the first eigenvalue is at least five times as large as the second eigenvalue for most of the grade/content areas. This substantial difference in magnitude indicates that one factor appears to be dominant and that the Communication Arts, Mathematics, and Science tests are essentially one-dimensional.

This evidence supports the claim that there is a dominant dimension underlying the items/tasks in each test and that scores from each test represent performance primarily determined by that ability. Construct-irrelevant variance such as factual knowledge

irrelevant for doing well in a subject does not appear to create significant nuisance factors.

9.4 Analyses by Content Standard

Two sets of analyses were conducted for the content standard level in another attempt to assess the construct validity of MAP. First, the reliability of each Content Standard was computed. Second, correlation coefficients that measure the relationship between the Content Standards were computed.

9.5 Reliability of Content Standards

Cronbach's (1951) Alpha was computed for each of the Content Standards by grade/content area using the census data. Tables 9.11 through 9.27 report the reliability statistics along the diagonal of each matrix for each grade/content area. Reliability indices, such as Cronbach's Alpha, are a function of the number of test items. It is expected that Cronbach's Alpha would be low for a Content Standard assessed by a small number of items (e.g., Writing Formally and Informally).

9.6 Correlations among Content Standard Subscores

In this section, we measure the strength of the interrelationships among the Content Standards by computing correlation between the content standards. Tables 9.11 through 9.27 report the uncorrected Pearson product-moment (PPM) correlation coefficients, the PPM corrected for attenuation (CAPP), in addition to the reliability coefficients described above. The PPM among the Content Standard subscores is presented below the diagonal portion of the matrix, the CAPP is presented above the diagonal portion of the matrix, and the reliability coefficients are shown on the diagonal in each table.

The uncorrected correlation coefficients (PPM) in Tables 9.11 to 9.27 should be interpreted in the context of the reliability coefficient. In general, we expect to see lower PPM coefficients between variables that are less reliable. Overall, the PPM coefficients show that performance on one Content Standard is moderately to strongly related to performance on another Content Standard within the same content area. As noted above, the value of the correlation coefficients will be affected by the limited number of items measuring each Content Standard. So, caution should be used when comparing the PPM coefficients measuring the relationships between Content Standards to those measuring the relationships between content areas (Table 9.28). We expect to see a more modest relationship (smaller correlation coefficients) reported between the Content Standards as a consequence of the lower number of items measuring each content standard (e.g., Writing Formally and Informally).

Indeed, the PPM between two content standard subscores may be artificially low because of measurement error. We can correct for the attenuation of the PPM statistically using Spearman's formula:

$$CAPP\text{M} = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}}$$

Where r_{xy} is the PPM between two content standards, r_{xx} is the reliability of one of those content standards, and r_{yy} is the reliability for the other content standard.

Across all tables, the CAPPM indicate strong relationships between the content standards. In some cases, the CAPPM is greater than 1.0. “Disattenuated values greater than 1.00 indicate that measurement error is not randomly distributed” (Schumacker, 1996). The strong relationships suggested by the CAPPM in Tables 9.11 to 9.27 are further evidence of the validity of the test construct. Since the overall content area is comprised of the content standard subscores and the content area is expected to measure a single dimension, then we would expect that these subscores are also highly related.

9.7 Divergent (Discriminant) Validity

Measures of different constructs should not be highly correlated with each other. Divergent validity is a subtype of construct validity that can be assessed by the extent to which measures of constructs that theoretically should not be related to each other are, in fact, observed as not related to each other. Typically, correlation coefficients among measures of unrelated or distantly related constructs are examined in support of divergent validity.

To assess the divergent validity of the MAP tests, correlations were computed between the Math and Communication Arts scale scores for students who took both MAP subject area tests in 2007. These correlations are based on the census data and the results are shown in Table 9.28. The correlation coefficients ranged from 0.74 (between Communication Arts and Mathematics in Grades 4) to 0.84 (between Mathematics and Science in Grade 8). The correlation coefficients suggest that individual student scores for Communication Arts and Mathematics are moderately- to highly-related. The correlation coefficients between Science and the other two content areas suggest that the Science MAP is highly related to the Communication Arts and Mathematics MAP. The tests are not perfectly related to each other, suggesting that different constructs are being tapped; however, the test scores do appear, at least, moderately related to one another, suggesting they are tapping into a similar knowledge base. This is especially true of the Science test. The Science MAP is comprised of many constructed-response items, which may help account for its relationship with the Communication Arts test. The Science MAP tests similar thinking skills and item types as are found in the Mathematics MAP, which may help account for the strong correlation between the Science and Mathematics test scores.

Table 9. 1: Reliability in Communication Arts

Grade	Number of Items	Number of Score Points	Cronbach's Alpha
3	58	68	0.91
4	56	63	0.91
5	56	63	0.91
6	55	62	0.90
7	62	73	0.92
8	61	69	0.91
11	63	73	0.92
11b	63	73	0.93

Table 9. 2: Reliability in Mathematics

Grade	Number of Items	Number of Score Points	Cronbach's Alpha
3	60	67	0.92
4	65	77	0.92
5	62	69	0.91
6	61	68	0.92
7	62	69	0.92
8	64	76	0.93
10	58	70	0.94

Table 9. 3: Reliability in Science

Grade	Number of Items	Number of Score Points	Cronbach's Alpha
5	53	80	0.91
8	58	87	0.93
11	58	92	0.93

Table 9. 4: Decision Accuracy and Consistency Conditioned on Level of Achievement

Content Area	Grade	Accuracy				Consistency			
		Below Basic	Basic	Prof.	Adv.	Below Basic	Basic	Prof.	Adv.
Communication Arts	3	0.83	0.84	0.64	0.85	0.72	0.80	0.53	0.75
	4	0.83	0.84	0.72	0.84	0.70	0.80	0.63	0.74
	5	0.82	0.84	0.70	0.85	0.69	0.79	0.60	0.76
	6	0.82	0.82	0.70	0.84	0.70	0.76	0.60	0.74
	7	0.84	0.83	0.75	0.85	0.73	0.78	0.66	0.76
	8	0.81	0.87	0.70	0.85	0.66	0.83	0.60	0.75
	11	0.85	0.83	0.71	0.85	0.76	0.77	0.61	0.75
	11b	0.90	0.80	0.67	0.84	0.85	0.73	0.55	0.73
Mathematics	3	0.82	0.87	0.79	0.84	0.69	0.83	0.71	0.74
	4	0.83	0.87	0.80	0.84	0.70	0.82	0.72	0.72
	5	0.82	0.85	0.74	0.84	0.69	0.80	0.65	0.73
	6	0.84	0.83	0.77	0.85	0.73	0.77	0.69	0.76
	7	0.85	0.79	0.76	0.85	0.76	0.72	0.67	0.76
	8	0.87	0.79	0.73	0.86	0.80	0.72	0.63	0.78
	10	0.89	0.77	0.79	0.87	0.83	0.69	0.71	0.78
	Science	5	0.83	0.83	0.67	0.85	0.72	0.78	0.57
8		0.87	0.80	0.80	0.84	0.80	0.73	0.72	0.74
11		0.89	0.69	0.79	0.85	0.83	0.58	0.71	0.76

Table 9. 5: Decision Accuracy and Consistency at Achievement Cut Points

Content Area	Grade	Accuracy			Consistency		
		Below Basic/ Basic	Basic/ Prof.	Prof./Adv.	Below Basic/ Basic	Basic/ Prof.	Prof./Adv.
Communication Arts	3	0.95	0.91	0.94	0.94	0.87	0.91
	4	0.96	0.91	0.94	0.94	0.87	0.92
	5	0.96	0.90	0.93	0.95	0.87	0.91
	6	0.95	0.90	0.94	0.93	0.86	0.91
	7	0.96	0.91	0.95	0.94	0.87	0.92
	8	0.97	0.91	0.94	0.96	0.87	0.91
	11	0.94	0.91	0.95	0.92	0.88	0.93
	11b	0.92	0.94	0.97	0.89	0.91	0.95
Mathematics	3	0.97	0.91	0.96	0.95	0.88	0.94
	4	0.97	0.91	0.96	0.95	0.88	0.94
	5	0.96	0.91	0.95	0.95	0.87	0.92
	6	0.96	0.91	0.95	0.94	0.87	0.92
	7	0.94	0.91	0.95	0.92	0.87	0.92
	8	0.94	0.92	0.95	0.91	0.88	0.93
	10	0.94	0.92	0.96	0.91	0.89	0.94
	Science	5	0.95	0.91	0.94	0.93	0.87
8		0.94	0.92	0.96	0.91	0.89	0.95
11		0.93	0.92	0.96	0.89	0.88	0.94

Table 9. 6: SEM by Subgroup

Grade	Category	Group	CA SEM	MA SEM	SC SEM
3	Overall		11.00	10.58	
	Ethnicity	White (not Hispanic)	10.90	10.69	
		Black (not Hispanic)	11.35	10.30	
		Hispanic	10.99	10.21	
		Asian/Pacific Islander	11.35	11.94	
Native American		11.12	10.53		
Gender	Male	11.07	10.66		
	Female	10.90	10.46		
Accommodations	No	10.73	10.59		
	Yes	13.05	10.76		
4	Overall		10.08	9.68	
	Ethnicity	White (not Hispanic)	10.06	9.69	
		Black (not Hispanic)	10.26	9.70	
		Hispanic	9.85	9.41	
		Asian/Pacific Islander	10.17	10.19	
Native American		10.26	9.15		
Gender	Male	10.15	9.77		
	Female	10.01	9.56		
Accommodations	No	9.79	9.62		
	Yes	12.28	10.13		
5	Overall		10.10	11.96	9.67
	Ethnicity	White (not Hispanic)	9.96	11.78	9.21
		Black (not Hispanic)	10.72	12.54	10.96
		Hispanic	9.84	11.74	9.93
		Asian/Pacific Islander	10.21	12.36	9.98
Native American		9.80	12.02	9.30	
Gender	Male	10.33	12.08	9.86	
	Female	9.82	11.79	9.43	
Accommodations	No	9.68	11.72	9.26	
	Yes	12.93	13.60	11.66	
6	Overall		10.44	11.95	
	Ethnicity	White (not Hispanic)	10.32	11.74	
		Black (not Hispanic)	11.01	12.82	
		Hispanic	10.44	11.67	
		Asian/Pacific Islander	10.29	12.56	
Native American		10.40	11.56		
Gender	Male	10.60	12.01		
	Female	10.23	11.87		
Accommodations	No	9.96	11.68		
	Yes	13.45	13.88		

Table 9. 6: SEM by Subgroup (Cont'd)

Grade	Category	Group	CA SEM	MA SEM	SC SEM
7	Overall		10.07	11.52	
	Ethnicity	White (not Hispanic)	9.92	11.17	
		Black (not Hispanic)	10.59	12.78	
		Hispanic	10.20	11.77	
		Asian/Pacific Islander	10.48	11.86	
		Native American	9.88	11.60	
	Gender	Male	10.28	11.74	
		Female	9.80	11.20	
	Accommodations	No	9.74	11.05	
		Yes	12.39	14.33	
8	Overall		9.90	10.74	8.01
	Ethnicity	White (not Hispanic)	9.63	10.26	7.61
		Black (not Hispanic)	10.69	12.39	9.00
		Hispanic	10.41	11.35	8.41
		Asian/Pacific Islander	10.23	10.73	8.16
		Native American	9.61	11.07	8.20
	Gender	Male	10.18	10.91	8.10
		Female	9.52	10.49	7.87
	Accommodations	No	9.46	10.20	7.71
		Yes	12.80	14.25	9.63
HS	Overall		10.24	12.31	10.51
	Ethnicity	White (not Hispanic)	10.12	11.86	9.96
		Black (not Hispanic)	10.75	14.13	12.59
		Hispanic	10.46	12.68	11.15
		Asian/Pacific Islander	10.70	12.35	10.57
		Native American	10.21	12.16	9.83
	Gender	Male	10.39	12.46	10.67
		Female	10.03	12.07	10.24
	Accommodations	No	10.00	11.86	10.06
		Yes	12.60	15.85	13.68
11 Breach	Overall		11.68	-	-
	Ethnicity	White (not Hispanic)	10.82	-	-
		Black (not Hispanic)	12.07	-	-
		Hispanic	11.40	-	-
		Asian/Pacific Islander	14.88	-	-
		Native American	-	-	-
	Gender	Male	12.33	-	-
		Female	10.89	-	-
	Accommodations	No	11.35	-	-
		Yes	17.07	-	-

Table 9. 7: Conditional Standard Error of Measurement at the Basic, Proficient, & Advanced Cut Scores

Content Area	Grade	Basic		Proficient		Advanced	
		Cut Score	SEM	Cut Score	SEM	Cut Score	SEM
Communication Arts	3	592	9	648	9	673	12
	4	612	8	662	9	691	12
	5	625	8	675	8	702	11
	6	631	9	676	8	704	11
	7	634	9	680	8	712	11
	8	639	9	696	8	723	9
	11	679	8	725	9	753	10
	11b	679	10	725	7	753	9
Mathematics	3	568	10	628	9	667	15
	4	596	8	651	8	688	12
	5	605	11	668	9	706	13
	6	628	11	681	9	721	13
	7	640	12	685	8	724	10
	8	670	10	710	8	741	8
	10	695	11	738	9	785	10
	Science	5	626	9	669	7	692
8		671	7	703	6	735	7
11		702	9	729	8	770	8

Table 9. 8: Principal Component Analysis for Communication Arts

Grade	Eigenvalue	Percent of Variance Explained	Cumulative Percent of Variance Explained
Grade 3			
First Component	10.87	18.74	18.74
Second Component	2.01	3.46	22.19
Ratio (First/Second)	5.42		
Grade 4			
First Component	10.79	19.26	19.26
Second Component	1.76	3.15	22.41
Ratio (First/Second)	6.11		
Grade 5			
First Component	10.52	18.78	18.78
Second Component	1.42	2.54	21.32
Ratio (First/Second)	7.40		
Grade 6			
First Component	9.81	17.83	17.83
Second Component	1.59	2.89	20.72
Ratio (First/Second)	6.18		
Grade 7			
First Component	11.29	18.21	18.21
Second Component	1.68	2.72	20.93
Ratio (First/Second)	6.71		
Grade 8			
First Component	11.00	18.03	18.03
Second Component	2.06	3.38	21.41
Ratio (First/Second)	5.33		
Grade 11			
First Component	11.65	18.49	18.49
Second Component	2.00	3.17	21.67
Ratio (First/Second)	5.83		
Grade 11 Breach			
First Component	12.42	19.71	19.71
Second Component	2.22	3.52	23.23
Ratio (First/Second)	5.59		

Table 9.9: Principal Component Analysis for Mathematics

Grade	Eigenvalue	Percent of Variance Explained	Cumulative Percent of Variance Explained
Grade 3			
First Component	11.19	18.66	18.66
Second Component	1.72	2.86	21.52
Ratio (First/Second)	6.52		
Grade 4			
First Component	11.89	18.30	18.30
Second Component	1.75	2.69	20.99
Ratio (First/Second)	6.80		
Grade 5			
First Component	10.69	17.24	17.24
Second Component	1.87	3.02	20.26
Ratio (First/Second)	5.71		
Grade 6			
First Component	10.90	17.87	17.87
Second Component	1.39	2.29	20.16
Ratio (First/Second)	7.82		
Grade 7			
First Component	11.66	18.81	18.81
Second Component	1.90	3.07	21.88
Ratio (First/Second)	6.13		
Grade 8			
First Component	12.11	18.93	18.93
Second Component	1.99	3.11	22.04
Ratio (First/Second)	6.08		
Grade 10			
First Component	13.51	23.29	23.29
Second Component	1.93	3.33	26.62
Ratio (First/Second)	6.99		

Table 9. 10: Principal Component Analysis for Science

Grade	Eigenvalue	Percent of Variance Explained	Cumulative Percent of Variance Explained
Grade 5			
First Component	9.89	18.66	18.66
Second Component	1.73	3.26	21.92
Ratio (First/Second)	5.73		
Grade 8			
First Component	12.52	21.58	21.58
Second Component	1.79	3.08	24.66
Ratio (First/Second)	7.01		
Grade 11			
First Component	12.38	21.35	21.35
Second Component	1.71	2.95	24.30
Ratio (First/Second)	7.24		

Table 9. 11: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 3

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	16	0.73	0.90	0.87	0.92	0.90
02	Reading Fiction/Poetry/Drama	23	0.70	0.84	0.93	0.91	1.11
03	Reading Nonfiction	15	0.64	0.74	0.75	0.98	1.10
04	Writing Formally & Informally	4	0.51	0.54	0.55	0.42	0.95
05	Combined Reading	38	0.72	0.96	0.90	0.58	0.89

Table 9. 12: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 4

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	10	0.58	0.89	0.84	0.61	0.89
02	Reading Fiction/Poetry/Drama	38	0.64	0.88	0.92	0.64	1.10
03	Reading Nonfiction	6	0.50	0.68	0.61	0.65	1.08
04	Writing Formally & Informally	2	0.33	0.43	0.36	0.51	0.66
05	Combined Reading	44	0.64	0.98	0.80	0.44	0.90

Table 9. 13: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 5

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	13	0.72	0.88	0.91	0.82	0.91
02	Reading Fiction/Poetry/Drama	19	0.67	0.79	0.95	0.83	1.09
03	Reading Nonfiction	22	0.70	0.76	0.81	0.94	1.13
04	Writing Formally & Informally	2	0.37	0.39	0.45	0.28	0.90
05	Combined Reading	41	0.73	0.92	0.96	0.45	0.89

Table 9. 14: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 6

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	13	0.67	0.91	0.94	-	0.93
02	Reading Fiction/Poetry/Drama	23	0.68	0.82	0.97	-	1.12
03	Reading Nonfiction	18	0.67	0.76	0.75	-	1.13
04	Writing Formally & Informally	1	0.23	0.30	0.27	-	-
05	Combined Reading	41	0.72	0.96	0.92	0.30	0.88

Table 9. 15: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 7

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	16	0.72	0.90	0.89	0.81	0.91
02	Reading Fiction/Poetry/Drama	33	0.71	0.86	0.99	0.82	1.12
03	Reading Nonfiction	10	0.62	0.75	0.67	0.80	1.12
04	Writing Formally & Informally	4	0.55	0.61	0.52	0.63	0.82
05	Combined Reading	42	0.73	0.98	0.86	0.61	0.89

Table 9. 16: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 8

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	15	0.68	0.84	0.89	0.72	0.87
02	Reading Fiction/Poetry/Drama	16	0.62	0.80	0.96	0.81	1.10
03	Reading Nonfiction	27	0.66	0.78	0.82	0.86	1.12
04	Writing Formally & Informally	3	0.39	0.48	0.51	0.44	0.85
05	Combined Reading	43	0.68	0.93	0.96	0.53	0.89

Table 9. 17: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Communication Arts, Grade 11

No.	Content Standard	Number of Items	01	02	03	04	05
01	Speaking/Writing Standard English	14	0.65	0.85	0.86	0.74	0.87
02	Reading Fiction/Poetry/Drama	11	0.59	0.74	0.92	0.81	1.08
03	Reading Nonfiction	32	0.65	0.73	0.86	0.76	1.10
04	Writing Formally & Informally	6	0.49	0.57	0.58	0.68	0.80
05	Combined Reading	43	0.67	0.88	0.97	0.62	0.90

Table 9. 18: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 3

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	22	0.81	0.98	0.92	1.06	1.01
02	Algebraic Relationship	12	0.78	0.77	0.92	1.02	0.97
03	Geometric and Spatial	11	0.66	0.64	0.63	0.99	0.96
04	Measurement	9	0.73	0.68	0.60	0.58	1.05
05	Data and Probability	6	0.61	0.57	0.51	0.53	0.44

Table 9. 19: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards Mathematics, Grade 4

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	19	0.82	0.99	0.87	0.90	0.93
02	Algebraic Relationship	10	0.75	0.70	0.91	0.98	0.98
03	Geometric and Spatial	12	0.61	0.59	0.60	0.92	0.91
04	Measurement	13	0.69	0.69	0.60	0.71	0.95
05	Data and Probability	11	0.67	0.65	0.56	0.64	0.64

Table 9. 20: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 5

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	16	0.77	0.98	0.81	0.99	0.95
02	Algebraic Relationship	11	0.72	0.70	0.86	1.03	0.97
03	Geometric and Spatial	11	0.59	0.60	0.69	0.90	0.88
04	Measurement	13	0.71	0.71	0.61	0.68	0.99
05	Data and Probability	11	0.64	0.63	0.56	0.62	0.59

Table 9. 21: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 6

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	17	0.78	0.97	0.89	0.97	0.96
02	Algebraic Relationship	11	0.69	0.65	0.96	1.01	1.02
03	Geometric and Spatial	11	0.61	0.61	0.61	0.96	0.96
04	Measurement	11	0.72	0.68	0.63	0.70	1.01
05	Data and Probability	11	0.69	0.67	0.61	0.69	0.66

Table 9. 22: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 7

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	16	0.76	0.97	0.95	0.97	0.98
02	Algebraic Relationship	11	0.68	0.65	1.01	0.97	1.01
03	Geometric and Spatial	12	0.68	0.67	0.68	1.00	1.00
04	Measurement	11	0.74	0.68	0.71	0.75	1.00
05	Data and Probability	12	0.70	0.67	0.68	0.71	0.67

Table 9. 23: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 8

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	14	0.72	0.95	0.92	0.94	0.99
02	Algebraic Relationship	16	0.70	0.77	0.98	1.00	1.03
03	Geometric and Spatial	13	0.65	0.72	0.70	1.00	0.98
04	Measurement	10	0.69	0.76	0.73	0.75	1.02
05	Data and Probability	11	0.67	0.72	0.66	0.71	0.64

Table 9. 24: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Mathematics, Grade 10

No.	Content Standard	Number of Items	01	02	03	04	05
01	Number and Operations	11	0.75	0.96	0.94	0.97	0.98
02	Algebraic Relationship	15	0.75	0.81	0.97	0.99	0.97
03	Geometric and Spatial	12	0.71	0.76	0.77	1.01	0.98
04	Measurement	9	0.70	0.74	0.75	0.70	0.99
05	Data and Probability	11	0.73	0.75	0.74	0.72	0.74

Table 9. 25: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Science, Grade 5

No.	Content Standard	Number of Items	01	02	03	04	05	06	07	08
01	Matter and Energy	6	0.56	1.02	0.91	0.89	1.04	0.99	0.96	0.96
02	Force and Motion	4	0.47	0.39	0.94	0.95	1.08	1.03	0.99	1.06
03	Characteristics of Living Organisms	6	0.56	0.47	0.67	0.95	0.95	0.95	0.86	0.95
04	Interactions of Organisms	6	0.49	0.44	0.58	0.56	0.95	0.91	0.86	0.96
05	Earth's Processes	6	0.53	0.46	0.54	0.49	0.47	1.03	0.98	1.05
06	The Universe	5	0.54	0.47	0.57	0.49	0.52	0.53	0.93	0.99
07	Scientific Inquiry	14	0.60	0.52	0.59	0.54	0.57	0.57	0.71	0.95
08	Technology and the Environment	6	0.49	0.45	0.53	0.49	0.49	0.49	0.54	0.46

Table 9. 26: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Science, Grade 8

No.	Content Standard	Number of Items	01	02	03	04	05	06	07	08
01	Matter and Energy	7	0.61	0.99	0.98	0.98	1.01	1.00	0.92	0.96
02	Force and Motion	5	0.56	0.53	0.95	0.95	0.96	0.98	0.96	0.96
03	Characteristics of Living Organisms	8	0.58	0.52	0.57	1.00	1.00	0.92	0.92	1.02
04	Interactions of Organisms	5	0.62	0.56	0.61	0.66	1.00	0.96	0.89	1.01
05	Earth's Processes	8	0.62	0.56	0.60	0.65	0.63	0.97	0.91	0.99
06	The Universe	4	0.62	0.57	0.55	0.62	0.62	0.63	0.90	0.93
07	Scientific Inquiry	17	0.64	0.62	0.62	0.65	0.64	0.63	0.79	0.94
08	Technology and the Environment	4	0.59	0.55	0.60	0.65	0.61	0.58	0.65	0.62

Table 9. 27: Reliability (Diagonal) of Each Content Standard, and Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) Among Content Standards: Science, Grade 11

No.	Content Standard	Number of Items	01	02	03	04	05	06	07	08
01	Matter and Energy	8	0.72	0.95	0.94	0.87	0.93	0.87	0.87	0.87
02	Force and Motion	6	0.62	0.59	0.96	0.98	1.01	1.00	0.93	1.02
03	Characteristics of Living Organisms	8	0.63	0.58	0.62	0.95	0.98	0.90	0.84	0.92
04	Interactions of Organisms	4	0.57	0.58	0.58	0.59	0.98	0.95	0.90	0.99
05	Earth's Processes	6	0.60	0.59	0.58	0.57	0.57	0.99	0.90	0.99
06	The Universe	4	0.52	0.54	0.50	0.52	0.53	0.50	0.93	1.00
07	Scientific Inquiry	18	0.67	0.65	0.60	0.63	0.62	0.60	0.83	0.94
08	Technology and the Environment	4	0.47	0.50	0.46	0.48	0.47	0.45	0.54	0.40

Table 9. 28: Inter-Correlation of Communication Arts, Mathematics, and Science Scale Scores

Grade	CA/MA	CA/SC	MA/SC
3	0.76	-	-
4	0.74	-	-
5	0.75	0.78	0.80
6	0.76	-	-
7	0.78	-	-
8	0.77	0.81	0.85
10	-	-	-
11	-	0.80	-

Figure 9. 1: SEM Plot Communication Arts, Grade 3

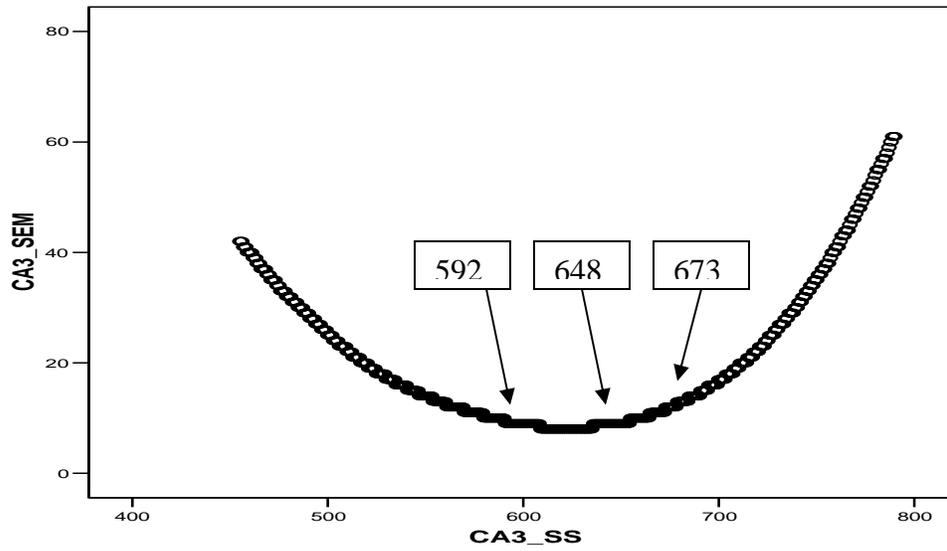


Figure 9. 2: SEM Plot Communication Arts, Grade 4

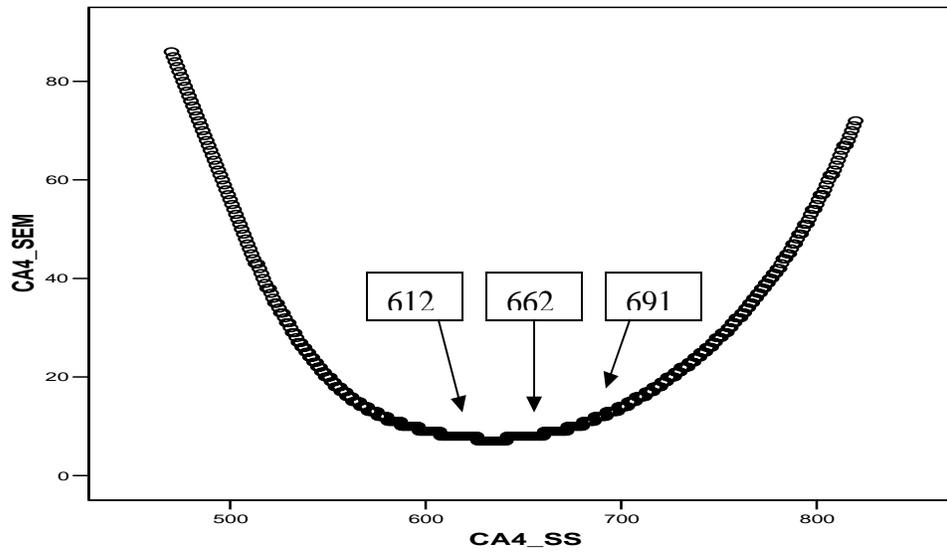


Figure 9. 3: SEM Plot Communication Arts, Grade 5

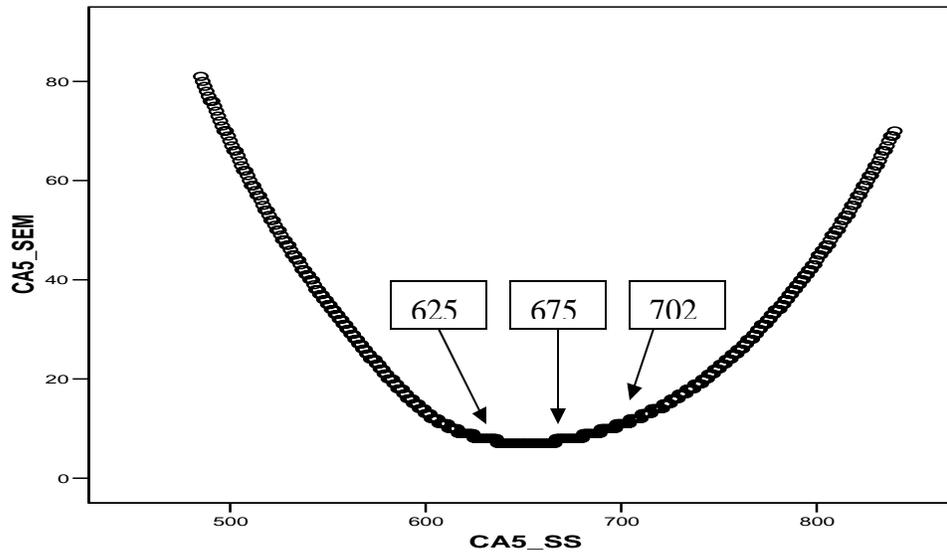


Figure 9. 4: SEM Plot Communication Arts, Grade 6

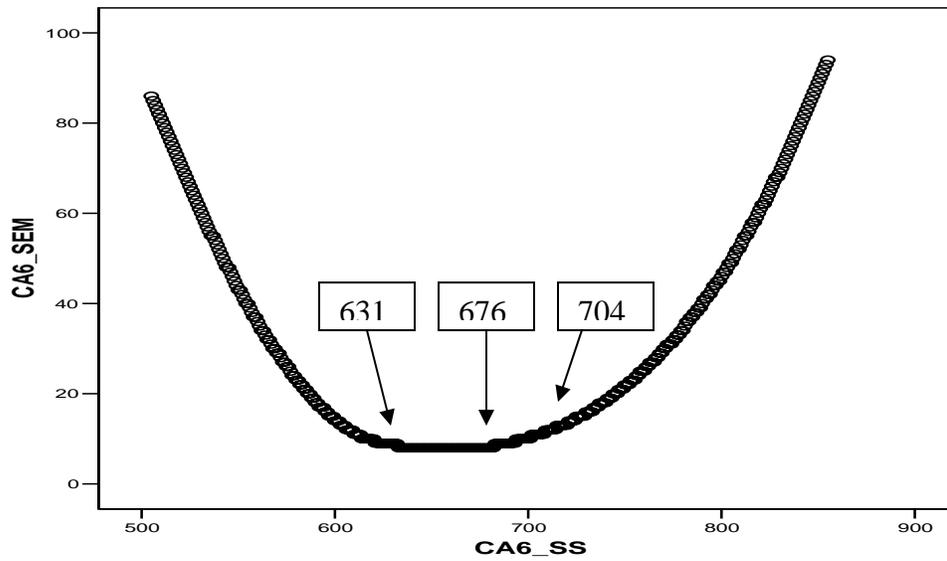


Figure 9. 5: SEM Plot Communication Arts, Grade 7

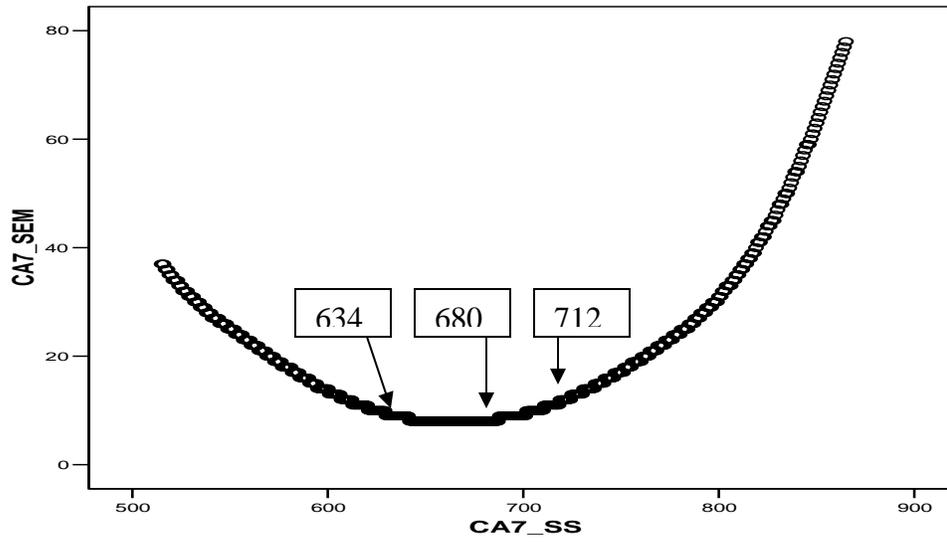


Figure 9. 6: SEM Plot Communication Arts, Grade 8

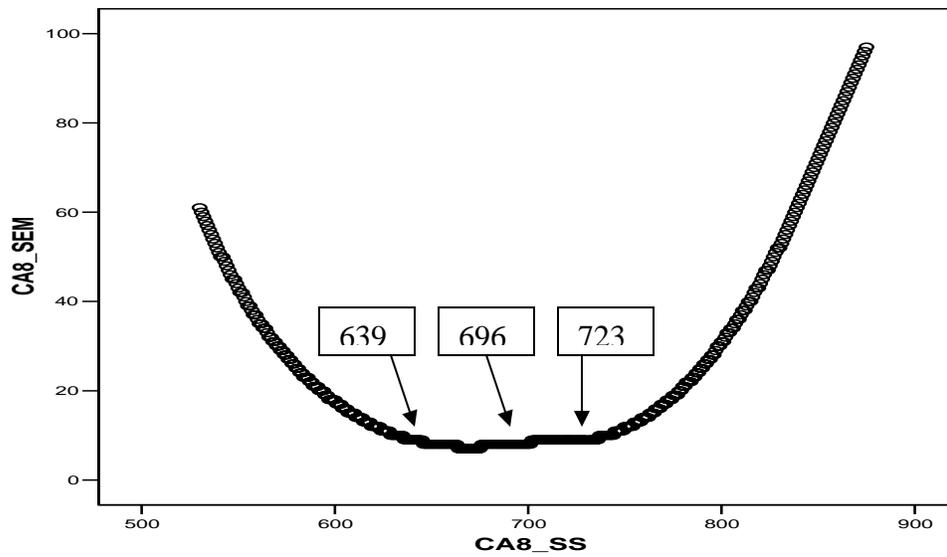


Figure 9. 7: SEM Plot Communication Arts, Grade 11

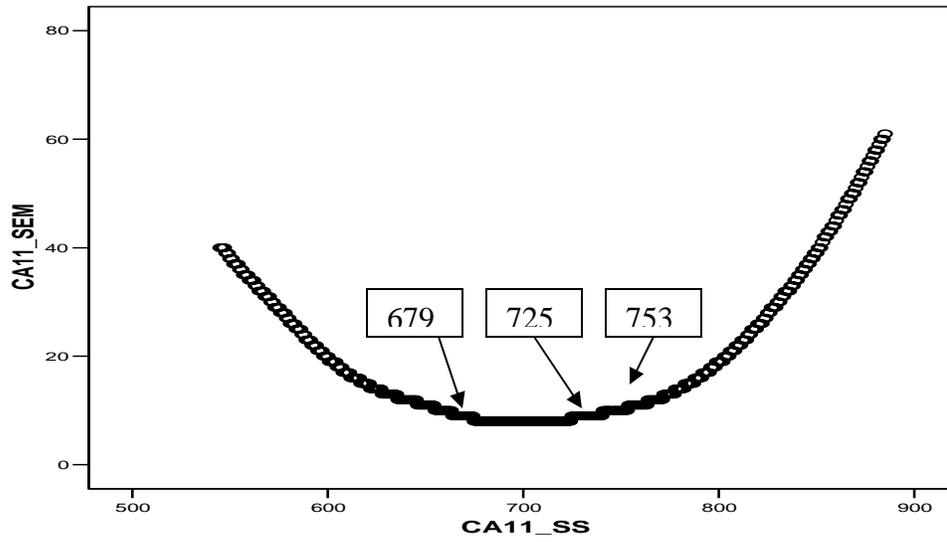


Figure 9. 8: SEM Plot, Communication Arts Grade 11 Breach

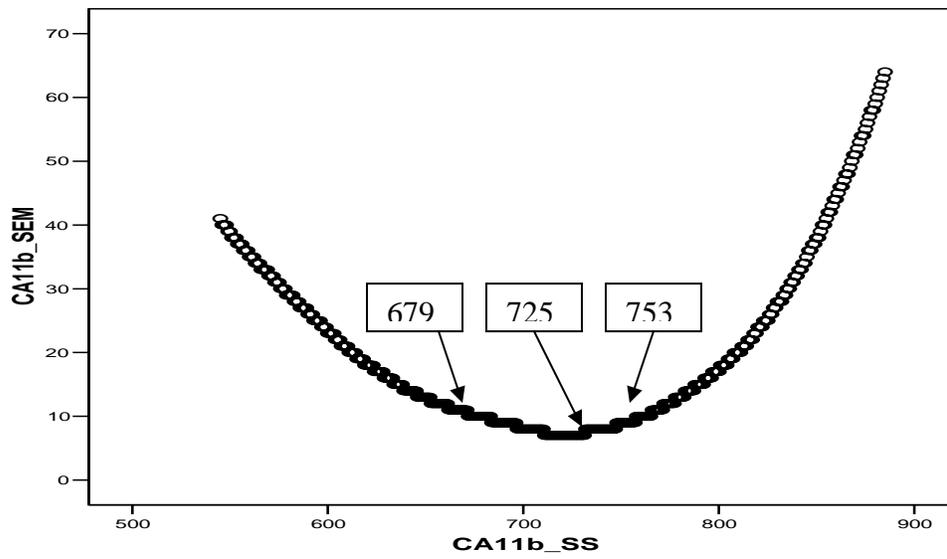


Figure 9. 9: SEM Plot Mathematics, Grade 3

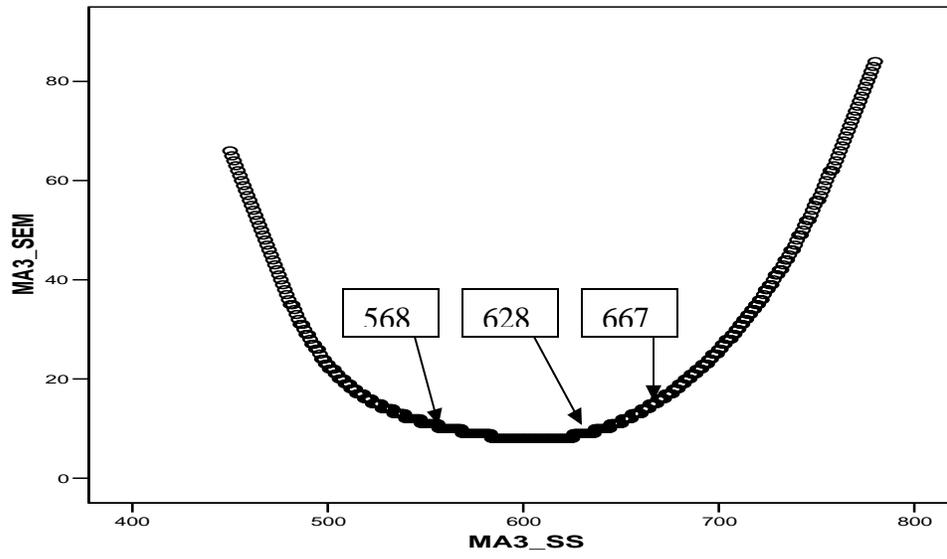


Figure 9. 10: SEM Plot Mathematics, Grade 4

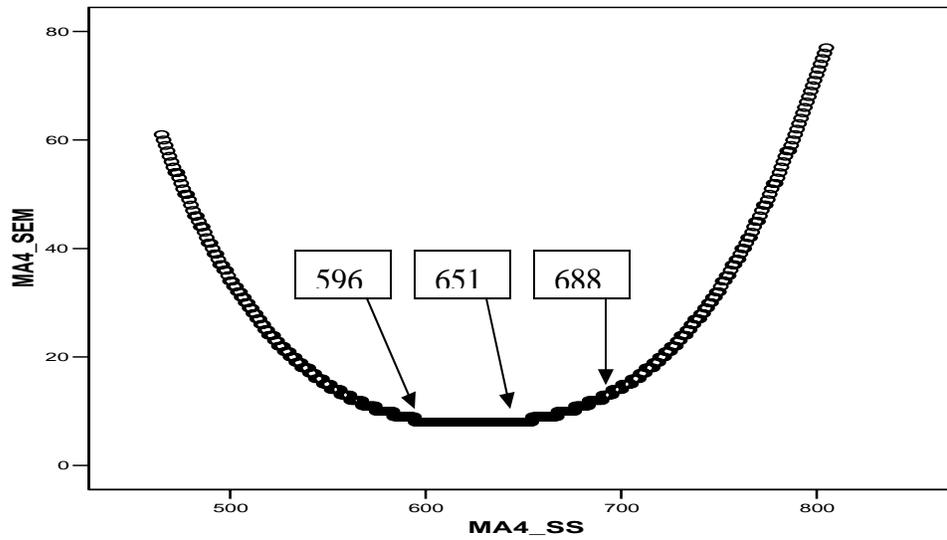


Figure 9.11: SEM Plot Mathematics, Grade 5

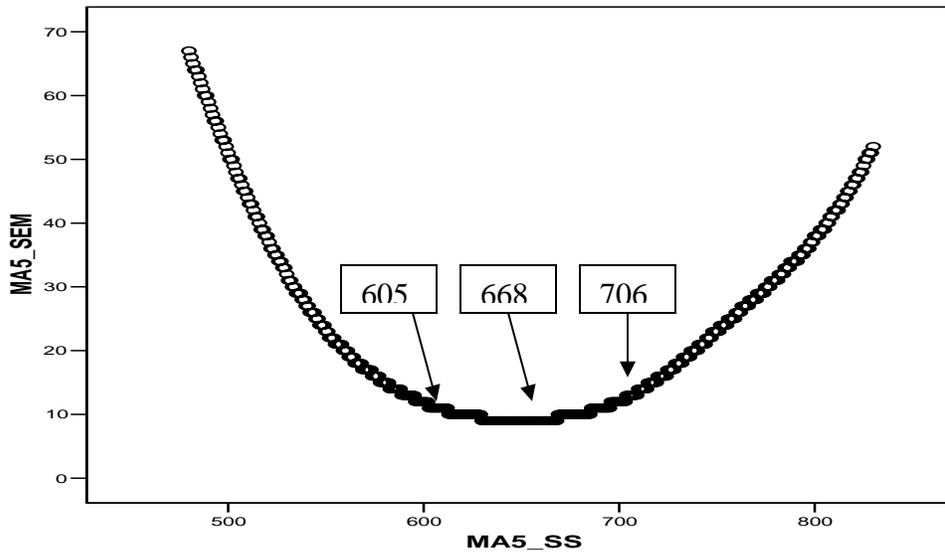


Figure 9.12: SEM Plot Mathematics, Grade 6

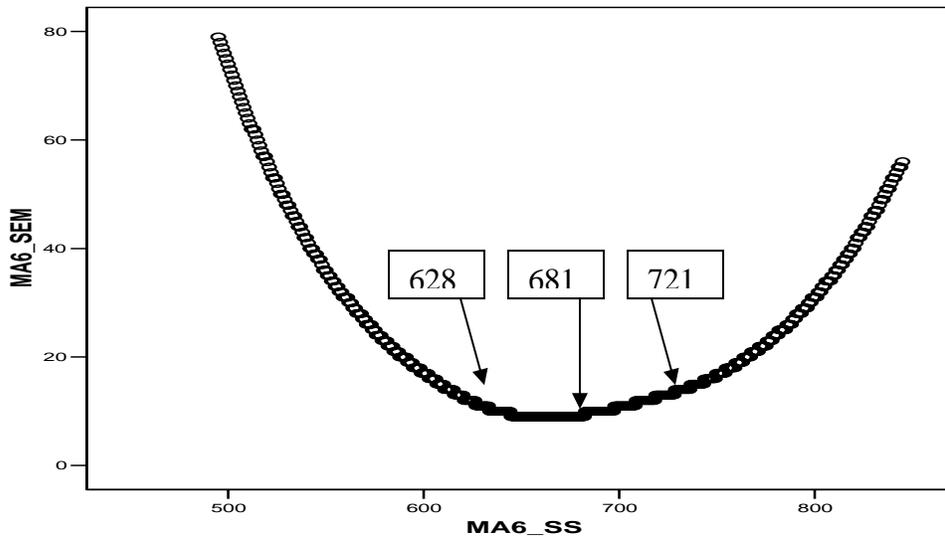


Figure 9.13: SEM Plot Mathematics, Grade 7

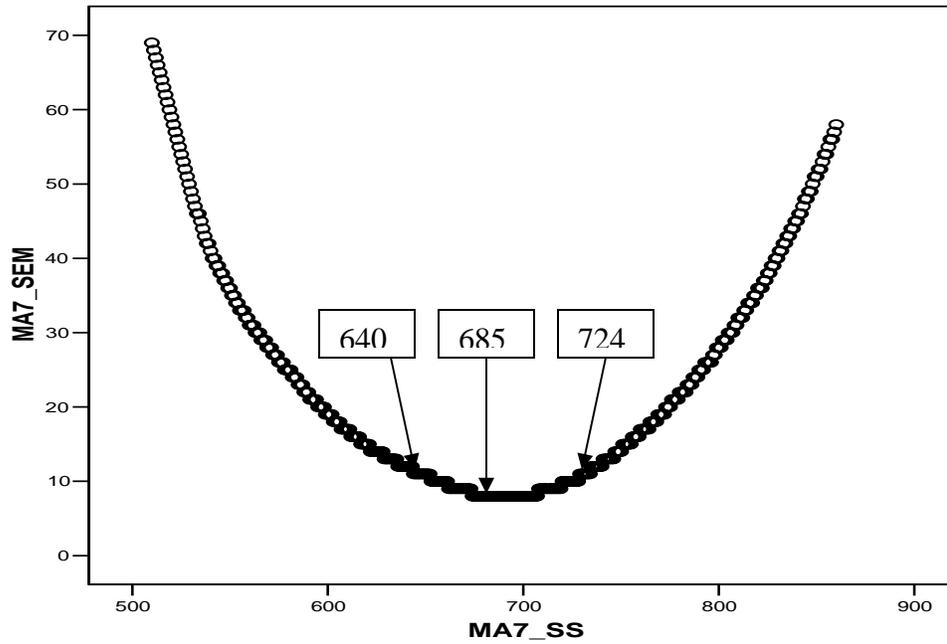


Figure 9.1: SEM Plot Mathematics, Grade 8

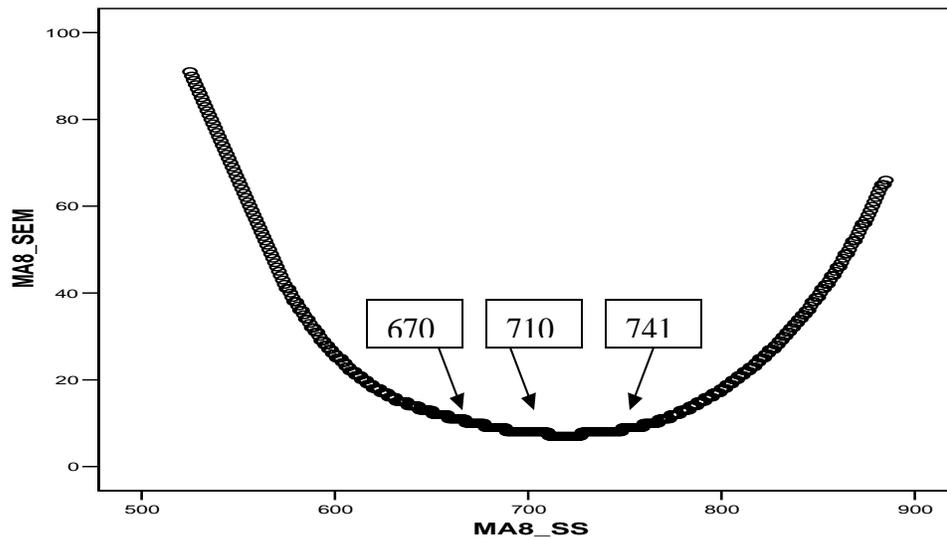


Figure 9. 14: SEM Plot Mathematics Grade, 10

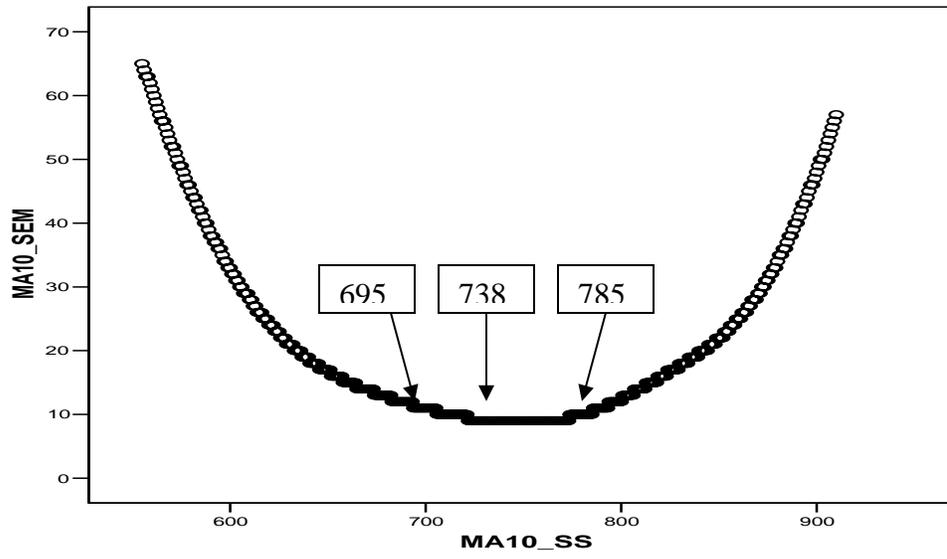


Figure 9. 15: SEM Plot Science, Grade 5

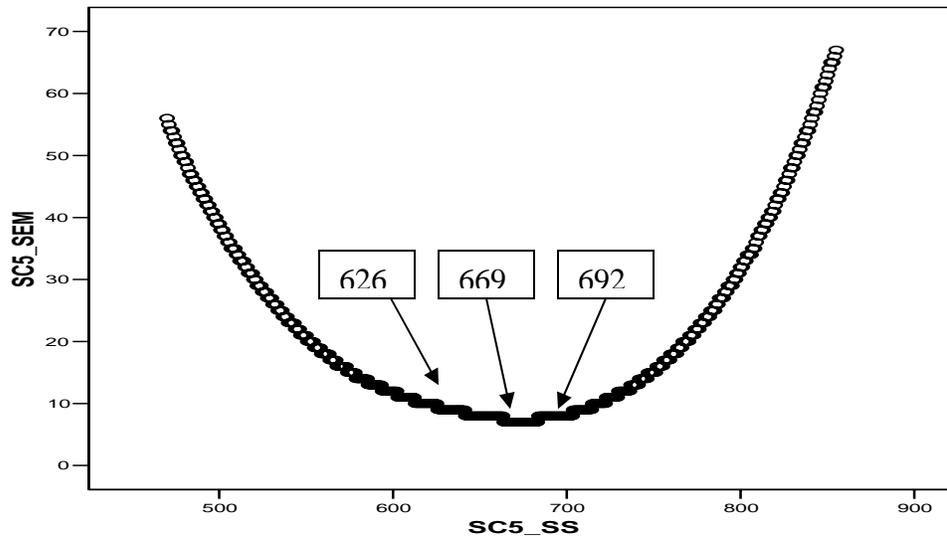


Figure 9. 16: SEM Plot Science, Grade 8

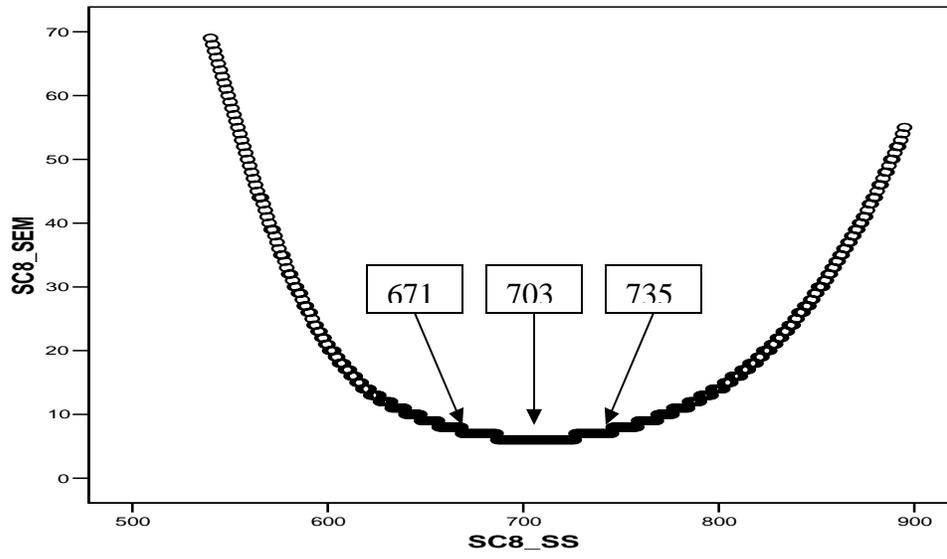


Figure 9. 17: SEM Plot Science, Grade 11

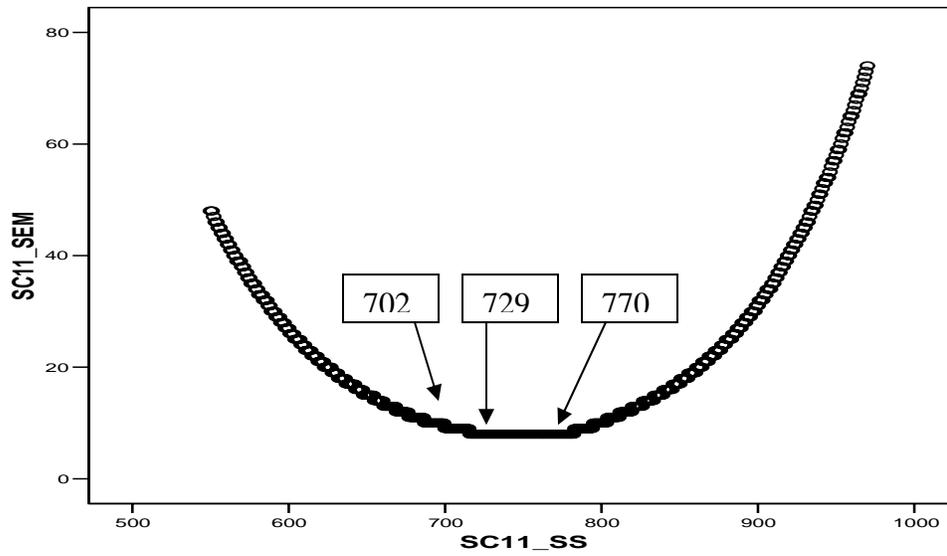


Figure 9. 18: Scree Plot Communication Arts, Grade 3

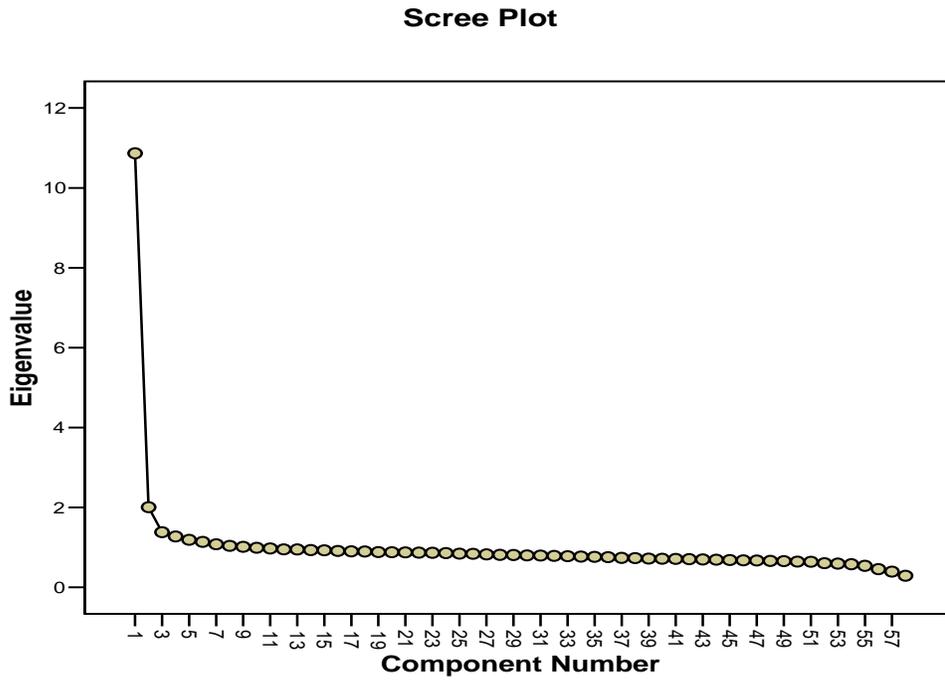


Figure 9. 19: Scree Plot Communication Arts, Grade 4

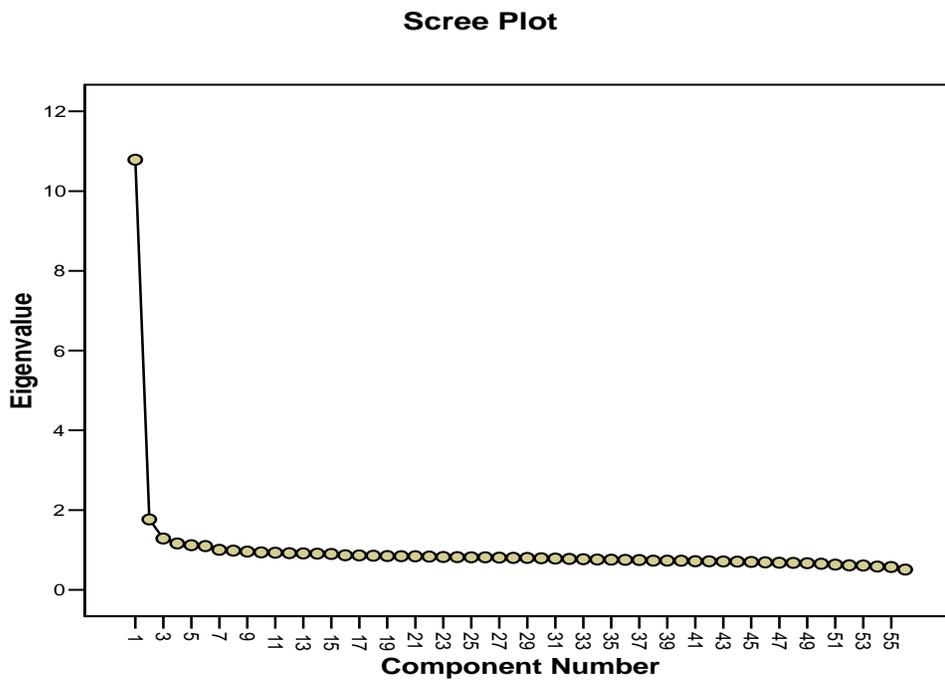


Figure 9. 20: Scree Plot Communication Arts, Grade 5

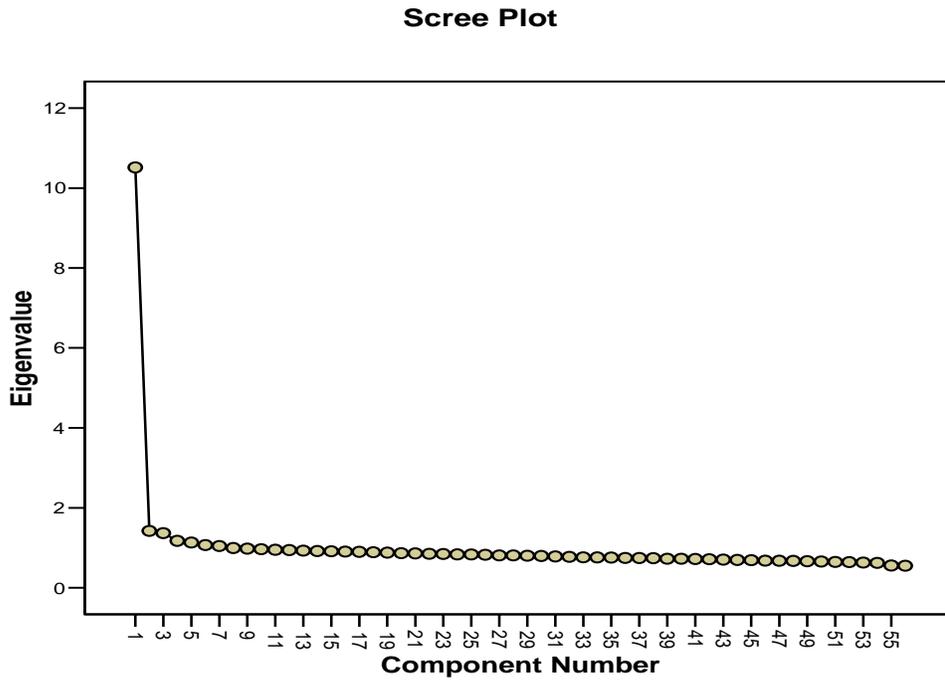


Figure 9. 21: Scree Plot Communication Arts, Grade 6

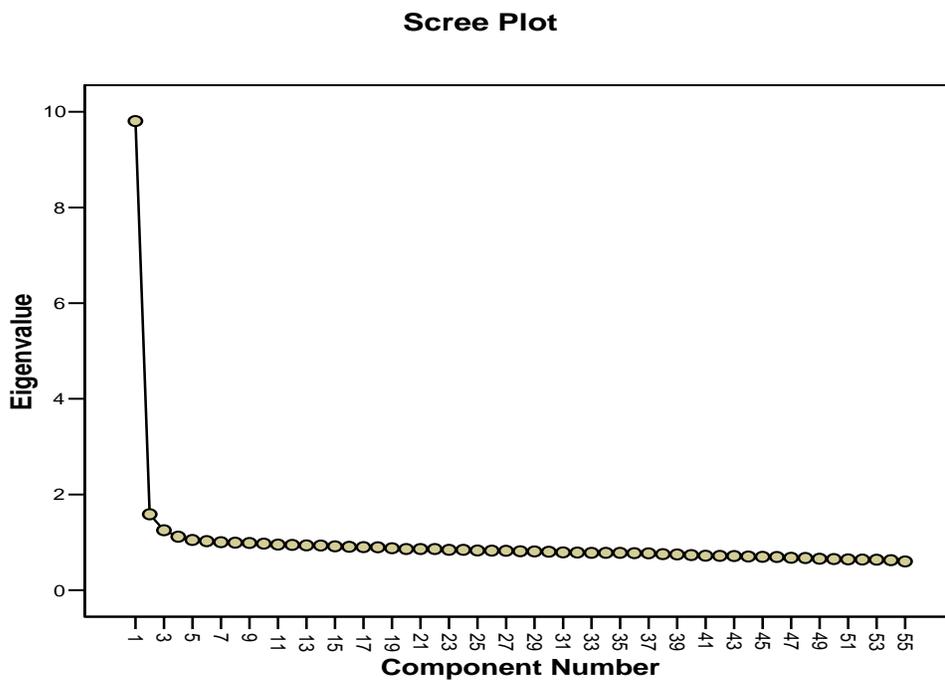


Figure 9.22: Scree Plot Communication Arts, Grade 7

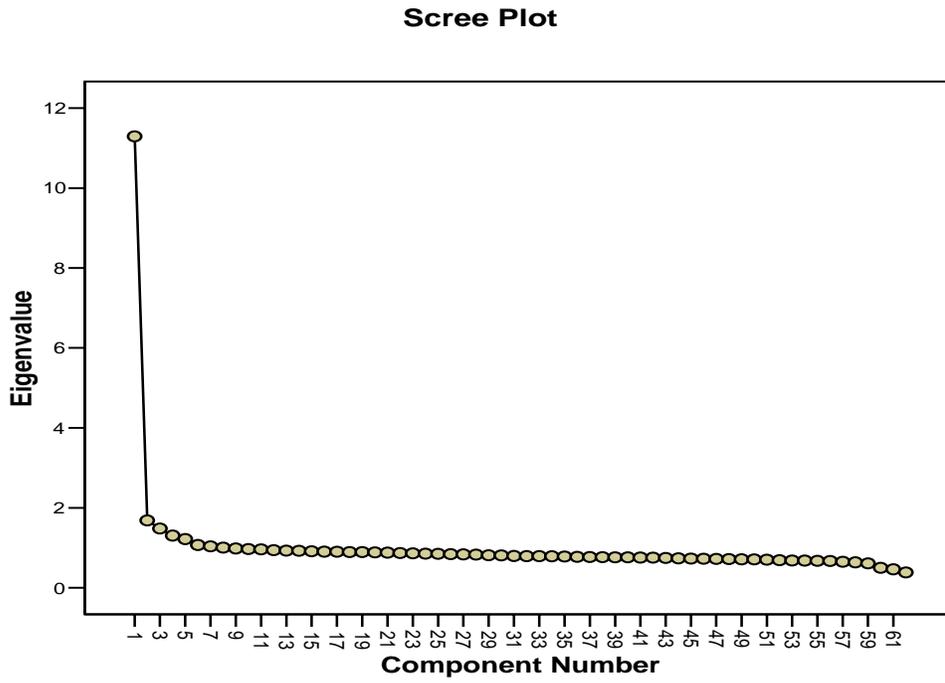


Figure 9.23: Scree Plot Communication Arts, Grade 8

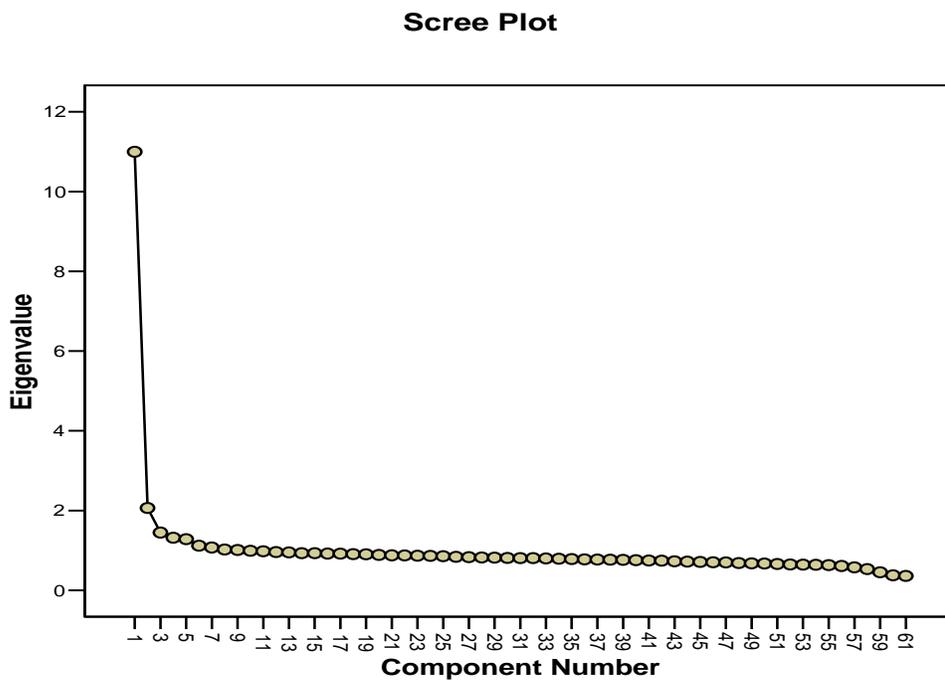


Figure 9. 24: Scree Plot Communication Arts, Grade 11

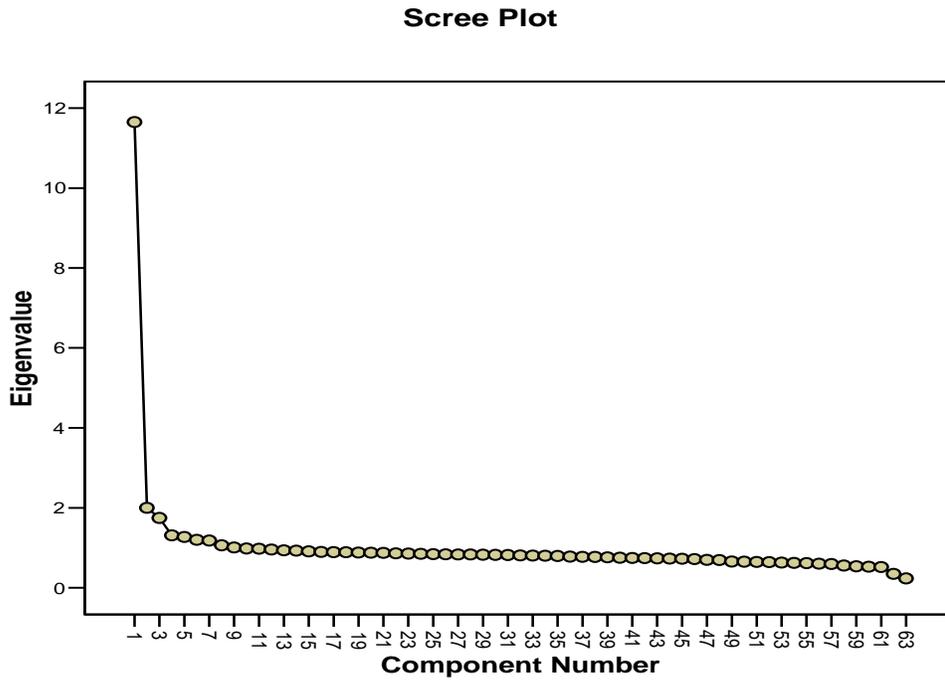


Figure 9. 25: Scree Plot Communication Arts, Grade 11 Breach

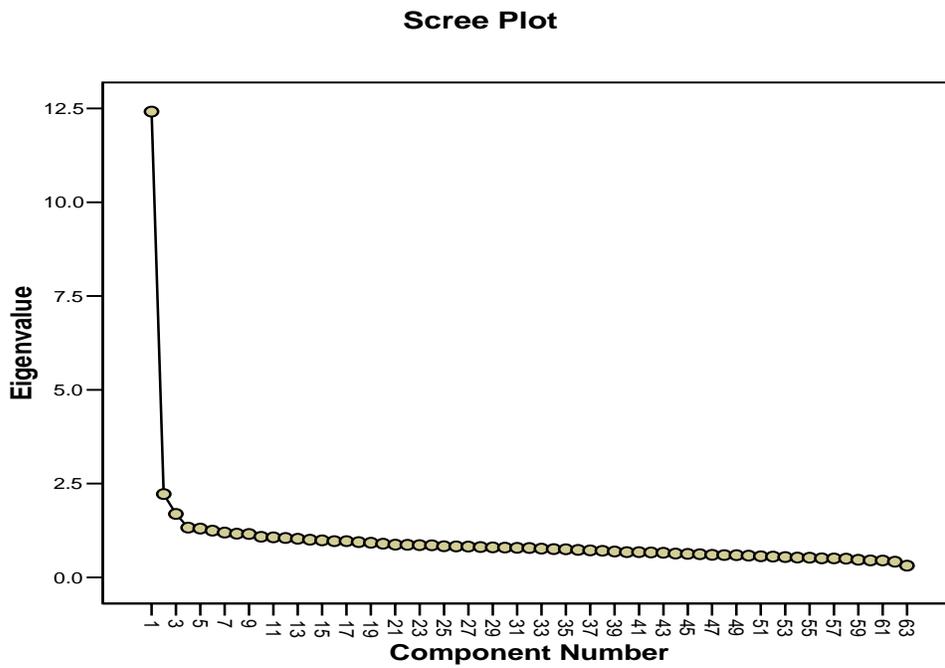


Figure 9. 26: Scree Plot Mathematics, Grade 3

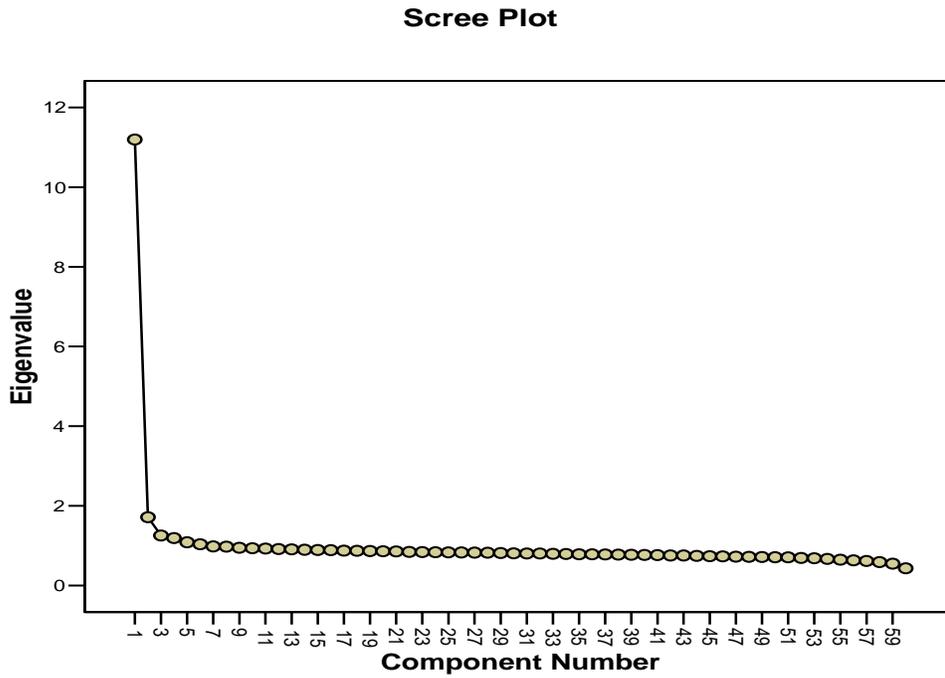


Figure 9. 27: Scree Plot Mathematics, Grade 4

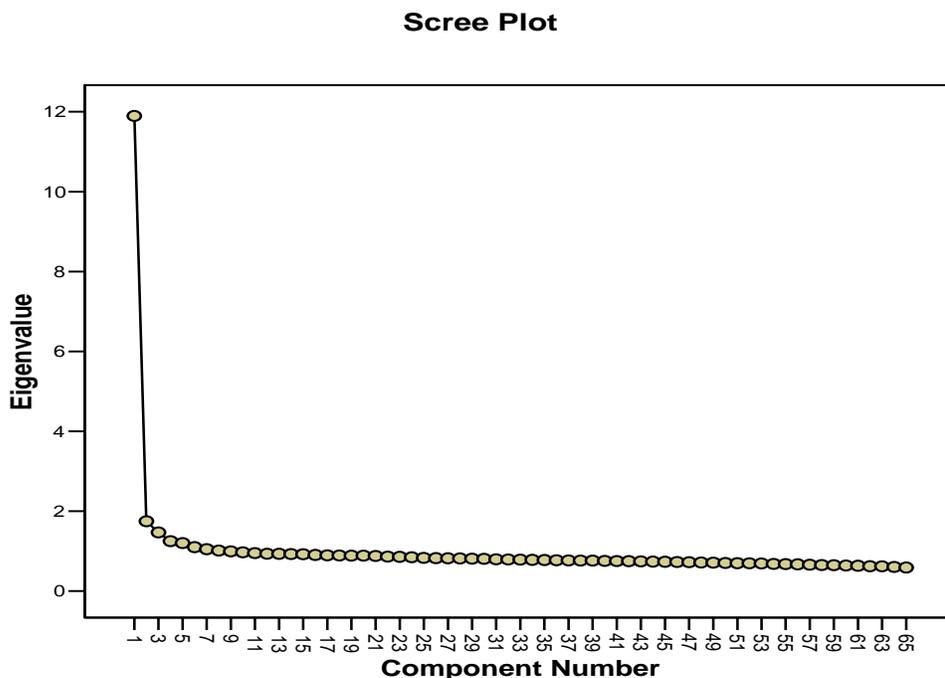


Figure 9. 28: Scree Plot Mathematics, Grade 5

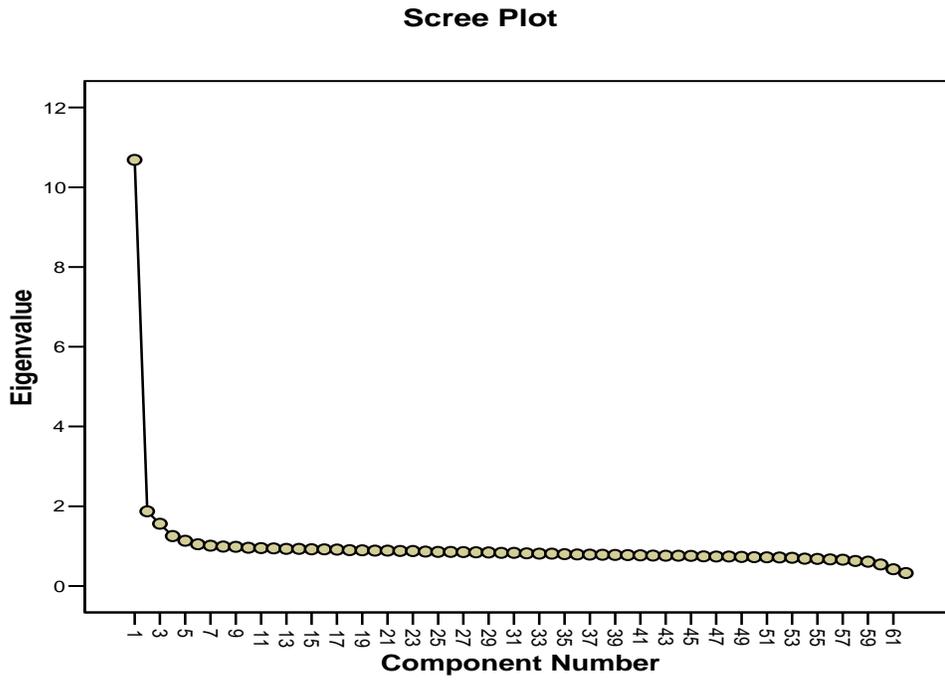


Figure 9. 29: Scree Plot Mathematics, Grade 6

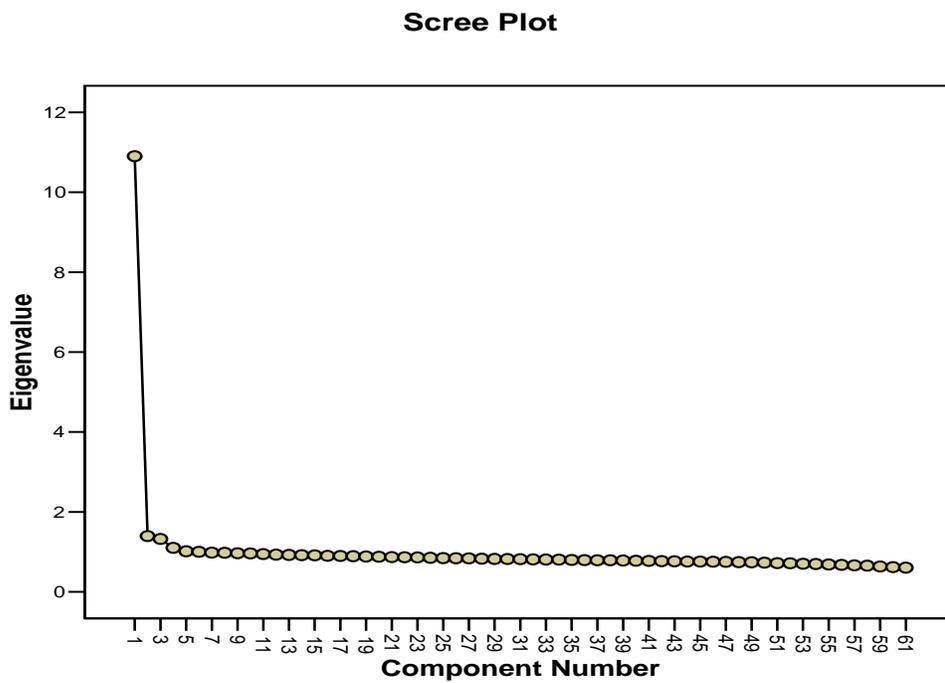


Figure 9.30: Scree Plot Mathematics, Grade 7

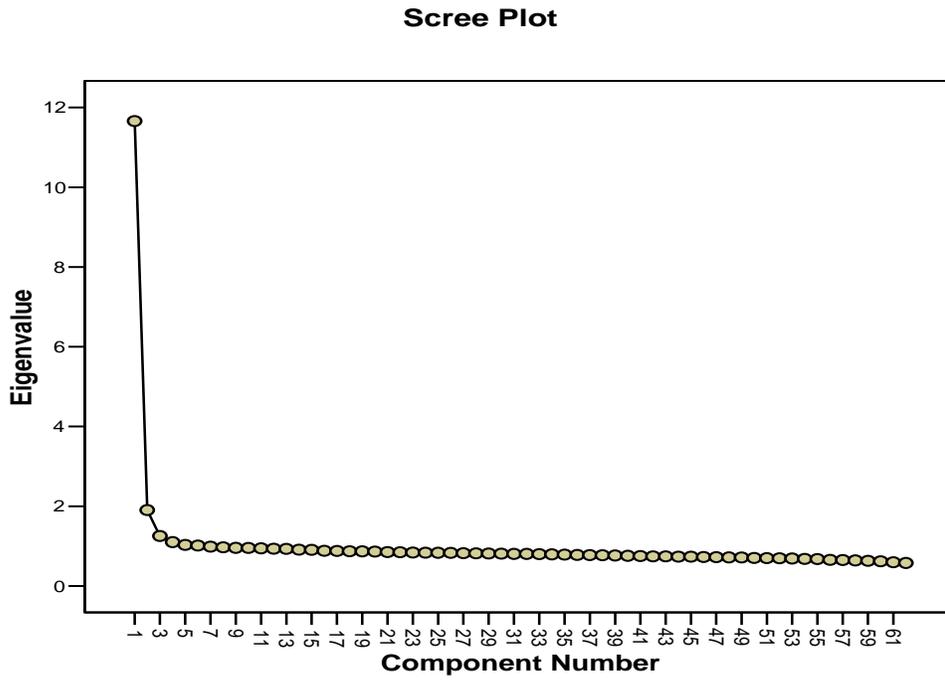


Figure 9.31: Scree Plot Mathematics, Grade 8

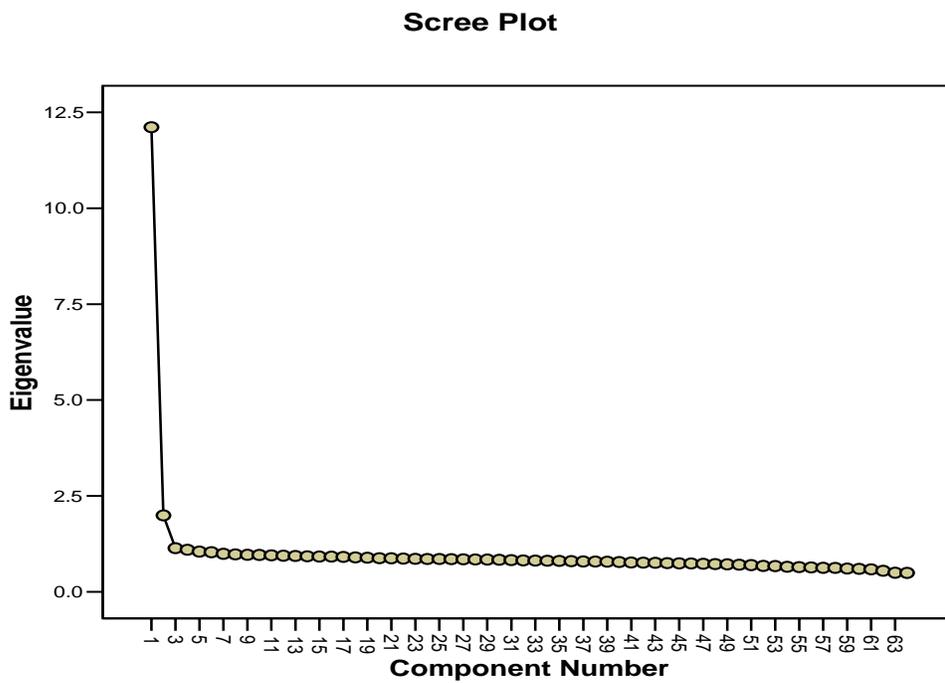


Figure 9.32: Scree Plot Mathematics, Grade 10

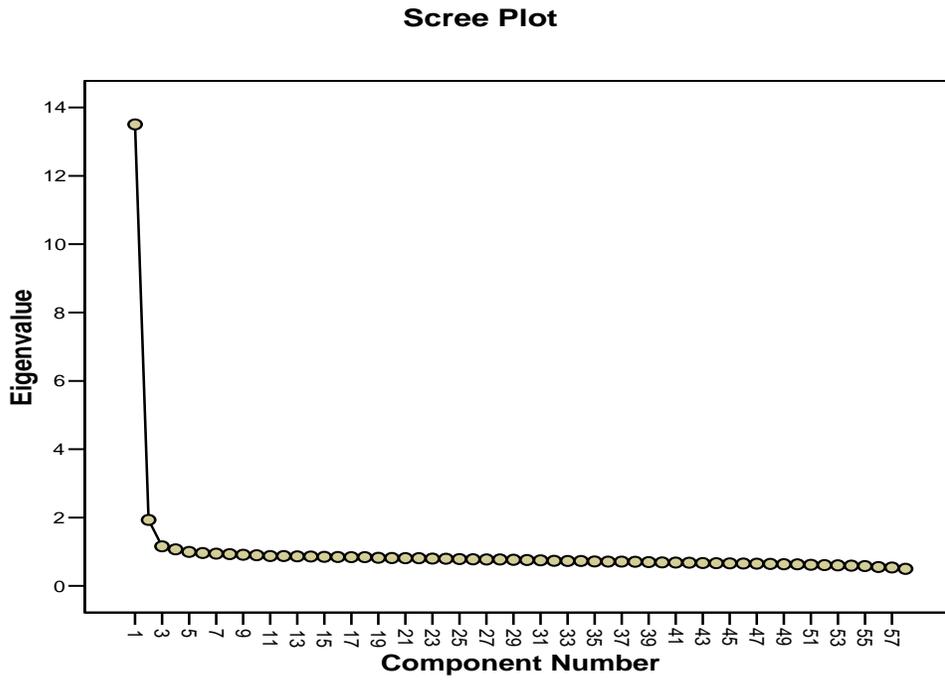


Figure 9.33: Scree Plot Science, Grade 5

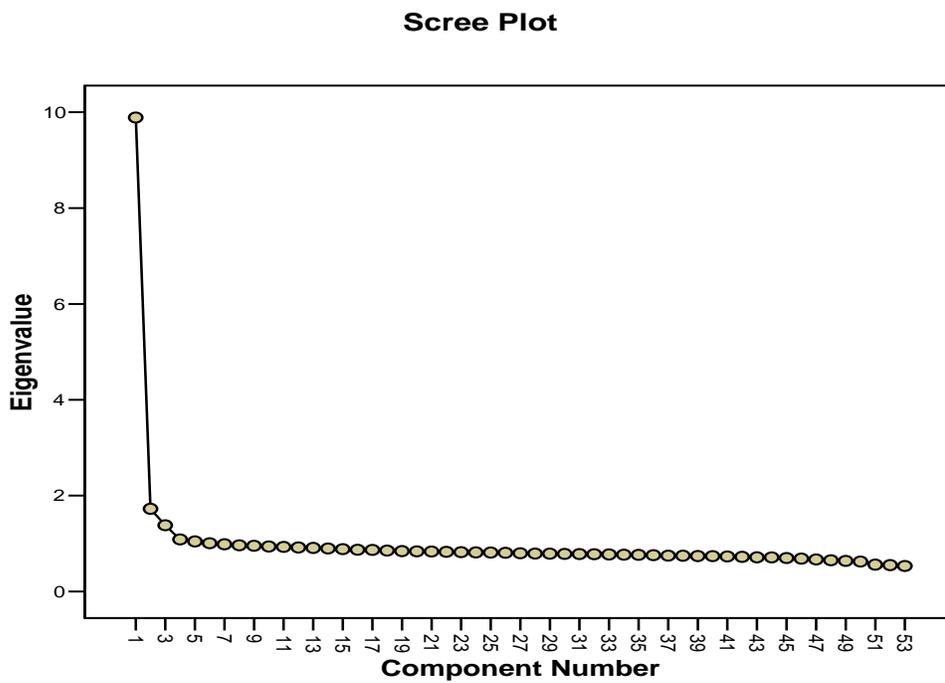


Figure 9.34: Scree Plot Science, Grade 8

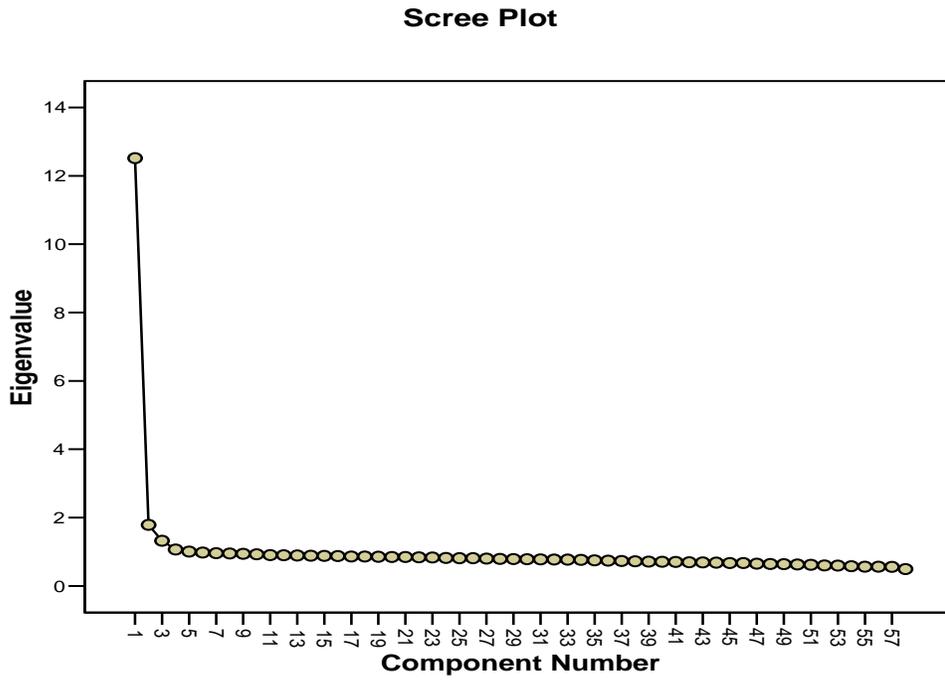
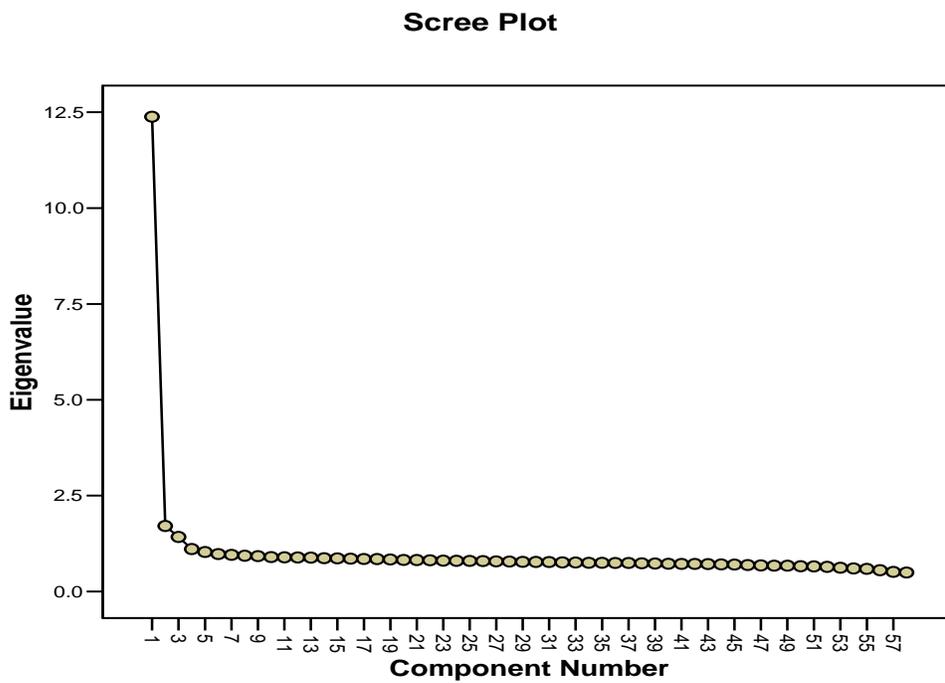


Figure 9.35: Scree Plot Science, Grade 11



CHAPTER 10: FAIRNESS

As noted in the *Standards*, there are varying definitions of fairness. In this chapter, we first examine fairness as it relates to minimizing bias on the test. We then look at test performance among varying subgroups assessed by MAP. It should be noted that differences in test performance among subgroups does not mean that a test is unfair—it simply means that groups perform differentially on the test. Even when a test is carefully and properly constructed, differences may exist among subgroups as a result of differences in curriculum or learning by students in the subgroup.

This chapter addresses AERA/APA/NCME standards 7.1, 7.2, 7.3, and 7.4.

10.1 Minimizing Bias through Careful Test Development

The development of a test that is fair for all examinees begins in the early stages of planning and development. The item and test development processes that are used to minimize bias are summarized below.

First, careful attention was paid to content validity during the item- development and item-selection processes. Bias can occur only if the test is measuring different things for different groups. By eliminating irrelevant skills or knowledge from the items, the possibility of bias is reduced.

Second, item writers and test developers followed several published guidelines for reducing or eliminating bias. These included Guidelines for Bias-Free Publishing (MacMillan/McGraw-Hill, 1993a) and Reflecting Diversity: Multicultural Guidelines for Educational Publishing Professionals (Macmillan/McGraw-Hill, 1993b). Test developers reviewed the items and other testing materials with these guidelines in mind. Internal editorial reviews were conducted by at least three different people: a content editor who directly supervised the item writers; a style editor; and a content supervisor. The final test was again reviewed by at least these same people, and was also subjected to an independent review by a quality assurance editor.

Third, careful attention is given to item statistics throughout the test development process. As part of the test assembly process, attempts are made to avoid using or reusing items with poor statistical fit or distractors with positive point biserial correlations, since this may indicate that an item is tapping an ability that is irrelevant to the construct being measured. Differential item functioning (DIF) statistics are also examined during test construction. Items that have exhibited significant DIF against one or more subgroups are removed from further consideration unless it is essential to include them in order to meet content specifications.

Additional steps to reduce bias, including the use of Bias Review committees comprised of Missouri participants, are described in more detail in Chapter 3 of this report.

10.2 Evaluating Bias through Differential Item Functioning Statistics

After administering the test, an empirical approach known as differential item functioning (DIF) was used to examine the items. The DIF statistics indicate the degree to which members of a particular subgroup performs better or worse than expected on each item as compared to the reference group. The DIF procedures used and the results of these analyses are detailed in this section.

The position of CTB/McGraw-Hill concerning test bias is based on two general propositions. First, students may differ in their background knowledge, cognitive and academic skills, language, attitudes, and values. To the degree that these differences are large, no one curriculum and no one set of instructional materials will be equally suitable for all. Therefore, no one test will be equally appropriate for all. Furthermore, it is difficult to specify what amount of difference can be called large and to determine how these differences will affect the outcome of a particular test. Second, schools have been assigned the tasks of developing certain basic cognitive skills and supporting development of these skills equitably among all students. Therefore, there is a need for tests that measure the common skills and bodies of knowledge that are common to all learners. The test publisher's task is to develop assessments that measure these key cognitive skills without introducing extraneous or construct-irrelevant elements in the performances on which the measurement is based. If these tests require that students have culturally-specific knowledge and skills not taught in school, differences in performance among students can occur because of differences in student background and out-of-school learning. Such tests are measuring different things for different groups and can be called biased (Camilli & Shepard, 1994; Green, 1975). In order to lessen this bias, CTB/McGraw-Hill strives to minimize the role of the extraneous elements, thereby, increasing the number of students for whom the test is appropriate. As discussed above and in Chapter 3 of this report, careful attention is given during the test development and test construction processes to lessen the influence of these elements for large numbers of students (including the use of Bias Review committees). Unfortunately, in some cases these elements may continue to play a substantial role. To assess the extent to which items may be performing differently for various subgroups of interest, DIF analyses are conducted after each operational test administration.

DIF statistics are used to quantify differences in item performance between two groups after controlling for examinees' overall achievement level. Two DIF statistics that are commonly used for this purpose are the Mantel-Haenszel statistics (1959) and the Standardized Mean Difference (SMD) between the reference and focal groups, proposed by Dorans and Schmitt (1991).

For selected-response items, the Mantel-Haenszel (χ^2_{MH}) statistic was used to evaluate potential DIF items. In the MH procedure, subgroups are matched by their raw total test score, using a contingency table with K ability levels. When applying the MH procedure, the log-odds ratio α is assumed to be constant across the K matched levels. The χ^2_{MH} , then, estimates a pooled common odds ratio. Taking the natural logarithm of the common odds ratio and its confidence limits and multiplying these with the constant -2.35 , the

resulting values may then be placed on the MH delta metric (Δ_{MH}) for interpretive purposes. Items were flagged for DIF using the following criteria:

- Moderate DIF: Absolute value of the Mantel-Haenszel (Δ_{MH}) is significantly greater than zero (at the .05 level) and $-1.5 \leq \Delta_{MH} \leq -1$ or $1 \leq \Delta_{MH} \leq 1.5$.
- Large DIF: Absolute value of the Mantel-Haenszel (Δ_{MH}) that is significantly greater than zero (at the .05 level) and $|\Delta_{MH}|$ exceeds 1.5.

For constructed-response items, an effect size (ES) statistic based on the Mantel χ^2 will be used. ES is obtained by dividing the standardized mean difference (SMD) statistics by the standard deviation of the item. (A detailed description of these procedures can be found in Zwick, et al., 1993). Items are flagged using the same rules that are used in NAEP:

- Moderate DIF: If the Mantel statistic is significant ($p < .05$) and $|ES|$ is between 0.17 and 0.25
- Large DIF: If the Mantel statistic is significant ($p < .05$) and $|ES| \geq 0.25$

A positive DIF value indicates that the item favors the focal group, while a negative value indicates that the item disadvantages the focal group. Tables 10.1, 10.2 and 10.3 show the DIF results for the following subgroups:

- **Gender:** Focal group is Females; Reference group is Males.
- **Ethnicity:** Focal groups are Black, Hispanic, Asian/Pacific Islander, Native American/Alaskan; Reference group is White.
- **Accommodations:** Focal group is students who received one or more testing accommodations; Reference group is all others.

A negative SMD value implies that the focal group has a lower mean item score than the reference group, whereas a positive value implies that the focal group has a higher mean item score than the reference group, conditioned on the matching test score.

The DIF analyses are not performed for subgroups of less than 100. In these cases, the statistical procedures do not have sufficient power to detect differences should they exist.

Tables 10.1, 10.2, and 10.3 summarize the number of DIF flags by grade for each focal group. For example, in Grade 6 Communication Arts, there were 2 items flagged for DIF for the female subgroup. In this case, the items were flagged in favor of female students. Two items were flagged for DIF for the Black subgroup: one item exhibited strong positive DIF while the other exhibited strong negative DIF. Finally, two items were flagged for DIF for students with accommodations and both items exhibited strong positive DIF.

Again, any items included on the MAP (including those items flagged for DIF) have been thoroughly reviewed for content and bias by Missouri teachers, DESE staff, and CTB Content Development staff. Further, these items were reviewed for possible DIF flags during the field test stage of test development. The DIF flags found on the operational assessment do not necessarily indicate that an item is biased; rather, DIF flags indicate that the item functions differently for members of different groups (Camilli & Shepard, 1994). All items flagged for DIF in the tables above had been thoroughly reviewed before inclusion on the operational MAP to insure that they do not tap knowledge or specific ability irrelevant to the construct the test intends to measure. Items are not suppressed from operational scoring if they are flagged for DIF.

10.3 Evaluating Bias through Impact Analysis

The impact of achievement testing on minorities can be determined and reported in the form of average scores and also in terms of test score reliability. Tables 10.4 through 10.11 present the scale score means and standard deviations, numbers of students, effect size (Cohen's D), and test form reliability statistics (Coefficient Alpha, see Chapter 9) for various subgroups of interest.

10.3.1 Reliability

Tables 10.4 through 10.11 show the test reliability for the various subgroups of interest. This analysis shows that the test reliability is of acceptable magnitude for all of the subgroups.

10.3.2 Effect Size

Some believe that fairness is an issue whenever the measured ability differences between subgroups are overly large; however, a criterion for large difference is lacking. One way to evaluate the magnitude of the differences is to calculate the effect size. Cohen's D was used to calculate the effect size. Cohen's D is given by the formula:

$$d = \frac{\bar{x}_a - \bar{x}_b}{\sqrt{\frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{(n_a + n_b) - 2}}}$$

where \bar{x}_a is the mean score of group A, \bar{x}_b is the mean score of group B, s_a^2 is the variance of group A, s_b^2 is the variance of group B, n_a is the number of students in group A, and n_b is the number of students in group B.

Cohen's d , then, expresses the difference in group means in terms of the standard deviation. For example if $d=.34$ for two groups, then it may be interpreted as the mean difference between the two groups is .34 of the pooled standard deviation. Cohen (1988) offered guidelines for interpreting the meaning of the d statistic: $d=.20$ is a small effect

size, $d=.50$ is a medium effect size, and $d=.80$ is a large effect size. Even with these guidelines, caution should be used when judging the differences between the groups compared as there is debate in the measurement field regarding the appropriateness of these guidelines for standardized testing (Holland 2000).

Using Cohen's (1988) guidelines, certain trends become apparent in Tables 8.3-8.16. On the Communication Arts test in all grades, except Grade 5, gender has a small effect on mean test scores where girls outperform boys. On the Communication Arts, Mathematics, and Science test in all grades, accommodations tend to have a large effect on the mean test scores where students in the accommodated groups underperform students who are not in those groups.

In terms of the race/ethnicity in all grades, there is a moderate difference in mean Communication Arts test scores of black students compared to white students, where black students underperform white students on average. There is a small effect on mean test scores where Hispanics underperform white students on the Communication Arts in Grades 3 through 7 and 11 (both the general and breach form). In Grade 8, there is a moderate effect for Hispanics compared to white students. There is a small effect on the mean test scores, where Native Americans underperform white students on Communication Arts in Grades 4 through 8 and 11.

There is a moderate difference in mean Mathematics tests scores of black students compared to white students in Grade 3. In Grades 4 through 8 and 10, there is a large difference in mean Mathematics test scores of black students compared to white students. There is a small difference in mean Mathematics test scores of Hispanic students compared to white students in Grades 3 through 7, and a moderate difference in mean Mathematics scores in Grades 8 and 10. There is a small effect on mean test scores of Native American students compared to white students, where Native Americans underperform white students in Grades 4, 6, 7, 8, and 10 in Mathematics. Finally, there is a small effect on mean test scores of Asian students, where Asian students outperform white students in Grades 3 through 7 in Mathematics.

There is a large effect on mean Science test scores of Black students compared to white students in Grades 5, 8, and 11, where Black students underperform white students. There is a moderate effect on mean Science test scores of Hispanic students compared to white students in Grades 8 and 11 and a small effect in Grade 5, where Hispanic students underperform white students.

Table 10. 1: 2008 MAP DIF Statistics: Number of Flagged Items, Communication Arts

Grade	Group	Sample Size	Moderate Positive	Moderate Negative	Large Positive	Large Negative
3	Female	26621	0	0	1	1
	Asian/Pacific Islander	1096	0	1	5	2
	Black	10318	0	0	2	1
	Hispanic	2284	0	0	1	0
	Native American/Alaskan	200	0	0	0	0
	Accommodations	5056	1	0	3	0
4	Female	27616	0	0	0	0
	Asian/Pacific Islander	1072	0	1	3	2
	Black	10677	0	0	0	1
	Hispanic	2377	0	0	2	0
	Native American/Alaskan	208	0	0	0	0
	Accommodations	6044	0	0	3	0
5	Female	25627	0	0	0	0
	Asian/Pacific Islander	1034	0	0	3	3
	Black	9849	2	0	0	2
	Hispanic	2069	1	0	1	1
	Native American/Alaskan	209	0	1	0	1
	Accommodations	6043	0	0	0	0
6	Female	26343	1	0	1	0
	Asian/Pacific Islander	1041	0	0	0	2
	Black	9871	0	0	1	1
	Hispanic	1903	0	0	0	0
	Native American/Alaskan	238	0	0	0	0
	Accommodations	5711	0	0	2	0
7	Female	28425	0	1	4	1
	Asian/Pacific Islander	1141	1	1	1	2
	Black	10553	0	1	0	1
	Hispanic	2092	0	0	1	0
	Native American/Alaskan	262	0	0	0	0
	Accommodations	6046	3	0	1	0
8	Female	27449	2	1	2	1
	Asian/Pacific Islander	919	2	0	3	3
	Black	10598	1	0	1	3
	Hispanic	1801	0	0	4	2
	Native American/Alaskan	283	0	0	0	1
	Accommodations	5321	0	0	1	0
11	Female	31140	1	0	6	3
	Asian/Pacific Islander	1145	3	1	4	3
	Black	9588	2	0	3	2
	Hispanic	1548	2	2	1	1
	Native American/Alaskan	275	0	0	1	0
	Accommodations	4729	1	2	1	1

Table 10. 2: 2008 MAP DIF Statistics: Number of Flagged Items, Mathematics

Grade	Group	Sample Size	Moderate Positive	Moderate Negative	Large Positive	Large Negative
3	Female	26683	0	0	1	0
	Asian/Pacific Islander	1128	1	0	3	2
	Black	10333	0	0	2	1
	Hispanic	2326	1	0	1	0
	Native American/Alaskan	200	0	0	0	0
	Accommodations	5278	1	1	1	0
4	Female	27664	0	0	1	0
	Asian/Pacific Islander	1097	0	1	4	1
	Black	10698	0	0	3	0
	Hispanic	2414	0	0	1	1
	Native American/Alaskan	209	0	0	2	1
	Accommodations	6301	0	0	3	0
5	Female	25676	0	0	2	0
	Asian/Pacific Islander	1071	0	0	1	0
	Black	9884	0	0	3	0
	Hispanic	2096	0	0	0	0
	Native American/Alaskan	210	0	0	0	0
	Accommodations	6295	0	0	0	1
6	Female	26375	1	0	0	1
	Asian/Pacific Islander	1070	0	0	0	0
	Black	9882	0	0	0	1
	Hispanic	1927	0	0	0	0
	Native American/Alaskan	237	0	0	0	0
	Accommodations	5921	0	0	0	1
7	Female	28455	1	1	4	2
	Asian/Pacific Islander	1167	0	0	1	1
	Black	10555	1	0	0	0
	Hispanic	2125	0	0	1	0
	Native American/Alaskan	261	0	0	0	1
	Accommodations	6295	0	1	1	0
8	Female	27484	0	0	2	1
	Asian/Pacific Islander	936	1	0	1	2
	Black	10605	1	0	1	0
	Hispanic	1829	0	0	1	0
	Native American/Alaskan	284	0	0	0	0
	Accommodations	5643	0	0	1	2
11	Female	32691	0	0	4	0
	Asian/Pacific Islander	1088	3	1	2	1
	Black	11115	0	0	0	0
	Hispanic	1914	0	0	2	0
	Native American/Alaskan	318	0	0	0	0
	Accommodations	5670	0	0	0	0

Table 10. 3: 2008 MAP DIF Statistics: Number of Flagged Items, Science

Grade	Group	Sample Size	Moderate Positive	Moderate Negative	Large Positive	Large Negative
5	Female	25655	1	0	1	4
	Asian/Pacific Islander	1070	1	0	2	2
	Black	9870	0	0	3	0
	Hispanic	2094	0	0	3	1
	Native American/Alaskan	210	1	0	0	0
	Accommodations	6032	0	0	1	0
8	Female	27443	0	0	4	3
	Asian/Pacific Islander	938	0	1	3	1
	Black	10565	0	1	1	0
	Hispanic	1820	0	0	1	2
	Native American/Alaskan	282	0	0	0	0
	Accommodations	5369	1	0	0	0
11	Female	25424	3	1	2	4
	Asian/Pacific Islander	874	1	1	0	3
	Black	8154	0	0	1	2
	Hispanic	1280	0	0	1	1
	Native American/Alaskan	203	0	0	0	0
	Accommodations	3794	0	0	0	1

Table 10. 4: Impact Analysis, Grade 3

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	49942	642.65	35.62		0.91
		Black (not Hispanic)	11935	618.26	38.45	0.67	0.91
		Hispanic	2690	625.26	36.19	0.49	0.91
		Asian/Pacific Islander	1231	649.02	37.62	-0.18	0.91
		Native American	258	635.88	36.64	0.19	0.91
	Gender	Male	33879	632.81	38.39		0.92
		Female	32180	642.70	35.84	-0.27	0.91
	Accommodations	No	60003	641.97	33.81		0.90
		Yes	6176	595.15	44.78	1.34	0.92
Mathematics	Ethnicity	White (not Hispanic)	49946	626.85	35.01		0.91
		Black (not Hispanic)	11938	600.39	36.32	0.75	0.92
		Hispanic	2731	611.70	34.52	0.43	0.91
		Asian/Pacific Islander	1264	640.10	39.81	-0.38	0.91
		Native American	258	620.59	34.69	0.18	0.91
	Gender	Male	33911	621.99	38.06		0.92
		Female	32230	621.36	35.60	0.02	0.91
	Accommodations	No	59873	625.01	35.21		0.91
		Yes	6385	590.20	37.75	0.98	0.92

Table 10. 5: Impact Analysis, Grade 4

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	50497	659.85	31.79		0.90
		Black (not Hispanic)	12164	639.61	35.97	0.62	0.92
		Hispanic	2659	645.53	32.28	0.45	0.91
		Asian/Pacific Islander	1187	664.44	32.83	-0.14	0.90
		Native American	262	650.60	33.57	0.29	0.91
	Gender	Male	34297	650.77	34.57		0.91
		Female	32476	660.78	31.80	-0.30	0.90
	Accommodations	No	59721	660.29	29.09		0.89
		Yes	7152	616.53	42.42	1.42	0.92
Mathematics	Ethnicity	White (not Hispanic)	50498	649.38	31.87		0.91
		Black (not Hispanic)	12172	623.26	35.24	0.80	0.92
		Hispanic	2697	635.57	31.98	0.43	0.91
		Asian/Pacific Islander	1211	658.49	36.32	-0.28	0.92
		Native American	261	640.31	29.05	0.28	0.90
	Gender	Male	34333	644.24	35.24		0.92
		Female	32510	644.15	33.05	0.00	0.92
	Accommodations	No	59566	647.97	31.78		0.91
		Yes	7378	613.56	37.46	1.06	0.93

Table 10. 6: Impact Analysis, Grade 5

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	49831	675.73	31.95		0.90
		Black (not Hispanic)	11699	654.32	35.39	0.66	0.91
		Hispanic	2444	662.18	32.16	0.42	0.91
		Asian/Pacific Islander	1175	682.38	33.60	-0.21	0.91
		Native American	280	672.38	30.84	0.10	0.90
	Gender	Male	33344	668.35	35.65		0.92
		Female	32099	674.78	31.21	-0.19	0.90
	Accommodations	No	58055	676.35	28.98		0.89
		Yes	7489	633.73	42.72	1.38	0.91
Mathematics	Ethnicity	White (not Hispanic)	49841	667.51	37.73		0.90
		Black (not Hispanic)	11707	636.04	42.40	0.81	0.91
		Hispanic	2474	651.04	38.49	0.44	0.91
		Asian/Pacific Islander	1218	679.94	44.20	-0.33	0.92
		Native American	281	661.18	41.85	0.17	0.92
	Gender	Male	33396	661.54	42.16		0.92
		Female	32140	661.39	39.13	0.00	0.91
	Accommodations	No	57880	666.70	37.14		0.90
		Yes	7756	622.17	44.65	1.17	0.91
Science	Ethnicity	White (not Hispanic)	49809	667.73	27.66		0.89
		Black (not Hispanic)	11692	637.77	34.41	1.03	0.90
		Hispanic	2473	649.71	31.20	0.65	0.90
		Asian/Pacific Islander	1217	667.03	35.13	0.03	0.92
		Native American	281	663.23	29.14	0.16	0.90
	Gender	Male	33371	662.96	32.38		0.91
		Female	32116	660.32	30.50	0.08	0.90
	Accommodations	No	58159	665.01	28.85		0.90
		Yes	7427	635.26	38.37	0.99	0.91

Table 10. 7: Impact Analysis, Grade 6

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	50403	675.23	31.73		0.89
		Black (not Hispanic)	11415	654.69	35.58	0.63	0.90
		Hispanic	2235	662.72	32.48	0.39	0.90
		Asian/Pacific Islander	1248	680.06	34.95	-0.15	0.91
		Native American	289	668.34	31.85	0.22	0.89
	Gender	Male	33419	667.63	35.11		0.91
		Female	32176	675.10	31.27	-0.22	0.89
	Accommodations	No	58601	676.07	28.87		0.88
		Yes	7071	631.45	41.73	1.46	0.90
Mathematics	Ethnicity	White (not Hispanic)	50392	684.34	38.15		0.91
		Black (not Hispanic)	11417	652.99	43.51	0.80	0.91
		Hispanic	2260	668.50	37.59	0.42	0.90
		Asian/Pacific Islander	1278	694.41	45.19	-0.26	0.92
		Native American	288	672.58	40.29	0.31	0.92
	Gender	Male	33445	677.75	42.46		0.92
		Female	32193	679.27	39.65	-0.04	0.91
	Accommodations	No	58463	683.63	37.41		0.90
		Yes	7253	636.82	45.75	1.22	0.91

Table 10. 8: Impact Analysis, Grade 7

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	50853	680.93	32.67		0.91
		Black (not Hispanic)	11873	655.41	36.41	0.76	0.92
		Hispanic	2303	666.60	35.34	0.44	0.92
		Asian/Pacific Islander	1298	684.37	39.54	-0.10	0.93
		Native American	291	670.53	35.11	0.32	0.92
	Gender	Male	34334	670.87	36.57		0.92
		Female	32266	681.28	32.51	-0.30	0.91
	Accommodations	No	59670	681.11	30.59		0.90
		Yes	7031	631.40	39.04	1.57	0.90
Mathematics	Ethnicity	White (not Hispanic)	50834	687.73	38.29		0.91
		Black (not Hispanic)	11859	653.69	41.36	0.88	0.90
		Hispanic	2336	669.84	39.57	0.47	0.91
		Asian/Pacific Islander	1325	698.43	48.70	-0.28	0.94
		Native American	291	673.72	41.46	0.37	0.92
	Gender	Male	34340	681.71	43.23		0.93
		Female	32287	680.67	39.24	0.03	0.92
	Accommodations	No	59480	686.68	37.46		0.91
		Yes	7247	635.79	44.06	1.33	0.89

Table 10. 9: Impact Analysis, Grade 8

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	51293	696.20	30.87		0.90
		Black (not Hispanic)	12200	671.45	35.32	0.78	0.91
		Hispanic	2149	678.61	35.65	0.57	0.91
		Asian/Pacific Islander	1175	696.69	38.18	-0.02	0.93
		Native American	330	687.63	32.76	0.28	0.91
	Gender	Male	34384	686.53	35.10		0.92
		Female	32766	695.91	31.02	-0.28	0.91
	Accommodations	No	60723	695.67	29.30		0.90
		Yes	6555	648.18	39.77	1.56	0.90
Mathematics	Ethnicity	White (not Hispanic)	51289	708.24	35.76		0.92
		Black (not Hispanic)	12198	673.27	40.07	0.95	0.90
		Hispanic	2175	690.11	38.98	0.51	0.92
		Asian/Pacific Islander	1192	714.34	46.80	-0.17	0.95
		Native American	331	697.35	35.51	0.30	0.90
	Gender	Male	34398	701.09	41.19		0.93
		Female	32789	701.64	37.32	-0.01	0.92
	Accommodations	No	60451	706.13	35.90		0.92
		Yes	6861	658.72	43.04	1.29	0.89
Science	Ethnicity	White (not Hispanic)	51238	700.71	26.69		0.92
		Black (not Hispanic)	12155	669.55	32.18	1.12	0.92
		Hispanic	2167	683.18	31.77	0.65	0.93
		Asian/Pacific Islander	1194	697.65	36.25	0.11	0.95
		Native American	329	691.02	29.49	0.36	0.92
	Gender	Male	34335	695.01	31.82		0.94
		Female	32750	693.75	29.34	0.04	0.93
	Accommodations	No	60649	697.77	28.15		0.92
		Yes	6560	662.76	34.69	1.21	0.92

Table 10. 10: Impact Analysis, Grade 11 (Communication Arts, Science), Grade 10 (Mathematics)

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	49255	717.58	34.39		0.91
		Black (not Hispanic)	8810	692.55	35.95	0.72	0.91
		Hispanic	1420	702.11	35.61	0.45	0.91
		Asian/Pacific Islander	1117	721.37	37.94	-0.11	0.92
		Native American	274	708.36	32.79	0.27	0.90
	Gender	Male	30370	708.61	37.61		0.92
		Female	30536	718.59	33.32	-0.28	0.91
	Accommodations	No	56267	717.51	32.77		0.91
		Yes	4774	667.08	38.36	1.52	0.89
Mathematics	Ethnicity	White (not Hispanic)	53731	737.67	45.61		0.93
		Black (not Hispanic)	11329	691.39	48.41	1.00	0.91
		Hispanic	1950	713.31	47.89	0.53	0.93
		Asian/Pacific Islander	1223	746.06	53.98	-0.18	0.95
		Native American	330	720.07	46.10	0.39	0.93
	Gender	Male	34672	728.95	52.08		0.94
		Female	33900	729.83	46.72	-0.02	0.93
	Accommodations	No	62862	734.94	45.73		0.93
		Yes	5914	669.24	48.74	1.43	0.89
Science	Ethnicity	White (not Hispanic)	49362	728.72	36.01		0.92
		Black (not Hispanic)	9634	691.03	43.51	1.01	0.92
		Hispanic	1553	708.32	40.70	0.56	0.92
		Asian/Pacific Islander	1155	727.97	45.14	0.02	0.95
		Native American	268	719.18	34.78	0.26	0.92
	Gender	Male	30834	724.07	42.45		0.94
		Female	31143	720.52	37.40	0.09	0.93
	Accommodations	No	57395	726.18	36.83		0.93
		Yes	4738	673.94	46.32	1.39	0.91

Table 10. 11: Impact Analysis, Grade 11 Breach

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size	Coefficient Alpha
Communication Arts	Ethnicity	White (not Hispanic)	273	712.19	40.72		0.93
		Black (not Hispanic)	853	686.35	42.06	0.62	0.92
		Hispanic	134	694.49	41.43	0.43	0.92
		Asian/Pacific Islander	31	713.48	51.66	-0.03	0.92
		Native American	1	650.00	-	-	-
	Gender	Male	586	686.61	46.73		0.93
		Female	702	699.33	38.82	-0.30	0.92
	Accommodations	No	1244	695.77	41.24		0.92
		Yes	55	633.82	46.10	1.49	0.86

REFERENCES

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association, Inc.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika*, 46, 443–459.
- Burket, G. R. (1995). PARDUX [Computer program]. Monterey, CA: CTB/McGraw-Hill.
- Camilli, G., & Shepard, A. L. (1994). *Methods for Identifying Biased Test Items*. Sage Publication.
- Candell, G.L. & Drasgow, F. (1988). An iterative procedure for linking metrics bias in item response theory. *Applied Psychological Measurement*, 12(3), 253-260.
- Cattell, R.B. (1966), "The Scree Test for the Number of Factors," *Multivariate Behavioral Research*, 1, 245 -276.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*. 16, 297-334.
- CTB/McGraw-Hill (2005). *Missouri Assessment Program Final Bookmark Standard Setting Technical Report*. Monterey, CA: Author.
- CTB/McGraw-Hill (2008). *Guide to Interpreting Results*. Author.
- CTB/McGraw-Hill (2008). *Missouri Assessment Program Bookmark Standard Setting Technical Report 2008 for Missouri Achievement-Level Setting Grades 5, 8, and 11 Science*. Author.
- Dorans, N.J., & Schmitt, M.P. (1991). *Constructed response and differential item functioning: A pragmatic approach*. Princeton: Educational Testing Service.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.

- Green, D.R. (1975). *Procedures for assessing bias in achievement tests*. Paper presented at the National Institute of Education Conference on Test Bias, Annapolis, MD.
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Hingham, MA: Kluwer-Nijhoff Pub.
- Karkee, T. & Choi, S. (2005). Impact of eliminating anchor items flagged from statistical criteria on test score classifications in common item equating. Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Canada, Apr 15, 2005)
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Macmillan/McGraw-Hill. (1993a). *Guidelines for Bias-Free Publish*. New York: Author.
- Macmillan/McGraw-Hill. (1993b). *Reflecting diversity: Multicultural guidelines for educational publishing professionals*. New York: Author.
- 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.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7*, 201–210.
- Thissen, D. (1982). Marginal maximum-likelihood estimation for the one-parameter logistic model. *Psychometrika, 47*, 175–186.
- Voelkle, M., Schwarz, R., Arenson, E., & Ito, K. (2002). An investigation of Factors Affecting Stocking & Lord Equating. (Paper in Progress.)
- Yen, W. M. (1981). Using simulation results to choose a latent trait model. *Applied Psychological Measurement, 5*, 245–262.
- Yen, W. M. (1993) . Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement, 30*, 187-213.
- Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30*, 233-251

Appendix A

<i>DESE Presentation on Test Coordinator's Manual.....</i>	<i>A-1</i>
<i>DESE Presentation on Test Examiner's Manual.....</i>	<i>A-11</i>
<i>Presentation on Test Administration.....</i>	<i>A-18</i>

DESE Presentation on Test Coordinator's Manual

Test Coordinator's Manual 2008



1

2008 MAP Test Coordinator's Manual (TCM)

The TCM is primarily focused on the responsibilities of:

- District Test Coordinators (DTC's); and
- School Test Coordinators (STC's).

The TCM also contains:

- An Overview;
- Timelines; and
- A Glossary of Terms.

2

Missouri Assessment Program

MAP 2008

Communication Arts.....Grades 3-8 & 11
Math.....Grades 3-8 & 10
Science.....Grades 5, 8 & 11

3

Dates to Remember:

[Arrival of Materials](#)
Feb. 20-27, 2008.....Test Coordinator packages to district;
March 3-14, 2008.....Test Materials arrive in districts;

[Testing Window](#)
March 31-April 25, 2008.... Administer Assessments;

[Testing Materials](#)
April 11, 2008.....Deadline for ordering additional testing materials without incurring additional shipping cost;

NOTE: If testing materials are ordered after April 11, the district will be responsible for **paying** the shipping costs.

April 21, 2008.....Final deadline for ordering additional testing materials;

[Return of Materials](#)
April 28, 2008.....Deadline to contact CTB/McGraw-Hill for pickup of testing materials;
May 1, 2008.....Deadline for testing materials to be picked up by CTB's transportation vendor;

[Test Results](#)
August-October 2008.....Reports shipped to districts.

4

DTC's Responsibilities:

- Checks, reviews and distributes testing materials to the STC's;
- Trains STC's on MAP processes;
- Assumes STC role when necessary;
- Guarantees security for all testing materials;
- Acts as sole channel for all communication between districts and CTB MAP Service Line; and
- Collects all testing materials after the test and returns to CTB/McGraw-Hill.

5

STC's Responsibilities

- Receives and checks all testing materials from the DTC;
- Assumes DTC role when necessary;
- Guarantees security of all testing materials;
- Disseminates Examiner's Manuals;
- Trains all Examiners;
- Checks [Group Information Sheets](#) (GISs);
- Completes [School/Group Lists](#); and
- Collects all testing materials after testing, checks and organizes materials for return to the DTC.

6

Examiner's Responsibility

- Examiner responsibilities are addressed in the 2008 MAP Test Examiner's Manual provided for each grade.

7

Summary of 2008 Changes

- 2008 MAP includes all required tests, bound together in one book:
 - Communication Arts; Mathematics; and Science.
- Content area labels are no longer required.
- Voluntary Social Studies assessment will not be offered in 2008.
- In 2009-2010 Government and American History required EOC Assessments.
- Scorable and unused test materials **MUST** be boxed separately.

8

- A Missing Test Materials Form is used to notify CTB of contaminated and/or destroyed test books.
- Form included in the Test Coordinator's Kit.
- Grade 10 has an individual Mathematics' test book.
- Grade 11 has individual Science and Communication Arts test books.
- Content areas are bound together in alphabetical order.

9

- A single School/Group List for all books is being used this year.
- Field test items are embedded throughout the operational form of the MAP this year.
- One Group Information Sheet (GIS) per grades 3 -10;
- Two GIS for 11th grade:
 - One GIS for C.A.; &
 - One GIS for Science.
- DTC's responsibilities changed for organizing and packing testing materials.

10

- STC's responsibilities now include:
 - Packing test materials in the return, shipping boxes; and
 - Affixing the return, shipping labels.
- Test materials are now required to be shipped in **GREEN** shaded boxes provided by CTB/McGraw-Hill.

IF	THEN
More boxes are needed --	The DTC will order them.
Boxes are used other than those provided by CTB/McGraw-Hill --	Green stock labels, which are included in the Test Coordinator's Kit, must be affixed to them.

11

- After testing, the STC should collect and securely destroy all Examiner's Manuals.
- A 5% overage of test books is included in each school's shipment.
- Plus, a 5% overage is also sent to the DTC.
- For students who can not use their student barcode labels, a new line is added to the test book front cover to allow students to add their district and school name.
- The Test Book Accountability Form replaces the Security Barcode Verification Form.

12

- Neither an SIS nor a blank test books will be returned for students taking the MAP-Alternate Assessment (MAP-A).

- A new bubble was added for the Blind/Visually Impaired student who does not read Braille:
04 Oral Reading – Blind/Partial Sight.

13

NCLB requires **all** Missouri students to take the MAP test.

Only two groups are exempt:

- Group 1: MAP-Alternate (MAP-A); and
- Group 2: ELL students in the United States, 12 months or fewer, may be exempt from taking the Communication Arts test.

14

Inclusion of Special Populations

- Accommodations for special populations can be found on the DESE website.
- The following are examples of special populations:
 - IEP students;
 - Individual Accommodation Program (IAP 504) students;
 - Students not tested:
 - MAP-A
 - ELL
 - Out-of-District students;
 - Homebound students;
 - Home-schooled students;
 - H.S. Career Education students;
 - Gifted Students; and
 - MoVIP students.

15

Students Testing Out-of-District

- The home DTC (where the student is enrolled) delivers the testing materials to the serving districts/agencies;
- After testing, the completed materials are returned back to the home DTC; and
- The GIS determines where students' result will be reported.

16

Homebound Students:

- Must take the test if they are receiving homebound services;
- Must have test delivered by a trained Examiner who guarantees the security of testing materials; and
- May be tested either at home or school.

17

Home-School Students

- May take part with the local district at the district's discretion;
- Must test if receiving services; and
- Must test at local school.

18

Student Make-up Sessions

IF	THEN
Absent during one or more sessions --	Schedule make-up session
Student is unable to test during make-up session --	Follow Student Absences Procedures

19

Student Absences

1. Write student's name on front of the unused test book;
2. Affix student barcode label if accurate;
3. Complete SIS if student barcode label is inaccurate;
4. Code SIS for absences; and
5. Return test book with **scorable** books.

20

Large Print/Braille Procedures

- Must be transcribed to a regular edition test book to receive score;
- Must be labeled, **“Contents transcribed to a regular test book -- DO NOT SCORE”**;
- Must use special handling and packaging instructions; and
- Must return with **unused** testing materials.

21

Translators

For Mathematics and Science MAP Assessments only:

- ELL students can use their native language to give oral or written responses to assessments;
- ELL students' responses must be translated into English and scribed verbatim into a regular test book;
- Translators must be trained in administering the MAP;
- Translators can review tests before administration in a secure environment;
- Translators must guarantee security of MAP testing materials; and
- ELL accommodation codes apply (see Examiner's Manual codes).

22

Invalidations

- Three categories for test invalidation:
 - Student discovered cheating;
 - Examiner paraphrased test questions in any content area; or
 - Examiner reads any part of the Communication Arts test to students.

23

Invalidation Procedures

The SCT must:

- Be in agreement with Examiner regarding the test invalidation;
- Complete the Teacher Invalidation grid on the SIS;
- Provide demographic information to the DTC; and
- Include invalidate test materials with the **scorable** testing materials.
- DTC sends written communication to the Director of Assessment **with a copy to Director of Accountability, Data and Accreditation.**

24

Check lists for DTC and STC

- Convenient check lists are provided to the DTC and STC on pages 7 and 8 of the TCM;
- DTC and STC must ensure all functions are completed; and
- DTC and STC are responsible for both check lists if either of the responsible parties/roles is not available.

25

Student Barcode Labels

- Student barcode labels are in the Test Coordinator's Kit;
- Student barcode labels include:
 - Information from MOSIS Phase III (pre-code system); and
 - Demographic information (but not all biographical data) from the pre-code file;
- CTB-McGraw-Hill prints and provides:
 - One biographical master label; and
 - Two student barcode labels; and
- DTC and STC do the following:
 - Check student barcode label against the student pre-code roster for accuracy;
 - Determine viability of labels; and
 - Handle exceptions appropriately.

26

Test Book Cover (student barcode label positioning)

AFFIX

27

Handling Student Barcode Labels

- Error in: *Student Name, Birth Date, Gender, Race/Ethnicity, and MOSIS ID* → **Don't use barcode; bubble in all info on SIS**
- If label is affixed and then found to be wrong → Place **two blank labels** over the incorrect label and then **bubble all** information the SIS
- Wrong student label is affixed → Place blank label over incorrect label; and **Then affix correct label**

28

Using Student Barcode Labels

No student barcode label

Notify local student data management person to enroll/submit the student in core data.
Bubble in all info on SIS
Leave barcode area blank!

29

Step 1: Review Testing Materials

- The District Test Coordinator's Kit includes two folders for **EACH** school
- One for the DTC; and
 - One for the STC.

30

DTC Folder

- District Packing List;
- District cover Letter;
- Materials Ordering and Inventory Information Flyer;
- Test Book Accountability Form;
- Add/Short Form – District;
- Missing Test Materials Form;
- Braille Omit Return Instruction Sheet;
- Student Barcode Label Instruction Sheet;
- Test Coordinator's Manuals; and
- Blank District Return Shipping Labels:
 - Blue for scorable materials; and
 - White for unused Inventory.

31

STC Folder

- School Packing List;
- School Cover letter;
- Add/Short Form – School;
- Braille Omit Return Instruction Sheet;
- School Group Lists (SGL's);
- Group Information Sheet (GIS) for each teacher and grade;
- Test Coordinator's Manuals;
- Blank Green stock labels;
- Return Shipping Labels:
 - Blue for scorable materials; and
 - White for unused Inventory;
- Student Barcode Label Instruction Sheet;
- Student Pre-code Roster;
- Blank Barcode Label Stock; and
- Student Barcode Labels.

32

Testing Materials

- Packaged by school;
- Shipped to the district's office or the designated address used in the online order;
- Include:
 - Examiner's Manuals;
 - Test books;
 - Large white envelopes; and
 - Ancillary testing materials.

33

Verify Shipment

- Compare packing list material against shipment; and
- Follow procedures for ordering more materials if needed, using Add/Short Form.

34

Map Spring 2008 Ordering Additional Testing Materials

Start Date	End Date	Shipping Mode	Shipping Costs
March 14	April 11	USP ground Service	CTB
April 14 th	April 21 st	Next-day or 2 nd day service	District

35

Step 2: Distribute Testing Materials

DTC's responsibilities:

- Maintains security;
- Tracks security barcodes to confirm start and end barcode numbers for each shrink-wrapped bundle;
- Matches numbers with packing list; and
- Reports discrepancies to CTB.

36

Test Book Accountability Form

- This form replaces the past Security Barcode Form;
- Deals with the security barcode information;
- One copy is needed for **each** school in the district;
- The form is used to ensure 100% accountability of test books;
- Instructions for the DTC and STC are outlined on page 13 - 15 in TCM;
- Pretest and post-test responsibilities regarding the form apply to both the DTC and STC; and
- This form or a copy of it is **never** given to the Examiners.

37

Security Barcode

- Used for test book security;
- Used for inventory – each book consecutively numbered;
- Used to ensure 100% accountability of test books;
- Used for missing inventory reports generated by CTB/McGraw-Hill;
- Used by DESE to track barcode numbers, district and school name; and
- Located on lower right-hand corner directly above "Spring 2008"

38

Security SHALL NOT'S

- Testing materials **shall not** be photocopied, duplicated, or made accessible to non-testing personnel;
- Testing materials **shall not** be viewed by Examiners before testing;
NOTE: The exceptions are ELL translators and Special Education teachers who are pre-selecting items.
- Testing materials **shall not** be left in an unsecured area at any time, for any reason -- must be locked in room or cabinet at school or district office before, between and after testing sessions; and
- Test books **shall not** be shared between schools.

39

7 Steps in the Life Cycle of Test Administration

1. **DTC** guarantees the security of the testing materials -- every school must have sufficient, satisfactory and locked security;
2. **DTC** houses materials at the district office if the school's security is insufficient, unsatisfactory or unlocked;
3. **DTC** distributes the materials to all **STC's** in the district;
4. **DTC** delivers appropriate testing materials for out-of-district students, prior to the first day of testing. Also, the **DTC** makes arrangements for returning materials after testing is complete;
5. **STC** distributes the 2008 MAP *Examiner's Manuals* to all who will administer the test as soon as possible ;
6. **STC** collects all student draft work and scratch paper and securely destroys after testing; and
7. **DTC** saves their folder and boxes for use after test administration.

40

Materials Needed for Each Examiner

- *Examiner's Manual* for appropriate grade level;
- Appropriate quantities of books, manipulatives and reference sheets;
- Student barcode labels for each student;
- Pre-coded Group Information Sheet (GIS):
 - One for each grade, 3-8;
 - One for Grade 10 Mathematics;
 - One for Grade 11 Communication Arts;
 - One for Grade 11 Science; and
 - No separate GIS for Special Education students.
- Appropriate quantity of large white envelopes (each holds approximately 5-10 books); and
- School/Group List for use after test administration.

41

Step 3: Collect Testing Materials

- All used and unused books must be accounted for and returned;
- One test book is returned for each student, except Grade 11;
- All manipulatives, reference sheets, scratch paper, extra envelopes, and contaminated test books are secured destroyed;
- Examiner's Manuals are collected and destroyed; and
- Bulleted items on pages 19 and 20 are used to ensure that all tasks are completed appropriately.

42

Markings in Test Books

In order to be scored properly, the following must take place:

1. Test books completed in ink are transcribed into another test book with a **non-mechanical #2 pencil**;
2. Student responses written on coding tracks/margins are erased and transcribed onto the response line or box;
3. All stray marks on the coding track are erased;
4. All underlining of text is erased around the answer choices; and
5. Test books marked with highlighters are transcribed into other test books for scoring.

43

Step 4: Check the Organization of Materials Collected

Collect the following after testing:

- Large white envelopes with all items in the exact order prescribed on page 21-22 and the illustration on page 34;
- All MAP test books, (used and unused books boxed separately), including student barcodes and/or completed SIS's for each students;

Unused books include:

- Test books which are damaged or have manufacturing errors;
 - Test books written in a language other than English;
 - Test books that are partially used; and
 - Test book that are Braille/Large Print editions with contents transcribed to a regular test book.
- Completed GIS.

44

Step 5: Check Student Information Sheet (SIS)

- Each Examiner's materials **MUST** be grouped together;
- Physical Conditions of SIS can affect/interfere with scoring. Those conditions are listed on pages 23-24;
- Biographical data must be checked on the barcode label & pre-code roster;
- SIS of the test book must be completed **only** when:
 - Pre-coded student barcode label can not be used; and
 - Pre-code roster information is inaccurate for a student.
- A completed SIS and/or a test book with a student barcode label must be received for every eligible student.

45

Sample SIS

- A picture of the SIS and explanations of the fields can be found on pages 25-26.
- Reminder: Identify Special Populations and Invalidations.

46

Step 6: Check the Teacher/Group Information Sheet (GIS)

- GIS provides data that is used on reports – notify the DTC if any errors exist on the GIS;
- GIS is submitted for each grade, except Grade 11 which requires two;
- GIS has both hand-entered and pre-coded information – **both must be accurate**;
- GIS is scannable and **can not** be photocopied; and
- GIS's are placed on top of test books whose scores are to be reported together.

47

Step 7: Complete School/Group List (S/GL)

- S/GL is used by CTB McGraw-Hill to inventory test books;
- S/GL can be photocopied;
- S/GL should have an entry for every GIS that was completed;
- S/GL information includes:
 - Pre-coded:
 - District Name and code number;
 - School Name and code number; and
 - Country Code Number.
 - Hand-written:
 - Contact person's phone number;
 - Group Name is identical to "Teacher Name" on GIS;
 - Grade Number (except Grade 11 only, indicate Communications Arts or Science);
 - Number of books being returned; and
 - Total number of students (tested, MAP- A, invalidated, absent all sessions, unused books).
- Sample S/GL on page 32; and
- Directions for completing fields on page 33.

48

Step 8: Organize Materials for the DTC

The STC will:

1. Reuse the CTB/McGraw-Hill green-shaded boxes in which testing materials arrived;
2. Package the following scorable materials:
 - School/Group Lists; and
 - Large white envelopes, organized by grade in ascending order, accompanied by GIS forms.
3. Package unused testing materials;
4. Affix shipping labels and number each set of boxes separately:
 - Blue labeled on scorable books, numbered 1 of X, 2 of X, etc.
 - White labels on unused books, numbered 1 of X, 2 of X, etc.
5. Send materials to the DTC in unsealed boxes.

49

Step 9: Package and Ship Testing Materials

The DTC will:

- Ensure all testing materials are received from each school in the district;
- Verify TCM instructions for the STC were followed exactly;
- Contact any STC who delays returning materials;
- Added packing material;
- Schedule testing material pickup online;
- Schedule pickup no later than **April 28, 2008**;
- Follow instructions outlined in TCM on page 37-38; and
- Fax Test Book Accountability Forms to CTB/McGraw-Hill.

50

Glossary of Terms

A glossary is listed at the end of the manual from pages 39-41. Most terms are the same as previous years. Two noteworthy changes are as follows:

- Level Not Determined; and
- Valid Attempt

51

Level Not Determined (LND)

This designation is for students who did not receive a MAP score for any one of the following three reasons:

1. An SIS is returned to CTB/McGraw-Hill with a blank test book;
2. A student does not attempt any items in one or more sessions; or
3. A student is absent all 3 sessions.

52

Valid Attempt

53

Questions???? Questions????



Missouri Department of Elementary
and Secondary Education
800-845-3545 (Assessment Section)

54

DESE Presentation on Test Examiner's Manual

MAP Examiner's Manual 2008



1

This presentation will include:

- An overview of MAP requirements;
- Index overview;
- Changes in the 2008 MAP; and
- Glossary overview.

2

Overview

- Missouri Assessment Program/Educational reform mandated by the Outstanding School Act of 1993
- Required: Communication Arts, Mathematics, and Science
- Types of items:
 - Selected Response
 - Constructed Response
 - Performance Events

DESE uses information obtained through MAP to:
–monitor the progress of students in meeting the Show-Me-Standards;
–inform the public and state legislature about students' performance; and
–help make informed decisions about educational issues.

3

Table of Contents

- Step 1 Plan Your Testing Schedule
- Step 2 Organize Your Classroom
- Step 3 Check Your Testing Materials
- Step 4 Before Testing (**Student Identification Information**)
- Step 5 Administer the Test
 - Communication Arts, Mathematics, and Science**
- Step 6 Invalidations and Make-ups
- Step 7 After Testing: Student Status Coding
- Step 8 Assemble Materials for Return
- Glossary

4

The Glossary is:

- the last 3 pages in Manual;
- an extensive list of terms associated with the MAP;
- contains terms that will be referred to while handling MAP materials and administering the assessment; and
- a helpful resource while reviewing the procedures/guidelines for administering the MAP.

5

Changes for 2008:

6

STEP 1

Be prepared for the test (pp.1-4)

1. Review testing schedules. (Page 1 for both content areas).
2. Review in advance the testing directions in the Examiner's Manual.
3. Mirror appropriate accommodations for MAP testing to students' current IEP or 504 accommodations -- more discussion to follow in step 7.
4. Provide advanced copies of the MAP test to Special Education teachers who are pre-selecting items for their qualifying students.
5. Provide advanced copies of the MAP test to translators who are translating the test to English Language Learners who qualifying for this accommodation.
6. Complete strictly-timed sessions within the designated, allotted time.
7. Allow completion of testing sessions to students who are making adequate progress, as determined by the examiner, for testing sessions which are **NOT** strictly timed.
8. Complete all testing sessions in one day.
9. Use proctors as needed -- approximately 1 per 50 students.

7

Use Standardized Testing Procedures (p.3)

- Following the instructions exactly ensures similar testing conditions in all classrooms.
 - Test directions should be read **exactly as written**.
 - Follow time allowances for strictly-timed sections of the test.
- Be sure students understand the directions and how to mark answers. Test directions can be paraphrased **but test items may not!**

8

Encourage all students to attempt every item

Don't return books to students who have completed the test

Don't ask individual students to finish incomplete or incorrect answers to items

9

Step 1

Large Print and Braille (p.4)

- Transcribe student responses into a regular edition of a test book.
- Mark the large print or Braille book and return to CTB/McGraw-Hill with the unused test books.
- Refer to *Braille Omit Instruction Sheet*

CONTENTS TRANSCRIBED
TO A REGULAR TEST BOOK.

DO NOT SCORE

10

Step 2

Organize Your Classroom (p.4)

Plan for following:

- Distributing and collecting materials;
- Seating arrangements;
- Eliminating noise distractions;
- Using a **"Do Not Disturb Sign"**;
- Removing all **content** or **process** information from classroom, i.e. the walls; and
- Noting **start** and **stop** times on the board for the strictly-timed sections of the test.

11

Check Your Testing Materials (p. 5)

- Examiner materials (11 items)
- Student materials
 - **Tools**
 - CA (only **standardized** books may be used at appropriate grade levels and sessions)
 - MA (only the manipulatives provided and **calculators** may be used at appropriate grade levels and sessions)
 - SC (manipulatives provided)

12

Security Barcode Instructions (p. 6)

- *Examiner's Manual* (p. 6)
- *Test Coordinator's Manual* (pp. 14 - 18)
- No separate instructions sent to districts
- Directions are exactly the same as last year

Security barcodes are located on the lower right-hand corner of the students' test books.

These should be checked by DTC and STC.
The examiner should count the books.

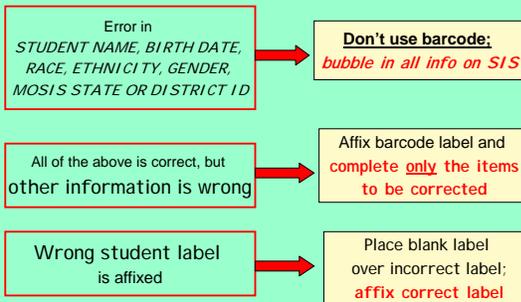
13

Use of Translators

- All tests **except** Communication Arts may be read to ELL students
- ELL students may give responses **orally** or in **writing**
- All written responses must be transcribed **verbatim** to another MAP test book
- **Translators must be trained in giving the MAP**

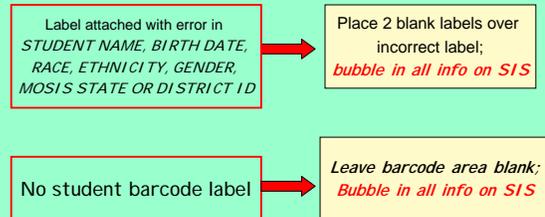
14

Step 4 Before Testing (p.8) Using Student Barcode Labels



15

Using Student Barcode Labels



16

How to Fill Out the SIS (p. 8)

1. Student Name
2. Birth Date
3. Race/Ethnicity
4. Gender
5. MOSIS State ID
6. District Student ID (Optional)

17

Step 5 Prepare Your Students (p.11)

- Help students approach testing in a relaxed positive way
- Explain the purpose of the test
- Point out that some items may be more difficult and may be new to students – *they are not expected to know all the answers but they are expected to do their best*



18

Step 5

Administer the Test (p.11)

- Examiner's instructions are arranged by grade in Communication Arts, Math and Science across content areas this year.
- Sample questions are illustrated and printed in the Examiner's Manual.
- A start/stop time graphic is printed for examiner's use.
- Sessions cannot be split over 2 days or over lunch periods
- Break times are printed in the Examiner's Manual

Have all manipulative(s) "punched-out" prior to testing.

19

Specific Instructions for Test



- Each grade has specific instructions
- Communication Arts is extensive
- Notice the wording about **not paraphrasing test questions** and pronouncing only **one** word per sentence

20

Released Item Responses

- **No test can be photocopied**
- Only released items may be captured and scored.
- A list of released items is on page 17 of SS and SC Examiner's Manuals (Intermediate and High School Only).
- Use of carbonless paper (only for SS and SC)
- Districts may use their own carbonless paper. **None** can be ordered this year.

21

Step 6 Invalidations

Tests are invalidated for the following three reasons:

- Student is discovered cheating
- Examiner paraphrases test question in **any** content area
- Examiner reads any part of the Communication Arts Test, other than students who are Blind/Visually Impaired who do not read Braille.

Examiner must provide **STC** with student information and the reason for the invalidation. **DTC** must send a letter with this information to the Director of Assessment at DESE.

22

Absences and Make-up Sessions

Make-up Sessions:

- Make-up sessions should be scheduled for students who miss one or more sessions of the MAP.
- **"Level Not Determined" (LND) will be ascribed to enrolled students who DO NOT participate in the MAP during either regular or makeup sessions.**
- Every enrolled student must have a test book returned to CTB.
- The absent student's test book must be appropriately labeled and also returned to CTB.

23

Step 7 Student Status Coding

- Student Information has been condensed to one page including *Student Information, Student Status, Accommodations, Disability Diagnosis, and Invalidations*
- Accommodations for ELL and IEP are listed and described on **two separate charts**
- Notice footnotes in *Accommodations*:
 - oral reading of CA,
 - paraphrasing of all tests,
 - extending time for Terra Nova, and
 - the use of former accommodations
- Note invalidations

24

How to fill in Student Status

- SES Bubble to be filled in by STC
- Note: New Title III bubble for ELL student receiving services through Title III
- Consult district Federal Programs Coordinator for Title III students



25

Student Status (cont.)

• H.S. Career Education

Approved Career Education programs can be viewed by accessing the Core Data web application system at:

<http://k12apps.dese.mo.gov/webapps/logon.asp>

- Fill in **ALL** appropriate bubbles that apply (necessary for AYP/MSIP information)



26

Students Not Tested in the Content Area

- All students must be accounted for when administering the MAP
- TWO categories of students are not required to be assessed by the MAP
 - **MAP Alternate (MAP-A) students**
 - **ELL (English Language Learner) in the U.S. less than 12 total months**
 - In **CA ONLY** - **note exact number of months (required on SIS)**
 - For **MAP A** or **ELL less than 12 total months (CA ONLY)** - **you need to return a completed SIS in a test book for these students.**

27

Step 8

Assemble Materials for Return

- After testing, check all SIS and test books for completion and correctness of information according to Step 7
- Complete and check Group Information Sheets (GIS)
 - **Organize test books whose scores are to be reported together**

28

Assemble Materials (cont.)

- Organize materials as shown on picture of envelope (possibly 4 categories)
 - GIS for class/group
 - Completed test books
 - Test books with bar code label or completed SIS for MAP-A students
 - **For CA only:** Test books with barcode label or completed SIS for ELL students in the U.S. 12 months or fewer.

29

Assemble Materials (cont.)

- Unused test books, Large Print, and Braille test books can be placed in a box together labeled "Unused Test Books"
- **Draft copies of writing prompt and scratch paper should be given to STC for secure destruction**



30

Contact Us

St. Louis Regional MAP Center

(314) 516-6628

(888) 627-8675

Eric Hadley mapscrif@umsl.edu

Pat Johnson mapsti@umsl.edu

Carolyn Weinzirl weinzirlc@umsl.edu

Pat Winkler maprf@umsl.edu

31

Presentation on Test Administration

Test Coordinator's Manual 2008



1

2008 MAP Test Coordinator's Manual (TCM)

The TCM is primarily focused on the responsibilities of:

- District Test Coordinators (DTC's); and
- School Test Coordinators (STC's).

The TCM also contains:

- An Overview;
- Timelines; and
- A Glossary of Terms.

2

Missouri Assessment Program

MAP 2008

Communication Arts.....Grades 3-8 & 11
Math.....Grades 3-8 & 10
Science.....Grades 5, 8 & 11

3

Dates to Remember:

[Arrival of Materials](#)
Feb. 20-27, 2008.....Test Coordinator packages to district;
March 3-14, 2008.....Test Materials arrive in districts;

[Testing Window](#)
March 31-April 25, 2008.... Administer Assessments;

[Testing Materials](#)
April 11, 2008.....Deadline for ordering additional testing materials without incurring additional shipping cost;

NOTE: If testing materials are ordered after April 11, the district will be responsible for **paying** the shipping costs.

April 21, 2008.....Final deadline for ordering additional testing materials;

[Return of Materials](#)
April 28, 2008.....Deadline to contact CTB/McGraw-Hill for pickup of testing materials;
May 1, 2008.....Deadline for testing materials to be picked up by CTB's transportation vendor;

[Test Results](#)
August-October 2008.....Reports shipped to districts.

4

DTC's Responsibilities:

- Checks, reviews and distributes testing materials to the STC's;
- Trains STC's on MAP processes;
- Assumes STC role when necessary;
- Guarantees security for all testing materials;
- Acts as sole channel for all communication between districts and CTB MAP Service Line; and
- Collects all testing materials after the test and returns to CTB/McGraw-Hill.

5

STC's Responsibilities

- Receives and checks all testing materials from the DTC;
- Assumes DTC role when necessary;
- Guarantees security of all testing materials;
- Disseminates Examiner's Manuals;
- Trains all Examiners;
- Checks [Group Information Sheets](#) (GISs);
- Completes [School/Group Lists](#); and
- Collects all testing materials after testing, checks and organizes materials for return to the DTC.

6

Examiner's Responsibility

- Examiner responsibilities are addressed in the 2008 MAP Test Examiner's Manual provided for each grade.

7

Summary of 2008 Changes

- 2008 MAP includes all required tests, bound together in one book:
- Communication Arts; Mathematics; and Science.
- Voluntary Social Studies assessment will not be offered in 2008.
- In 2009-2010 Government and American History required EOC Assessments.
- Content area labels are no longer required.
- Scorable and unused test materials **MUST** be boxed separately.

8

- A Missing Test Materials Form is used to notify CTB of contaminated and/or destroyed test books.
- Form included in the Test Coordinator's Kit.
- Grade 10 has an individual Mathematics' test book.
- Grade 11 has individual Science and Communication Arts test books.
- Content areas are bound together in alphabetical order.

9

- A single School/Group List for all books is being used this year.
- Field test items are embedded throughout the operational form of the MAP this year.
- One Group Information Sheet (GIS) per grades 3 -10;
- Two GIS for 11th grade:
 - One GIS for C.A.; &
 - One GIS for Science.
- DTC's responsibilities changed for organizing and packing testing materials.

10

- STC's responsibilities now include:
 - Packing test materials in the return, shipping boxes; and
 - Affixing the return, shipping labels.
- Test materials are now required to be shipped in **GREEN** shaded boxes provided by CTB/McGraw-Hill.

IF	THEN
More boxes are needed --	The DTC will order them.
Boxes are used other than those provided by CTB/McGraw-Hill --	Green stock labels, which are included in the Test Coordinator's Kit, must be affixed to them.

11

- After testing, the STC should collect and securely destroy all Examiner's Manuals.
- A 5% overage of test books is included in each school's shipment.
- Plus, a 5% overage is also sent to the DTC.
- For students who can not use their student barcode labels, a new line is added to the test book front cover to allow students to add their district and school name.
- The Test Book Accountability Form replaces the Security Barcode Verification Form.

12

- Neither an SIS nor a blank test books will be returned for students taking the MAP-Alternate Assessment (MAP-A).

- A new bubble was added for the Blind/Visually Impaired student who does not read Braille:
04 Oral Reading – Blind/Partial Sight.

13

NCLB requires **all** Missouri students to take the MAP test.

Only two groups are exempt:

- Group 1: MAP-Alternate (MAP-A); and
- Group 2: ELL students in the United States, 12 months or fewer, may be exempt from taking the Communication Arts test.

14

Inclusion of Special Populations

- Accommodations for special populations can be found on the DESE website.
- The following are examples of special populations:
 - IEP students;
 - Individual Accommodation Program (IAP 504) students;
 - Students not tested:
 - MAP-A
 - ELL
 - Out-of-District students;
 - Homebound students;
 - Home-schooled students;
 - H.S. Career Education students;
 - Gifted Students; and
 - MoVIP students.

15

Students Testing Out-of-District

- The home DTC (where the student is enrolled) delivers the testing materials to the serving districts/agencies;
- After testing, the completed materials are returned back to the home DTC; and
- The GIS determines where students' result will be reported.

16

Homebound Students:

- Must take the test if they are receiving homebound services;
- Must have test delivered by a trained Examiner who guarantees the security of testing materials; and
- May be tested either at home or school.

17

Home-School Students

- May take part with the local district at the district's discretion;
- Must test if receiving services; and
- Must test at local school.

18

Student Make-up Sessions

IF	THEN
Absent during one or more sessions --	Schedule make-up session
Student is unable to test during make-up session --	Follow Student Absences Procedures

19

Student Absences

1. Write student's name on front of the unused test book;
2. Affix student barcode label if accurate;
3. Complete SIS if student barcode label is inaccurate;
4. Code SIS for absences; and
5. Return test book with **scorable** books.

20

Large Print/Braille Procedures

- Must be transcribed to a regular edition test book to receive score;
- Must be labeled, **“Contents transcribed to a regular test book -- DO NOT SCORE”**;
- Must use special handling and packaging instructions; and
- Must return with **unused** testing materials.

21

Translators

For Mathematics and Science MAP Assessments only:

- ELL students can use their native language to give oral or written responses to assessments;
- ELL students' responses must be translated into English and scribed verbatim into a regular test book;
- Translators must be trained in administering the MAP;
- Translators can review tests before administration in a secure environment;
- Translators must guarantee security of MAP testing materials; and
- ELL accommodation codes apply (see Examiner's Manual codes).

22

Invalidations

- Three categories for test invalidation:
 - Student discovered cheating;
 - Examiner paraphrased test questions in any content area; or
 - Examiner reads any part of the Communication Arts test to students.

23

Invalidation Procedures

The SCT must:

- Be in agreement with Examiner regarding the test invalidation;
- Complete the Teacher Invalidation grid on the SIS;
- Provide demographic information to the DTC; and
- Include invalidate test materials with the **scorable** testing materials.
- DTC sends written communication to the Director of Assessment **with a copy to Director of Accountability, Data and Accreditation.**

24

Check lists for DTC and STC

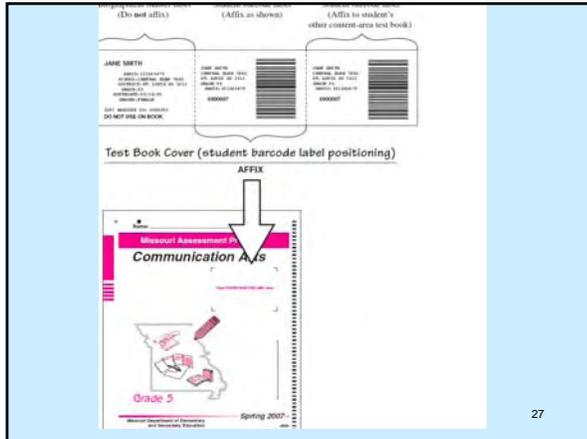
- Convenient check lists are provided to the DTC and STC on pages 7 and 8 of the TCM;
- DTC and STC must ensure all functions are completed; and
- DTC and STC are responsible for both check lists if either of the responsible parties/roles is not available.

25

Student Barcode Labels

- Student barcode labels are in the Test Coordinator's Kit;
- Student barcode labels include:
 - Information from MOSIS Phase III (pre-code system); and
 - Demographic information (but not all biographical data) from the pre-code file;
- CTB-McGraw-Hill prints and provides:
 - One biographical master label; and
 - Two student barcode labels; and
- DTC and STC do the following:
 - Check student barcode label against the student pre-code roster for accuracy;
 - Determine viability of labels; and
 - Handle exceptions appropriately.

26



27

Handling Student Barcode Labels

- Error in:
Student Name, Birth Date,
Gender, Race/Ethnicity, and
MOSIS ID

→ **Don't use barcode; bubble in all info on SIS**
- If label is affixed and then found to be wrong

→ Place **two blank labels** over the incorrect label and then **bubble all** information the SIS
- Wrong student label is affixed

→ Place blank label over incorrect label; and **Then affix correct label**

28

Using Student Barcode Labels

No student barcode label

→ Notify local student data management person to enroll/submit the student in core data.
Bubble in all info on SIS
Leave barcode area blank!

29

Step 1: Review Testing Materials

- The District Test Coordinator's Kit includes two folders for **EACH** school
- One for the DTC; and
 - One for the STC.

30

DTC Folder

- District Packing List;
- District cover Letter;
- Materials Ordering and Inventory Information Flyer;
- Test Book Accountability Form;
- Add/Short Form – District;
- Missing Test Materials Form;
- Braille Omit Return Instruction Sheet;
- Student Barcode Label Instruction Sheet;
- Test Coordinator’s Manuals; and
- Blank District Return Shipping Labels:
 - Blue for scorable materials; and
 - White for unused Inventory.

31

STC Folder

- School Packing List;
- School Cover letter;
- Add/Short Form – School;
- Braille Omit Return Instruction Sheet;
- School Group Lists (SGL’s);
- Group Information Sheet (GIS) for each teacher and grade;
- Test Coordinator’s Manuals;
- Blank Green stock labels;
- Return Shipping Labels:
 - Blue for scorable materials; and
 - White for unused Inventory;
- Student Barcode Label Instruction Sheet;
- Student Pre-code Roster;
- Blank Barcode Label Stock; and
- Student Barcode Labels.

32

Testing Materials

- Packaged by school;
- Shipped to the district’s office or the designated address used in the online order;
- Include:
 - Examiner’s Manuals;
 - Test books;
 - Large white envelopes; and
 - Ancillary testing materials.

33

Verify Shipment

- Compare packing list material against shipment; and
- Follow procedures for ordering more materials if needed, using Add/Short Form.

34

Map Spring 2008 Ordering Additional Testing Materials

Start Date	End Date	Shipping Mode	Shipping Costs
March 14	April 11	USP ground Service	CTB
April 14 th	April 21 st	Next-day or 2 nd day service	District

35

Step 2: Distribute Testing Materials

DTC’s responsibilities:

- Maintains security;
- Tracks security barcodes to confirm start and end barcode numbers for each shrink-wrapped bundle;
- Matches numbers with packing list; and
- Reports discrepancies to CTB.

36

Test Book Accountability Form

- This form replaces the past Security Barcode Form;
- Deals with the security barcode information;
- One copy is needed for **each** school in the district;
- The form is used to ensure 100% accountability of test books;
- Instructions for the DTC and STC are outlined on page 13 - 15 in TCM;
- Pretest and post-test responsibilities regarding the form apply to both the DTC and STC; and
- This form or a copy of it is **never** given to the Examiners.

37

Security Barcode

- Used for test book security;
- Used for inventory – each book consecutively numbered;
- Used to ensure 100% accountability of test books;
- Used for missing inventory reports generated by CTB/McGraw-Hill;
- Used by DESE to track barcode numbers, district and school name; and
- Located on lower right-hand corner directly above "Spring 2008"

38

Security SHALL NOT'S

- Testing materials **shall not** be photocopied, duplicated, or made accessible to non-testing personnel;
- Testing materials **shall not** be viewed by Examiners before testing;
NOTE: The exceptions are ELL translators and Special Education teachers who are pre-selecting items.
- Testing materials **shall not** be left in an unsecured area at any time, for any reason -- must be locked in room or cabinet at school or district office before, between and after testing sessions; and
- Test books **shall not** be shared between schools.

39

7 Steps in the Life Cycle of Test Administration

1. **DTC** guarantees the security of the testing materials -- every school must have sufficient, satisfactory and locked security;
2. **DTC** houses materials at the district office if the school's security is insufficient, unsatisfactory or unlocked;
3. **DTC** distributes the materials to all **STC's** in the district;
4. **DTC** delivers appropriate testing materials for out-of-district students, prior to the first day of testing. Also, the **DTC** makes arrangements for returning materials after testing is complete;
5. **STC** distributes the 2008 MAP *Examiner's Manuals* to all who will administer the test as soon as possible ;
6. **STC** collects all student draft work and scratch paper and securely destroys after testing; and
7. **DTC** saves their folder and boxes for use after test administration.

40

Materials Needed for Each Examiner

- *Examiner's Manual* for appropriate grade level;
- Appropriate quantities of books, manipulatives and reference sheets;
- Student barcode labels for each student;
- Pre-coded Group Information Sheet (GIS):
 - One for each grade, 3-8;
 - One for Grade 10 Mathematics;
 - One for Grade 11 Communication Arts;
 - One for Grade 11 Science; and
 - No separate GIS for Special Education students.
- Appropriate quantity of large white envelopes (each holds approximately 5-10 books); and
- School/Group List for use after test administration.

41

Step 3: Collect Testing Materials

- All used and unused books must be accounted for and returned;
- One test book is returned for each student, except Grade 11;
- All manipulatives, reference sheets, scratch paper, extra envelopes, and contaminated test books are secured destroyed;
- Examiner's Manuals are collected and destroyed; and
- Bulleted items on pages 19 and 20 are used to ensure that all tasks are completed appropriately.

42

Markings in Test Books

In order to be scored properly, the following must take place:

1. Test books completed in ink are transcribed into another test book with a **non-mechanical #2 pencil**;
2. Student responses written on coding tracks/margins are erased and transcribed onto the response line or box;
3. All stray marks on the coding track are erased;
4. All underlining of text is erased around the answer choices; and
5. Test books marked with highlighters are transcribed into other test books for scoring.

43

Step 4: Check the Organization of Materials Collected

Collect the following after testing:

- Large white envelopes with all items in the exact order prescribed on page 21-22 and the illustration on page 34;
- All MAP test books, (used and unused books boxed separately), including student barcodes and/or completed SIS's for each student;

Unused books include:

- Test books which are damaged or have manufacturing errors;
 - Test books written in a language other than English;
 - Test books that are partially used; and
 - Test book that are Braille/Large Print editions with contents transcribed to a regular test book.
- Completed GIS.

44

Step 5: Check Student Information Sheet (SIS)

- Each Examiner's materials **MUST** be grouped together;
- Physical Conditions of SIS can affect/interfere with scoring. Those conditions are listed on pages 23-24;
- Biographical data must be checked on the barcode label & pre-code roster;
- SIS of the test book must be completed **only** when:
 - Pre-coded student barcode label can not be used; and
 - Pre-code roster information is inaccurate for a student.
- A completed SIS and/or a test book with a student barcode label must be received for every eligible student.

45

Sample SIS

- A picture of the SIS and explanations of the fields can be found on pages 25-26.
- Reminder: Identify Special Populations and Invalidations.

46

Step 6: Check the Teacher/Group Information Sheet (GIS)

- GIS provides data that is used on reports – notify the DTC if any errors exist on the GIS;
- GIS is submitted for each grade, except Grade 11 which requires two;
- GIS has both hand-entered and pre-coded information – **both must be accurate**;
- GIS is scannable and **can not** be photocopied; and
- GIS's are placed on top of test books whose scores are to be reported together.

47

Step 7: Complete School/Group List (S/GL)

- S/GL is used by CTB McGraw-Hill to inventory test books;
- S/GL can be photocopied;
- S/GL should have an entry for every GIS that was completed;
- S/GL information includes:
 - Pre-coded:
 - District Name and code number;
 - School Name and code number; and
 - Country Code Number.
 - Hand-written:
 - Contact person's phone number;
 - Group Name is identical to "Teacher Name" on GIS;
 - Grade Number (except Grade 11 only, indicate Communications Arts or Science);
 - Number of books being returned; and
 - Total number of students (tested, MAP- A, invalidated, absent all sessions, unused books).
- Sample S/GL on page 32; and
- Directions for completing fields on page 33.

48

Step 8: Organize Materials for the DTC

The STC will:

1. Reuse the CTB/McGraw-Hill green-shaded boxes in which testing materials arrived;
2. Package the following scorable materials:
 - School/Group Lists; and
 - Large white envelopes, organized by grade in ascending order, accompanied by GIS forms.
3. Package unused testing materials;
4. Affix shipping labels and number each set of boxes separately:
 - Blue labeled on scorable books, numbered 1 of X, 2 of X, etc.
 - White labels on unused books, numbered 1 of X, 2 of X, etc.
5. Send materials to the DTC in unsealed boxes.

49

Step 9: Package and Ship Testing Materials

The DTC will:

- Ensure all testing materials are received from each school in the district;
- Verify TCM instructions for the STC were followed exactly;
- Contact any STC who delays returning materials;
- Added packing material;
- Schedule testing material pickup online;
- Schedule pickup no later than **April 28, 2008**;
- Follow instructions outlined in TCM on page 37-38; and
- Fax Test Book Accountability Forms to CTB/McGraw-Hill.

50

Glossary of Terms

A glossary is listed at the end of the manual from pages 39-41. Most terms are the same as previous years. Two noteworthy changes are as follows:

- Level Not Determined; and
- Valid Attempt

51

Level Not Determined (LND)

This designation is for students who did not receive a MAP score for any one of the following three reasons:

1. An SIS is returned to CTB/McGraw-Hill with a blank test book;
2. A student does not attempt any items in one or more sessions; or
3. A student is absent all 3 sessions.

52

Valid Attempt

53

Questions???? Questions????



Missouri Department of Elementary
and Secondary Education
800-845-3545 (Assessment Section)

54

Appendix B

*Administration of the Breach Form of Grade 11 Communication Arts
MAP.....B-1*

Missouri Assessment Program Linking Design.....B-8

Missouri Assessment Program Anchor Evaluation.....B-21

Administration of the Breach Form of Grade 11 Communication Arts MAP

Administration of the Breach Form of Grade 11 Communication Arts MAP

Karla Egan
CTB/McGraw-Hill
August 12, 2008

In 2008, four Missouri districts were administered a breach form of TerraNova in Grade 11 after it was discovered that these districts had access to the TerraNova form embedded in the 2008 MAP. Table 1 lists the affected districts.

Table 1. List of Districts who Administered Alternate Form of MAP

Charleston
Clark County
Kansas City
Maplewood

In this document, CTB describes the breach form of TerraNova, the calibration and linking for the breach form, the analyses of person fit conducted after the calibration and linking, and the district-level data.

Breach Form

The MAP assessment is an augmented NRT. Items from CTB's norm-reference product, TerraNova, are aligned to the Missouri content standards and test design blueprints. Custom items are used to fill any gaps in this alignment. In 2008, the Grade 11 Communication Arts MAP was comprised of TerraNova Form A and 20 custom items.

The breach form was comprised of TerraNova Form C as well as the same custom items embedded in 2008. TerraNova Form C was administered in 2007. Form C had already been aligned to the MO content standards. It was also determined by CTB's Content Development team that Form C could be substituted with Form A and adequate coverage of the test blueprint would still exist. See Table 2 for the distribution of items by content standard for TerraNova Form A versus TerraNova Form C.

Table 2. Number of Items Measuring each Content Standard for TerraNova Forms A and C

Content Standard	Number of Items	
	Form A	Form C
Writing Standard English	4	5
Reading (Fiction and Non-Fiction)	34	34
Writing Formally	1	0

Administration. Students taking the breach form were administered items in two test books. Students were administered the custom items in Sessions 1 and 2 in the 2008 test book. They were then administered the Session 3 TerraNova items (Parts 1 and 3) from the 2007 test book. Finally, they were administered the remaining custom items in Session 4 in the 2008 test book.

Calibrating and Linking

The items from the 2008 breach form were concurrently calibrated with the items from the 2008 regular MAP. Items were calibrated using the 3-parameter model for MC items and the 2-parameter partial credit model for CR items. The items were linked to the MAP scale using the Stocking and Lord (1983) procedure using the TerraNova Form A items. A detailed description of the calibration and linking procedures will be included in the Technical Report.

Residual Analyses

Following the implementation of the concurrent calibration, it was important to check that the models were functioning in the same manner for the students administered the regular form and students administered the breach form. To do this, we analyzed the residual item score distributions for the multiple-choice and constructed response items held in common by the two groups of students. If the residual patterns are similar between the students who were administered the two forms, then we can be confident that the calibration and linking analyses were appropriate for both groups of students.

MC Items: To find the residual score distributions for MC items, the following equation was used

$$\text{Residual} = \text{observed score} - P, \text{ where } P \text{ is the probability of a correct response.}$$

CR Items: For the CR items, we first found the expected score:

$$\sum (x P(x)),$$

where x is the score point and $P(x)$ is the probability of the student obtaining that score point.

To find the residual score distributions for CR items, the following equation was used

$$\text{Residual} = \text{observed score} - \text{expected score.}$$

For each combination of person and item, the response probability was calculated using the 3PL/2PPC model with the item and person parameters replaced by their estimates.

QQ Plots: Quantile-quantile (QQ) plots of the residuals from the MC items and from the CR items for the two groups of students were created. The QQ plot allowed us to

graphically compare the similarity of the distributions of residuals for the two groups of students. If the distributions are similar, then they should fall along a 45-degree line on the plot.

Figure 1 shows the absolute values of the residuals from the students taking the breach form compared to the students taking the regular form for the MC items taken by both groups. The QQ plot shows that the distribution of residuals is very similar for the two groups indicating that the calibration results were appropriate for both groups of students.

Figure 2 shows the absolute values of the residuals from the students taking the breach form compared to the students taking the regular form for the CR items taken by both groups. Again, the QQ plot shows that the distribution of residuals is very similar for the two groups. There appears to be one outlier at the very tail of the distribution; however, this does not undermine the basic finding that the calibration results were appropriate for both groups of students.

District-level Data

The district-level data for the four affected districts was further analyzed to look at changes between 2007 and 2008. Table 3 reports the mean MAP scores and standard deviations for these districts from 2006, 2007 and 2008. The final row of this table reports the mean results for the rest of the state (not including these four districts). Table 3 shows that the overall mean scale score of the four breach districts is fairly stable from 2006 to 2008, decreasing by approximately 2 points from 2007 to 2008. A similar trend was observed for the mean scale score for the rest of the state from 2007 to 2008.

Table 4 shows the mean TerraNova scores and standard deviations for these districts from 2006, 2007 and 2008. The final row of this table reports the mean results for the rest of the state (not including these four districts). Table 4 shows that the overall mean scale score for the breach districts and the rest of the state declined by approximately 2 points from 2007 to 2008. It should be noted here that the same TerraNova form was administered to the breach districts in 2006, 2007, and 2008 while the rest of the state took the same form in 2006 and 2007 and a different form in 2008.

Summary

The items from the regular and breach forms were put on scale through a concurrent calibration using 3PL/2PPC IRT models and linked to the Missouri scale through the TerraNova items on the regular form. In doing this, it was important to ascertain if the IRT models used were appropriate for both the students taking the regular form and the breach form. The results of the analyses confirmed that the models were appropriate.

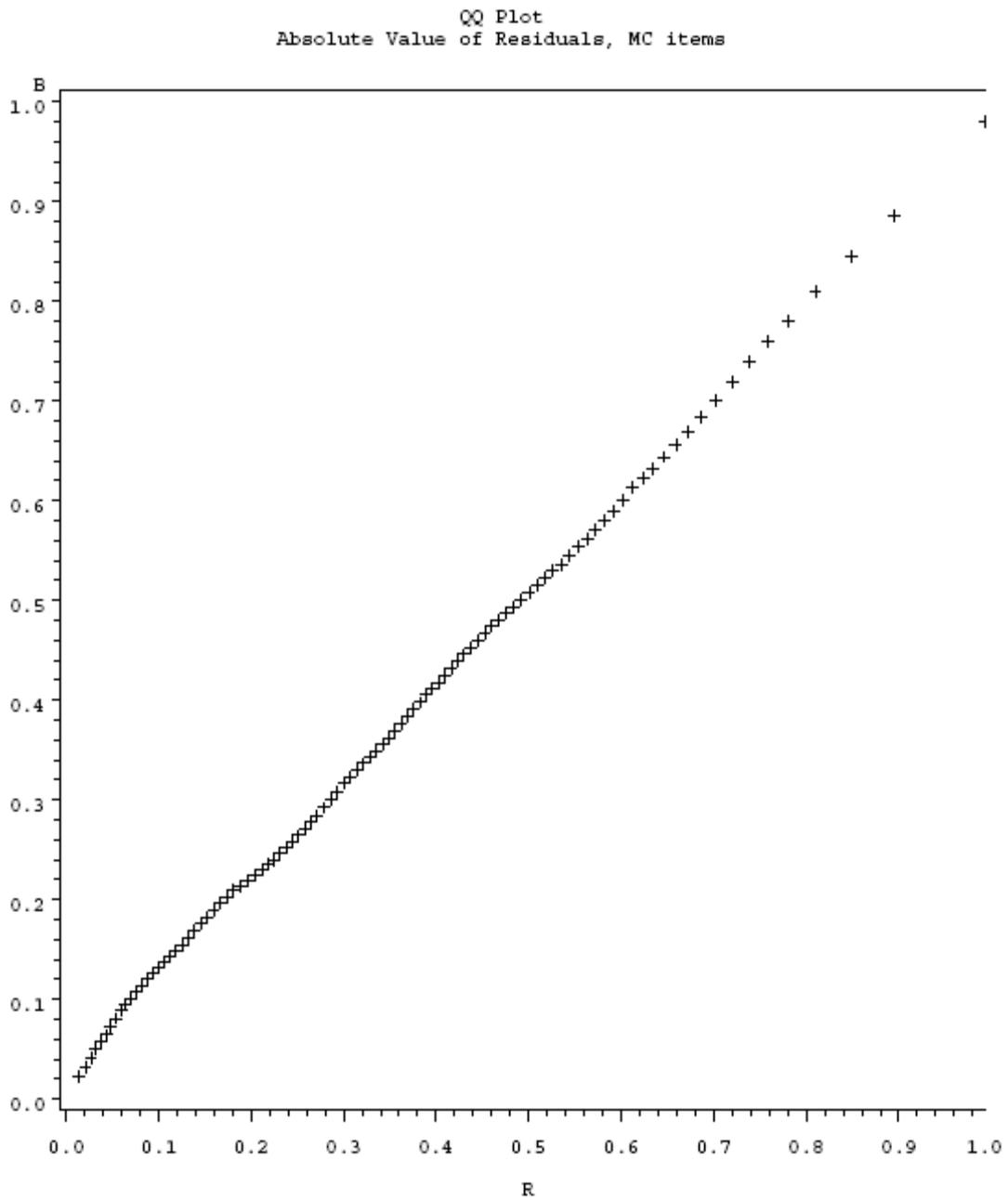


Figure 1. QQ Plot of the absolute value of the residuals from the breach form (y-axis) and regular form (x-axis) for the MC items.

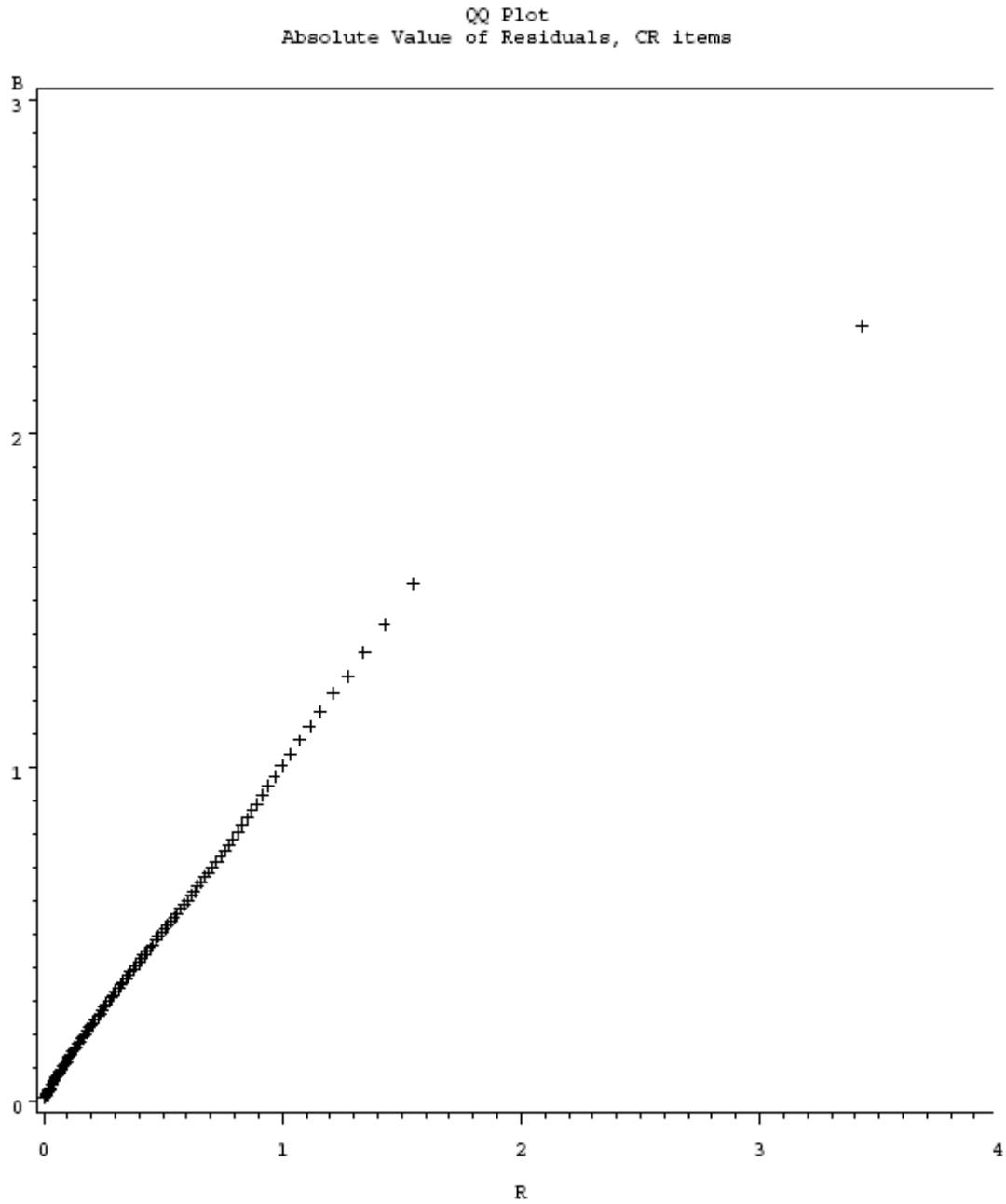


Figure 2. QQ Plot of the absolute value of the residuals from the breach form (y-axis) and regular form (x-axis) for the CR items.

Table 3. Mean MAP scores (standard deviations) and n-counts for 2006, 2007, and 2008

District	2006		2007		2008	
	N count	MAP Mean (SD)	N count	MAP Mean (SD)	N count	MAP Mean (SD)
CHARLESTON	61	708.25 (36.35)	75	708.01 (34.14)	62	688.94 (43.60)
CLARK CO RI	77	720.78 (24.04)	81	714.81 (31.11)	69	716.49 (29.87)
KANSAS CITY	1065	691.24 (35.71)	1122	692.50 (33.54)	1036	690.39 (43.37)
MAPLEWOOD	72	709.01 (35.78)	71	706.31 (30.84)	89	711.43 (41.29)
OVERALL Breach Districts	1275	694.84 (36.12)	1349	695.43 (33.93)	1256	693.14 (43.28)
STATE (minus Breach Districts)	60004	716.69 (31.42)	61482	715.86 (31.32)	61041	713.56 (35.89)

Table 4. Mean NRT scores (standard deviations) and n-counts for 2006, 2007, and 2008

District	2006		2007		2008	
	N count	MAP Mean (SD)	N count	MAP Mean (SD)	N count	MAP Mean (SD)
CHARLESTON	61	710.18 (37.32)	75	707.51 (40.54)	62	682.42 (49.25)
CLARK CO RI	77	724.21 (26.23)	81	718.96 (31.02)	69	710.84 (39.02)
KANSAS CITY	1098	691.22 (42.34)	1144	693.64 (37.60)	1036	684.96 (45.43)
MAPLEWOOD	72	711.04 (31.84)	71	707.85 (30.86)	89	713.70 (36.70)
OVERALL Breach Districts	1308	695.14 (41.84)	1371	696.63 (37.73)	1256	688.29 (45.62)
STATE (minus Breach Districts)	60522	718.12 (37.26)	61884	718.38 (36.76)	60716	716.30 (38.24)

Missouri Assessment Program Linking Design

Missouri Assessment Program Linking Design

Karla Egan

August 19, 2008

The Grades 3 through 8 MAP are linked to the MAP scale through the embedded *TerraNova* NRT test. In past years, this link has been accomplished using the *TerraNova* parameters from the national norming study. In 2008, this design was altered so that the MAP tests were again linked back to the MAP scale through the *TerraNova* items; however, in 2008, the anchor parameters were based on Missouri students. In this paper, we first explicate the original linking design and the results of this linking design, as well as the alternate linking design and the results from this linking design. Finally, we examine the linking design for future MAP forms.

Original Linking Design

The Grades 3 through 8 MAP are comprised of the *TerraNova* NRT plus additional custom items. The tests are constructed to meet a test blueprint designed to measure Missouri's Grade-level Expectations. Every two years the embedded *TerraNova* form is changed for an alternate *TerraNova* form to minimize exposure of the NRT items. The custom items were changed ever year; thus, there is 100% replacement of items every two years.

In 2006 and 2007, *TerraNova* Form C was embedded in Grades 3 through 8 MAP. In 2008, *TerraNova* Form D was embedded in Grades 3 through 8 MAP. The embedded *TerraNova* form is used to maintain the test scale; in other words, the MAP is linked to the MAP scale through the embedded *TerraNova* form. Since both Form C and Form D are alternate forms, it should be possible to substitute one form for another and maintain the link to the MAP scale. The fundamental assumption underlying this approach is that the relationship between the 2007 and 2008 MAP forms for Missouri students is accurately established on the basis of the nationally-derived relationship between Forms C and D of *TerraNova*.

After the initial IRT item calibration, items parameters were linked to the *TerraNova* scale using the Stocking & Lord (1983) test characteristic curve (TCC) procedure. For the linking, the intact *TerraNova* Form D Survey items served as anchors. The anchor parameters were derived from the *TerraNova* norming study.

The MAP calibrations and linking were conducted using a sample of Missouri students. This sample comprised, at least, 80% of the total student population. All results reported in this paper are based in the 80% calibration sample.

Results

The anchor items were evaluated for performance after the initial linking. The anchor evaluation is discussed later in Appendix B. No Communication Arts anchor items were

flagged for poor performance. Table 1 shows the number of Mathematics items that were turned off as anchors due to poor performance. The results of the anchor analyses will be discussed in more detail in the *2008 Missouri Technical Report*.

Table 1. Number of Mathematics Anchor Items Turned Off as Anchors by Grade

Grade	Number of Flagged Items
3	1
4	1
5	2
6	0
7	1
8	2
10	3

The items in Table 1 were turned off as anchor items after examining the effect that their removal had on the content structure of the anchor. It should be noted that students were still scored on these items.

The paper on Anchor Evaluation presented later in Appendix B provide plots of the input TCCs and the transformed estimated TCCs for the Communication Arts and Mathematics anchor items by grade. Based on these plots, the linking appeared to work well.

The results of the original linking showed larger-than expected variability in MAP results across grades. Table 2 presents the mean scale score and impact data once students were scored following the implementation of the original linking. The results are based on the 80% calibration sample. Upon examination of Table 2, there are clear areas that are cause for concern. For example, there is an almost 14 percentage point decrease in the percent of students at or above Proficiency in Grade 5 Mathematics. This type of decline, while possible, is probably unlikely.

The larger-than expected variability first led us to analyze the data for possible processing and/or sampling errors. A thorough investigation revealed no processing or sampling errors. Next, the variability in cross-year MAP results led us to investigate the data for possible form effects.

Table 2. Mean and Standard Deviation (SD) of 2006, 2007, and 2008 MAP Scale Scores (SS) and Percent at or above Proficient (% P+A) for Communication Arts and Mathematics, using the Original Linking Design to derive the 2008 results.

Grade	Year	Communication Arts				Mathematics			
		N	Mean SS	SD SS	% P+A	N	Mean SS	SD SS	% P+A
3	2006	64486	639.86	36.84	43.0	64763	621.59	39.1	43.7
	2007	66345	639.58	38.04	43.2	66610	622.39	38.7	45.4
	2008	54794	635.07	35.69	36.4	54907	622.37	37.4	44.8
4	2006	65179	654.55	38.56	44.3	65306	643.88	37.1	43.8
	2007	65250	656.11	39.51	45.6	65363	644.47	36.6	44.9
	2008	56866	650.15	34.48	37.9	56976	640.63	35.8	40.4
5	2006	66007	668.18	37.09	45.5	66123	660.06	40.0	43.8
	2007	65461	671.01	37.14	48.3	65498	663.21	41.5	47.0
	2008	52300	665.25	34.85	40.2	52429	651.72	36.2	33.2
6	2006	66948	666.85	33.70	42.7	67017	673.30	39.8	44.3
	2007	66247	667.99	34.63	44.1	66332	676.31	41.8	48.3
	2008	53590	670.29	35.51	46.5	53666	677.99	41.2	50.6
7	2006	70290	671.63	37.06	43.5	70698	675.38	41.3	43.5
	2007	67167	672.11	36.26	45.2	67554	677.41	42.6	45.5
	2008	58764	678.55	37.59	52.4	58830	678.58	37.2	46.2
8	2006	72483	686.85	37.87	42.1	72542	697.73	40.4	40.4
	2007	70175	686.89	37.54	42.2	70204	698.33	42.0	41.2
	2008	56312	681.49	38.99	38.1	56387	697.43	41.8	40.6

Form Effect

To explore the possibility of a form effect, we computed the *TerraNova* scale scores from Form D and compared those to the 2007 Form C scale score. Table 3 presents the mean *TerraNova* scale scores from the 2007 administration of *TerraNova* and the 2008 operational administration of *TerraNova* Form D. It also shows the difference between the Form D mean scale score and the Form C mean scale score. The magnitude of the difference between the scale score indicates the possible presence of a form effect.

Table 3 shows that students had higher mean NRT scale scores in 2007 than in 2008 in all grades in both content areas, except Grades 6 and 7 Communication Arts and Mathematics. Table 3 also shows that the form effect is strongest in Grade 5 Mathematics, where there was a 10 point drop in the mean scale score. There are also possible form effects in Grades 5, 7, and 8 Communication Arts; however, the magnitude of the effect in these grades is smaller than in Grade 5 Mathematics.

The results of this analysis led us to believe that a form effect may be artificially deflating student scores in certain grade/content areas (e.g. Grade 5 Mathematics, Grade 8 Communication Arts) and inflating student scores in other grade/content areas (e.g. Grade 7 Communication Arts). The results of these analyses led us to look for an alternate means of linking the 2008 MAP to the MAP scale.

Table 3. Mean NRT Scale Scores from 2007 Operational Administration, 2008 Operational Administration, and Difference between the Mean Scale Scores (Form D minus Form C)

		2007 NRT Form C			2008 NRT Form D			Difference (Form D - Form C)
		N	Mean	SD	N	Mean	SD	Mean
Communication Arts	3	66574	639.38	42.14	66153	635.77	42.22	-3.60
	4	65308	655.02	42.48	66826	653.02	43.10	-2.00
	5	65481	670.99	42.18	65499	666.86	40.82	-4.12
	6	66295	670.49	41.91	65658	671.58	40.35	1.09
	7	67522	673.98	42.86	66664	679.61	43.90	5.63
	8	70277	689.65	43.88	67229	683.05	42.48	-6.60
Mathematics	3	66653	623.86	46.92	66271	622.74	45.28	-1.12
	4	65382	646.92	45.44	66949	644.92	46.61	-2.00
	5	65516	666.08	51.45	65633	655.76	46.57	-10.32
	6	66354	676.67	46.87	65725	679.22	47.74	2.55
	7	67592	680.23	52.28	66737	682.40	43.65	2.17
	8	70258	700.68	48.49	67308	699.53	48.40	-1.16

Alternate Linking Design

The alternate linking design was comprised of four steps. Each of these steps will be detailed below.

Step 1. Estimate new parameters for the Survey Form D items using the students who took both the 2007 Field test and the 2007 OP test (approximately 12,000 students per grade/content area).

Step 2. Link 2008 MAP items to the MAP scale through Survey Form D parameters that were based on Missouri students. The fundamental assumption underlying this approach is that students had the same level of proficiency on the field and operational tests.

Step 3. Adjust for a motivation effect using information from both linking methods.

Step 4. Apply new scaling constants to untransformed parameters to put 2008 MAP items onto the MAP scale.

Step 1. Missouri-based Parameters for *TerraNova*

In 2007, Missouri administered a standalone field test to approximately 12,000 students in Grades 3 through 8. Table 4 shows the number of students who took the standalone field test in 2007.

Table 4. Missouri 2008 Alternate Linking - # of Students included in the linking between the 2007 FT and 2007 OP tests

Content	Grade	# of Students
Communication Arts	3	12584
	4	12500
	5	12411
	6	12006
	7	11463
	8	11873
Mathematics	3	12565
	4	12724
	5	12851
	6	12606
	7	13077
	8	14550

This field test was administered approximately two weeks after the close of the operational testing window. *TerraNova* Form D was embedded in the standalone field test.

The *TerraNova* Form D items were put onto the MAP scale using the approximately 12,000 students who took both the standalone form and the operational form. By doing this, we were able to estimate Missouri-based parameters for the Form D items. We could then use the Missouri-based parameters as the anchor parameters to link the 2008 operational to the MAP scale.

Special note on Grade 8 Communication Arts. For Grade 8 Communication Arts, one passage and related items was replaced with a passage from Form A *TerraNova* because the author of the original Form D passage would no longer grant permission to use it. Because the items from Form A were not field-tested in 2007, it was not possible to estimate Missouri-based parameters for these items (n=6). For Grade 8 Communication Arts, new parameters were found for the remaining Form D items. The remaining items (n=28) were used to link the Grade 8 Communication Arts MAP to the MAP scale.

Step 2. Linking with Missouri-based Anchor Parameters

Using the same methodology as the original linking, the 2008 items parameters were linked to the *TerraNova* scale using the Stocking & Lord (1983) test characteristic curve (TCC) procedure. For the linking, the intact *TerraNova* Form D Survey items served as anchors; however, this time the Missouri-derived parameters were used instead of the

parameters derived from the *TerraNova* norming study. No anchor items were flagged for poor performance during the analysis of the anchor items.

The paper on Anchor Evaluation presented later Appendix B provide plots of the input TCCs and the transformed estimated TCCs for the Communication Arts and Mathematics anchor items by grade for the linking with Missouri-based parameters. Based on these plots, the linking appeared to work well.

Table 5 shows the results from the alternate linking. Again the results are based on the 80% calibration sample. In this case, student performance increases across all grade/content areas. In some cases, the percent at or above Proficient increases by almost 10 percentage points from 2007 (see Grade 8 Communication Arts).

Ideally, the results would have been stable across the two linking methodologies. If both methodologies would have shown a large decrease in Grade 5 Mathematics, then this result was likely true. For example, in Grade 7 Communication Arts, both linking methods resulted in fairly large gains over the 2007 scores.

As noted above, the fundamental assumption underlying this approach is that students had the same level of proficiency on the field and operational tests. We studied this assumption by comparing the 2007 NRT field-test data to the 2008 operational data. We explain this study in the next section.

Table 5. Mean and Standard Deviation (SD) of 2006, 2007, and 2008 MAP Scale Scores (SS) and Percent at or above Proficient (% P+A) for Communication Arts and Mathematics, using the Missouri-based parameters to derive the 2008 results.

Grade	Year	Communication Arts				Mathematics			
		N	Mean SS	SD SS	% P+A	N	Mean SS	SD SS	% P+A
3	2006	64486	639.86	36.84	43.0	64763	621.59	39.1	43.7
	2007	66345	639.58	38.04	43.2	66610	622.39	38.7	45.4
	2008	54794	641.44	37.32	45.0	54907	624.81	37.13	47.6
4	2006	65179	654.55	38.56	44.3	65306	643.88	37.1	43.8
	2007	65250	656.11	39.51	45.6	65363	644.47	36.6	44.9
	2008	56866	658.68	33.30	49.3	56976	646.76	34.18	47.6
5	2006	66007	668.18	37.09	45.5	66123	660.06	40.0	43.8
	2007	65461	671.01	37.14	48.3	65498	663.21	41.5	47
	2008	52300	674.56	33.26	52.4	52429	664.40	40.77	49.2
6	2006	66948	666.85	33.70	42.7	67017	673.30	39.8	44.3
	2007	66247	667.99	34.63	44.1	66332	676.31	41.8	48.3
	2008	53590	674.17	33.42	51.5	53666	681.30	41.16	53.9
7	2006	70290	671.63	37.06	43.5	70698	675.38	41.3	43.5
	2007	67167	672.11	36.26	45.2	67554	677.41	42.6	45.5
	2008	58764	679.46	34.62	53.6	58830	684.71	41.36	53.4
8	2006	72483	686.85	37.87	42.1	72542	697.73	40.4	40.4
	2007	70175	686.89	37.54	42.2	70204	698.33	42.0	41.2
	2008	56312	694.24	33.29	52.4	56387	704.65	39.41	47.8

Step 3. Motivation Effect

Before starting any of these analyses, we were concerned about the motivation of students who took the standalone field test. These students would have known that it was a field test and it was administered during the last few weeks of the school year for many of the students.

Table 6 shows the mean NRT scale scores from the 2007 field-test administration of Form D and the 2008 operational administration of Form D. It also shows the difference between the mean scale scores from the two forms. Table 6 shows that the mean scale scores from the operational administration were routinely higher than the field-test administration. The results in Table 6 suggest that a motivation effect is present in the data. This meant that it would be necessary to account for the motivation of students during the linking process.

Table 6. Mean NRT Scale Scores from 2007 Field-test Administration, 2008 Operational Administration, and Difference between Mean Scale Scores (2008 minus 2007)

		2007 FT NRT Form D			2008 NRT Form D			Difference (08OP-07FT)
		N	Mean	SD	N	Mean	SD	Mean
Communication Arts	3	12980	634.91	41.94	66153	635.77	42.22	0.86
	4	12934	645.88	46.60	66826	653.02	43.10	7.15
	5	12923	661.68	43.93	65499	666.86	40.82	5.19
	6	12371	664.42	40.61	65658	671.58	40.35	7.16
	7	12074	674.15	44.02	66664	679.61	43.90	5.47
	8	12375	680.06	47.52	67229	683.05	42.48	2.98
Mathematics	3	13008	619.12	44.39	66271	622.74	45.28	3.62
	4	13174	643.32	46.02	66949	644.92	46.61	1.60
	5	13150	653.43	45.38	65633	655.76	46.57	2.33
	6	13245	672.56	48.32	65725	679.22	47.74	6.65
	7	13573	675.09	43.25	66737	682.40	43.65	7.31
	8	15069	696.16	49.31	67308	699.53	48.40	3.37

Adjusting for the Motivation Effect

To adjust for the motivation effect, we used information from both linking procedures. There are many ways to adjust for the motivation effect. Since grade-level information was only available for a single data point within a content area, the results from all grades within a content area were summarized to enhance the stability of the results.

To adjust for the motivation effect, the mean standardized growth (z) from 2007 was calculated for both linking procedures.

$$(1) \quad z_{ij} = \frac{\bar{x}_{ij} - \bar{x}_{2007}}{\sigma_{2007}},$$

where \bar{x}_{ij} is the mean MAP scale score for linking method i at grade j , \bar{x}_{2007} is the mean MAP scale score from 2007, and σ_{2007} is the standard deviation from 2007.

The standardized growth was then averaged across all grades,

$$(2) \quad \bar{z}_i = \frac{\sum_{j=1}^n (z_{ij})}{n},$$

where \bar{z}_i is the mean standardized growth across all grades within a content area for the 2008 linking method i ; z_{ij} is the standardized growth for linking method i at grade j ; and n is the total number of grades.

Then the grand mean was found for mean standardized growth of the two linking methods,

$$(3) \quad Z = \frac{\sum_{i=1}^2 (\bar{z}_i)}{2},$$

where Z is the grand mean.

Finally, the motivation offset was calculated found,

$$(4) \quad M = Z - \bar{z}_{alt},$$

where M is motivation offset and \bar{z}_{alt} is the mean standardized growth across all grades for the alternate linking method.

Step 4. Finding New Scaling Constants

Once the motivation offset was found, it was possible to find new scaling constants to apply to the theta-metric item parameters to place them on the Missouri scale.

First, the desired mean MAP scale score was found for each grade,

$$(5) \quad \mu_j^* = \bar{x}_j^A - (\sigma_j^A * M),$$

where μ_j^* is the desired mean for grade j , \bar{x}_j is the mean for grade j using the alternate linking method, σ_j^A is the standard deviation for grade j from the alternate linking A , and M is the motivation offset.

Next, the new scaling constants were computed,

$$(5) \quad M_1 = \frac{\sigma_j^A}{\sigma_j^T},$$

$$(6) \quad M_2 = \mu_j^* - \frac{\sigma_j^A}{\sigma_j^T} * \mu_j^T,$$

where σ_j^A is the standard deviation for grade j from the alternate linking A , σ_j^T is the standard deviation of the unscaled thetas T for grade j , μ_j^* is the desired mean for grade j , and μ_j^T is the unscaled theta, T , mean for grade j .

Final Results

Once the new scaling constants were applied, results were tabulated and are summarized in Table 7. Again, these results are based on the 80% calibration sample. In Grades 3, 4, and 5 in both content areas, the performance of students decreased or maintained, in terms of the percent of students at or above Proficient compared to 2007. In Grades 6 through 8 in both content areas, the percent of students at or above Proficient increased from 2007.

Table 7. Mean and Standard Deviation (SD) of 2006, 2007, and 2008 MAP Scale Scores (SS) and Percent at or above Proficient (% P+A) for Communication Arts and Mathematics, using the New Scaling Constants to derive the 2008 results.

Grade	Year	Communication Arts				Mathematics			
		N	Mean SS	SD SS	% P+A	N	Mean SS	SD SS	% P+A
3	2006	64486	639.86	36.84	43.0	64763	621.59	39.1	43.7
	2007	66345	639.58	38.04	43.2	66610	622.39	38.7	45.4
	2008	54794	638.00	37.31	40.7	54907	621.94	37.2	44.1
4	2006	65179	654.55	38.56	44.3	65306	643.88	37.1	43.8
	2007	65250	656.11	39.51	45.6	65363	644.47	36.6	44.9
	2008	56866	655.61	33.22	45.1	56976	644.11	34.2	44.2
5	2006	66007	668.18	37.09	45.5	66123	660.06	40.0	43.8
	2007	65461	671.01	37.14	48.3	65498	663.21	41.5	47
	2008	52300	671.50	33.18	48.2	52429	661.25	40.7	45.7
6	2006	66948	666.85	33.70	42.7	67017	673.30	39.8	44.3
	2007	66247	667.99	34.63	44.1	66332	676.31	41.8	48.3
	2008	53590	671.10	33.3	47.1	53666	678.11	41.1	50.4
7	2006	70290	671.63	37.06	43.5	70698	675.38	41.3	43.5
	2007	67167	672.11	36.26	45.2	67554	677.41	42.6	45.5
	2008	58764	676.26	34.6	49.5	58830	681.52	41.3	50.0
8	2006	72483	686.85	37.87	42.1	72542	697.73	40.4	40.4
	2007	70175	686.89	37.54	42.2	70204	698.33	42.0	41.2
	2008	56312	691.18	33.2	48.2	56387	701.61	39.4	44.4

Future Linking of MAP to MAP Scale

In order to continue to link MAP to the MAP scale, it will be necessary to continue to use Missouri-based parameters for the anchor items. This is a straightforward process in 2009 because the same *TerraNova* form will be embedded as the anchor set as was used in 2008. Thus, the 2008 scaled parameters will be used as the anchor set.

In 2010, the current plan is to use *TerraNova* Form E as the anchor. In order to obtain Missouri-based parameters for Form E items, the Form E items will be embedded in the 2009 form. Form E will be divided in 6 parts, and the items will be distributed across each of the 6 MAP forms. The Form E items will be administered in the same section as the 2009 *TerraNova* form immediately after the completion of the 2009 *TerraNova* form. Form E was divided among the 6 MAP forms to minimize testing time for students. This design allows us to continue to spiral the multiple forms within a classroom in all grades, except Grade 3 Communication Arts.

For Grade 3 Communication Arts, it will be necessary to spiral the forms by school instead of within a classroom. The Grade 3 Communication Arts test is administered with oral directions. When Form E is split among the six MAP forms, it requires a different set of directions for each set of items. It is not logistically feasible for a third grade teacher to orally administer six different test forms within the same classroom. For this reason, the Grade 3 Communication Arts form will be spiraled by school.

Conclusions

The original linking design links 2008 MAP forms to the MAP scale using *TerraNova* parameters established on a standardization sample. We did this using Form C in 2006 & 2007. We are using Form D in 2008 MAP. The fundamental assumption underlying this approach is that the relationship between the 2007 and 2008 MAP forms for Missouri students is accurately established on the basis of the nationally-derived relationship between Forms C and D of *TerraNova*. The results of the Form D linking showed larger-than expected variability in MAP results across grades.

As a standard approach to investigating large score changes in equating, we examined an alternative defensible equating approach to see if evidence would reinforce or conflict with results from the initial analyses.

Under this alternative linking design Form D anchor parameters were put on the MAP scale based on data collected during the 2007 operational and field test (FT). The link was based on the approximately 11,000 MO students who took both Form C (operationally) and Form D (FT) in 2007. The fundamental assumption underlying this approach is that students had the same level of proficiency on the field and operational tests.

Evidence from our analyses suggests that the assumptions underlying each approach hold fairly well for some tests but not for all tests. In particular, the national and MO-based

relationships between the Forms C and D items appear to differ for some tests, and there is evidence to suggest that students may have had higher motivation on the operational versus field test situations in 2007.

Each approach has advantages and we found that an approach that combines information from both leads to the most defensible and reasonable results. In this combined approach:

- the overall cross-grade growth was determined by the simple average of the growth calculated independently for each of the two methods.
- the pattern of grade-to-grade growth was based on the alternate Missouri-based, linking method. This was done because of the significantly greater stability and less variability arising from this approach.
- Final scoring tables under this approach are obtained by applying a simple linear transformation to the parameters from the alternate, Missouri-based calibration, to achieve total cross-grade growth from the composite

Since the operational 2008 and 2009 forms will have items in common, we do not expect to face this challenge next year. In 2010 *TerraNova* Form E items will be introduced, and we will modify research analysis plans to support the stability and reliability of the linking .

Finally, Missouri norms are based on nationally derived parameters and are unaffected by the MAP linking approach described above.

Missouri Assessment Program Anchor Evaluation

Missouri Assessment Program: Anchor Evaluation for Communication Arts, Mathematics, and Science

The anchor items were evaluated immediately following the calibration and equating of the Missouri Assessment Program (MAP). This report outlines the methods used to evaluate anchor items for the MAP and the results of the analyses.

Two different linkings, original and alternate, were completed for the Grades 3 through 8 MAP in Communication Arts and Mathematics. The linking methodologies are described in *Linking Design for the 2008 Grades 3–8 MAP*, also found in this Appendix. The anchor evaluation was conducted after each linking.

Methods Used to Evaluate Anchor Items

For the Missouri assessments, two statistical methods are used to evaluate anchor items: (1) iterative linking (Candell & Drasgow, 1988) using Stocking and Lord's (SL;1983) test characteristic curve method; and (2) differences between the item-ability regression curves.

Test Characteristic Curve Method¹

The Stocking and Lord (1983) procedure, also called the test characteristic curve (TCC) method, minimizes the mean squared difference between the two TCCs, one based on estimates from the previous calibration and the other on transformed estimates from the current calibration. Let $\hat{\psi}_j$ be the test characteristic curve based on estimates from a previous calibration and $\hat{\psi}_j^*$ be the test characteristic curve based on transformed estimates from the current calibration.

$$\hat{\psi}_j = \hat{\psi}(\theta_j) = \sum_{i=1}^n P_i(\theta_j; a_i, b_i, c_i)$$

$$\hat{\psi}_j^* = \hat{\psi}(\theta_j) = \sum_{i=1}^n P_i(\theta_j; \frac{a_i}{M_1}, M_1 b_i + M_2, c_i)$$

The TCC method determines the scaling constants (M_1 and M_2) by minimizing the following quadratic loss function (F):

$$F = \frac{1}{N} \sum_{a=1}^N (\hat{\psi}_j - \hat{\psi}_j^*)^2$$

Differential item functioning was evaluated by examining previous (input) and transformed (estimated) item parameters. The item with an absolute difference of parameters greater than two times the root mean square deviation was flagged. The difference was also monitored by plotting input and estimated item parameters.

¹ Text explaining the Test Characteristic Curve Method, Delta-Plot Method, and Lord's Chi Square is taken from Karkee and Choi (2005). *Impact of Eliminating Anchor Items Flagged from Statistical Criteria on Test Score Classifications in Common Item Equating*. Paper presented at the American Educational Research Association, Montreal, Canada.

IRT Item-Ability Regression Curves

We will also compute differences between the item-ability regression curves of the anchor items for the *TerraNova* and 2008 calibrations. The differences between the curves will be evaluated using the following statistics:

- UnWtd Mean = Average signed difference in estimated probability.
- UnWtd Mean Abs Dif = Average Absolute (unsigned) difference in estimated probability.
- UnWtd RMSD = Root mean squared difference.

Both unweighted and weighted versions of these statistics will be calculated. Unweighted differences give equal weight to differences across the ability spectrum. Weighted differences assign weights according to the number of test-takers that are impacted.

The weighted versions of these differences are:

- Wtd Mean = Weighted average signed difference in estimated probability.
- Wtd Mean Abs = Weighted average Absolute (unsigned) difference in estimated probability.
- WtdRMSD = Weighted Root mean squared difference.

For the six statistics listed above, differences greater than $+0.10$ are considered large, and differences between $+0.07$ and 0.10 are considered moderate.

Additionally, the Maximum Absolute difference (MaxAbsDifPC) will be identified. For MaxAbsDIFPC, large differences are those greater than $+0.15$, and moderate differences are all differences between $+0.125$ and 0.15 .

While dropping an anchor item flagged based solely on statistical criteria has its simplicity, this option may change the content coverage and equating constants, shift scale score distributions, and affect the classification of students by moving them into different proficiency levels. Before an anchor item may be dropped from an anchor set, the adequacy of the content coverage must be evaluated.

As stated above, an item is removed from the anchor set only if it adversely affects quality of scaling, not desirability of results. As such, CTB will not consider how the removal of an item affects the overall mean scale score or the impact data (percent of students in each achievement level) when recommending items for removal.

Items removed from the anchor set are still scored as part of the whole test. Anchor items are considered for exclusion from the MAP under the following conditions:

1. Items flagged using the TCC method are considered for exclusion when the correlation between the input and equated item parameters is below $.90$ for the b -parameter or below $.80$ for the a -parameter. If the exclusion of an outlying anchor item increases the

correlation to above .90 for the b -parameter or above .80 for the a -parameter, then the anchor is a candidate for removal.

2. An item is a candidate for removal when it is flagged for large differences on four of the seven statistics considered when examining the differences between the IRT regression curves.
3. Removal of the item will only be considered after alternative explanations have been considered that may explain shifts in performance. For example, performance on the anchor item may improve because of a statewide initiative emphasizing instruction on a particular set of skills. In this case, improved performance on the item represents true growth in that area. Removing the anchor item may artificially lower test scores.
4. Removal of the item may not significantly alter the content distribution of the anchor set. The distribution of the anchor items across the content standards must remain within 10% of the 2008 test blueprint.
5. The number of remaining items will remain at an acceptable level of anchor set reliability. Operationally, this means the anchor set will still be representative of the total test blueprint and that the anchor may not be less than 20% of the total test length.

Results of Analyses

Table 1 lists the flagged items using the two methods. Table 1 shows that only Mathematics items were only flagged for the original linking. In Grades 3, 4, 5, 7, and 8, only two items were flagged for further investigation. In Grade 10, seven items were flagged for further investigation.

Grade 3 Mathematics: Item 28 was flagged using the IRT Curve method. Further examination of the item revealed no content-based reason (e.g., the item's wording changed or a distractor was changed) for the change in performance. The removal of the item did not alter the content distribution of the anchor set, nor did it affect the quality of the overall equating. Anchor item 28 was removed from the Grade 3 Mathematics anchor.

Grade 4 Mathematics: Item 54 was flagged using the IRT Curve method. Further examination of the item revealed no content-based reason (e.g., the item's wording changed or a distractor was changed) for the change in performance. The removal of the items did not alter the content distribution of the anchor set, nor did it affect the quality of the overall equating. Anchor item 54 was removed from the Grade 4 Mathematics anchor.

Grade 5 Mathematics: Item 55 was flagged as an outlier on the b -parameter using the Stocking and Lord TCC method. It was also flagged using the IRT curve method. Item 43 was also flagged using the IRT curve method. Both items were easier and better discriminated in 2008 than they did on the *TerraNova* test. The removal of these items did not alter the content distribution of the anchor set. The removal of these items positively affected the quality of the overall equating. Anchor items 43 and 55 were removed from the Grade 5 Mathematics anchor.

Grade 7 Mathematics: Items 43 and 49 were flagged using the IRT curve method. Both items were easier than they had been on the *TerraNova* assessment. The removal of the item did not alter the content distribution of the anchor set. The removal of the items did not alter the content

distribution of the anchor set, nor did it affect the quality of the overall equating. Anchor item 43 and 49 were removed from the Grade 5 Mathematics anchor.

Grade 8 Mathematics: Items 43 and 47 were flagged using the IRT curve method. Item 43 better discriminated and was easier than it had been on the *TerraNova* assessment. Item 47 was easier than it had been on the *TerraNova* assessment. The removal of the item did not alter the content distribution of the anchor set. The removal of the items did not alter the content distribution of the anchor set, nor did it affect the quality of the overall equating. Anchor item 43 and 49 were removed from the Grade 5 Mathematics anchor.

Grade 10 Mathematics: Items 32 and 43 were flagged as outliers on the *b*-parameter using the Stocking and Lord TCC method. Items 37 and 44 were flagged as outliers on the *a*-parameter using the Stocking and Lord TCC method. Items 26, 37, 39, 42, and 43 were flagged using the IRT curve method. Of these items, three items were removed from the anchor set: 37, 43, and 44. Items 37 and 43 were identified as outliers using both methods. The removal of item 44 helped improved the quality of the overall equating as measured by the correlation coefficient of the input and estimated *a*-parameters to be greater than the desired level ($r > .80$). The removal of the items did not alter the content distribution of the anchor set.

Detailed Results of the Test Characteristic Curve Method

Tables 1 through 3 provides results for the TCC method for both the original and alternate (Alt) linking methods. These tables summarize the following information for each grade content area: grade level, number of iterations, scaling constants (M1 and M2), and quadratic loss function (F). Within each grade level, the following information is summarized for each item parameter estimate: difference (Diff), root mean square difference (RMSD), ratio of the standard deviation (SD Ratio), correlation (*r*) between input (*TerraNova*) and estimated (2008) anchor parameters, and the outlying anchor items. When applicable, the results of the recommended equating are also displayed in the table. Please note that the actual test characteristic curves are shown above in Figures 1–35 for the original and alternate linking.

These plots are used to assess the quality of the linking results. The light blue TCC lines in the plots are the TCCs for the input anchor items. The dark blue lines are the TCCs from the 2008 MAP parameter estimates transformed to the *TerraNova* scale². The closer the two TCCs are to each other at all ability levels, the more confidence we have in the equating result.

² The *c*-parameters for the MAP test data were fixed to the original *TerraNova* *c*-parameters in order to provide more accurate equating results (Voelkle, Schwarz, Arenson, & Ito 2002).

Detailed Results Comparing the IRT Anchor Regression Curves

Tables 5 through 33 present the detailed results for both the original and alternate linking when the IRT Anchor Regression method is used. These tables summarize the seven statistics examined using this method. The headers in the tables are abbreviated as follows:

- UnWtd RMSD = unweighted root mean squared difference
- UnWtd Mean Abs Difference = average absolute difference in estimated probability.
- UnWtd Max = maximum absolute difference.
- UnWtd Mean = average signed difference in estimated probability.
- Wtd RMSD = weighted root mean squared difference.
- Wtd Mean Abs Difference = weighted average absolute difference in estimated probability.
- Wtd Mean = weighted average signed difference in estimated probability.

Again, for six of the statistics listed above (except the maximum absolute difference), differences greater than $\pm.10$ were considered large, and differences greater than $\pm.07$ were considered moderate. For maximum absolute difference, large differences were those greater than $\pm.15$, and moderate differences were all differences greater than $\pm.125$.

Table 1. Anchor Items Flagged Using Different Methods

Content	Grade	Methods for Flagging	
		Stocking and Lord	IRT Curve
Mathematics	3		28
	4		54
	5	55 (b)	43, 55
	6		
	7		43, 49
	8		43, 47
	10	37(a), 44 (a), 32 (b), 43 (b)	26, 37, 39, 42, 43

Table 2. Detailed Results from the Test Characteristic Curve Method, Communication Arts

Grade	Iterations	M1	M2	F	Par	Diff	RMSD	SD Ratio	r	Items far from est. (RMS)
3	26	31.184	636.03	0.239	a b	0.001 0.989	0.004 5.944	1.164 0.948	0.876 0.978	Items 16, 21 Items 20, 28
3 Alt	19	32.583	643.16	0.124	a b	-0.001 0.376	0.002 3.524	0.910 0.970	0.969 0.993	Item 49 Items 24, 50
4	31	29.992	651.70	0.143	a b	0.001 0.472	0.004 8.300	1.017 0.988	0.849 0.955	Item 50 Item 15
4 Alt	30	28.457	660.33	0.092	a b	0.001 0.372	0.003 5.211	1.068 1.004	0.960 0.980	Item 31
5	19	29.453	667.96	0.094	a b	0.001 -0.098	0.005 9.176	0.976 0.996	0.881 0.932	Items 36, 40 Item 25
5 Alt	25	27.650	677.10	0.055	a b	0.001 -0.268	0.002 3.702	1.030 1.014	0.976 0.988	Items 40, 47 Items 24, 25
6	23	30.063	671.63	0.071	a b	0.001 -0.871	0.004 7.019	1.215 1.046	0.920 0.964	Items 16, 51 Item 42
6 Alt	12	27.446	675.84	0.079	a b	0.000 0.019	0.002 3.424	1.027 1.005	0.976 0.989	Item 29 Items 40, 41
7	23	32.852	680.59	0.161	a b	0.000 -0.596	0.004 10.920	1.006 1.153	0.831 0.942	Item 56 Items 23, 45
7 Alt	20	30.149	681.50	0.077	a b	0.000 -0.290	0.002 4.634	0.963 1.047	0.966 0.988	Items 34, 52, 53 Items 23, 45
8	19	34.049	682.60	0.211	a b	0.001 0.323	0.005 8.567	0.988 1.046	0.896 0.976	Items 24, 46
8 Alt	19	28.743	696.34	0.220	a b	0.000 -0.421	0.003 4.779	0.942 1.022	0.963 0.990	Item 51
11	16	30.829	715.74	0.234	a b	0.001 1.321	0.004 10.114	1.076 1.114	0.905 0.962	Items 20, 24 Item 50

Table 3. Detailed Results from the Test Characteristic Curve Method, Mathematics

Grade	Iterations	M1	M2	F	Par	Diff	RMSD	SD Ratio	r	Items far from est. (RMS)
3	47	32.366	622.34	0.193	a	0.000	0.004	0.851	0.847	Item 34
					b	1.030	13.003	0.948	0.923	Item 28
3 Ancs Removed	51	32.210	623.86	0.269	a	0.000	0.004	0.849	0.865	Item 34
					b	1.546	11.445	0.938	0.943	Item 33
3 Alt	28	32.056	626.18	0.116	a	0.000	0.001	0.959	0.980	Item 40
					b	-0.168	3.767	0.998	0.993	
4	39	31.253	643.59	0.395	a	0.003	0.004	1.092	0.910	Items 34, 54
					b	-1.179	13.074	1.147	0.935	Items 40, 54
4 Ancs Removed	37	31.854	642.95	0.336	a	0.002	0.003	1.101	0.937	Items 34, 55
					b	-1.313	11.339	1.122	0.951	Items 40, 55
4 Alt	30	30.552	648.15	0.222	a	0.001	0.002	1.074	0.971	Items 42, 44
					b	-0.432	2.686	1.000	0.997	Item 41
5	29	32.605	656.11	0.352	a	0.000	0.005	0.863	0.900	Items 32, 43
					b	-2.615	20.040	1.036	0.776	Item 55
5 Ancs Removed	29	31.353	652.83	0.300	a	0.000	0.004	0.946	0.917	Item 32
					b	-1.206	11.739	1.138	0.922	Items 44, 50
5 Alt	31	35.513	665.81	0.085	a	0.000	0.001	0.974	0.988	Items 32, 38
					b	-0.773	4.556	1.075	0.992	Items 26, 44
6	27	35.842	681.24	0.093	a	-0.001	0.003	0.987	0.919	Item 45
					b	-0.672	9.107	0.956	0.938	Item 53
6 Alt	25	35.499	684.78	0.012	a	0.000	0.001	0.969	0.986	Item 43
					b	-0.019	5.224	1.017	0.974	Item 53
7	13	33.257	681.33	0.039	a	0.001	0.004	1.201	0.874	Items 39, 54
					b	-0.336	10.983	1.132	0.933	Items 26, 49
7 Ancs Removed	14	32.708	679.53	0.043	a	0.002	0.004	1.238	0.898	Items 39, 54
					b	-0.521	9.779	1.163	0.952	Items 26, 33
7 Alt	8	36.513	686.02	0.060	a	0.000	0.003	1.025	0.938	Items 45, 51
					b	0.056	3.701	1.013	0.994	
8	7	37.583	701.54	0.199	a	0.000	0.005	1.138	0.809	Items 43, 48
					b	0.149	12.720	0.992	0.945	Items 47, 48
8 Ancs Removed	6	36.603	698.97	0.083	a	0.000	0.004	1.114	0.872	Item 48
					b	-0.091	10.657	0.989	0.962	Items 25, 48
8 Alt	7	34.322	706.17	0.032	a	0.001	0.002	1.068	0.968	Items 43, 53, 54
					b	-0.185	3.610	1.011	0.995	Items 26, 46
10	12	45.582	735.26	0.247	a	-0.001	0.006	0.831	0.644	Items 37, 44
					b	-1.715	20.586	1.110	0.824	Items 32, 43
10 Ancs Removed	14	43.886	731.08	0.232	a	-0.001	0.004	1.031	0.829	Item 32
					b	-1.912	17.662	1.143	0.863	

Table 4. Detailed Results from the Test Characteristic Curve Method, Science

Grade	Iterations	M1	M2	F	Par	Diff	RMSD	SD Ratio	r	Items far from est. (RMS)
5	12	27.927	662.70	0.345	a	0.000	0.004	1.090	0.917	Item 20
					b	-0.535	7.345	0.989	0.981	Item 13
8	24	26.847	696.23	0.456	a	0.001	0.004	1.109	0.906	Item 34
					b	-2.393	11.512	0.902	0.931	Item 36
11	5	33.812	724.69	0.243	a	0.000	0.004	1.012	0.924	Item 52
					b	-1.583	8.011	1.067	0.968	Items 43, 50

Table 5. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 3 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
16	0.03	0.02	0.08	-0.01	0.02	0.02	0.01
17	0.02	0.01	0.04	0.00	0.02	0.01	0.01
18	0.02	0.01	0.04	0.01	0.02	0.02	0.01
19	0.01	0.01	0.01	0.00	0.01	0.01	0.00
20	0.05	0.03	0.12	-0.03	0.02	0.01	0.00
21	0.04	0.02	0.10	-0.02	0.02	0.01	0.00
22	0.02	0.01	0.04	0.00	0.02	0.02	0.01
23	0.01	0.01	0.02	0.01	0.02	0.01	0.01
24	0.03	0.03	0.06	0.02	0.04	0.03	0.03
27	0.03	0.03	0.06	-0.03	0.05	0.05	-0.05
28	0.06	0.04	0.10	-0.04	0.08	0.08	-0.08
29	0.02	0.02	0.05	-0.01	0.03	0.02	0.00
30	0.02	0.01	0.05	-0.01	0.02	0.02	0.00
31	0.03	0.02	0.05	0.02	0.04	0.04	0.04
32	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
33	0.04	0.02	0.08	0.02	0.05	0.05	0.05
37	0.02	0.01	0.05	-0.01	0.02	0.01	-0.01
38	0.02	0.01	0.06	-0.01	0.04	0.03	-0.03
39	0.04	0.03	0.07	0.03	0.03	0.02	0.02
40	0.03	0.02	0.07	-0.02	0.05	0.04	-0.04
41	0.01	0.00	0.02	0.00	0.01	0.01	0.01
44	0.01	0.01	0.02	0.00	0.01	0.01	0.01
45	0.01	0.01	0.02	0.00	0.02	0.01	0.00
46	0.01	0.01	0.03	0.01	0.02	0.02	0.02
47	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
48	0.02	0.01	0.05	0.01	0.03	0.03	0.03
49	0.01	0.01	0.03	0.01	0.02	0.01	0.01
50	0.04	0.03	0.07	-0.03	0.05	0.05	-0.05
51	0.04	0.03	0.08	0.01	0.03	0.02	-0.01
52	0.01	0.01	0.02	0.00	0.02	0.02	-0.01

Table 6. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 4 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
15	0.10	0.07	0.18	-0.07	0.08	0.06	-0.06
16	0.03	0.01	0.08	-0.01	0.01	0.00	0.00
17	0.03	0.03	0.06	0.00	0.04	0.04	0.03
18	0.07	0.05	0.15	-0.05	0.05	0.03	-0.03
19	0.05	0.04	0.11	-0.02	0.04	0.03	0.00
20	0.02	0.01	0.02	0.00	0.02	0.02	0.01
21	0.02	0.01	0.04	-0.01	0.01	0.00	0.00
22	0.02	0.01	0.06	0.01	0.04	0.03	0.03
23	0.02	0.01	0.06	0.01	0.04	0.03	0.02
24	0.01	0.01	0.02	0.00	0.01	0.01	0.01
25	0.04	0.03	0.07	0.01	0.03	0.03	0.00
26	0.00	0.00	0.01	0.00	0.00	0.00	0.00
29	0.01	0.01	0.03	0.01	0.01	0.01	0.01
30	0.03	0.02	0.06	0.02	0.02	0.02	0.02
31	0.08	0.05	0.19	0.05	0.06	0.03	0.03
32	0.02	0.01	0.06	0.01	0.03	0.02	0.02
33	0.07	0.04	0.17	0.04	0.06	0.04	0.04
34	0.06	0.04	0.13	-0.03	0.09	0.08	-0.08
35	0.05	0.03	0.09	0.03	0.04	0.03	0.03
36	0.02	0.02	0.05	0.01	0.03	0.03	0.02
37	0.02	0.01	0.05	-0.01	0.02	0.01	-0.01
38	0.02	0.01	0.04	-0.01	0.02	0.01	-0.01
39	0.02	0.02	0.06	0.01	0.02	0.01	0.01
40	0.01	0.01	0.01	0.00	0.01	0.01	0.00
41	0.02	0.02	0.04	0.01	0.03	0.03	0.02
42	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
45	0.01	0.01	0.03	0.00	0.01	0.01	0.00
46	0.05	0.04	0.10	0.03	0.04	0.03	0.02
47	0.05	0.04	0.09	-0.04	0.07	0.07	-0.07
48	0.03	0.02	0.07	-0.02	0.04	0.03	-0.03
49	0.02	0.01	0.05	-0.01	0.03	0.03	-0.03
50	0.04	0.03	0.09	0.01	0.04	0.04	-0.02
51	0.00	0.00	0.01	0.00	0.01	0.00	0.00
52	0.02	0.01	0.05	0.01	0.03	0.03	0.03
53	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02

Table 7. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 5 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
15	0.04	0.02	0.08	-0.02	0.02	0.01	-0.01
16	0.07	0.05	0.16	-0.04	0.06	0.04	-0.03
17	0.00	0.00	0.01	0.00	0.01	0.01	0.00
18	0.04	0.03	0.08	-0.03	0.03	0.02	-0.02
19	0.07	0.05	0.13	-0.05	0.09	0.08	-0.08
20	0.01	0.01	0.02	0.00	0.02	0.01	0.01
21	0.03	0.03	0.06	0.03	0.05	0.04	0.04
22	0.03	0.02	0.06	0.02	0.03	0.03	0.03
23	0.02	0.01	0.03	0.01	0.02	0.02	0.02
24	0.07	0.05	0.14	0.05	0.08	0.06	0.06
25	0.11	0.08	0.18	0.07	0.10	0.08	0.07
26	0.03	0.02	0.08	0.02	0.03	0.02	0.02
27	0.04	0.03	0.05	0.03	0.04	0.04	0.04
28	0.05	0.04	0.09	0.04	0.06	0.06	0.06
29	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
30	0.05	0.03	0.14	0.03	0.07	0.05	0.05
36	0.03	0.02	0.07	0.00	0.03	0.02	0.02
37	0.05	0.03	0.12	-0.03	0.06	0.04	-0.04
38	0.01	0.01	0.02	0.01	0.02	0.02	0.02
39	0.05	0.03	0.10	-0.03	0.05	0.04	-0.03
40	0.03	0.02	0.06	0.00	0.04	0.03	0.00
41	0.03	0.02	0.05	0.00	0.03	0.03	-0.02
44	0.02	0.01	0.03	0.00	0.02	0.02	-0.02
45	0.03	0.02	0.08	-0.02	0.04	0.02	-0.02
46	0.03	0.02	0.07	0.02	0.05	0.04	0.04
47	0.01	0.01	0.03	0.00	0.01	0.01	0.01
48	0.05	0.03	0.13	-0.03	0.07	0.05	-0.05
49	0.01	0.01	0.02	0.00	0.02	0.01	0.00
50	0.04	0.03	0.07	-0.03	0.05	0.05	-0.05
51	0.03	0.02	0.09	-0.01	0.05	0.04	-0.04
52	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
53	0.03	0.02	0.07	-0.02	0.05	0.05	-0.05

Table 8. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 6 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
14	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
15	0.02	0.02	0.03	0.00	0.02	0.02	-0.02
16	0.03	0.02	0.08	-0.01	0.02	0.01	0.00
17	0.02	0.01	0.04	0.01	0.01	0.01	-0.01
18	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
19	0.04	0.03	0.07	-0.02	0.04	0.03	-0.03
20	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
21	0.02	0.01	0.03	0.01	0.01	0.01	0.01
22	0.02	0.01	0.02	0.00	0.02	0.02	0.00
23	0.03	0.02	0.06	0.02	0.04	0.04	0.04
24	0.02	0.01	0.06	-0.01	0.02	0.01	0.00
28	0.01	0.01	0.02	0.00	0.01	0.01	0.00
29	0.06	0.04	0.12	0.04	0.09	0.08	0.08
30	0.05	0.05	0.08	0.00	0.04	0.03	-0.02
31	0.03	0.02	0.06	0.02	0.03	0.03	0.03
34	0.01	0.01	0.02	0.00	0.01	0.01	0.00
35	0.05	0.05	0.08	0.02	0.05	0.05	0.03
36	0.03	0.02	0.06	-0.02	0.03	0.02	-0.02
37	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
38	0.01	0.01	0.03	0.00	0.02	0.01	-0.01
39	0.03	0.02	0.07	-0.02	0.05	0.04	-0.04
40	0.03	0.03	0.05	0.00	0.03	0.03	-0.02
41	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
42	0.09	0.07	0.15	0.07	0.08	0.07	0.07
43	0.02	0.01	0.03	-0.01	0.03	0.02	-0.02
47	0.02	0.02	0.05	0.02	0.02	0.02	0.02
48	0.02	0.01	0.05	0.01	0.03	0.02	0.02
49	0.04	0.03	0.07	-0.03	0.05	0.05	-0.05
50	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
51	0.03	0.02	0.07	0.01	0.04	0.04	0.03
52	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01

Table 9. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 7 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
19	0.01	0.01	0.03	0.01	0.01	0.01	0.01
20	0.05	0.04	0.09	-0.04	0.07	0.07	-0.07
21	0.02	0.02	0.05	-0.02	0.02	0.01	-0.01
22	0.04	0.02	0.08	-0.02	0.04	0.03	-0.02
23	0.08	0.05	0.14	0.05	0.05	0.04	0.04
24	0.01	0.00	0.01	0.00	0.01	0.01	0.01
25	0.07	0.06	0.11	-0.06	0.09	0.09	-0.09
26	0.03	0.02	0.05	0.02	0.02	0.01	0.01
27	0.04	0.02	0.08	0.02	0.02	0.01	0.00
28	0.03	0.02	0.07	-0.01	0.02	0.02	0.01
33	0.05	0.03	0.11	-0.03	0.05	0.03	-0.03
34	0.04	0.04	0.06	0.02	0.04	0.04	0.03
35	0.07	0.04	0.15	-0.04	0.08	0.06	-0.06
36	0.05	0.03	0.11	-0.03	0.06	0.04	-0.04
37	0.03	0.02	0.05	-0.02	0.03	0.03	-0.03
38	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
39	0.03	0.02	0.07	0.02	0.04	0.04	0.04
40	0.02	0.01	0.05	-0.01	0.04	0.03	-0.03
41	0.03	0.02	0.06	0.02	0.04	0.04	0.04
42	0.02	0.01	0.03	0.00	0.02	0.02	0.01
43	0.02	0.02	0.05	-0.01	0.03	0.03	-0.03
44	0.03	0.02	0.06	-0.02	0.04	0.04	-0.04
45	0.10	0.07	0.19	0.06	0.07	0.05	0.03
46	0.03	0.02	0.07	-0.02	0.05	0.05	-0.05
47	0.02	0.01	0.04	-0.01	0.02	0.02	-0.02
48	0.03	0.03	0.05	0.00	0.03	0.02	0.01
51	0.07	0.05	0.17	0.05	0.11	0.09	0.09
52	0.05	0.03	0.10	0.03	0.06	0.05	0.05
53	0.03	0.02	0.07	0.02	0.04	0.03	0.03
54	0.02	0.01	0.03	0.00	0.02	0.02	0.01
55	0.04	0.02	0.11	0.02	0.06	0.04	0.04
56	0.05	0.04	0.10	0.01	0.06	0.06	0.02
57	0.01	0.01	0.03	0.00	0.02	0.02	0.01

Table 10. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 8 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
19	0.01	0.01	0.03	0.00	0.01	0.01	0.01
20	0.04	0.03	0.08	0.03	0.03	0.02	0.02
21	0.02	0.01	0.04	0.01	0.02	0.02	0.02
22	0.02	0.01	0.04	0.00	0.01	0.01	0.01
23	0.02	0.02	0.06	-0.01	0.03	0.02	-0.01
24	0.05	0.04	0.10	0.04	0.03	0.02	0.02
25	0.02	0.02	0.03	-0.02	0.02	0.02	-0.02
26	0.03	0.02	0.08	-0.01	0.01	0.01	0.00
27	0.08	0.05	0.19	-0.05	0.08	0.05	-0.05
28	0.03	0.03	0.08	-0.02	0.03	0.03	-0.01
29	0.02	0.02	0.03	-0.01	0.02	0.02	-0.02
30	0.03	0.02	0.07	-0.01	0.03	0.03	-0.02
31	0.06	0.04	0.11	-0.03	0.05	0.03	-0.01
32	0.03	0.02	0.06	0.02	0.03	0.03	0.03
33	0.06	0.04	0.12	-0.03	0.05	0.03	-0.03
34	0.06	0.04	0.16	-0.04	0.05	0.02	-0.02
35	0.06	0.04	0.12	-0.04	0.07	0.05	-0.05
39	0.02	0.01	0.03	0.01	0.02	0.02	0.02
40	0.05	0.03	0.13	0.03	0.04	0.03	0.03
41	0.01	0.01	0.02	0.01	0.01	0.01	0.01
42	0.02	0.02	0.03	0.00	0.02	0.02	0.01
45	0.01	0.01	0.02	0.00	0.01	0.01	0.01
46	0.09	0.06	0.19	0.06	0.08	0.05	0.05
47	0.02	0.02	0.04	0.01	0.02	0.02	0.01
48	0.03	0.02	0.08	0.01	0.03	0.02	0.01
49	0.04	0.02	0.10	-0.02	0.07	0.05	-0.05
50	0.02	0.02	0.03	0.01	0.02	0.02	0.02
51	0.04	0.03	0.11	0.03	0.04	0.03	0.02
52	0.00	0.00	0.01	0.00	0.00	0.00	0.00
53	0.04	0.02	0.09	0.02	0.05	0.04	0.04
54	0.02	0.01	0.03	0.01	0.02	0.02	0.02
55	0.02	0.01	0.06	-0.01	0.03	0.02	-0.02
56	0.04	0.02	0.12	-0.02	0.06	0.04	-0.04
57	0.04	0.02	0.10	-0.02	0.06	0.04	-0.04

Table 11. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 11 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
20	0.02	0.01	0.06	-0.01	0.03	0.03	-0.01
21	0.04	0.02	0.10	-0.02	0.02	0.01	0.00
22	0.05	0.04	0.10	-0.04	0.06	0.05	-0.05
23	0.03	0.02	0.05	-0.02	0.02	0.02	-0.01
24	0.06	0.04	0.15	-0.03	0.09	0.08	-0.08
25	0.04	0.03	0.07	-0.03	0.04	0.04	-0.04
26	0.03	0.02	0.07	-0.01	0.03	0.02	-0.01
27	0.05	0.03	0.10	-0.03	0.07	0.07	-0.07
28	0.04	0.02	0.08	-0.02	0.05	0.03	-0.03
29	0.01	0.01	0.02	0.00	0.01	0.01	0.00
30	0.04	0.03	0.06	-0.02	0.03	0.02	-0.02
31	0.05	0.03	0.11	0.03	0.03	0.01	0.01
33	0.05	0.04	0.08	-0.01	0.04	0.04	0.01
35	0.03	0.01	0.07	0.01	0.04	0.03	0.03
36	0.04	0.02	0.10	0.02	0.05	0.04	0.04
37	0.03	0.02	0.08	0.02	0.05	0.03	0.03
38	0.03	0.01	0.07	0.01	0.04	0.03	0.03
39	0.04	0.03	0.08	0.02	0.04	0.03	0.00
40	0.05	0.03	0.09	0.03	0.05	0.04	0.04
41	0.01	0.01	0.04	0.01	0.02	0.01	0.01
42	0.02	0.02	0.05	-0.01	0.03	0.03	-0.02
43	0.02	0.01	0.07	0.01	0.03	0.02	0.02
44	0.02	0.02	0.04	0.00	0.03	0.02	0.00
45	0.04	0.03	0.08	-0.03	0.06	0.05	-0.05
47	0.05	0.04	0.11	0.04	0.08	0.07	0.07
48	0.04	0.03	0.08	0.03	0.04	0.03	0.02
49	0.01	0.01	0.02	0.01	0.01	0.01	0.01
50	0.14	0.09	0.28	-0.09	0.06	0.04	-0.04
52	0.00	0.00	0.01	0.00	0.00	0.00	0.00
53	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
54	0.03	0.02	0.05	0.02	0.04	0.03	0.03
55	0.03	0.02	0.06	0.02	0.04	0.03	0.03
56	0.02	0.01	0.03	0.01	0.02	0.02	0.02
57	0.01	0.01	0.01	0.01	0.01	0.01	0.01

**Table 12. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 3
(Original Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.04	0.02	0.08	-0.02	0.02	0.01	-0.01
25	0.04	0.03	0.08	0.00	0.03	0.03	0.02
26	0.07	0.06	0.13	-0.06	0.08	0.07	-0.07
27	0.04	0.03	0.08	-0.03	0.06	0.05	-0.05
28	0.16	0.11	0.32	-0.11	0.17	0.14	-0.14
29	0.02	0.01	0.06	0.01	0.01	0.01	0.00
30	0.05	0.04	0.08	-0.04	0.07	0.06	-0.06
31	0.01	0.00	0.01	0.00	0.01	0.01	0.01
32	0.06	0.04	0.13	0.03	0.04	0.02	0.01
33	0.07	0.04	0.14	-0.04	0.03	0.02	-0.01
34	0.03	0.02	0.08	0.01	0.02	0.01	0.01
35	0.04	0.03	0.10	-0.02	0.02	0.01	0.00
36	0.06	0.04	0.12	-0.04	0.03	0.02	-0.02
37	0.03	0.02	0.08	0.02	0.02	0.01	0.00
38	0.06	0.04	0.14	0.04	0.09	0.08	0.08
39	0.02	0.01	0.06	-0.01	0.02	0.01	-0.01
40	0.09	0.06	0.17	0.06	0.07	0.05	0.05
41	0.04	0.03	0.09	0.03	0.04	0.03	0.03
42	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
43	0.05	0.03	0.12	0.03	0.03	0.01	0.01
44	0.04	0.03	0.08	0.01	0.05	0.04	0.02
45	0.02	0.01	0.02	0.01	0.02	0.02	0.02
46	0.06	0.05	0.12	0.05	0.08	0.07	0.07
47	0.02	0.01	0.04	0.00	0.02	0.02	-0.01
48	0.06	0.04	0.11	0.04	0.08	0.07	0.07
49	0.04	0.03	0.07	0.02	0.03	0.02	0.02
50	0.03	0.02	0.05	0.02	0.03	0.02	0.02
51	0.02	0.01	0.06	-0.01	0.04	0.03	-0.03
52	0.05	0.03	0.11	-0.03	0.07	0.06	-0.06
53	0.03	0.02	0.07	0.01	0.04	0.03	0.03

Table 13. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 4 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
25	0.03	0.02	0.06	-0.02	0.02	0.01	-0.01
26	0.05	0.03	0.11	-0.03	0.06	0.06	-0.06
27	0.04	0.03	0.08	-0.02	0.04	0.03	-0.02
28	0.04	0.03	0.08	-0.03	0.05	0.04	-0.04
29	0.02	0.01	0.05	-0.01	0.02	0.01	-0.01
30	0.04	0.03	0.10	-0.02	0.06	0.04	-0.04
31	0.02	0.01	0.04	-0.01	0.03	0.03	-0.03
32	0.04	0.03	0.09	-0.03	0.04	0.02	-0.02
33	0.03	0.02	0.05	0.01	0.03	0.03	0.03
34	0.05	0.03	0.11	-0.03	0.06	0.05	-0.04
35	0.00	0.00	0.01	0.00	0.01	0.01	0.00
36	0.01	0.01	0.02	0.01	0.02	0.02	0.02
37	0.05	0.03	0.10	-0.03	0.02	0.01	-0.01
38	0.00	0.00	0.01	0.00	0.00	0.00	0.00
39	0.04	0.03	0.07	0.02	0.06	0.05	0.05
40	0.07	0.06	0.11	0.06	0.06	0.06	0.06
41	0.04	0.03	0.07	0.02	0.05	0.05	0.04
42	0.01	0.01	0.03	0.00	0.02	0.01	0.01
43	0.05	0.03	0.11	-0.03	0.06	0.05	-0.05
44	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
45	0.04	0.03	0.06	0.00	0.04	0.04	-0.03
46	0.03	0.02	0.06	-0.01	0.02	0.01	0.00
47	0.05	0.03	0.12	-0.03	0.07	0.06	-0.06
48	0.02	0.02	0.04	0.02	0.02	0.02	0.02
49	0.03	0.02	0.06	-0.02	0.03	0.03	-0.03
50	0.08	0.05	0.16	0.05	0.04	0.02	0.02
51	0.03	0.02	0.08	-0.02	0.04	0.03	-0.03
52	0.06	0.04	0.12	0.04	0.09	0.08	0.08
53	0.03	0.02	0.07	-0.02	0.03	0.02	-0.02
54	0.17	0.13	0.29	0.13	0.17	0.15	0.15
55	0.10	0.07	0.21	0.07	0.11	0.09	0.09
56	0.06	0.05	0.11	0.05	0.09	0.09	0.09

Table 14. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 5 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.01	0.01	0.02	-0.01	0.02	0.02	-0.02
25	0.03	0.03	0.06	-0.02	0.02	0.02	-0.01
26	0.05	0.03	0.09	-0.03	0.04	0.03	-0.03
27	0.08	0.05	0.18	-0.05	0.07	0.04	-0.04
28	0.05	0.04	0.10	-0.03	0.05	0.04	-0.04
29	0.03	0.02	0.08	-0.02	0.04	0.04	-0.04
30	0.01	0.01	0.03	0.01	0.01	0.01	0.00
31	0.01	0.01	0.01	0.01	0.01	0.01	0.01
32	0.04	0.02	0.10	-0.02	0.05	0.04	-0.03
33	0.03	0.02	0.06	0.01	0.02	0.02	-0.01
34	0.05	0.03	0.15	-0.03	0.07	0.05	-0.05
35	0.03	0.02	0.07	-0.02	0.04	0.03	-0.03
36	0.02	0.01	0.05	0.00	0.02	0.02	-0.01
37	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
38	0.02	0.01	0.05	-0.01	0.03	0.02	-0.02
39	0.08	0.06	0.17	-0.06	0.10	0.08	-0.08
40	0.04	0.03	0.08	0.03	0.04	0.03	0.03
41	0.07	0.05	0.14	-0.05	0.05	0.04	-0.03
42	0.04	0.03	0.07	0.03	0.05	0.05	0.05
43	0.12	0.10	0.22	0.09	0.18	0.17	0.17
44	0.09	0.07	0.17	0.07	0.06	0.05	0.05
45	0.04	0.02	0.07	-0.02	0.04	0.03	-0.03
46	0.02	0.01	0.04	-0.01	0.02	0.01	-0.01
47	0.05	0.03	0.10	-0.03	0.07	0.07	-0.07
48	0.05	0.04	0.09	0.02	0.03	0.02	0.01
49	0.05	0.04	0.10	0.03	0.05	0.03	0.02
50	0.12	0.08	0.24	0.08	0.09	0.06	0.06
51	0.03	0.02	0.08	-0.02	0.05	0.04	-0.04
52	0.02	0.01	0.06	-0.01	0.04	0.03	-0.03
53	0.02	0.01	0.05	-0.01	0.03	0.02	-0.02
54	0.04	0.04	0.08	0.03	0.06	0.06	0.06
55	0.25	0.22	0.37	0.22	0.32	0.31	0.31

Table 15. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 6 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.03	0.02	0.06	0.02	0.02	0.01	0.01
25	0.05	0.04	0.09	-0.04	0.07	0.07	-0.07
26	0.03	0.03	0.05	0.03	0.04	0.04	0.04
27	0.06	0.04	0.12	-0.04	0.06	0.05	-0.05
28	0.03	0.02	0.05	-0.02	0.03	0.03	-0.03
29	0.06	0.04	0.12	-0.04	0.09	0.08	-0.08
30	0.03	0.02	0.07	-0.02	0.05	0.04	-0.04
31	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
32	0.04	0.02	0.09	0.02	0.02	0.01	0.00
33	0.02	0.01	0.03	-0.01	0.03	0.02	-0.02
34	0.03	0.03	0.05	0.02	0.03	0.03	0.03
35	0.02	0.02	0.05	-0.02	0.03	0.03	-0.03
36	0.03	0.02	0.07	-0.02	0.04	0.03	-0.03
37	0.02	0.02	0.04	0.02	0.03	0.03	0.03
38	0.03	0.02	0.08	-0.02	0.05	0.04	-0.04
39	0.02	0.01	0.04	0.01	0.02	0.01	0.01
40	0.04	0.03	0.05	0.00	0.04	0.04	-0.02
41	0.03	0.02	0.07	-0.01	0.03	0.02	-0.01
42	0.01	0.01	0.01	0.00	0.01	0.01	0.01
43	0.03	0.02	0.08	0.02	0.05	0.04	0.04
44	0.05	0.03	0.13	-0.03	0.09	0.07	-0.07
45	0.03	0.02	0.06	0.01	0.03	0.03	0.01
46	0.02	0.01	0.04	-0.01	0.03	0.02	-0.02
47	0.02	0.01	0.07	0.01	0.04	0.02	0.02
48	0.03	0.02	0.06	0.01	0.03	0.03	0.01
49	0.05	0.03	0.12	0.03	0.08	0.07	0.07
50	0.01	0.00	0.02	0.00	0.01	0.01	0.01
51	0.05	0.03	0.09	0.03	0.07	0.07	0.07
52	0.06	0.04	0.14	0.04	0.09	0.08	0.08
53	0.07	0.06	0.11	0.05	0.08	0.07	0.07
54	0.01	0.00	0.01	0.00	0.01	0.01	-0.01

**Table 16. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 7
(Original Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.02	0.02	0.03	-0.01	0.02	0.02	-0.02
25	0.05	0.04	0.11	0.04	0.07	0.06	0.06
26	0.09	0.07	0.15	-0.07	0.13	0.12	-0.12
27	0.03	0.01	0.07	-0.01	0.04	0.03	-0.03
28	0.03	0.03	0.06	-0.02	0.04	0.04	-0.04
29	0.04	0.03	0.08	-0.03	0.06	0.05	-0.05
30	0.02	0.01	0.05	-0.01	0.03	0.03	-0.03
31	0.04	0.04	0.07	-0.04	0.06	0.06	-0.06
32	0.01	0.01	0.02	0.00	0.01	0.01	0.00
33	0.09	0.04	0.23	0.04	0.03	0.01	0.01
34	0.03	0.02	0.07	-0.01	0.02	0.01	-0.01
35	0.04	0.02	0.08	0.02	0.03	0.02	0.01
36	0.03	0.02	0.06	-0.02	0.02	0.01	-0.01
37	0.04	0.03	0.07	0.03	0.05	0.05	0.05
38	0.03	0.03	0.06	-0.03	0.05	0.04	-0.04
39	0.04	0.03	0.08	0.01	0.05	0.05	0.04
40	0.05	0.03	0.11	0.03	0.04	0.02	0.02
41	0.05	0.03	0.09	-0.03	0.07	0.06	-0.06
42	0.04	0.03	0.08	0.03	0.05	0.04	0.04
43	0.08	0.05	0.18	0.05	0.12	0.11	0.11
44	0.09	0.06	0.18	-0.06	0.14	0.12	-0.12
45	0.03	0.02	0.05	0.01	0.03	0.02	0.01
46	0.01	0.01	0.04	0.01	0.02	0.02	0.02
47	0.01	0.01	0.03	0.00	0.02	0.02	-0.01
48	0.02	0.02	0.03	0.00	0.02	0.02	0.01
49	0.10	0.06	0.22	0.06	0.16	0.14	0.14
50	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
51	0.05	0.03	0.12	0.03	0.07	0.06	0.06
52	0.05	0.03	0.11	-0.03	0.08	0.07	-0.07
53	0.05	0.04	0.11	0.04	0.08	0.07	0.07
54	0.03	0.02	0.08	0.01	0.04	0.04	0.03
55	0.01	0.01	0.03	0.00	0.02	0.01	0.01

Table 17. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 8 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
25	0.07	0.05	0.13	-0.05	0.05	0.04	-0.04
26	0.02	0.02	0.05	-0.02	0.03	0.03	-0.03
27	0.06	0.04	0.11	-0.04	0.05	0.04	-0.04
28	0.00	0.00	0.01	0.00	0.00	0.00	0.00
29	0.02	0.02	0.03	0.02	0.03	0.03	0.03
30	0.07	0.04	0.16	-0.04	0.10	0.09	-0.09
31	0.06	0.04	0.11	-0.04	0.08	0.07	-0.07
32	0.04	0.04	0.07	0.01	0.03	0.03	-0.01
33	0.03	0.02	0.06	-0.02	0.04	0.04	-0.04
34	0.04	0.02	0.08	-0.02	0.03	0.02	-0.02
35	0.07	0.06	0.11	-0.05	0.09	0.09	-0.09
36	0.02	0.02	0.04	0.02	0.03	0.03	0.03
37	0.02	0.01	0.03	-0.01	0.03	0.02	-0.02
38	0.04	0.04	0.07	0.01	0.04	0.04	0.02
39	0.06	0.03	0.14	0.03	0.08	0.06	0.06
40	0.04	0.03	0.07	0.00	0.03	0.03	-0.02
41	0.04	0.02	0.07	0.02	0.03	0.02	0.02
42	0.02	0.01	0.04	0.01	0.03	0.02	0.02
43	0.11	0.07	0.26	0.07	0.18	0.15	0.14
44	0.04	0.04	0.07	0.03	0.05	0.04	0.04
45	0.02	0.01	0.03	0.01	0.02	0.01	0.01
46	0.03	0.03	0.06	0.01	0.03	0.02	-0.01
47	0.10	0.07	0.22	0.07	0.16	0.14	0.14
48	0.08	0.05	0.17	0.05	0.11	0.09	0.09
49	0.02	0.02	0.06	-0.01	0.02	0.02	0.00
50	0.04	0.02	0.11	-0.02	0.06	0.05	-0.05
51	0.06	0.04	0.11	-0.04	0.08	0.08	-0.08
52	0.01	0.01	0.04	0.00	0.02	0.02	-0.01
53	0.04	0.02	0.07	-0.02	0.05	0.04	-0.04
54	0.03	0.02	0.05	0.00	0.03	0.03	0.02
55	0.01	0.01	0.03	-0.01	0.02	0.01	-0.01

Table 18. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 10 (Original Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.07	0.05	0.15	-0.05	0.08	0.07	-0.07
25	0.06	0.05	0.12	-0.05	0.09	0.09	-0.09
26	0.11	0.08	0.25	-0.08	0.18	0.17	-0.17
27	0.03	0.03	0.05	0.00	0.04	0.03	-0.03
28	0.03	0.02	0.08	-0.01	0.05	0.04	-0.04
29	0.03	0.02	0.06	0.01	0.04	0.03	0.02
30	0.05	0.03	0.13	-0.03	0.08	0.06	-0.06
31	0.06	0.04	0.15	0.04	0.09	0.07	0.07
32	0.12	0.08	0.22	0.07	0.07	0.04	0.04
33	0.04	0.02	0.10	-0.02	0.06	0.04	-0.04
34	0.02	0.01	0.04	0.00	0.02	0.02	-0.01
35	0.02	0.01	0.05	0.01	0.02	0.01	0.01
36	0.03	0.02	0.09	-0.02	0.05	0.04	-0.04
37	0.09	0.05	0.22	0.05	0.13	0.11	0.10
38	0.01	0.01	0.02	-0.01	0.02	0.01	-0.01
39	0.08	0.06	0.17	0.06	0.12	0.11	0.11
40	0.04	0.03	0.09	-0.02	0.06	0.06	-0.05
41	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
42	0.08	0.05	0.19	0.05	0.13	0.11	0.11
43	0.16	0.12	0.34	0.12	0.25	0.24	0.24
44	0.04	0.03	0.10	0.00	0.06	0.05	0.01
45	0.02	0.01	0.04	0.01	0.03	0.03	0.02
46	0.07	0.04	0.17	-0.04	0.11	0.09	-0.09
47	0.04	0.03	0.07	-0.03	0.06	0.05	-0.05
48	0.03	0.01	0.08	0.01	0.04	0.03	0.03

**Table 19. Statistics Comparing IRT Item-Ability Regression Curves, Science, Grade 5
(Original Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
12	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
13	0.06	0.04	0.12	0.04	0.05	0.04	0.04
14	0.03	0.02	0.06	-0.02	0.03	0.03	-0.03
15	0.01	0.01	0.02	-0.01	0.02	0.02	-0.02
16	0.02	0.02	0.05	-0.02	0.02	0.02	-0.02
17	0.02	0.01	0.04	0.01	0.01	0.01	0.00
18	0.04	0.03	0.08	0.03	0.03	0.02	0.02
19	0.03	0.02	0.07	0.02	0.02	0.01	0.01
20	0.07	0.05	0.15	-0.04	0.04	0.03	-0.01
21	0.04	0.04	0.05	-0.01	0.03	0.03	-0.02
22	0.08	0.07	0.11	-0.05	0.09	0.08	-0.08
23	0.00	0.00	0.01	0.00	0.00	0.00	0.00
24	0.02	0.01	0.05	-0.01	0.03	0.03	-0.03
25	0.02	0.01	0.04	0.00	0.03	0.02	0.01
26	0.02	0.01	0.06	-0.01	0.03	0.03	-0.03
27	0.02	0.02	0.05	0.02	0.04	0.04	0.04
28	0.05	0.04	0.12	0.03	0.09	0.08	0.08
29	0.07	0.05	0.14	0.05	0.10	0.09	0.09
30	0.02	0.02	0.04	-0.01	0.03	0.03	-0.02
31	0.02	0.01	0.04	0.00	0.02	0.02	-0.01
32	0.06	0.03	0.15	-0.03	0.08	0.06	-0.06
33	0.03	0.02	0.08	0.02	0.05	0.05	0.05

**Table 20. Statistics Comparing IRT Item-Ability Regression Curves, Science, Grade 8
(Original Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
13	0.02	0.01	0.03	-0.01	0.02	0.01	-0.01
14	0.04	0.03	0.08	0.03	0.04	0.03	0.03
15	0.04	0.02	0.10	-0.02	0.04	0.03	-0.03
16	0.01	0.01	0.02	0.00	0.01	0.01	0.01
17	0.04	0.02	0.08	0.02	0.02	0.01	0.01
18	0.01	0.00	0.01	0.00	0.01	0.00	0.00
19	0.02	0.01	0.03	0.01	0.02	0.02	0.02
20	0.06	0.05	0.11	-0.05	0.08	0.07	-0.07
21	0.06	0.04	0.11	-0.04	0.05	0.04	-0.04
22	0.05	0.03	0.10	0.03	0.05	0.04	0.04
23	0.03	0.03	0.06	-0.03	0.04	0.04	-0.04
24	0.02	0.02	0.05	-0.02	0.02	0.02	-0.02
25	0.05	0.03	0.12	-0.03	0.07	0.06	-0.05
26	0.03	0.03	0.04	0.00	0.04	0.03	0.02
27	0.04	0.02	0.09	-0.02	0.05	0.04	-0.04
28	0.02	0.01	0.05	0.01	0.03	0.03	0.03
29	0.05	0.04	0.08	0.04	0.05	0.04	0.04
30	0.03	0.02	0.06	-0.02	0.04	0.04	-0.04
31	0.06	0.05	0.08	0.05	0.07	0.07	0.07
32	0.02	0.01	0.04	0.01	0.03	0.03	0.03
33	0.04	0.03	0.08	0.03	0.06	0.06	0.06
34	0.05	0.03	0.12	-0.02	0.08	0.07	-0.06
35	0.09	0.06	0.16	0.06	0.07	0.05	0.05
36	0.12	0.09	0.20	0.08	0.09	0.08	0.07

**Table 21. Statistics Comparing IRT Item-Ability Regression Curves, Science, Grade 11
(Original Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
30	0.03	0.02	0.07	-0.02	0.03	0.02	-0.01
31	0.04	0.03	0.07	0.03	0.06	0.06	0.06
32	0.02	0.02	0.04	0.01	0.02	0.02	0.01
33	0.02	0.02	0.04	-0.01	0.02	0.02	0.00
34	0.01	0.01	0.01	0.00	0.01	0.01	0.00
35	0.03	0.02	0.06	-0.01	0.04	0.03	-0.01
36	0.04	0.03	0.09	0.02	0.03	0.02	0.00
37	0.02	0.02	0.04	0.02	0.02	0.02	0.02
38	0.02	0.01	0.03	0.01	0.02	0.02	0.02
39	0.03	0.02	0.04	0.01	0.03	0.03	0.03
40	0.05	0.03	0.12	0.03	0.06	0.04	0.04
41	0.03	0.03	0.06	0.01	0.03	0.02	0.00
42	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
43	0.08	0.08	0.14	0.01	0.05	0.04	0.00
44	0.05	0.04	0.08	-0.04	0.06	0.05	-0.05
45	0.05	0.03	0.12	-0.03	0.08	0.07	-0.07
46	0.01	0.00	0.01	0.00	0.01	0.01	0.01
47	0.01	0.00	0.01	0.00	0.01	0.01	0.00
48	0.04	0.02	0.10	-0.02	0.05	0.03	-0.03
49	0.02	0.01	0.05	0.01	0.02	0.01	0.01
50	0.08	0.06	0.16	0.06	0.11	0.10	0.10
51	0.01	0.01	0.02	0.00	0.01	0.01	0.00
52	0.06	0.05	0.11	-0.01	0.08	0.07	-0.06

Table 22. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 3 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
16	0.02	0.01	0.06	-0.01	0.02	0.01	-0.01
17	0.01	0.00	0.02	0.00	0.01	0.00	0.00
18	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
19	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
20	0.03	0.01	0.06	-0.01	0.01	0.01	0.00
21	0.01	0.01	0.03	0.01	0.01	0.00	0.00
22	0.01	0.01	0.03	-0.01	0.01	0.01	-0.01
23	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
24	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
27	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
28	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
29	0.02	0.01	0.04	0.01	0.03	0.02	0.02
30	0.01	0.01	0.03	0.00	0.01	0.01	0.00
31	0.02	0.01	0.04	0.01	0.03	0.03	0.03
32	0.01	0.01	0.01	0.00	0.01	0.01	0.01
33	0.01	0.01	0.03	0.00	0.02	0.02	0.01
37	0.01	0.01	0.02	0.01	0.01	0.01	0.01
38	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
39	0.01	0.01	0.02	0.01	0.02	0.02	0.02
40	0.01	0.01	0.03	0.01	0.02	0.02	0.02
41	0.01	0.01	0.02	0.00	0.01	0.01	0.00
44	0.02	0.01	0.03	0.01	0.02	0.02	0.01
45	0.01	0.00	0.01	0.00	0.01	0.01	0.01
46	0.00	0.00	0.01	0.00	0.00	0.00	0.00
47	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
48	0.01	0.01	0.02	0.01	0.01	0.01	0.01
49	0.01	0.01	0.04	0.00	0.02	0.02	0.00
50	0.03	0.02	0.05	-0.02	0.03	0.02	-0.02
51	0.02	0.01	0.03	0.01	0.02	0.02	0.02
52	0.01	0.01	0.03	0.01	0.02	0.01	0.01

Table 23. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 4 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
15	0.03	0.02	0.06	-0.02	0.02	0.02	-0.02
16	0.03	0.01	0.07	-0.01	0.01	0.00	0.00
17	0.03	0.02	0.07	-0.02	0.02	0.01	-0.01
18	0.04	0.02	0.08	-0.02	0.03	0.02	-0.02
19	0.05	0.04	0.11	-0.03	0.04	0.03	-0.02
20	0.04	0.03	0.08	-0.03	0.03	0.02	-0.02
21	0.03	0.01	0.06	-0.01	0.01	0.00	0.00
22	0.01	0.01	0.03	0.00	0.02	0.01	0.00
23	0.01	0.01	0.02	0.00	0.01	0.01	0.00
24	0.02	0.01	0.04	-0.01	0.02	0.02	-0.02
25	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
26	0.03	0.02	0.07	0.02	0.03	0.02	0.02
29	0.03	0.02	0.07	0.02	0.03	0.02	0.02
30	0.02	0.02	0.05	0.02	0.02	0.02	0.02
31	0.07	0.04	0.16	0.04	0.05	0.03	0.03
32	0.01	0.00	0.02	0.00	0.01	0.01	0.01
33	0.04	0.02	0.09	0.02	0.03	0.02	0.02
34	0.01	0.01	0.02	0.01	0.02	0.01	0.01
35	0.04	0.03	0.09	0.03	0.04	0.03	0.03
36	0.01	0.00	0.01	0.00	0.01	0.01	0.00
37	0.01	0.01	0.02	0.01	0.02	0.02	0.02
38	0.00	0.00	0.01	0.00	0.01	0.01	0.01
39	0.01	0.01	0.02	0.00	0.01	0.00	0.00
40	0.01	0.01	0.01	0.00	0.01	0.01	0.01
41	0.01	0.01	0.03	0.00	0.01	0.01	0.00
42	0.01	0.01	0.01	0.00	0.01	0.01	0.01
45	0.00	0.00	0.01	0.00	0.01	0.01	-0.01
46	0.03	0.02	0.05	0.02	0.02	0.01	0.01
47	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
48	0.01	0.01	0.02	0.00	0.01	0.01	0.00
49	0.02	0.02	0.05	-0.01	0.03	0.03	-0.03
50	0.02	0.02	0.05	0.01	0.02	0.01	0.01
51	0.01	0.01	0.03	-0.01	0.01	0.01	-0.01
52	0.01	0.00	0.01	0.00	0.01	0.01	0.00
53	0.00	0.00	0.01	0.00	0.01	0.01	-0.01

Table 24. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 5 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
15	0.01	0.01	0.01	-0.01	0.01	0.01	-0.01
16	0.03	0.02	0.06	-0.01	0.02	0.01	-0.01
17	0.01	0.01	0.03	0.00	0.02	0.01	-0.01
18	0.02	0.01	0.03	-0.01	0.01	0.01	-0.01
19	0.02	0.02	0.05	-0.02	0.03	0.03	-0.03
20	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
21	0.01	0.01	0.02	0.01	0.01	0.01	0.00
22	0.04	0.03	0.07	0.03	0.05	0.04	0.04
23	0.01	0.01	0.02	0.01	0.01	0.01	0.01
24	0.04	0.02	0.08	0.02	0.04	0.03	0.03
25	0.03	0.02	0.05	0.02	0.03	0.02	0.02
26	0.00	0.00	0.00	0.00	0.00	0.00	0.00
27	0.03	0.03	0.05	0.02	0.03	0.03	0.03
28	0.01	0.01	0.02	0.01	0.01	0.01	0.01
29	0.00	0.00	0.01	0.00	0.00	0.00	0.00
30	0.02	0.02	0.06	0.02	0.03	0.02	0.02
36	0.01	0.01	0.03	0.01	0.01	0.01	0.01
37	0.02	0.01	0.05	-0.01	0.02	0.02	-0.02
38	0.01	0.01	0.02	-0.01	0.02	0.01	-0.01
39	0.01	0.01	0.02	-0.01	0.02	0.01	-0.01
40	0.02	0.01	0.04	-0.01	0.03	0.02	-0.02
41	0.02	0.02	0.03	-0.01	0.02	0.02	-0.02
44	0.02	0.01	0.04	-0.01	0.02	0.01	-0.01
45	0.02	0.01	0.04	-0.01	0.02	0.01	-0.01
46	0.02	0.01	0.04	0.01	0.03	0.02	0.02
47	0.02	0.01	0.03	0.00	0.01	0.01	0.01
48	0.01	0.01	0.04	-0.01	0.02	0.02	-0.02
49	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.02	0.02	0.04	-0.01	0.02	0.02	-0.02
51	0.00	0.00	0.01	0.00	0.01	0.01	-0.01
52	0.00	0.00	0.01	0.00	0.01	0.01	0.01
53	0.01	0.01	0.02	0.01	0.01	0.00	0.00

Table 25. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 6 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
14	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
15	0.01	0.01	0.01	0.00	0.00	0.00	0.00
16	0.01	0.00	0.01	0.00	0.01	0.00	0.00
17	0.01	0.01	0.02	0.01	0.01	0.01	0.01
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
20	0.01	0.01	0.02	-0.01	0.02	0.02	-0.02
21	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
22	0.03	0.02	0.05	-0.02	0.03	0.03	-0.03
23	0.02	0.02	0.03	0.00	0.02	0.02	-0.01
24	0.02	0.01	0.05	-0.01	0.01	0.01	-0.01
28	0.01	0.01	0.03	-0.01	0.01	0.01	-0.01
29	0.02	0.02	0.04	-0.01	0.03	0.02	0.00
30	0.01	0.01	0.01	-0.01	0.01	0.01	-0.01
31	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
34	0.01	0.00	0.02	0.00	0.01	0.01	-0.01
35	0.01	0.00	0.01	0.00	0.01	0.00	0.00
36	0.00	0.00	0.00	0.00	0.00	0.00	0.00
37	0.01	0.01	0.01	0.00	0.01	0.01	-0.01
38	0.01	0.01	0.04	-0.01	0.02	0.02	-0.02
39	0.00	0.00	0.01	0.00	0.01	0.01	-0.01
40	0.03	0.03	0.06	0.02	0.03	0.02	0.02
41	0.03	0.03	0.07	0.03	0.05	0.05	0.05
42	0.01	0.01	0.02	0.01	0.01	0.01	0.01
43	0.01	0.01	0.03	0.00	0.01	0.01	0.00
47	0.03	0.02	0.06	0.02	0.03	0.02	0.02
48	0.04	0.02	0.09	0.02	0.04	0.03	0.03
49	0.01	0.01	0.03	-0.01	0.01	0.01	0.00
50	0.02	0.02	0.03	0.02	0.03	0.03	0.03
51	0.01	0.00	0.02	0.00	0.01	0.01	0.01
52	0.01	0.01	0.02	0.00	0.01	0.01	0.01

Table 26. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 7 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
19	0.01	0.00	0.02	0.00	0.00	0.00	0.00
20	0.03	0.02	0.04	-0.02	0.04	0.03	-0.03
21	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
22	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
23	0.02	0.02	0.06	0.01	0.01	0.01	-0.01
24	0.01	0.00	0.01	0.00	0.00	0.00	0.00
25	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
26	0.02	0.01	0.04	0.01	0.02	0.01	0.01
27	0.00	0.00	0.01	0.00	0.00	0.00	0.00
28	0.01	0.01	0.03	-0.01	0.01	0.01	-0.01
33	0.02	0.01	0.05	-0.01	0.02	0.02	-0.02
34	0.04	0.04	0.06	0.00	0.04	0.03	0.00
35	0.04	0.02	0.09	-0.02	0.05	0.04	-0.04
36	0.03	0.02	0.07	-0.02	0.04	0.04	-0.04
37	0.03	0.02	0.05	-0.02	0.03	0.03	-0.03
38	0.00	0.00	0.01	0.00	0.00	0.00	0.00
39	0.03	0.02	0.07	0.02	0.04	0.03	0.03
40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
41	0.03	0.02	0.07	0.02	0.04	0.03	0.03
42	0.01	0.01	0.04	-0.01	0.02	0.01	-0.01
43	0.01	0.01	0.02	0.01	0.01	0.01	0.01
44	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
45	0.04	0.03	0.06	0.03	0.03	0.02	0.02
46	0.01	0.01	0.01	0.00	0.01	0.01	0.00
47	0.01	0.01	0.02	0.00	0.01	0.01	0.01
48	0.02	0.02	0.05	-0.01	0.02	0.02	-0.01
51	0.04	0.02	0.08	0.02	0.05	0.04	0.04
52	0.03	0.02	0.07	0.02	0.04	0.03	0.03
53	0.02	0.01	0.04	0.01	0.02	0.02	0.01
54	0.01	0.01	0.02	-0.01	0.02	0.01	-0.01
55	0.00	0.00	0.01	0.00	0.00	0.00	0.00
56	0.03	0.02	0.05	0.01	0.03	0.03	0.03
57	0.01	0.01	0.03	0.01	0.02	0.02	0.02

Table 27. Statistics Comparing IRT Item-Ability Regression Curves, Communication Arts, Grade 8 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
19	0.01	0.01	0.02	0.01	0.01	0.00	0.00
20	0.01	0.01	0.02	0.01	0.01	0.01	0.01
21	0.01	0.01	0.02	0.00	0.01	0.01	0.00
22	0.01	0.00	0.01	0.00	0.00	0.00	0.00
23	0.04	0.02	0.09	-0.02	0.05	0.03	-0.03
24	0.03	0.02	0.05	0.02	0.02	0.01	0.01
25	0.03	0.03	0.06	0.02	0.05	0.05	0.05
26	0.02	0.01	0.06	-0.01	0.01	0.00	0.00
27	0.05	0.03	0.14	-0.03	0.06	0.04	-0.04
28	0.05	0.03	0.11	-0.03	0.05	0.04	-0.04
29	0.02	0.02	0.02	0.02	0.02	0.02	0.02
39	0.03	0.02	0.06	-0.02	0.03	0.02	-0.01
40	0.02	0.01	0.05	0.01	0.02	0.01	0.01
41	0.00	0.00	0.00	0.00	0.00	0.00	0.00
42	0.01	0.01	0.01	0.00	0.01	0.01	0.01
45	0.02	0.01	0.03	0.00	0.02	0.02	-0.01
46	0.04	0.03	0.10	0.03	0.04	0.02	0.02
47	0.00	0.00	0.00	0.00	0.00	0.00	0.00
48	0.01	0.01	0.03	0.01	0.01	0.01	0.00
49	0.01	0.01	0.03	0.00	0.01	0.01	0.00
50	0.02	0.01	0.03	-0.01	0.01	0.01	-0.01
51	0.03	0.02	0.08	0.01	0.03	0.02	0.02
52	0.01	0.01	0.03	0.01	0.02	0.02	0.02
53	0.00	0.00	0.01	0.00	0.00	0.00	0.00
54	0.03	0.02	0.07	0.02	0.05	0.05	0.05
55	0.01	0.01	0.04	-0.01	0.02	0.01	-0.01
56	0.02	0.01	0.06	-0.01	0.03	0.02	-0.02
57	0.01	0.01	0.03	-0.01	0.02	0.01	-0.01

**Table 28. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 3
(Alternate Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.01	0.01	0.01	0.01	0.01	0.01	0.01
25	0.01	0.01	0.02	0.01	0.01	0.01	0.01
26	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
27	0.02	0.01	0.04	0.01	0.03	0.03	0.03
28	0.03	0.02	0.08	-0.02	0.05	0.04	-0.04
29	0.03	0.02	0.06	0.02	0.02	0.01	0.01
30	0.01	0.00	0.01	0.00	0.00	0.00	0.00
31	0.01	0.01	0.03	0.01	0.02	0.02	0.02
32	0.00	0.00	0.01	0.00	0.01	0.01	0.01
33	0.02	0.01	0.03	0.01	0.02	0.02	0.02
34	0.01	0.00	0.01	0.00	0.00	0.00	0.00
35	0.02	0.01	0.05	0.01	0.01	0.01	0.01
36	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
37	0.01	0.01	0.02	0.00	0.01	0.01	0.01
38	0.00	0.00	0.01	0.00	0.00	0.00	0.00
39	0.03	0.02	0.06	-0.02	0.02	0.01	-0.01
40	0.02	0.01	0.03	0.01	0.02	0.01	0.01
41	0.01	0.01	0.02	0.01	0.02	0.01	0.01
42	0.01	0.01	0.01	0.00	0.01	0.01	0.00
43	0.00	0.00	0.01	0.00	0.00	0.00	0.00
44	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
45	0.00	0.00	0.01	0.00	0.00	0.00	0.00
46	0.03	0.02	0.06	-0.02	0.04	0.03	-0.03
47	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
48	0.02	0.02	0.04	-0.02	0.02	0.02	-0.02
49	0.02	0.01	0.03	-0.01	0.01	0.01	-0.01
50	0.02	0.02	0.04	0.02	0.03	0.03	0.03
51	0.01	0.01	0.03	0.00	0.02	0.01	-0.01
52	0.00	0.00	0.01	0.00	0.00	0.00	0.00
53	0.03	0.02	0.07	0.02	0.04	0.04	0.04

Table 29. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 4 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
25	0.01	0.01	0.01	0.01	0.01	0.01	0.01
26	0.03	0.02	0.05	-0.02	0.03	0.03	-0.03
27	0.01	0.01	0.02	0.00	0.01	0.01	0.01
28	0.01	0.01	0.02	0.00	0.01	0.01	0.01
29	0.01	0.00	0.02	0.00	0.01	0.00	0.00
30	0.03	0.02	0.07	-0.02	0.04	0.03	-0.03
31	0.02	0.02	0.04	0.01	0.02	0.02	0.01
32	0.01	0.01	0.03	0.01	0.01	0.01	0.00
33	0.00	0.00	0.01	0.00	0.00	0.00	0.00
34	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
35	0.01	0.01	0.03	0.01	0.02	0.01	0.01
36	0.00	0.00	0.01	0.00	0.00	0.00	0.00
37	0.00	0.00	0.01	0.00	0.00	0.00	0.00
38	0.01	0.01	0.01	0.01	0.01	0.01	0.01
39	0.01	0.01	0.02	0.00	0.01	0.01	0.00
40	0.02	0.01	0.02	0.01	0.02	0.02	0.02
41	0.03	0.02	0.07	0.02	0.03	0.02	0.02
42	0.02	0.01	0.05	-0.01	0.03	0.02	-0.02
43	0.01	0.01	0.02	0.00	0.01	0.01	0.00
44	0.02	0.01	0.04	0.00	0.02	0.02	0.01
45	0.02	0.01	0.04	0.01	0.02	0.01	0.01
46	0.01	0.00	0.01	0.00	0.01	0.01	0.01
47	0.01	0.01	0.02	0.00	0.01	0.01	0.01
48	0.01	0.01	0.01	0.00	0.01	0.01	0.01
49	0.01	0.01	0.03	0.01	0.01	0.01	0.01
50	0.02	0.01	0.03	0.01	0.01	0.01	0.01
51	0.01	0.01	0.02	0.00	0.01	0.01	0.01
52	0.02	0.01	0.04	-0.01	0.03	0.02	-0.02
53	0.01	0.01	0.03	-0.01	0.01	0.01	0.00
54	0.01	0.00	0.03	0.00	0.00	0.00	0.00
55	0.01	0.01	0.02	0.01	0.01	0.01	0.01
56	0.02	0.02	0.04	0.01	0.03	0.02	0.02

**Table 30. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 5
(Alternate Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.01	0.01	0.01	0.00	0.01	0.00	0.00
25	0.01	0.01	0.02	0.01	0.01	0.01	0.01
26	0.04	0.03	0.06	-0.03	0.04	0.03	-0.03
27	0.02	0.01	0.05	-0.01	0.02	0.01	-0.01
28	0.01	0.00	0.01	0.00	0.01	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.02	0.01	0.04	0.01	0.02	0.01	0.01
31	0.01	0.00	0.01	0.00	0.01	0.01	-0.01
32	0.01	0.01	0.02	0.00	0.01	0.01	0.00
33	0.02	0.01	0.04	0.01	0.02	0.01	0.01
34	0.01	0.01	0.03	-0.01	0.02	0.01	-0.01
35	0.01	0.01	0.02	0.00	0.01	0.01	0.00
36	0.00	0.00	0.01	0.00	0.00	0.00	0.00
37	0.02	0.02	0.05	-0.02	0.03	0.03	-0.03
38	0.01	0.01	0.03	0.00	0.02	0.02	-0.01
39	0.01	0.01	0.02	0.01	0.01	0.01	0.01
40	0.02	0.01	0.05	0.01	0.01	0.01	0.01
41	0.01	0.01	0.02	0.01	0.01	0.01	0.01
42	0.01	0.01	0.01	0.00	0.01	0.00	0.00
43	0.01	0.01	0.03	0.00	0.01	0.01	0.00
44	0.03	0.02	0.07	0.02	0.01	0.01	0.01
45	0.00	0.00	0.01	0.00	0.01	0.01	-0.01
46	0.01	0.01	0.03	0.01	0.02	0.02	0.02
47	0.03	0.02	0.06	0.02	0.04	0.04	0.04
48	0.01	0.01	0.02	0.01	0.01	0.01	0.01
49	0.00	0.00	0.01	0.00	0.00	0.00	0.00
50	0.02	0.02	0.04	0.02	0.02	0.01	0.01
51	0.00	0.00	0.01	0.00	0.01	0.01	-0.01
52	0.02	0.01	0.04	-0.01	0.03	0.02	-0.02
53	0.00	0.00	0.01	0.00	0.00	0.00	0.00
54	0.02	0.02	0.03	0.00	0.02	0.02	0.00
55	0.00	0.00	0.01	0.00	0.01	0.00	0.00

**Table 31. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 6
(Alternate Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.01	0.01	0.02	0.01	0.01	0.01	0.01
25	0.01	0.01	0.02	0.00	0.01	0.01	-0.01
26	0.01	0.01	0.01	-0.01	0.01	0.01	-0.01
27	0.02	0.01	0.03	-0.01	0.02	0.02	-0.02
28	0.01	0.01	0.02	0.00	0.01	0.00	0.00
29	0.02	0.02	0.04	-0.02	0.03	0.02	-0.02
30	0.01	0.01	0.02	0.00	0.01	0.01	0.00
31	0.01	0.01	0.02	0.00	0.01	0.01	0.00
32	0.01	0.01	0.03	0.01	0.01	0.01	0.01
33	0.02	0.01	0.03	0.01	0.01	0.01	0.01
34	0.02	0.01	0.03	0.01	0.02	0.01	0.01
35	0.01	0.01	0.02	0.01	0.02	0.01	0.01
36	0.02	0.01	0.04	-0.01	0.03	0.02	-0.02
37	0.02	0.01	0.03	0.01	0.02	0.01	0.01
38	0.03	0.02	0.08	0.02	0.04	0.03	0.03
39	0.01	0.01	0.02	0.00	0.01	0.01	0.01
40	0.01	0.01	0.03	-0.01	0.01	0.01	-0.01
41	0.00	0.00	0.01	0.00	0.00	0.00	0.00
42	0.02	0.01	0.03	0.01	0.02	0.02	0.02
43	0.01	0.01	0.03	0.00	0.02	0.01	-0.01
44	0.00	0.00	0.00	0.00	0.00	0.00	0.00
45	0.02	0.01	0.04	0.01	0.03	0.02	0.02
46	0.00	0.00	0.01	0.00	0.01	0.00	0.00
47	0.01	0.01	0.02	0.01	0.01	0.01	0.01
48	0.01	0.01	0.02	0.00	0.01	0.01	0.01
49	0.01	0.00	0.01	0.00	0.01	0.01	0.00
50	0.02	0.01	0.05	-0.01	0.02	0.02	-0.02
51	0.01	0.01	0.02	0.01	0.02	0.02	0.02
52	0.01	0.01	0.02	0.01	0.01	0.01	0.01
53	0.06	0.05	0.09	-0.05	0.07	0.06	-0.06
54	0.01	0.01	0.01	0.00	0.01	0.01	0.00

**Table 32. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 7
(Alternate Linking)**

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
24	0.02	0.01	0.03	0.01	0.02	0.01	0.01
25	0.02	0.01	0.04	-0.01	0.02	0.02	-0.02
26	0.02	0.01	0.04	-0.01	0.03	0.02	-0.02
27	0.01	0.00	0.02	0.00	0.01	0.01	-0.01
28	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
29	0.02	0.01	0.03	0.00	0.02	0.02	-0.01
30	0.01	0.00	0.02	0.00	0.01	0.01	-0.01
31	0.01	0.01	0.02	-0.01	0.02	0.02	-0.01
32	0.02	0.01	0.03	-0.01	0.03	0.02	-0.02
33	0.01	0.00	0.02	0.00	0.00	0.00	0.00
34	0.01	0.01	0.04	-0.01	0.02	0.01	-0.01
35	0.02	0.02	0.04	0.02	0.02	0.02	0.02
36	0.02	0.01	0.03	0.01	0.02	0.01	0.01
37	0.01	0.00	0.01	0.00	0.00	0.00	0.00
38	0.03	0.03	0.06	-0.02	0.03	0.03	-0.02
39	0.01	0.01	0.02	0.01	0.01	0.01	0.01
40	0.02	0.02	0.05	0.02	0.02	0.01	0.01
41	0.01	0.00	0.01	0.00	0.01	0.01	0.00
42	0.01	0.01	0.02	0.01	0.01	0.01	0.01
43	0.02	0.02	0.05	0.01	0.03	0.02	0.02
44	0.02	0.01	0.03	0.01	0.02	0.01	0.01
45	0.03	0.02	0.07	0.01	0.04	0.04	0.04
46	0.02	0.01	0.05	0.01	0.03	0.02	0.02
47	0.02	0.02	0.05	-0.01	0.03	0.03	-0.02
48	0.03	0.02	0.05	-0.01	0.03	0.02	-0.02
49	0.02	0.01	0.04	0.01	0.03	0.03	0.03
50	0.01	0.01	0.02	0.01	0.01	0.01	0.01
51	0.01	0.01	0.03	0.00	0.02	0.02	0.00
52	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
53	0.02	0.01	0.03	0.00	0.02	0.02	0.00
54	0.01	0.01	0.04	-0.01	0.02	0.02	-0.02
55	0.01	0.01	0.03	0.01	0.02	0.01	0.01

Table 33. Statistics Comparing IRT Item-Ability Regression Curves, Mathematics, Grade 8 (Alternate Linking)

Anchor Item Position	UnWtd RMSD	UnWtd Mean Abs Difference	UnWtd Max	UnWtd Mean	Wtd RMSD	Wtd Mean Abs Difference	Wtd Mean
25	0.00	0.00	0.01	0.00	0.00	0.00	0.00
26	0.03	0.02	0.05	-0.02	0.04	0.03	-0.03
27	0.01	0.01	0.03	-0.01	0.01	0.01	-0.01
28	0.01	0.01	0.02	0.01	0.01	0.01	0.01
29	0.02	0.01	0.03	0.01	0.02	0.01	0.01
30	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
31	0.01	0.01	0.02	-0.01	0.01	0.01	-0.01
32	0.00	0.00	0.01	0.00	0.01	0.00	0.00
33	0.02	0.02	0.04	0.00	0.02	0.02	0.00
34	0.01	0.01	0.03	0.01	0.02	0.02	0.02
35	0.03	0.03	0.05	0.01	0.02	0.02	0.00
36	0.01	0.01	0.03	0.01	0.01	0.01	0.01
37	0.02	0.01	0.03	0.01	0.02	0.02	0.02
38	0.02	0.01	0.04	-0.01	0.03	0.03	-0.03
39	0.02	0.01	0.04	0.01	0.02	0.02	0.02
40	0.02	0.01	0.03	0.00	0.01	0.01	-0.01
41	0.02	0.01	0.04	-0.01	0.01	0.01	0.00
42	0.01	0.01	0.02	0.01	0.02	0.01	0.01
43	0.01	0.01	0.03	0.00	0.02	0.02	0.00
44	0.02	0.02	0.04	0.01	0.02	0.02	0.01
45	0.01	0.01	0.02	-0.01	0.01	0.00	0.00
46	0.03	0.02	0.05	0.02	0.03	0.03	0.03
47	0.02	0.01	0.03	0.01	0.02	0.02	0.02
48	0.00	0.00	0.00	0.00	0.00	0.00	0.00
49	0.01	0.01	0.03	-0.01	0.02	0.02	-0.02
50	0.01	0.01	0.03	-0.01	0.02	0.01	-0.01
51	0.02	0.02	0.04	-0.02	0.03	0.03	-0.03
52	0.01	0.01	0.03	0.01	0.02	0.01	0.01
53	0.03	0.02	0.05	0.01	0.02	0.02	-0.01
54	0.03	0.02	0.06	-0.01	0.03	0.02	-0.01
55	0.02	0.01	0.05	0.01	0.02	0.01	0.01

Figure 1. Communication Arts, Grade 3 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

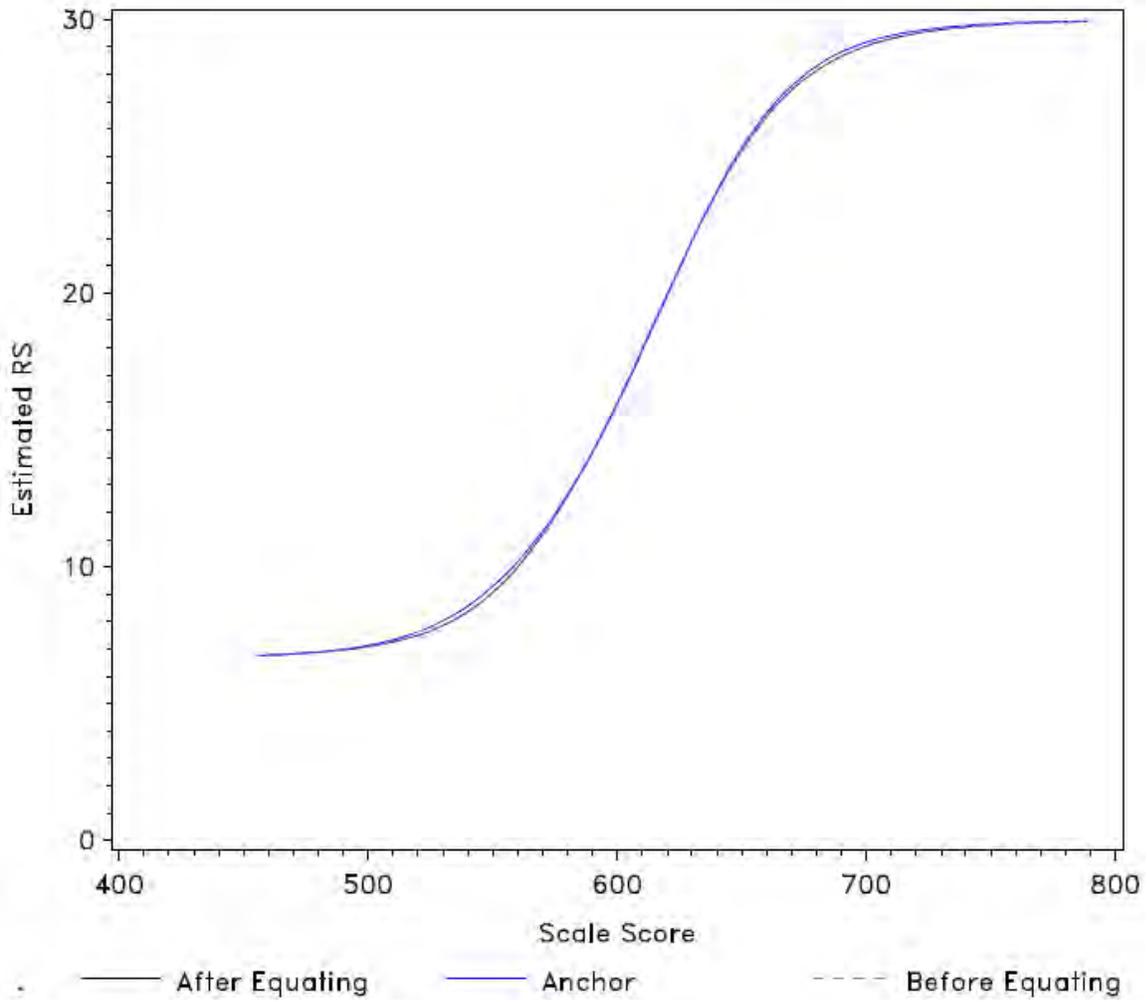


Figure 2. Communication Arts, Grade 4 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

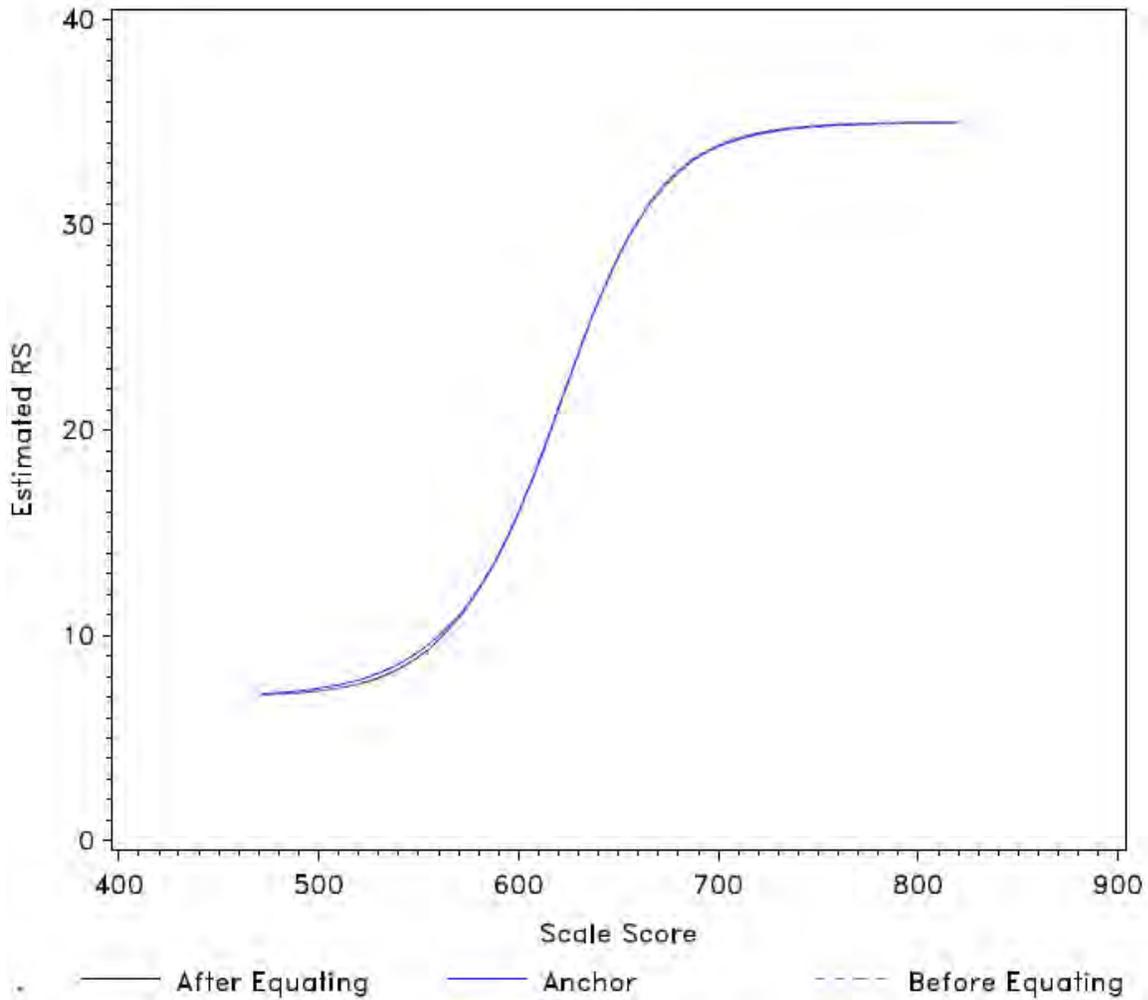


Figure 3. Communication Arts, Grade 5 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

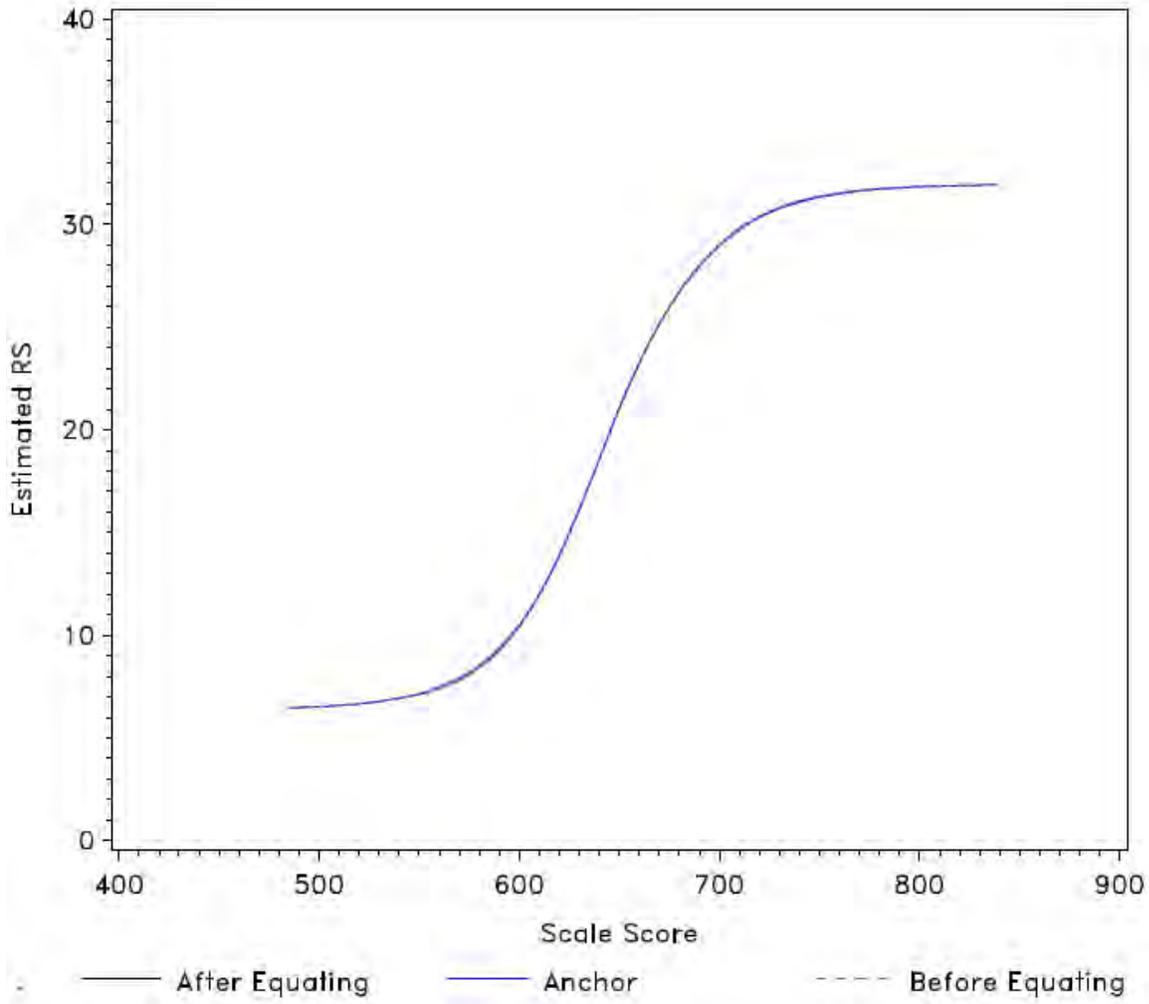


Figure 4. Communication Arts, Grade 6 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

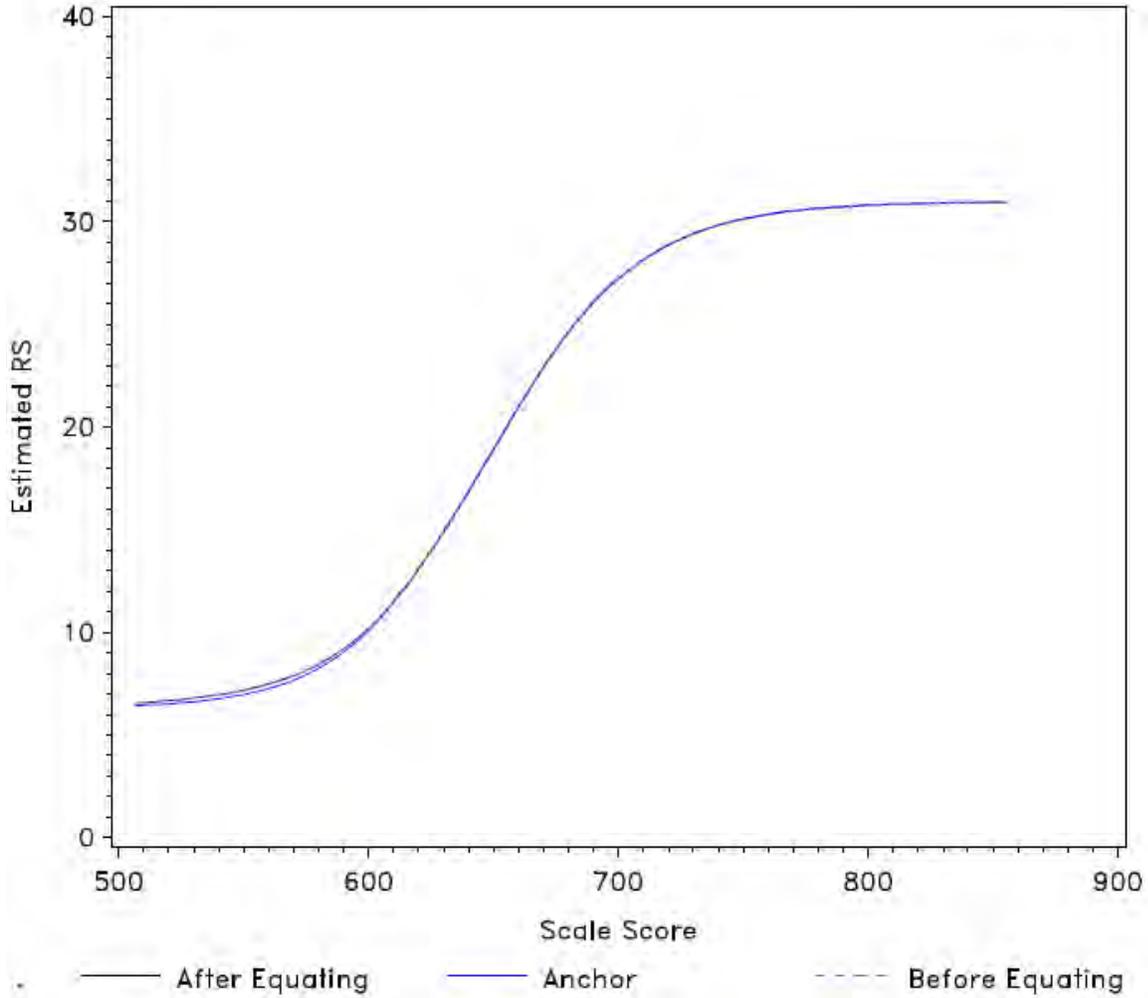


Figure 5. Communication Arts, Grade 7 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

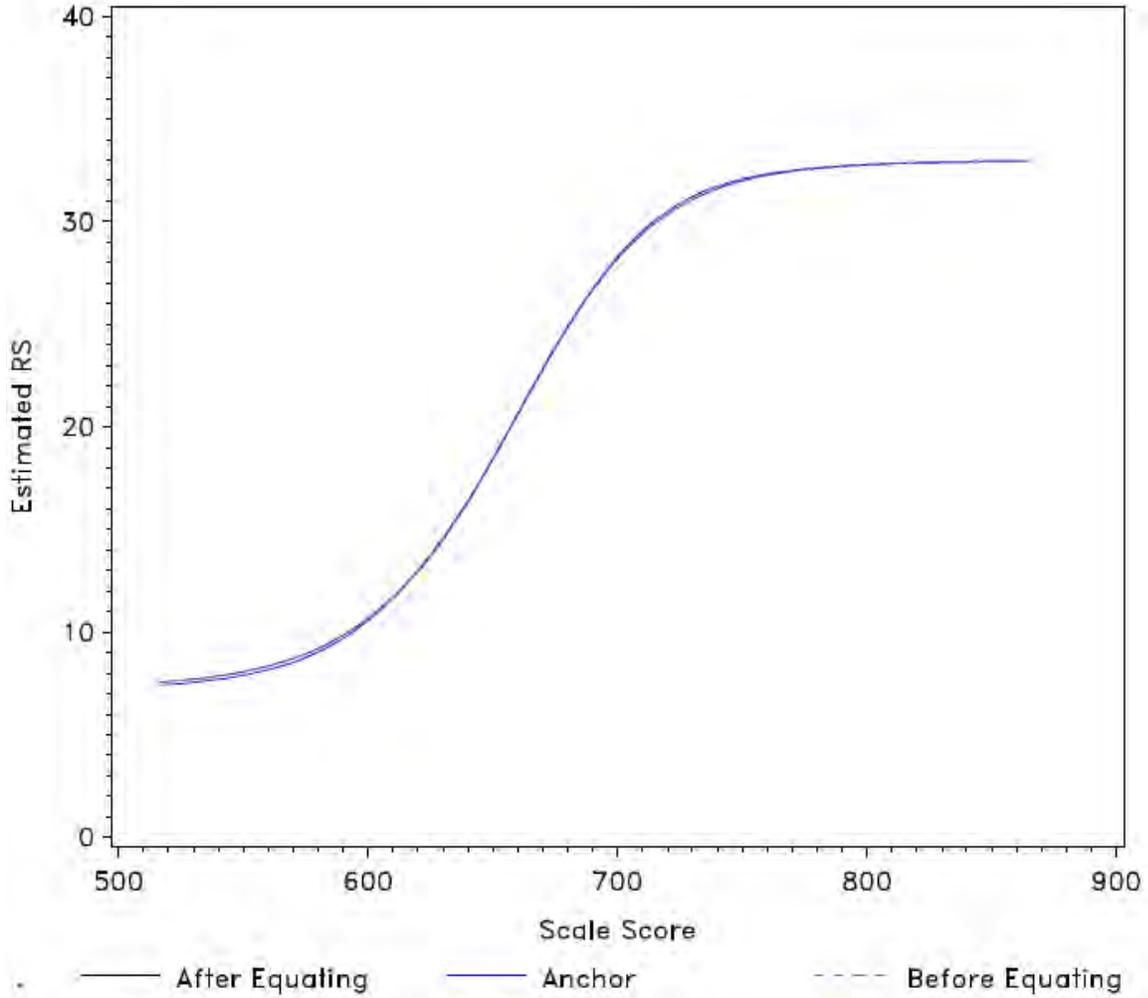


Figure 6. Communication Arts, Grade 8 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

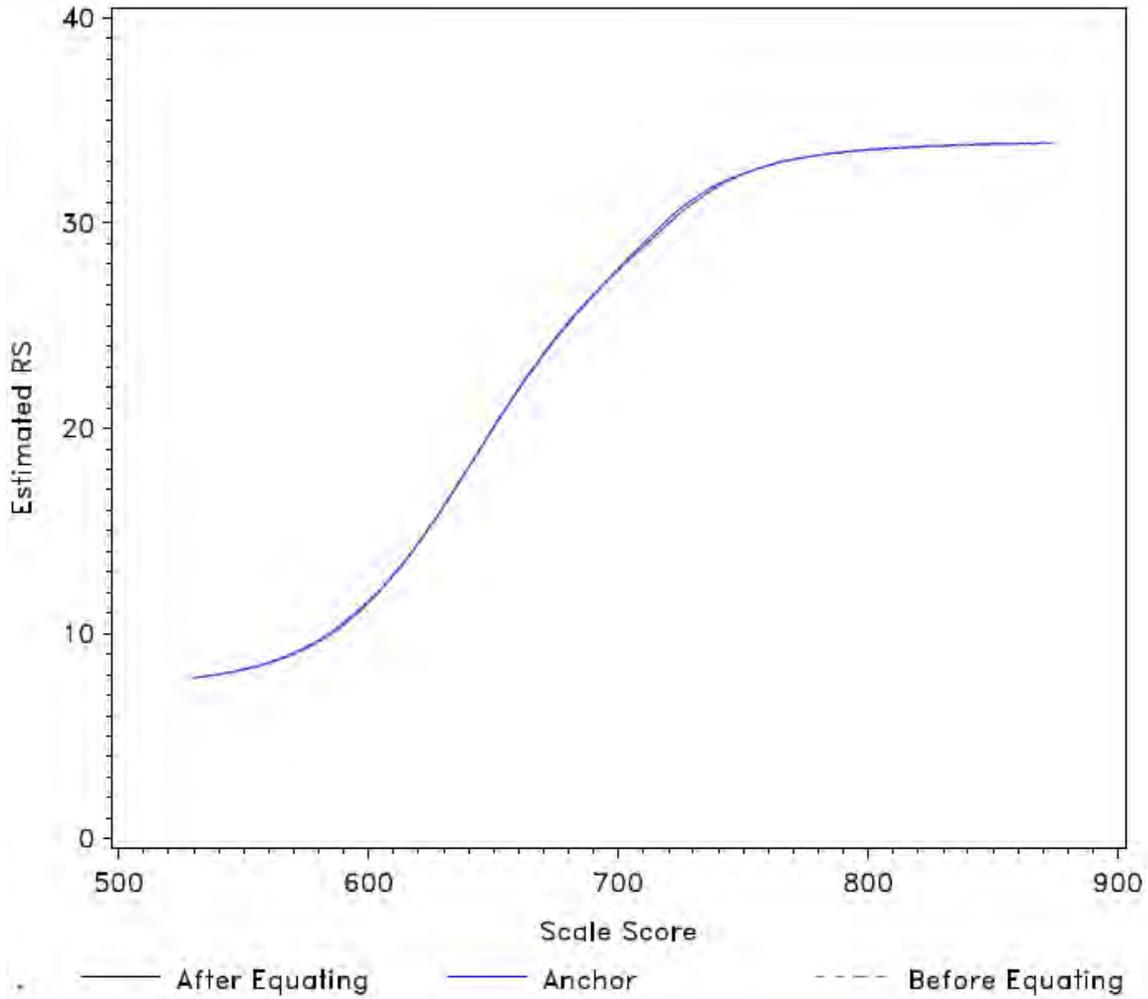


Figure 7. Communication Arts, Grade 11 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

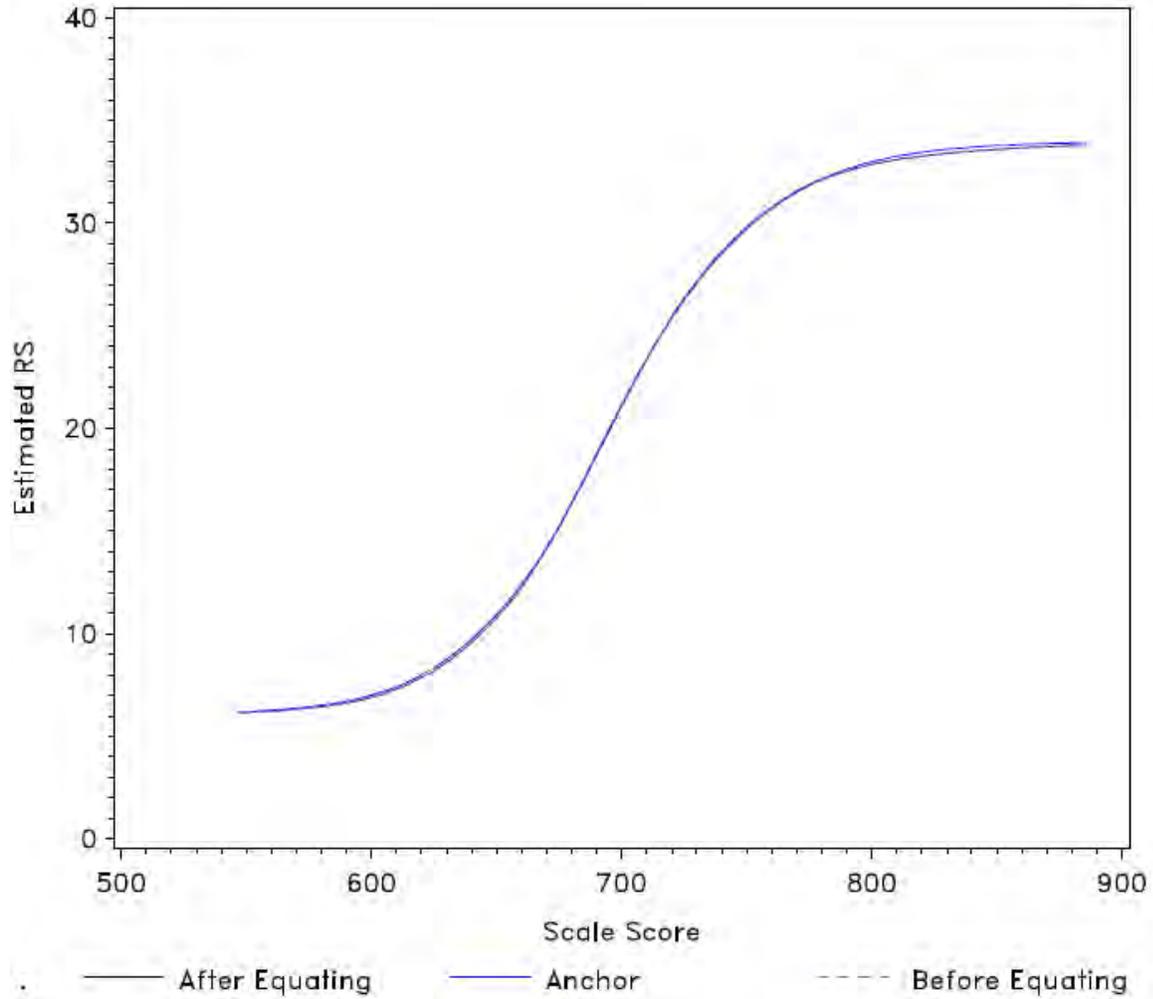


Figure 8. Mathematics, Grade 3 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

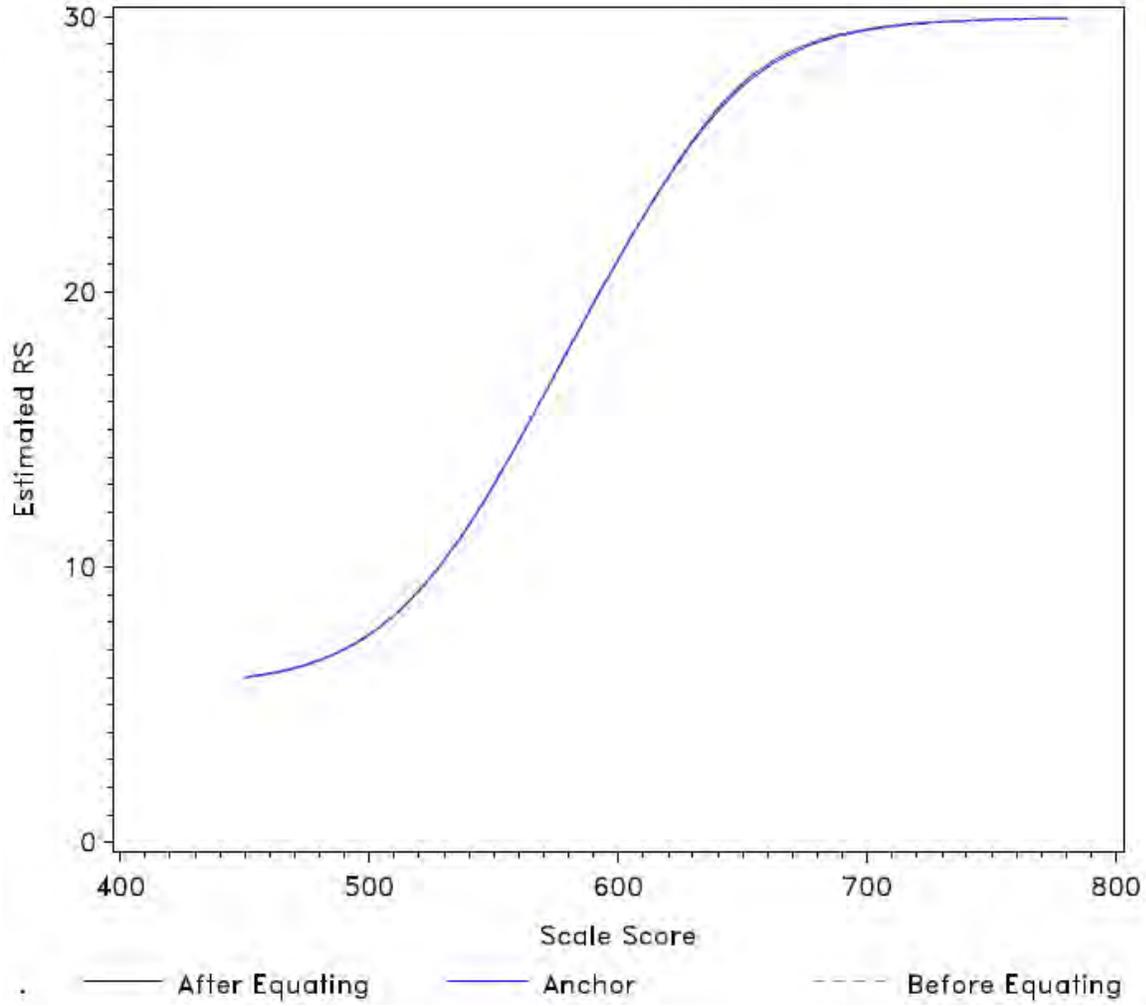


Figure 9. Mathematics, Grade 3 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking with Anchors Removed)

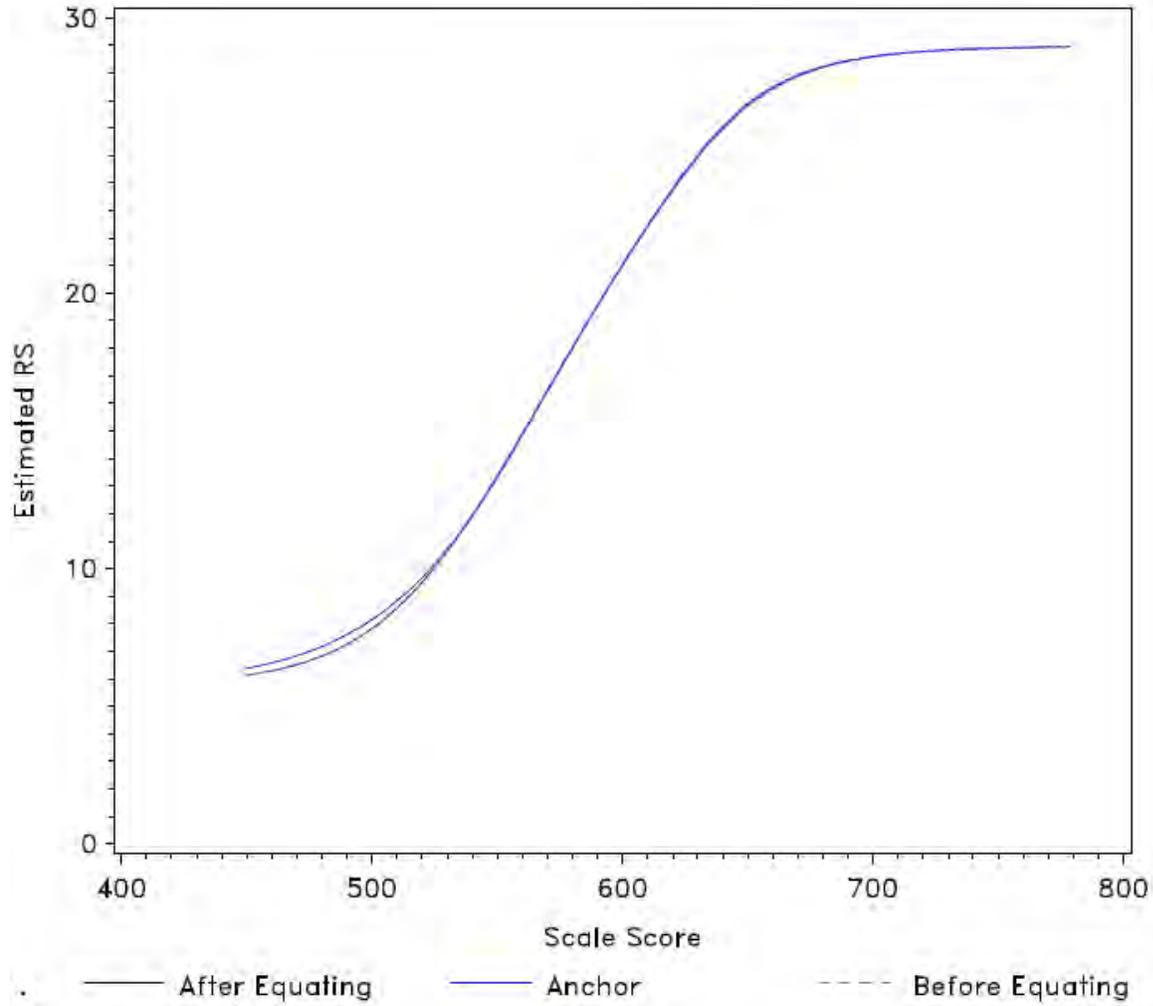


Figure 10. Mathematics, Grade 4 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

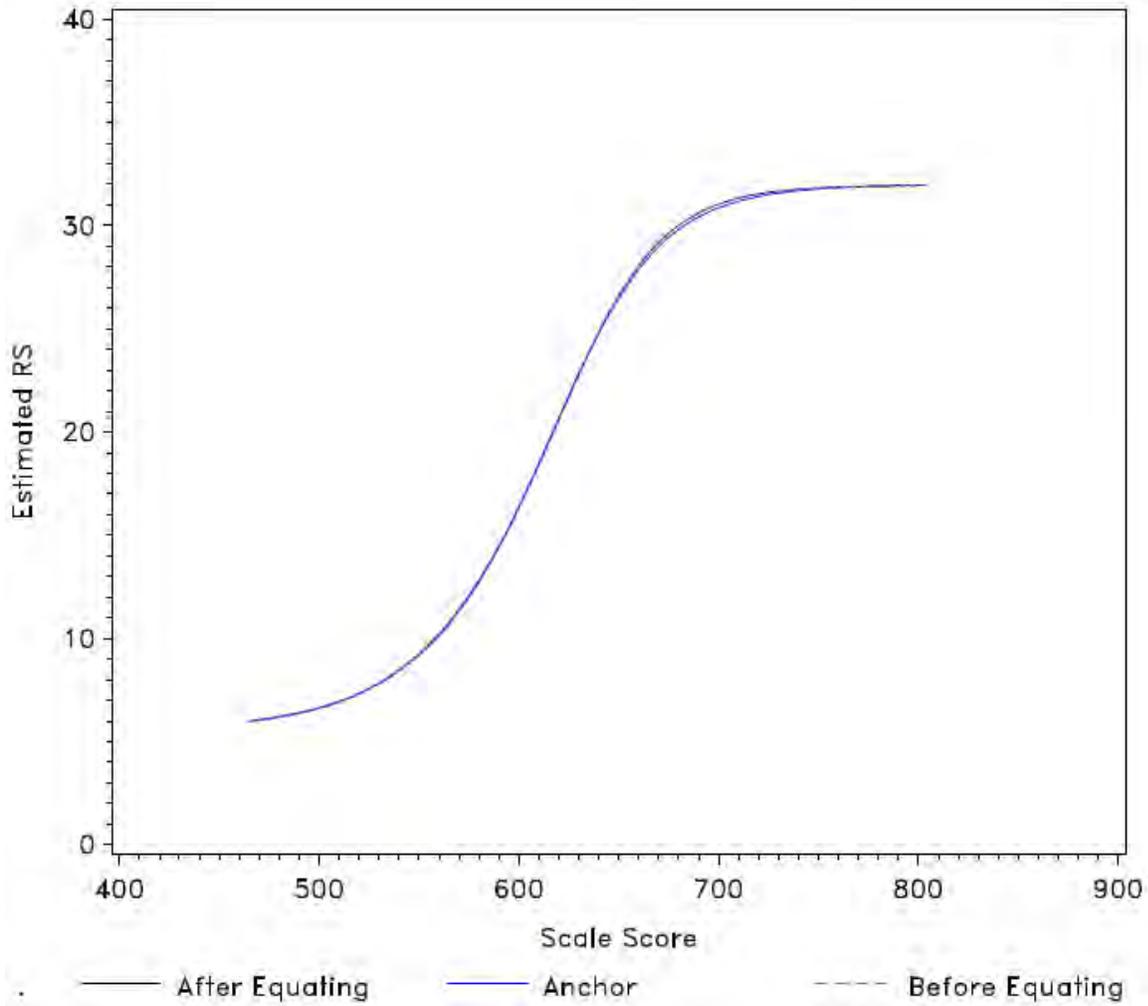


Figure 11. Mathematics, Grade 4 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking with Anchors Removed)

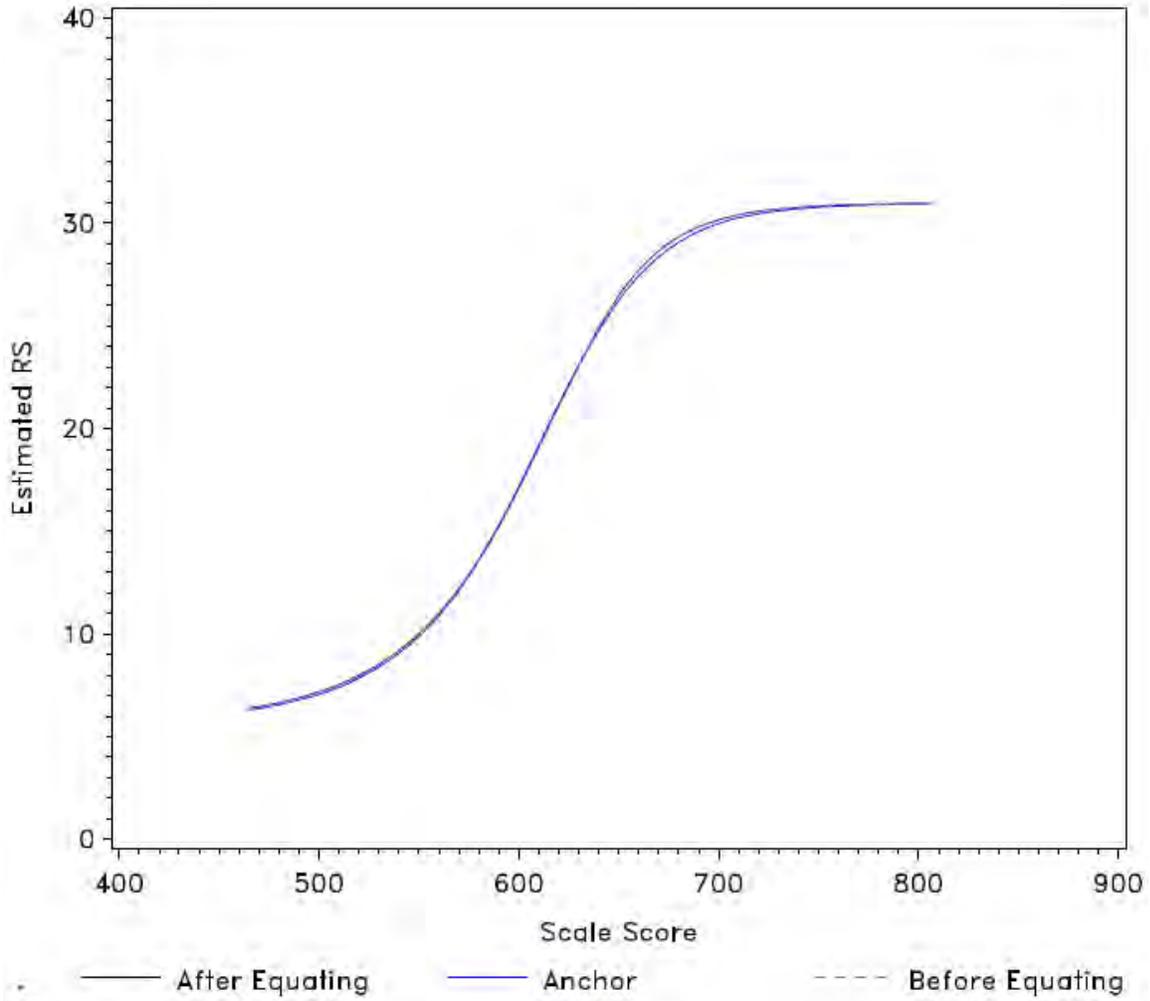


Figure 12. Mathematics, Grade 5 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

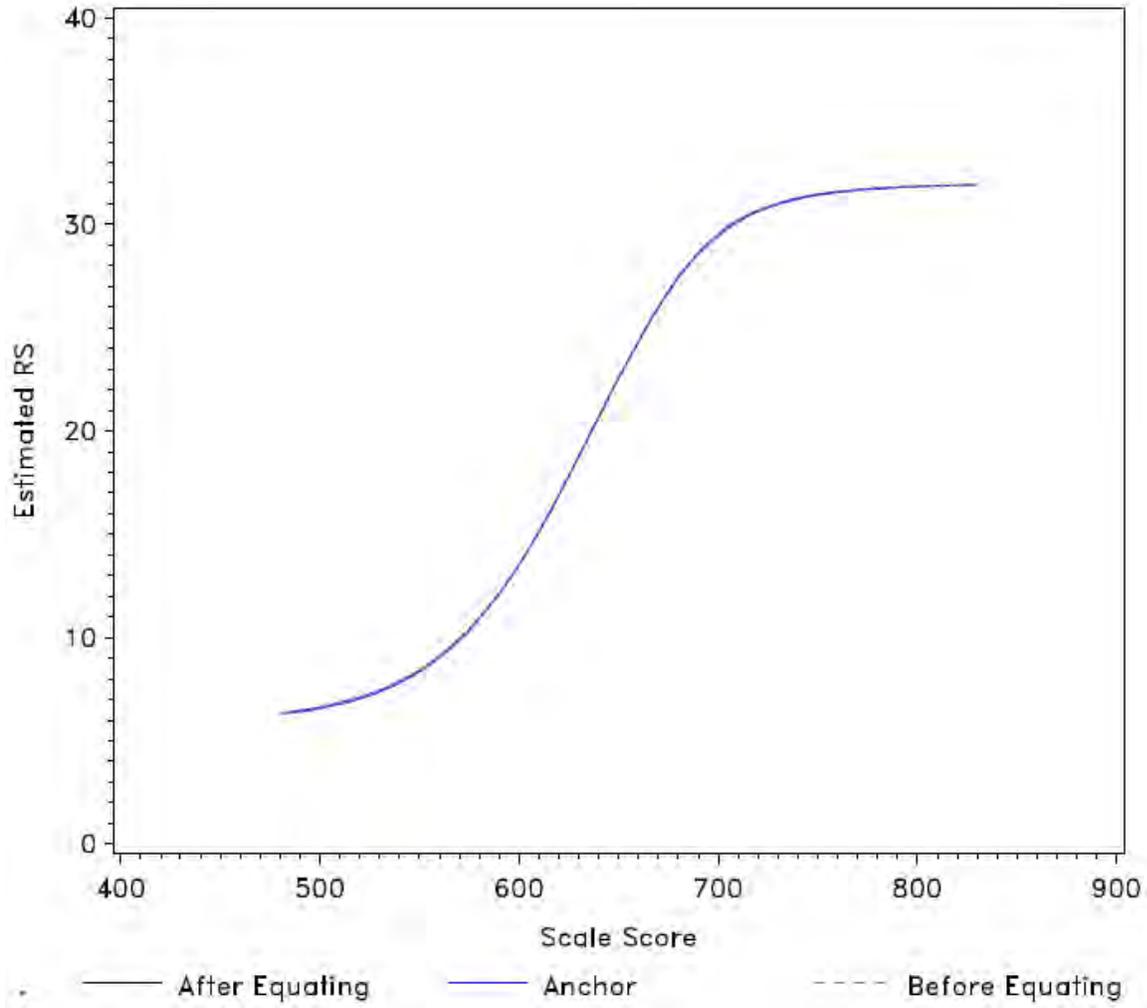


Figure 13. Mathematics, Grade 5 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking with Anchors Removed)

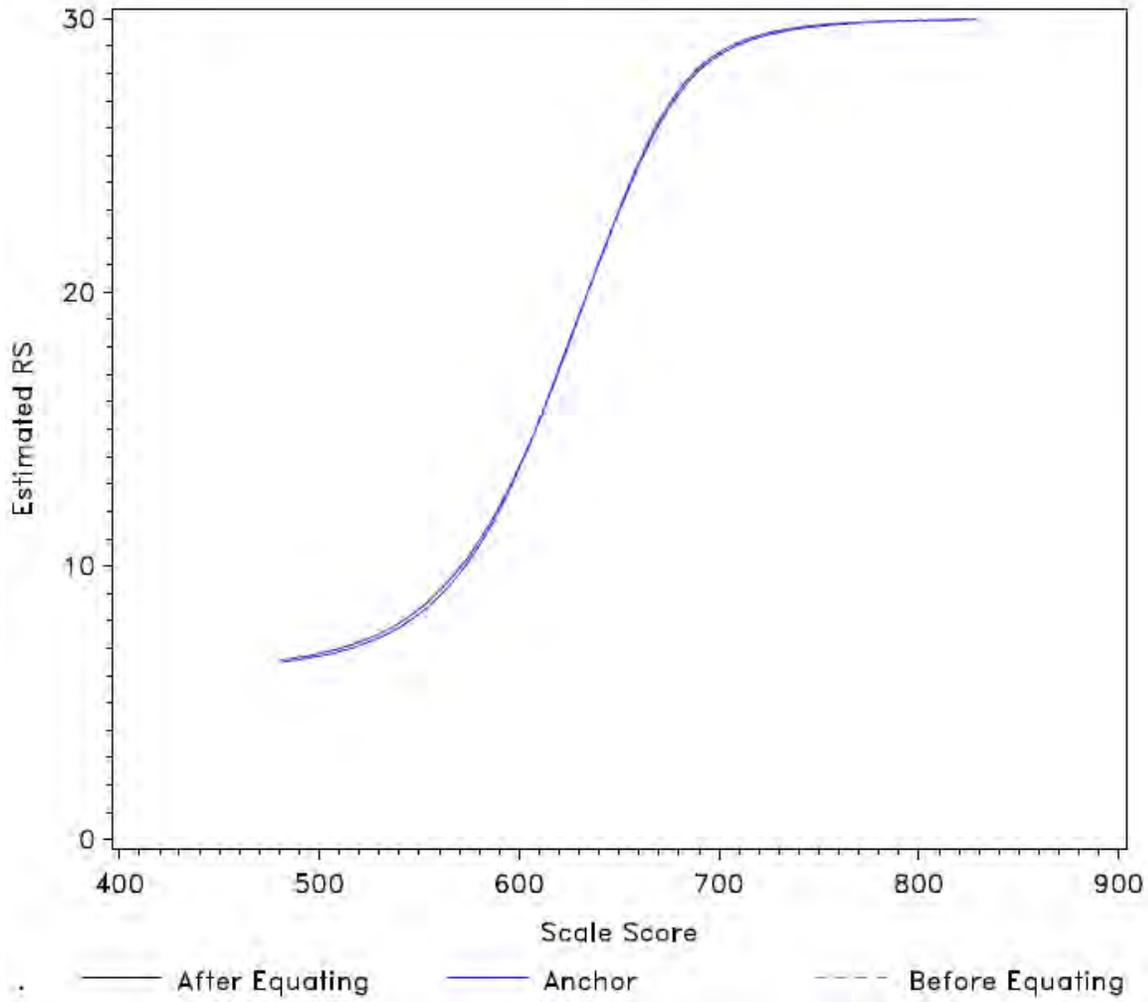


Figure 14. Mathematics, Grade 6 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

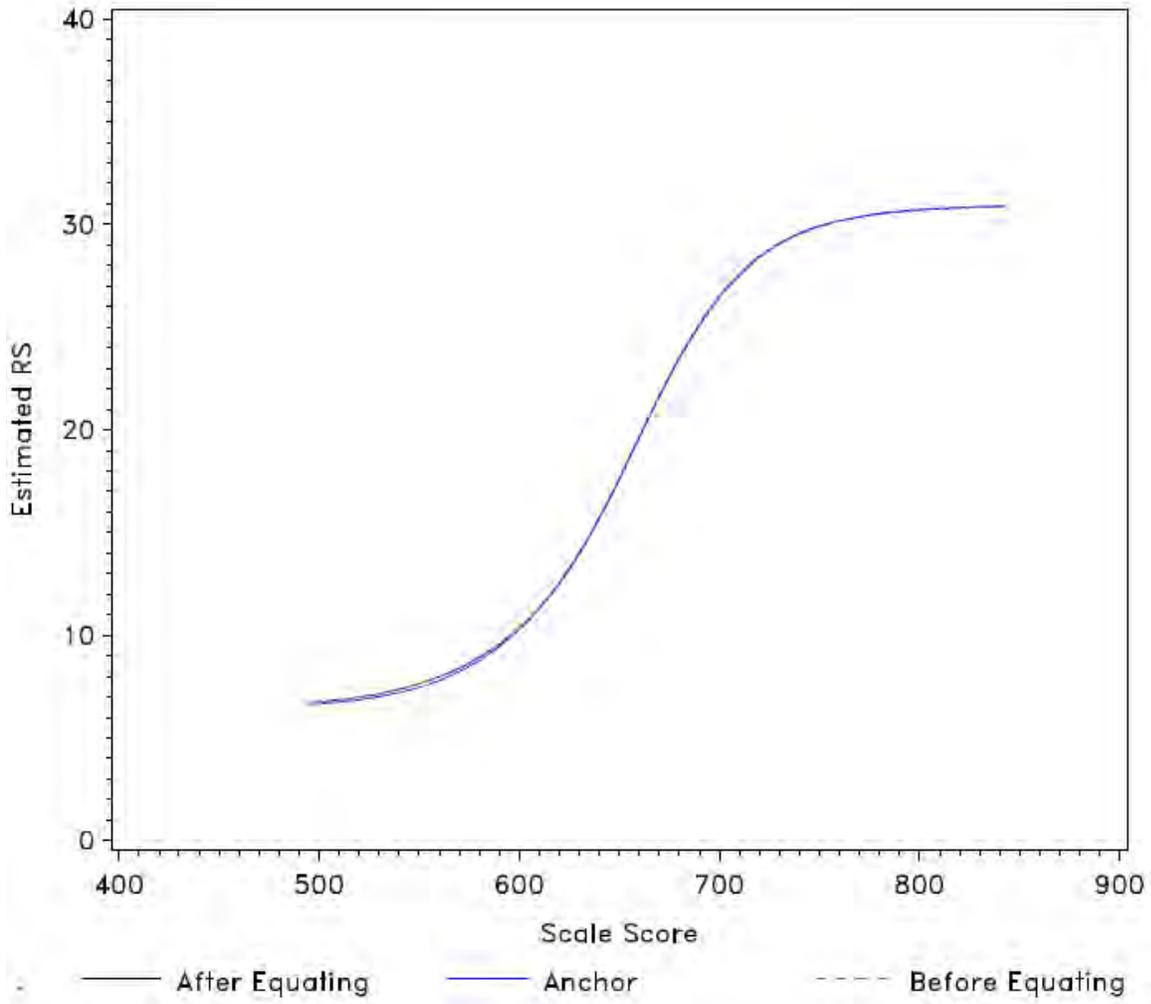


Figure 15. Mathematics, Grade 7 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

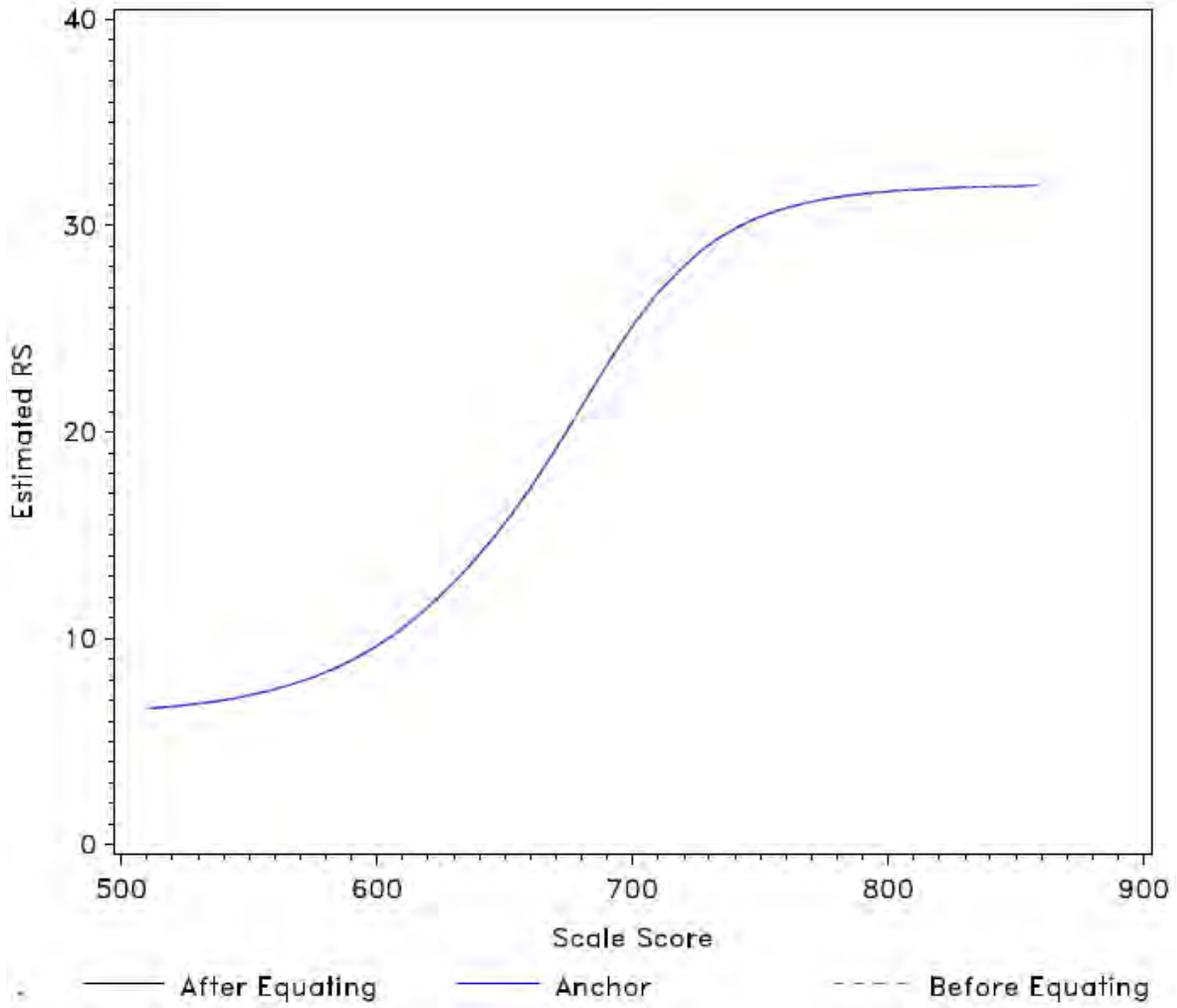


Figure 16. Mathematics, Grade 7 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking with Anchors Removed)

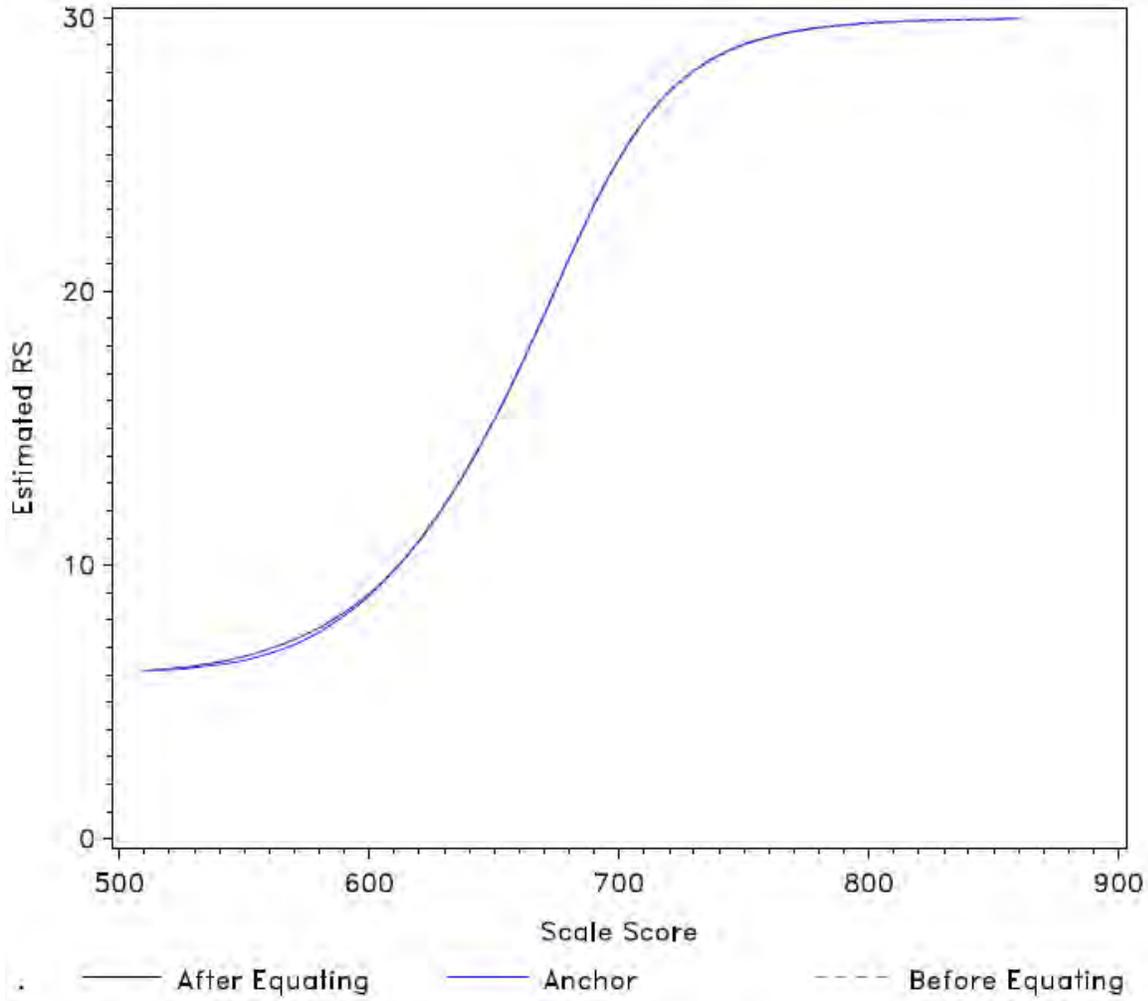


Figure 17. Mathematics, Grade 8 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

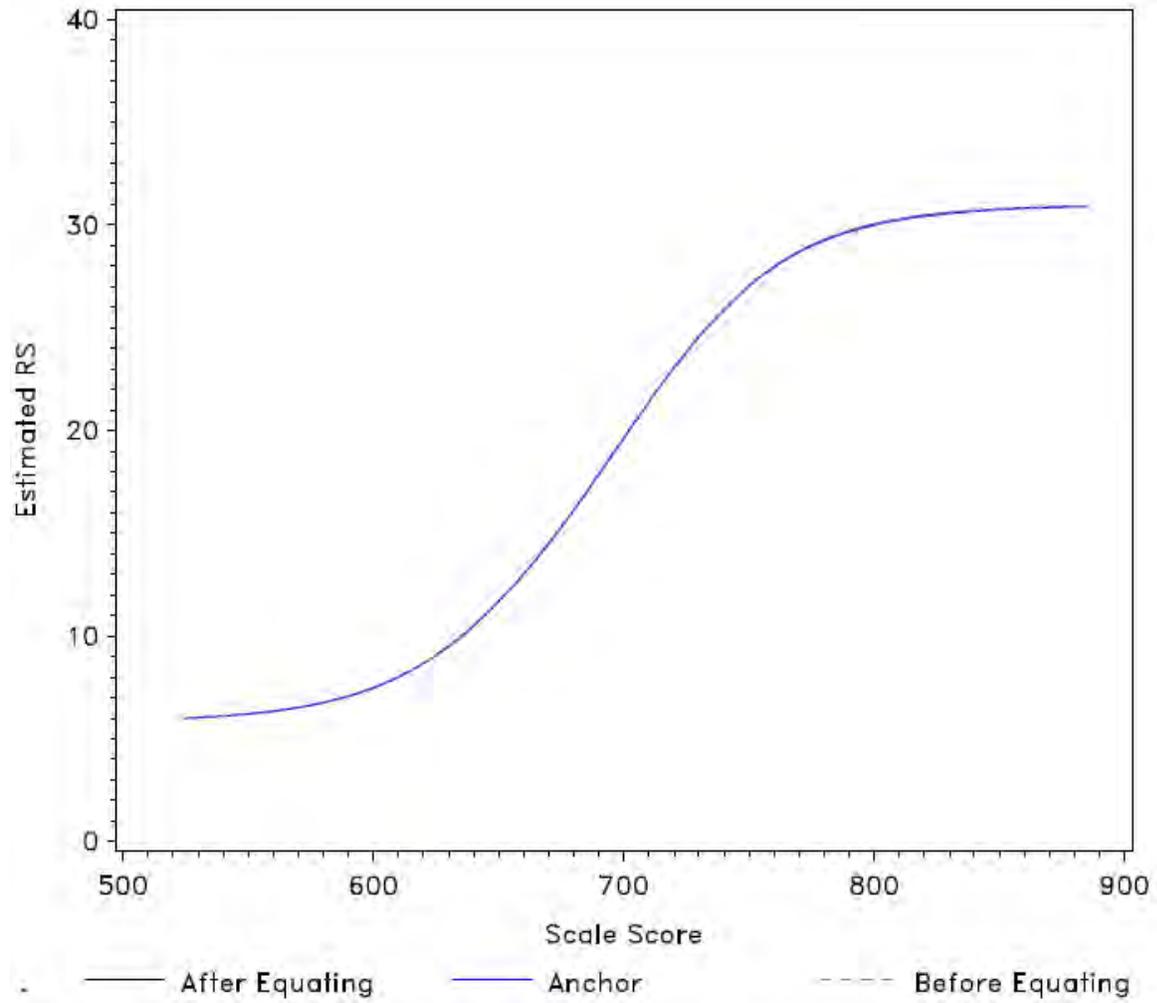


Figure 18. Mathematics, Grade 8 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking with Anchors Removed)

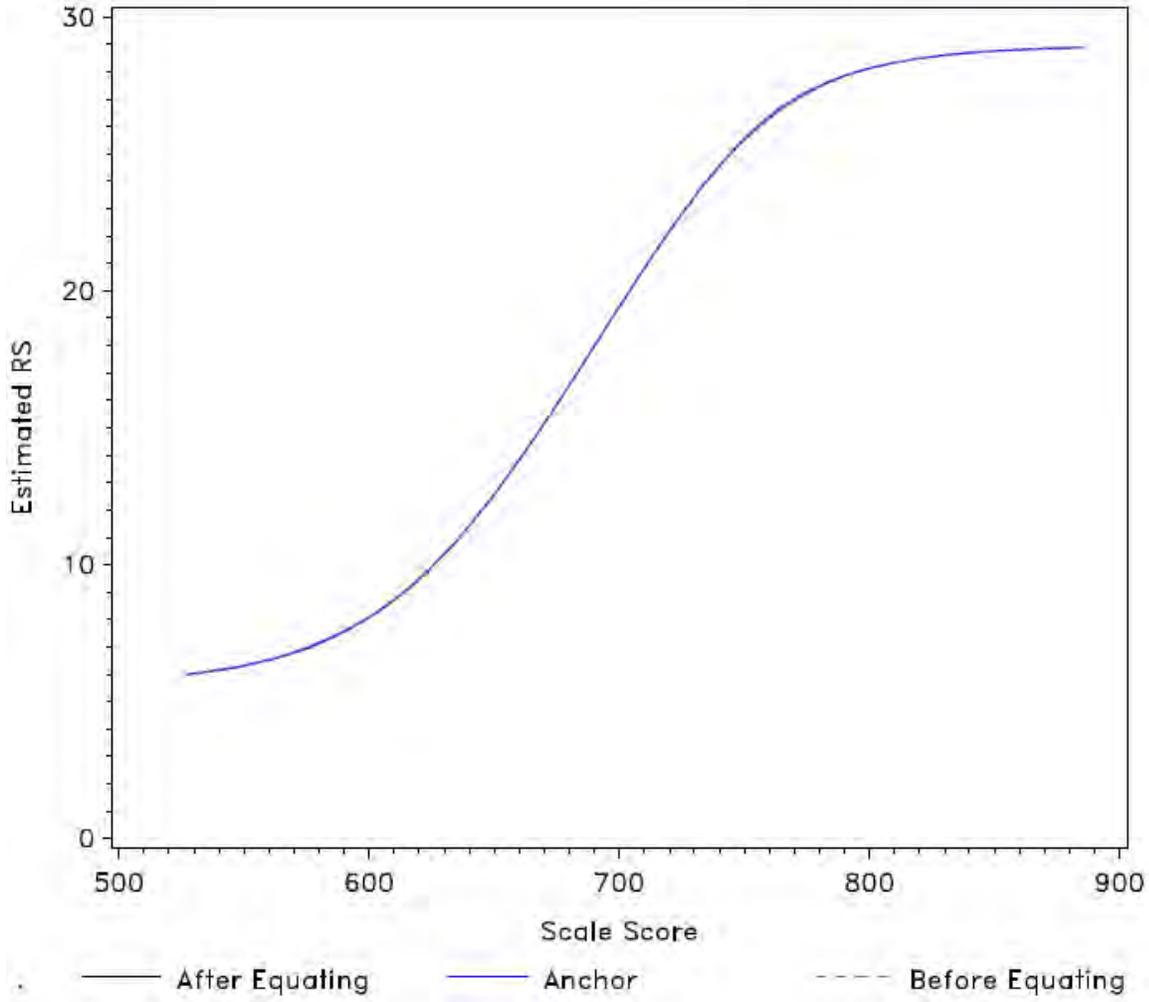


Figure 19. Mathematics, Grade 10 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

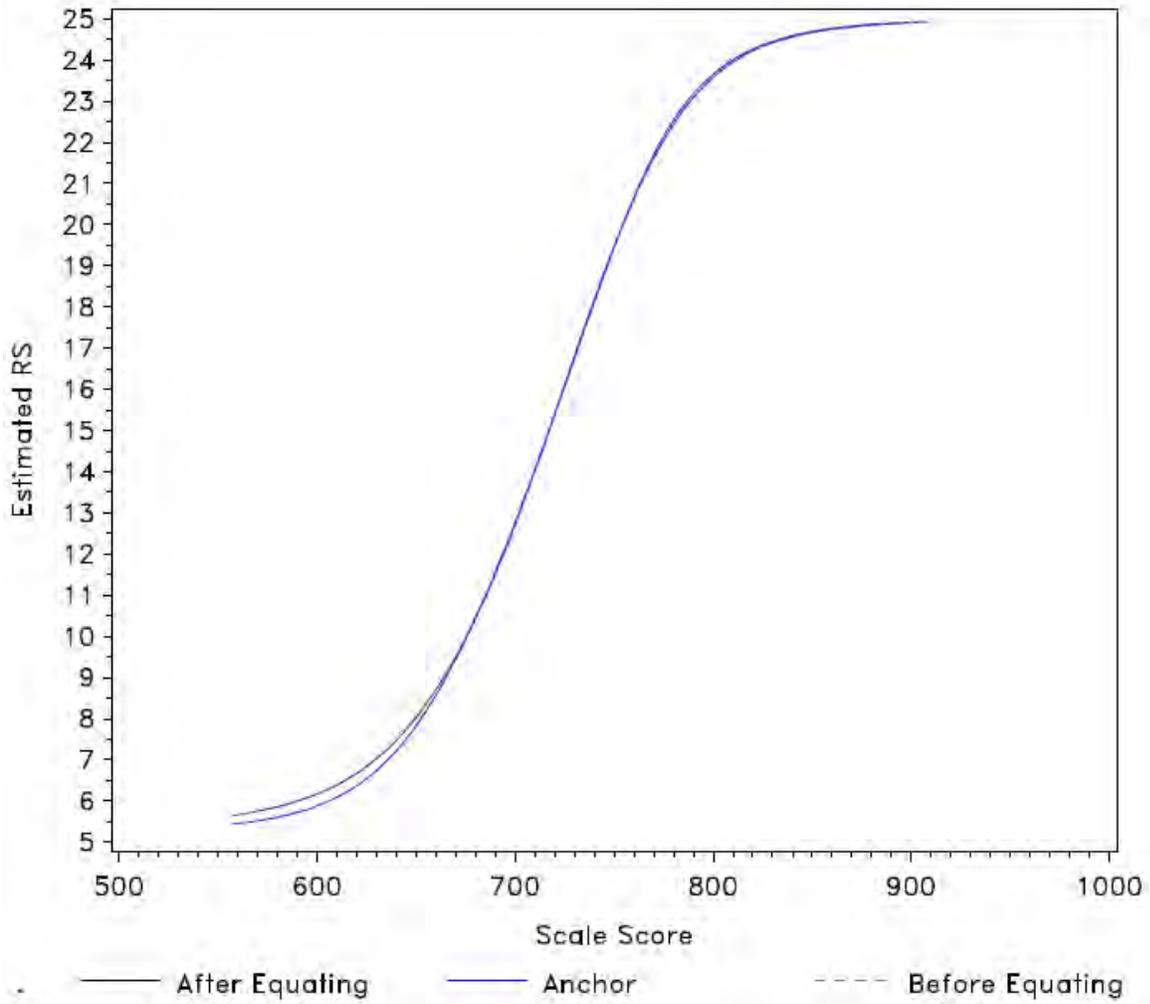


Figure 20. Mathematics, Grade 10 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking with Anchors Removed)

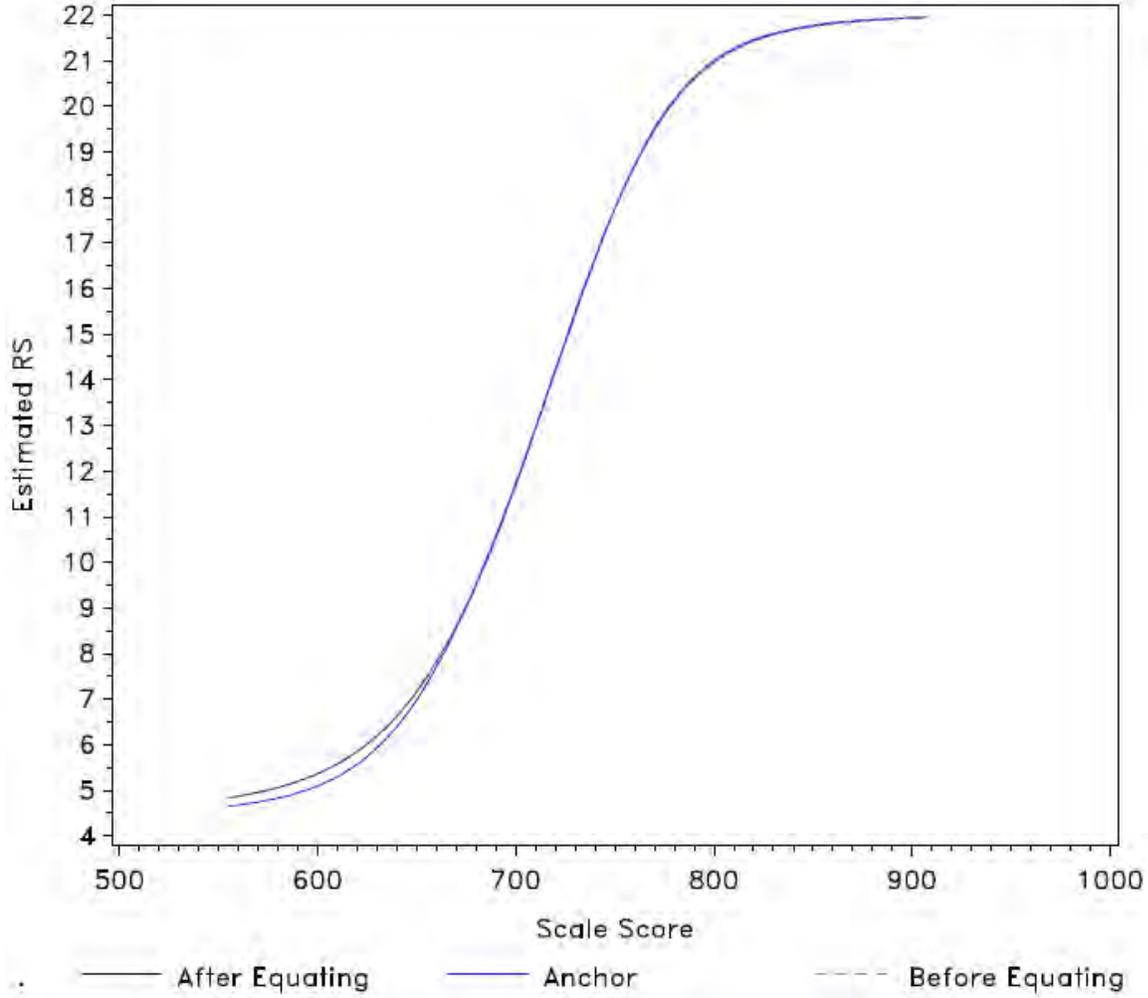


Figure 21. Science, Grade 5 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

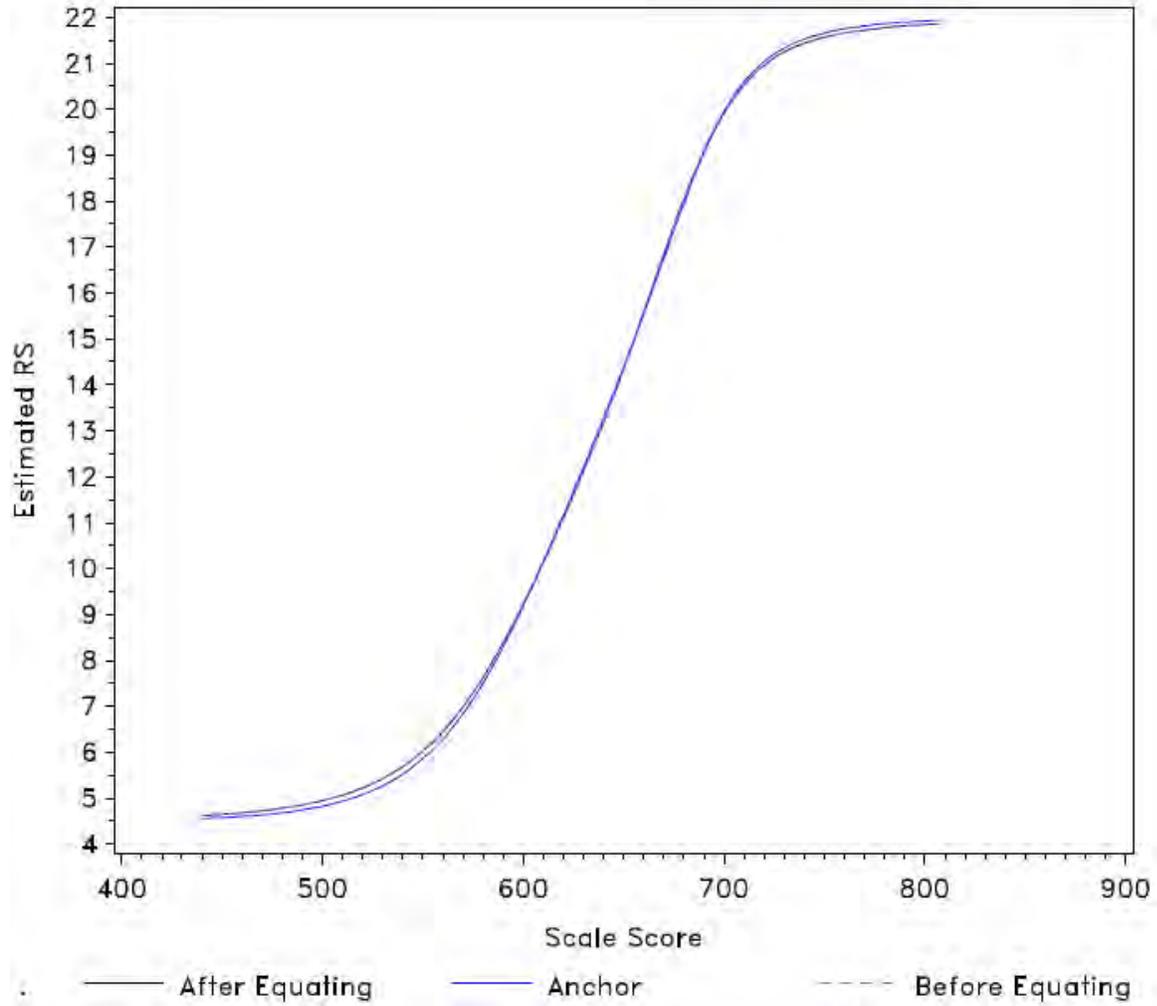


Figure 22. Science, Grade 8 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

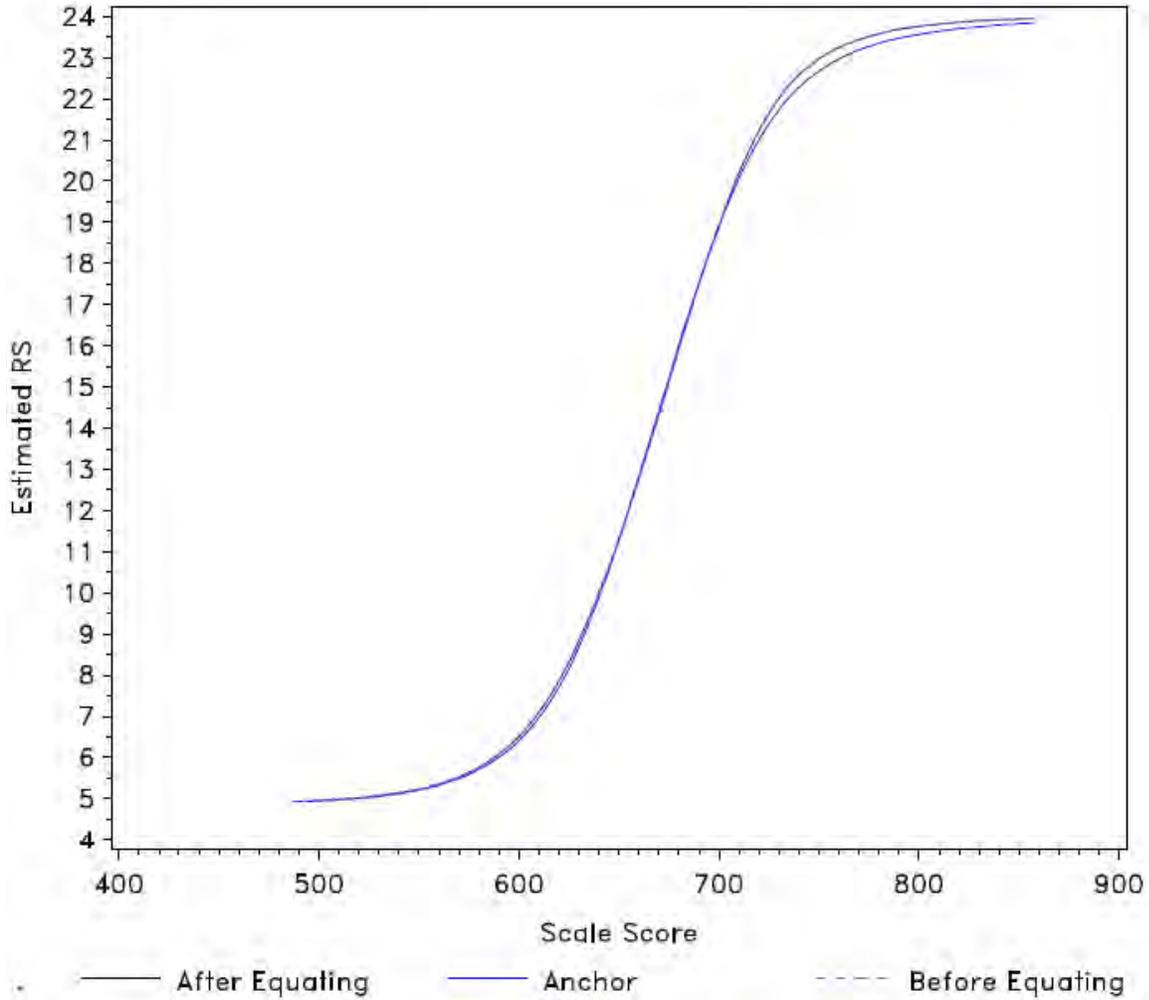


Figure 23. Science, Grade 11 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Original Linking)

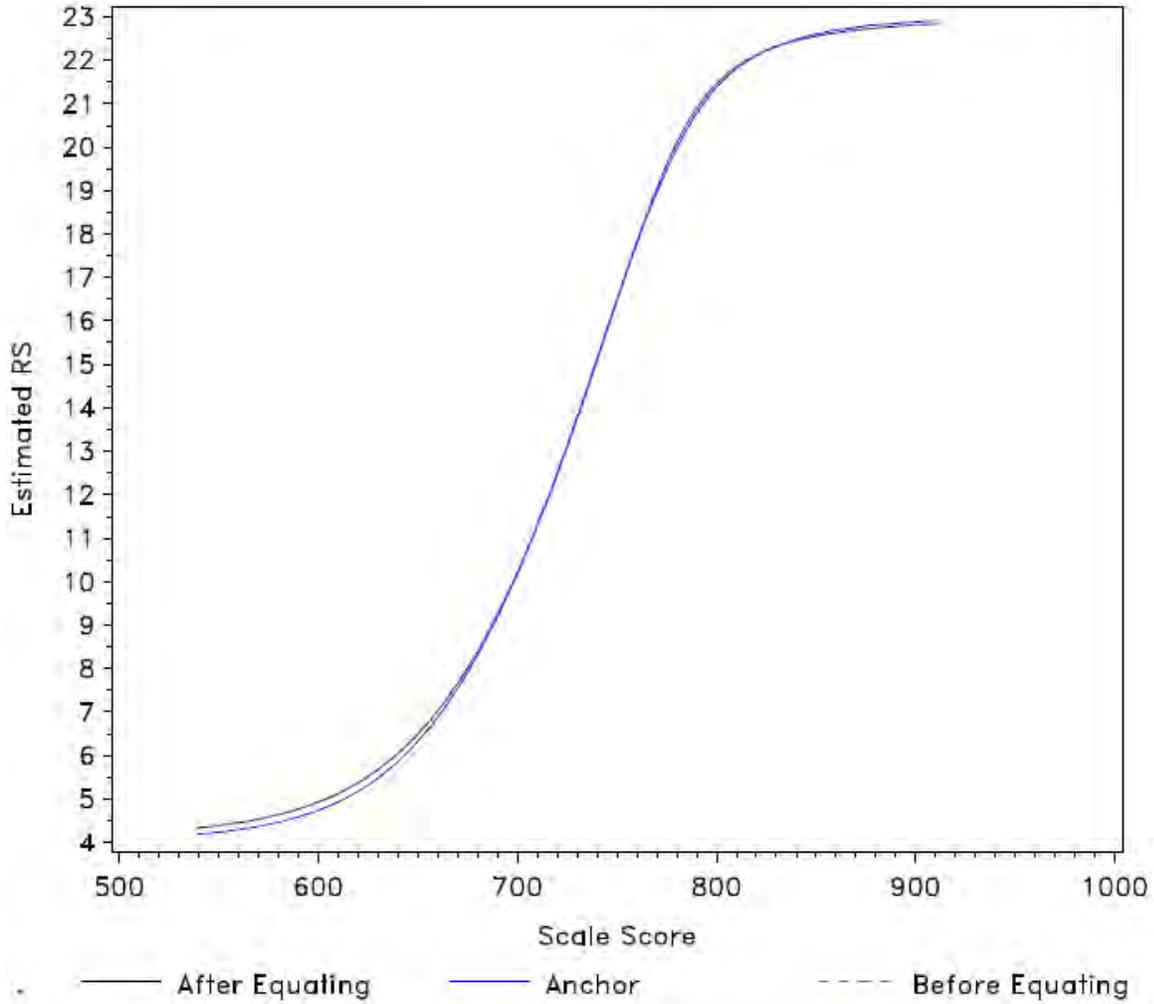


Figure 24. Communication Arts, Grade 3 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

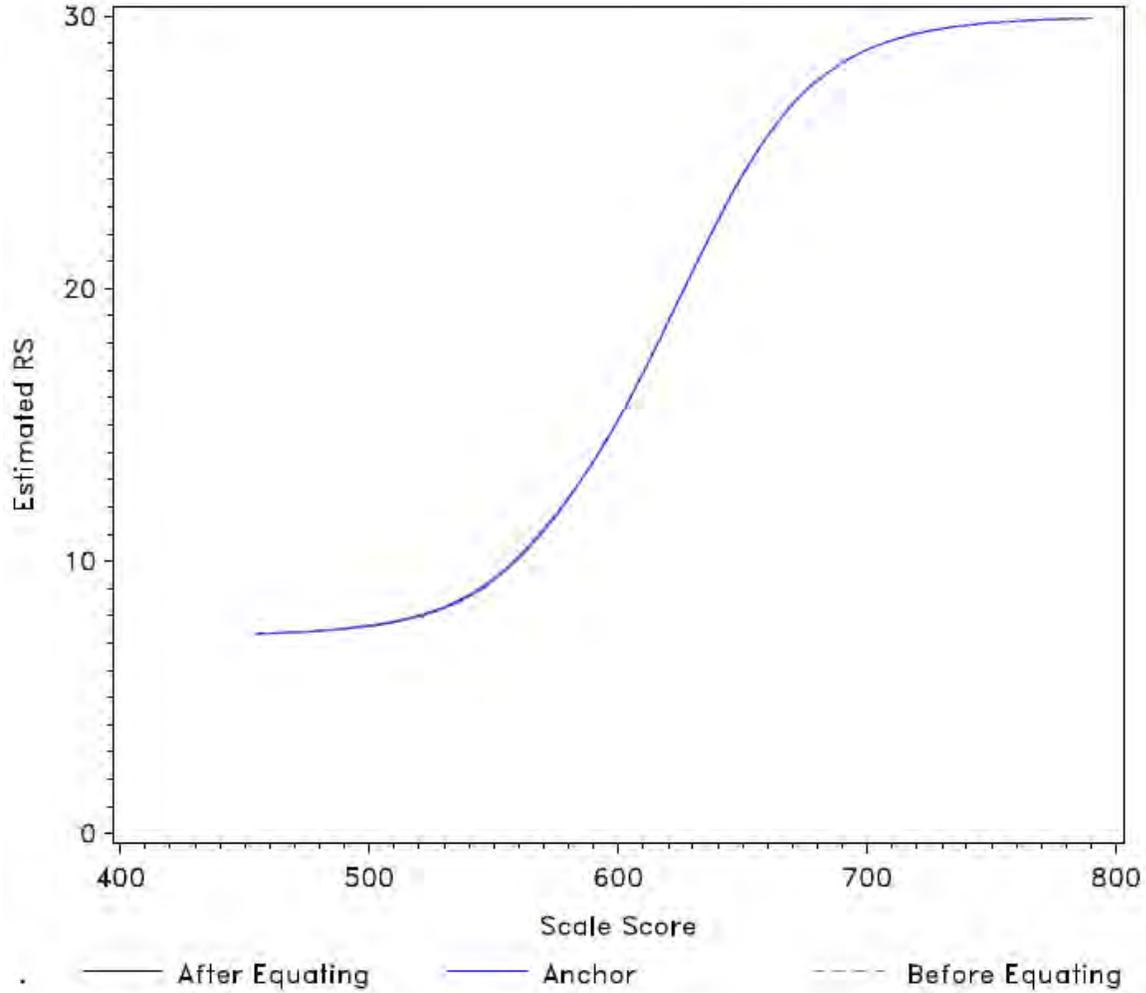


Figure 25. Communication Arts, Grade 4 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

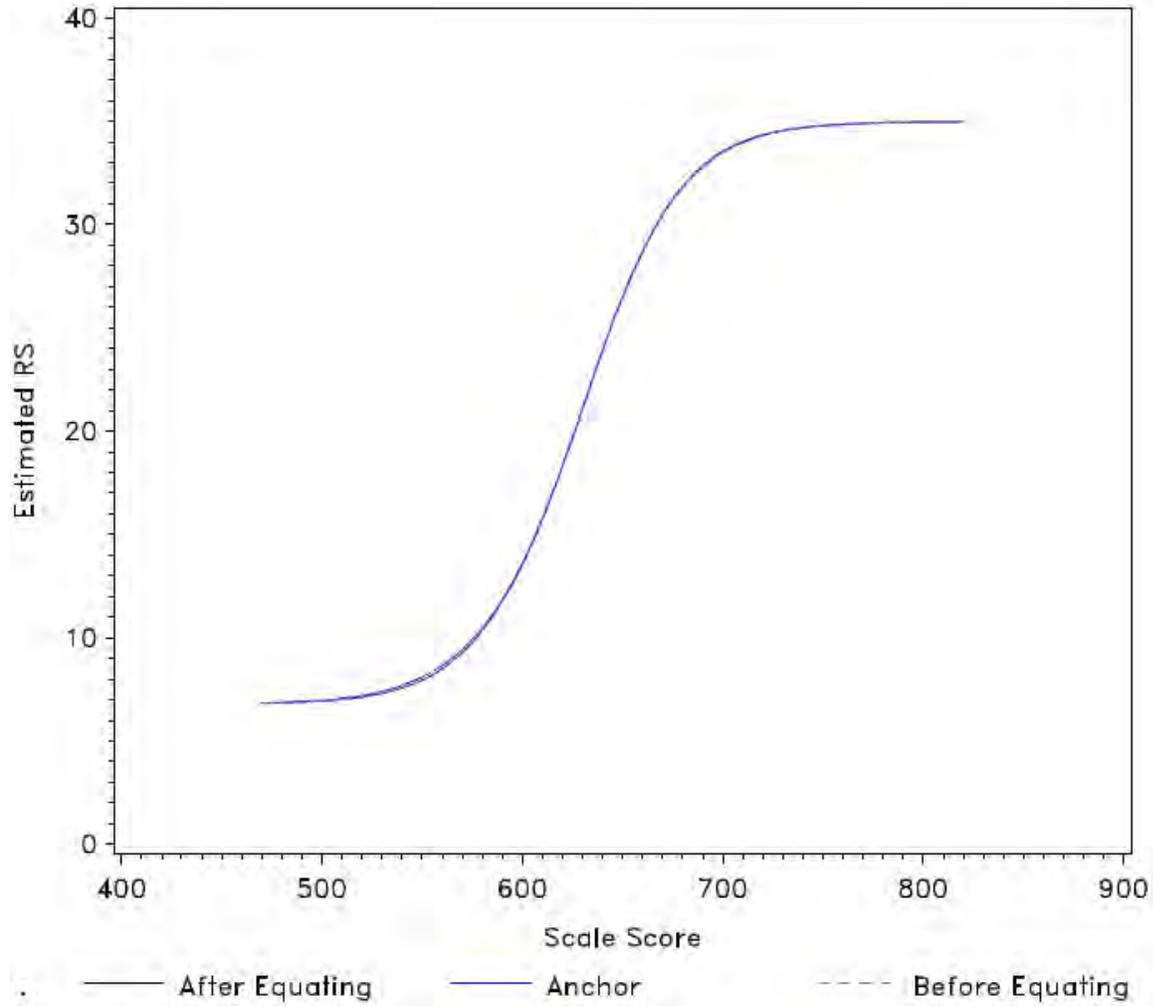


Figure 26. Communication Arts, Grade 5 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

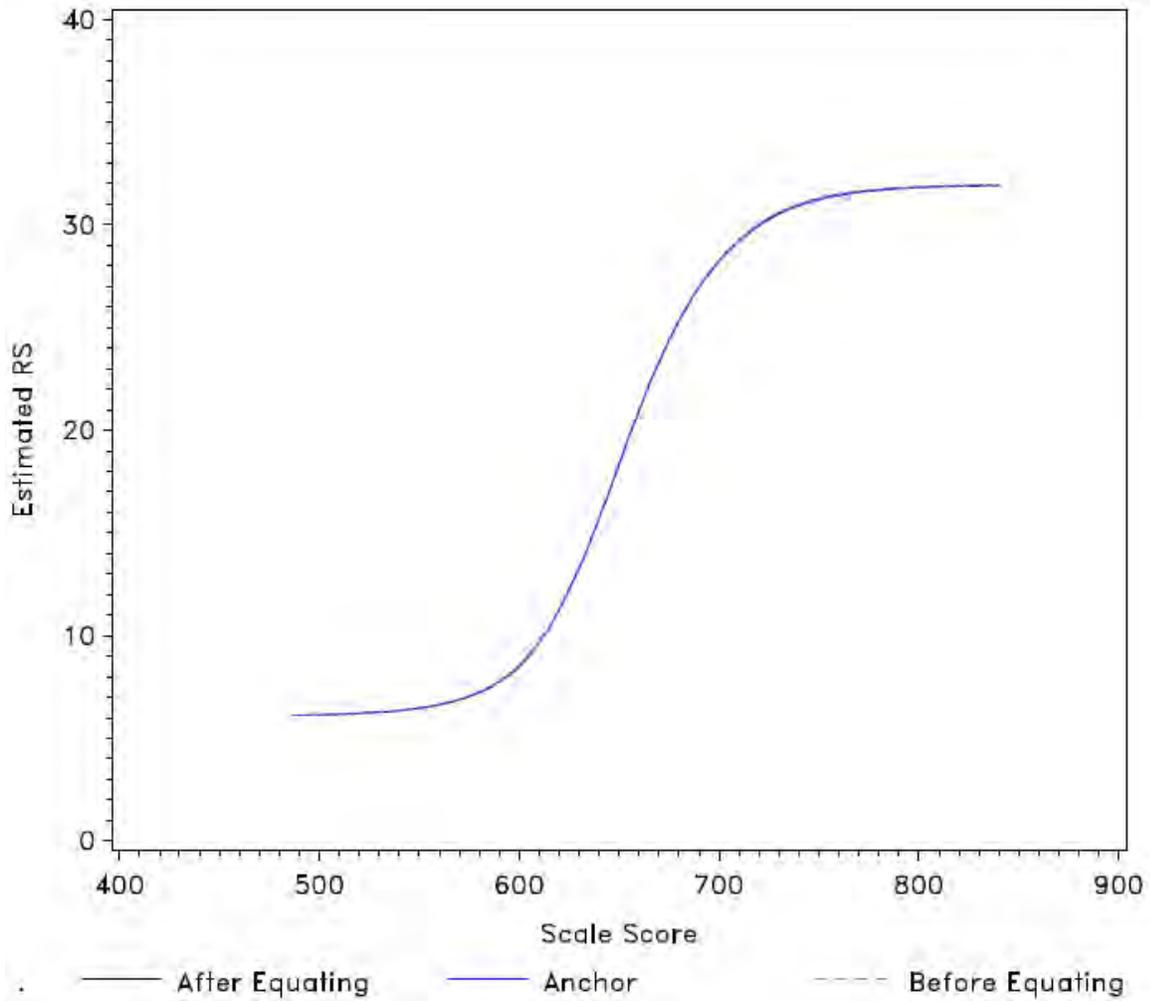


Figure 27. Communication Arts, Grade 6 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

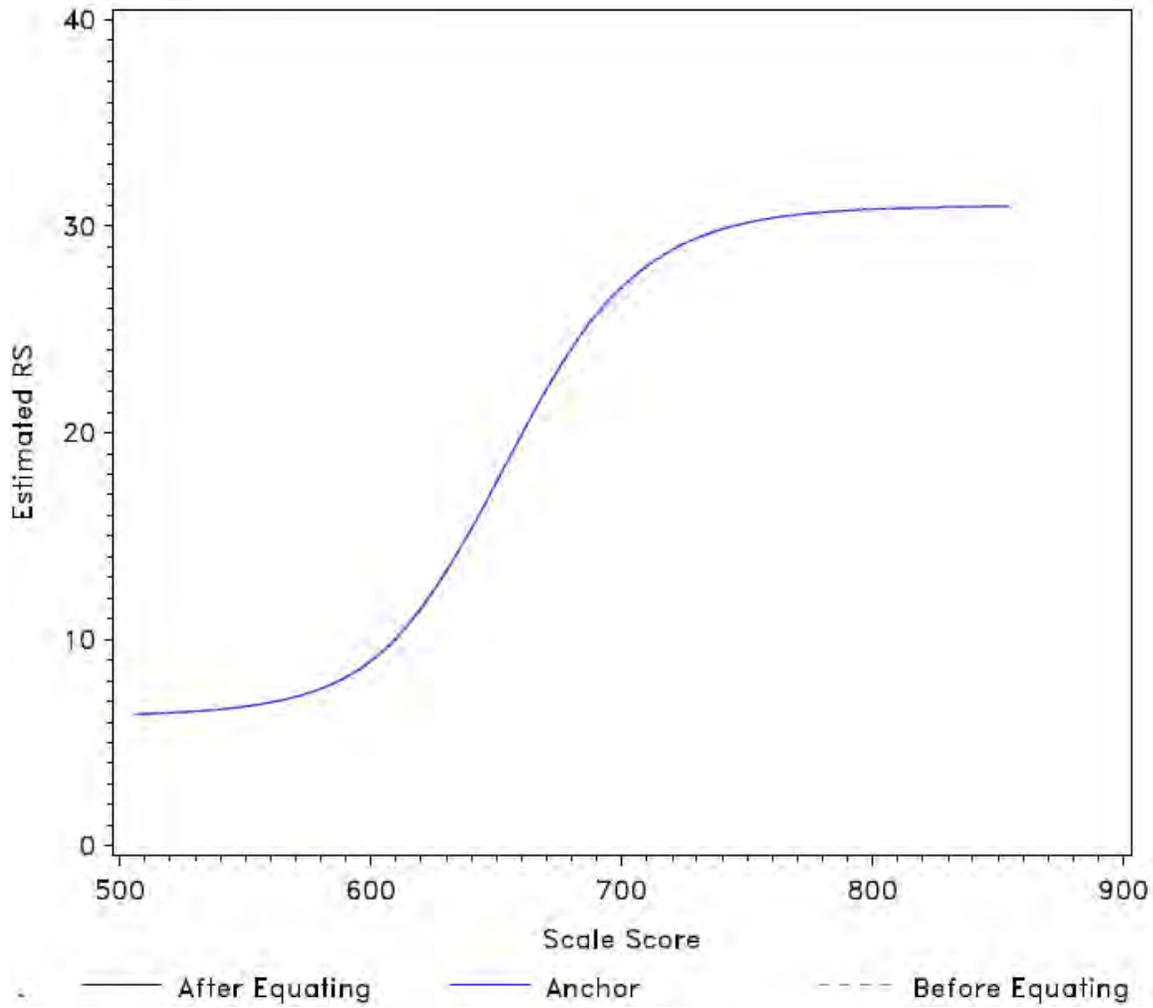


Figure 28. Communication Arts, Grade 7 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

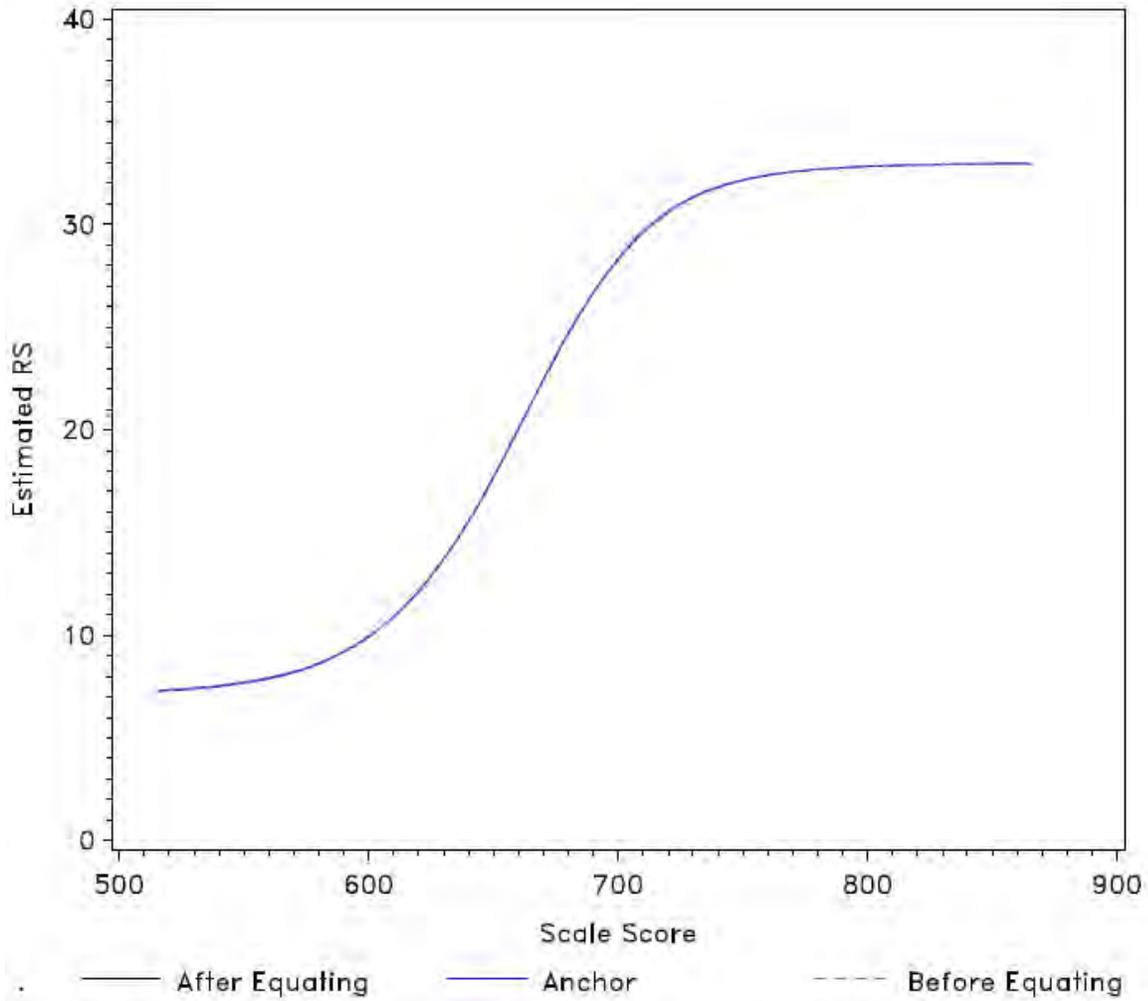


Figure 29. Communication Arts, Grade 8 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

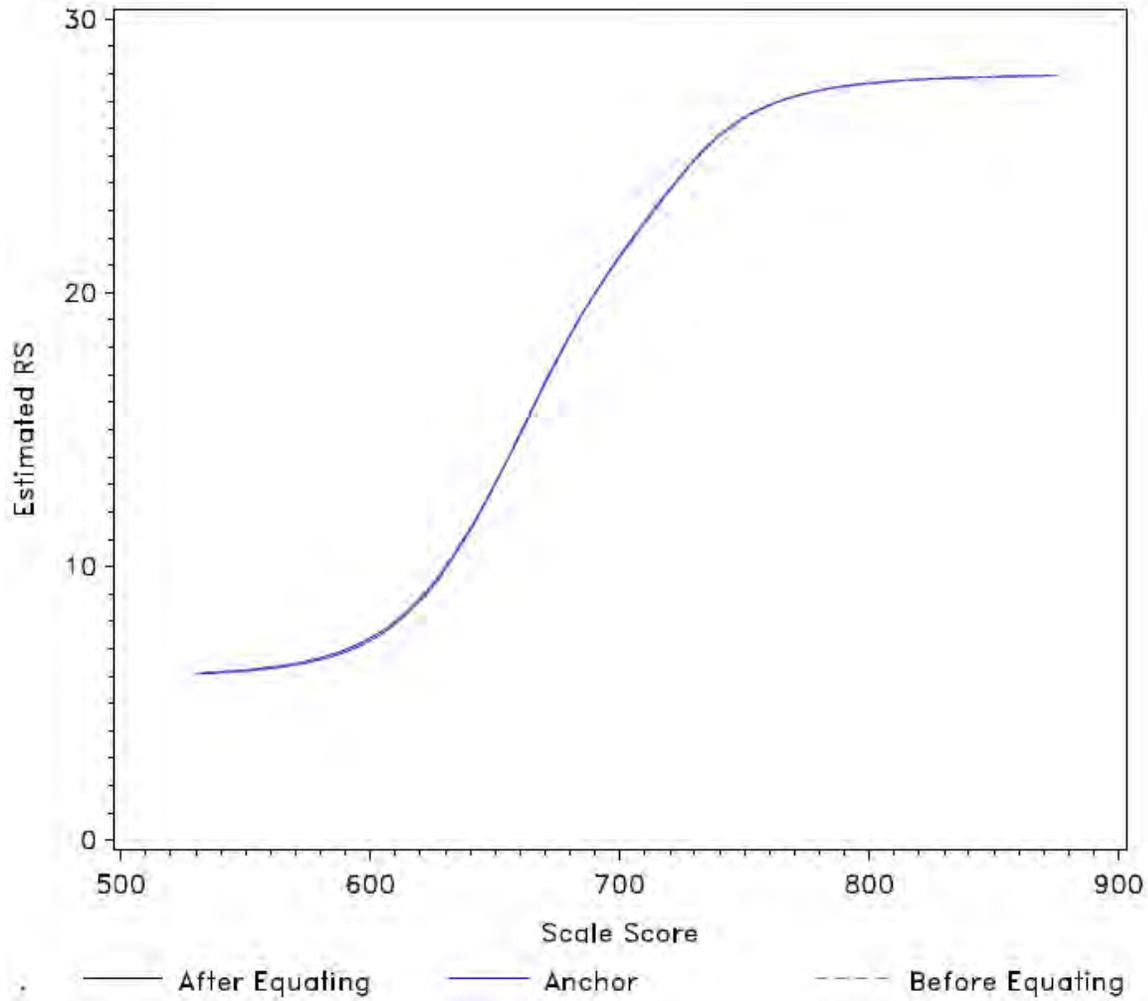


Figure 30. Mathematics, Grade 3 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

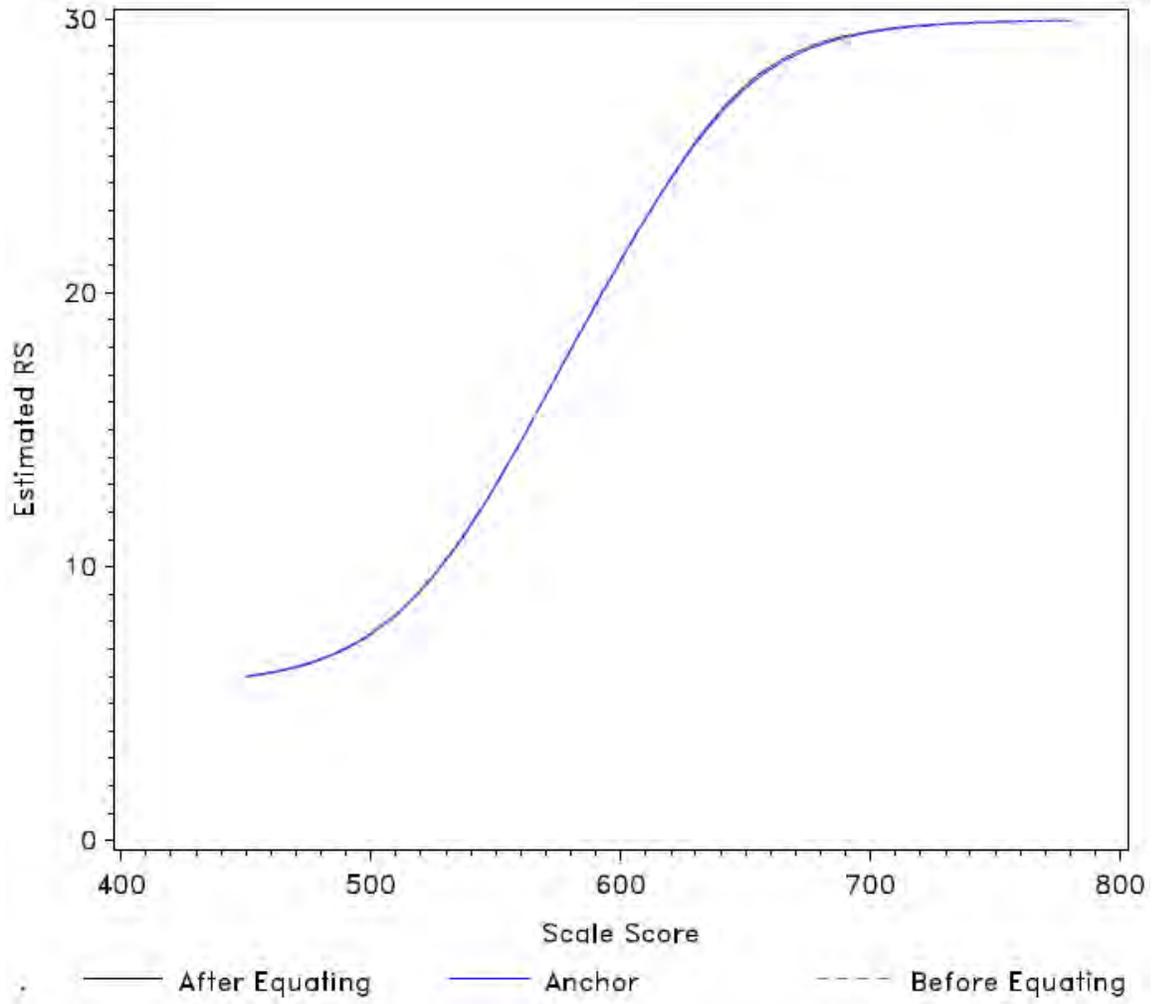


Figure 31. Mathematics, Grade 4 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

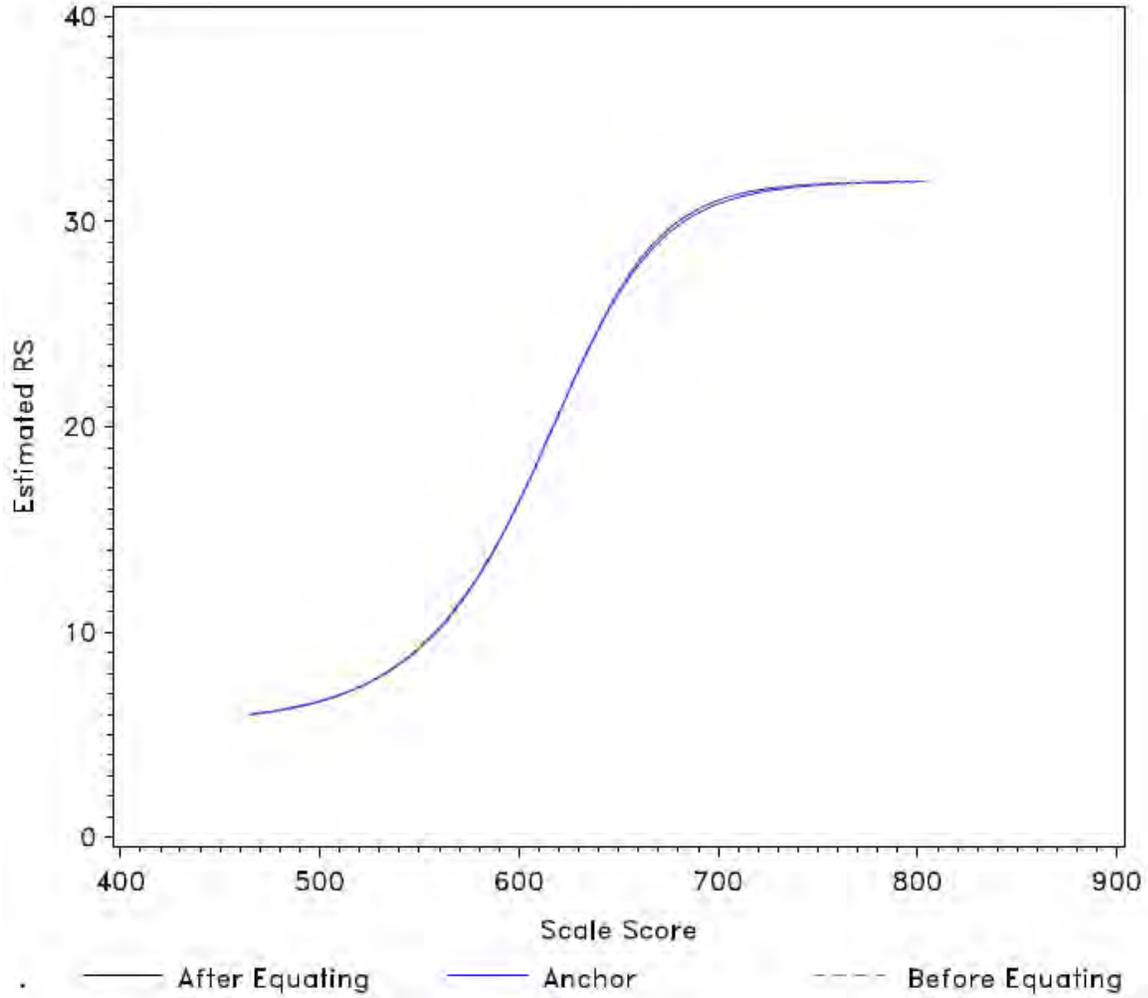


Figure 32. Mathematics, Grade 5 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

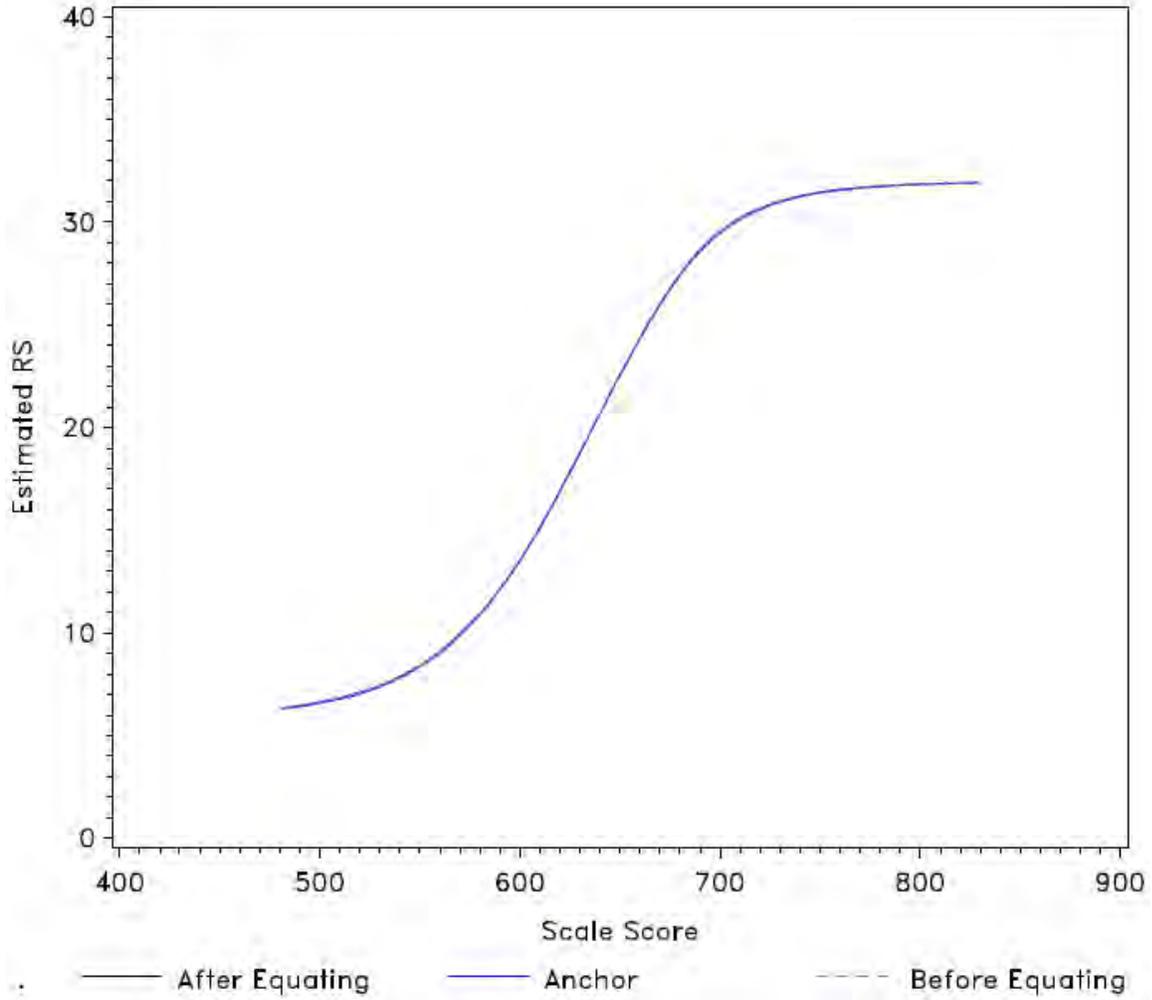


Figure 33. Mathematics, Grade 6 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

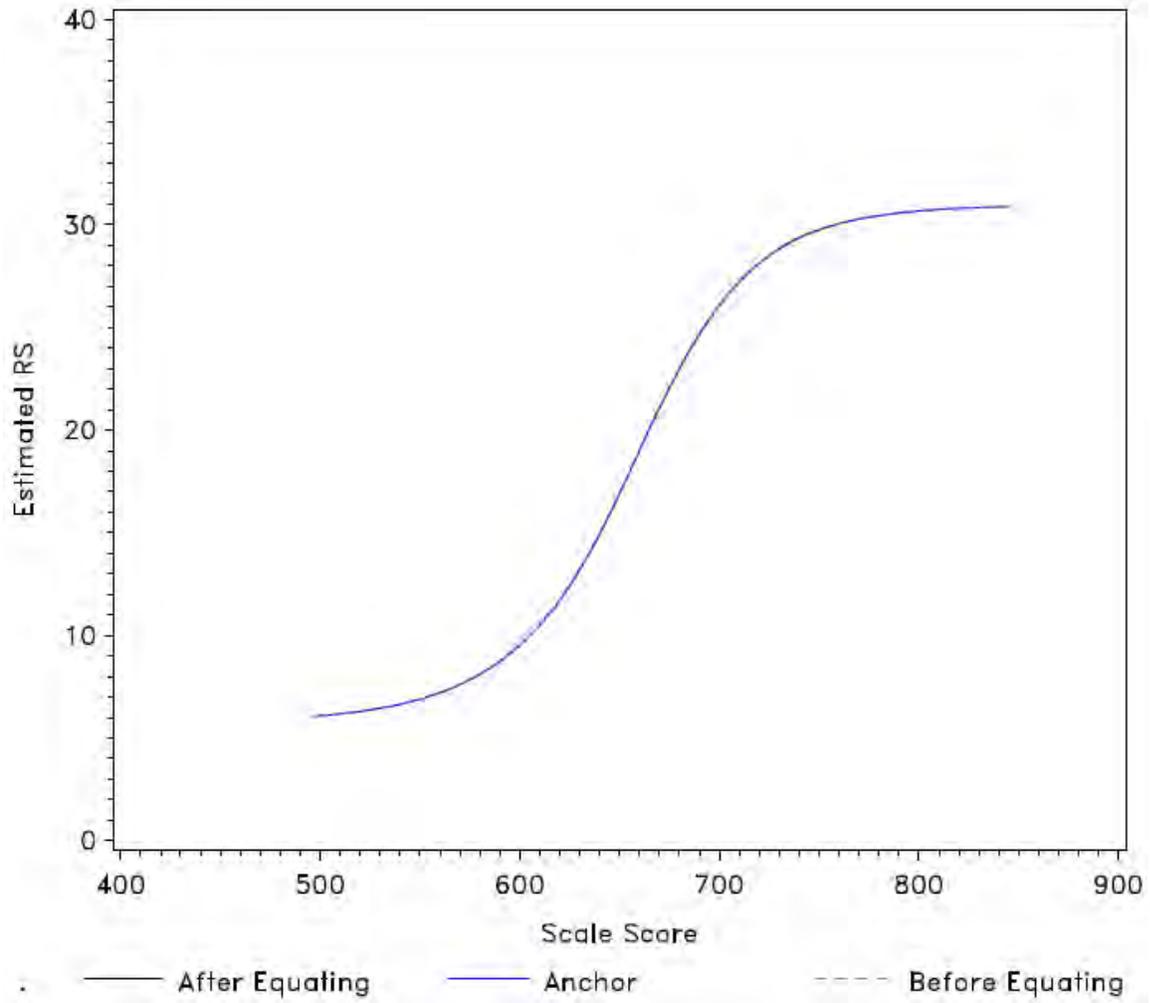


Figure 34. Mathematics, Grade 7 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)

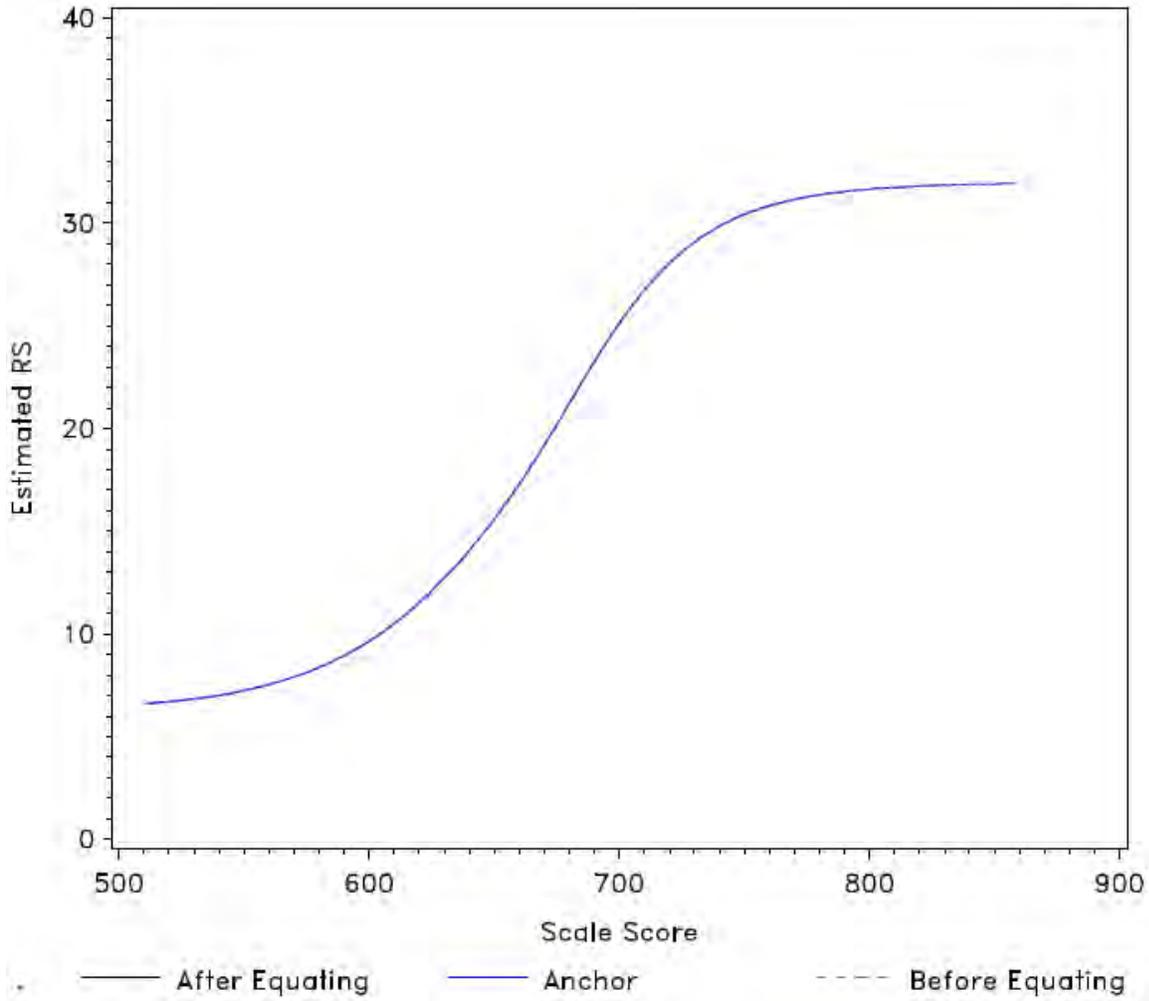
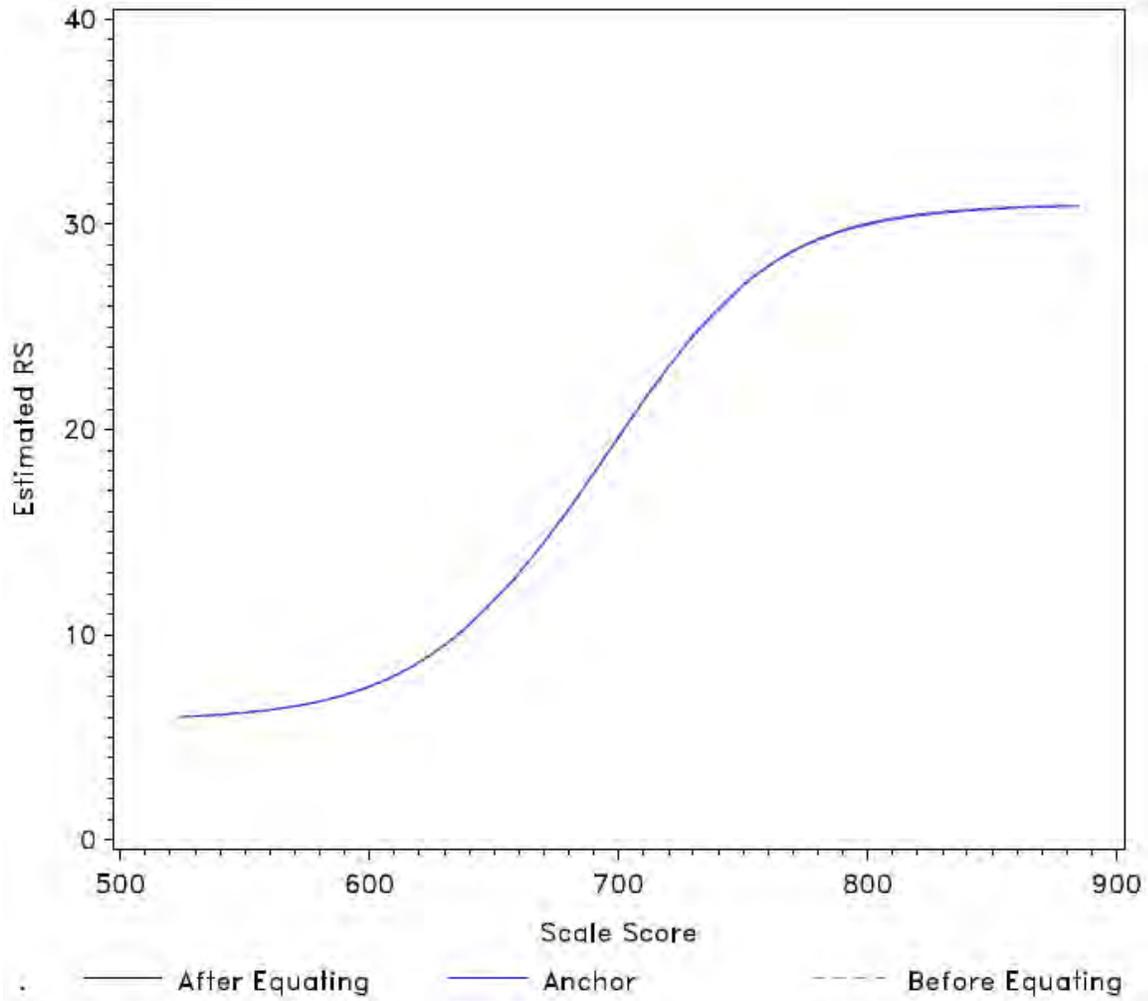


Figure 35. Mathematics, Grade 8 Test Characteristic Curves (TCC) for the Inputted Anchor Items and for the Estimated Anchor Items (Alternate Linking)



Appendix C

<i>Example of Presentation on Score Reports.....</i>	<i>C-1</i>
<i>Examples of Score Reports.....</i>	<i>C-12</i>
<i>Layout of General Research File.....</i>	<i>C-32</i>

Example of Presentation on Score Reports

What Is the Data Telling Us?

Ready to **analyze!**
Ready to **act!**

Objectives:

1. To review resources to interpret MAP data
2. To practice a process for examining, analyzing and responding to:
 - MAP Data
 - Classroom Assessments
 - Classroom Instruction
3. To remember that it's all about instruction and that Map Data is just one snapshot

Data?

It's not always black and white.

Change:

When We Focus on
Something,
We Make Progress

That is if . . .

- We know where we are
- We know where we want to go
- We know what it takes for us to get there

Give Change

a

Chance!

Before analyzing the data, let's
agree to . . .

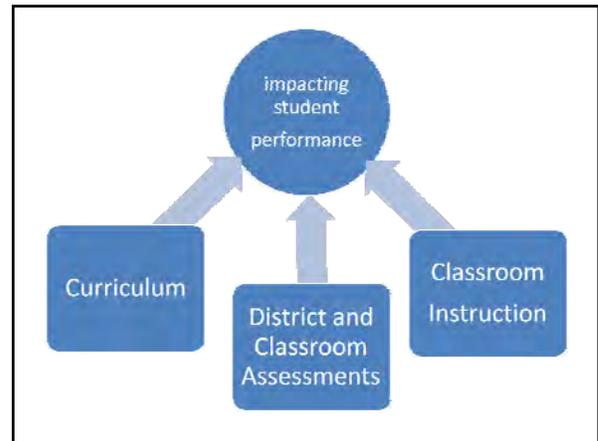
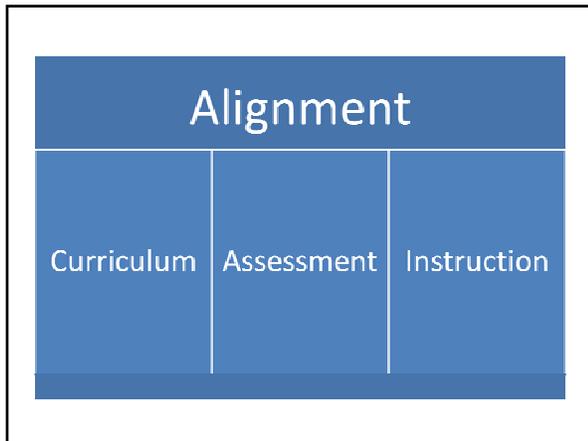
Focus on what we **can**
do, rather than what we
can't

Not leave anything about
teaching and learning to
chance

Set **high** goals not only for
our students, but for
ourselves.

The Numbers and
the **Stories** they Tell . . .

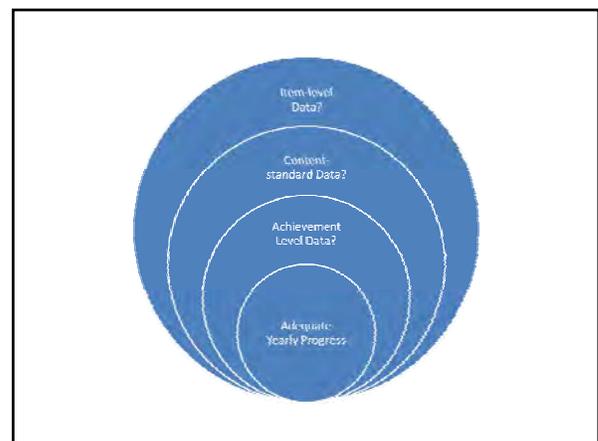




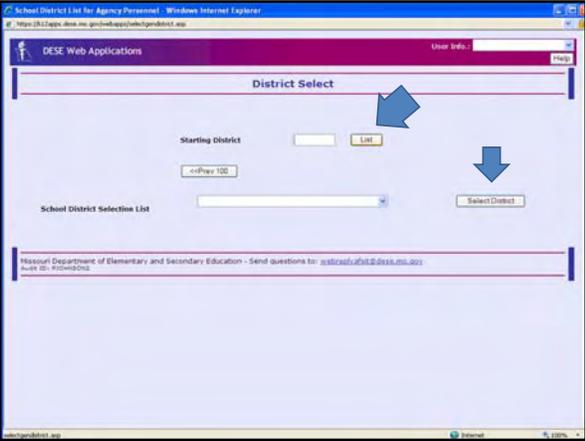
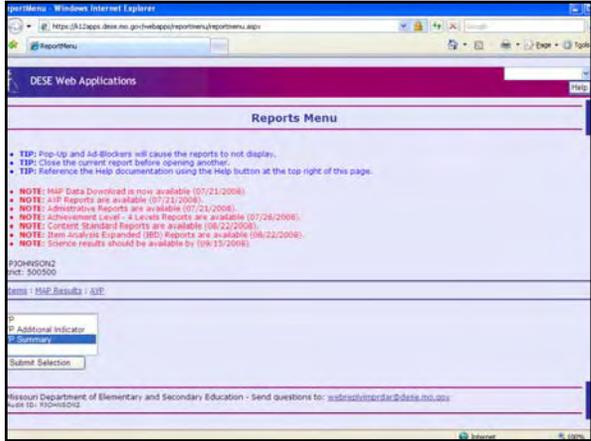
- Which reports can we use and how will we use them?
- **Adequate Yearly Progress** report to verify met or not met.
 - **Achievement-Level Descriptors** to examine students grade level learning targets.
 - **Achievement-Level** report allows for an "at a glance" comparison of data by groups and to gain a "top/bottom" view of percentages

- Which reports can we use and how will we use them?
- **Content Standard** report to analyze performance by strands across multiple years.
 - **Test blue print** to examine the focus of learning targets by strands.
 - **Item Benchmark Description** report to determine "learning targets" for improvement or recognition of strengths.

- Data by . . .
- **State**
 - **District**
 - Totals
 - Sub-groups
 - **Schools**
 - Totals
 - Sub-groups



Those are the **Reports**, so how do you retrieve them?



Missouri Department of Elementary and Secondary Education
 Adequate Yearly Progress - Summary Report
 2008: The AYP Data is based on the Adequate Yearly Progress Report

Name	Communication Arts			Mathematics			Science			All	Overall	Required
	Year	Groups	Met	Year	Groups	Met	Year	Groups	Met			
3000	2009	9	2	1	2	10	4	Yes	Yes	Yes	Yes	Yes
3000	2007	9	1	1	10	2	Yes	Yes	Yes	Yes	Yes	Yes
3000	2008	9	2	1	10	2	Yes	Yes	Yes	Yes	Yes	Yes
1000	2009	4	2	3	1	7	3	Yes	Yes	Yes	Yes	Yes
1000	2007	3	0	4	0	7	3	Yes	Yes	Yes	Yes	Yes
1000	2008	4	0	4	0	8	0	Yes	Yes	Yes	Yes	Yes
2000	2008	9	4	9	4	10	6	Yes	Yes	Yes	Yes	Yes
2000	2007	9	1	1	10	2	Yes	Yes	Yes	Yes	Yes	Yes
2000	2008	9	1	1	10	2	Yes	Yes	Yes	Yes	Yes	Yes
4000	2008	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2007	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2008	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2007	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2008	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2007	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2008	3	3	2	2	2	2	Yes	Yes	Yes	Yes	Yes
4000	2007	4	3	4	3	4	4	Yes	Yes	Yes	Yes	Yes
4000	2008	4	3	4	3	4	4	Yes	Yes	Yes	Yes	Yes
4000	2007	4	3	4	3	4	4	Yes	Yes	Yes	Yes	Yes
4000	2008	4	4	4	4	4	4	Yes	Yes	Yes	Yes	Yes
4000	2007	4	4	4	4	4	4	Yes	Yes	Yes	Yes	Yes
4000	2008	5	2	3	1	6	1	Yes	Yes	Yes	Yes	Yes

Annual Performance Report

Print: The AYP printout shows your location in red in the results. To change this setting, select 'Yes' from the location column. Also select 'Yes' from the 'Results' column.

Communication Arts	2008		Groups*	Met
	Year	Groups		
Mathematics	2	3	3	0
Additional Indicator	None			
Required Action: School Improvement Level 2				

Annual Report	2008	2009	2010	2011	2012	2013	2014
COMMUNICATIVE ARTS	3902	3903	3904	3905	3906	3907	3908
Annual Performance Target	18.0	18.0	20.0	20.0	20.0	20.0	20.0
School Total (All Kids)	PRCFC	33.3*	32.2*	32.4*	32.9*	24.4*	24.1*
Group	PRCFC	0	2.3	1.9	0	0	0
Group	PRCFC	33.3*	31.1*	30.8*	31.9*	24.4*	24.6*
Group	PRCFC	0	2.3	1.9	0	0	0
Group	PRCFC	0	2.3	1.9	0	0	0
Group	PRCFC	0	2.3	1.9	0	0	0
Group	PRCFC	0	2.3	1.9	0	0	0

- Adequate Yearly Progress**
- NCLB
 - Met or not met
 - Sub-groups – minimum of 30 students
 - Some of the terms: eg. – “safe harbor”
 - Updates/changes: **Growth Model**

Achievement Level Reports

- Show the percentage of students at each achievement level
- Compare data from multiple years to determine progress

Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
MAP Achievement Level 4 Report

Gender	Content Area	Grade	Year	ACC	REP	LMB	LMD%	BB	% BB	Basic	% Basic	Prof	% Prof	Adv	% Adv	MAP Score
Female	Mathematics	03	2006	24	24	0	0.0	2	8.3	14	58.3	7	29.2	1	4.2	729.58
Female	Mathematics	03	2007	27	27	0	0.0	5	22.2	17	63.0	4	14.8	0	0.0	802.58
Female	Mathematics	03	2008	12	12	0	0.0	0	0.0	5	41.7	7	58.3	0	0.0	748.58
Female	Mathematics	04	2006	27	27	0	0.0	10	37.0	15	55.6	2	7.4	0	0.0	875.43
Female	Mathematics	04	2007	27	27	0	0.0	4	14.8	18	66.7	3	11.1	0	0.0	753.79
Female	Mathematics	04	2008	19	19	0	0.0	5	26.3	13	68.4	1	5.3	0	0.0	876.94
Female	Mathematics	05	2006	24	24	0	0.0	3	12.5	18	75.0	3	12.5	0	0.0	700.00
Female	Mathematics	05	2007	21	21	0	0.0	0	0.0	20	95.2	0	0.0	0	0.0	723.71
Female	Mathematics	05	2008	23	23	0	0.0	3	13.0	17	73.9	3	13.0	0	0.0	700.00
Female	Mathematics	06	2006	22	22	0	0.0	2	9.1	13	59.1	7	31.8	0	0.0	722.71
Female	Mathematics	06	2007	25	25	0	0.0	1	4.0	15	60.0	4	16.0	0	0.0	715.00
Female	Mathematics	06	2008	23	23	0	0.0	3	13.0	12	52.2	4	17.4	0	0.0	748.58
Male	Mathematics	03	2006	21	21	0	0.0	3	14.3	14	66.7	3	14.3	1	4.8	708.58
Male	Mathematics	03	2007	27	27	0	0.0	4	14.8	17	63.0	3	11.1	0	0.0	705.43
Male	Mathematics	03	2008	17	17	0	0.0	0	0.0	10	60.2	2	11.8	0	0.0	711.88
Male	Mathematics	04	2006	23	23	0	0.0	2	8.7	20	87.0	1	4.3	0	0.0	805.79
Male	Mathematics	04	2007	11	11	0	0.0	1	9.1	10	90.9	1	9.1	0	0.0	728.58
Male	Mathematics	04	2008	14	14	0	0.0	3	21.4	20	142.9	1	7.1	0	0.0	728.58
Male	Mathematics	05	2006	32	32	0	0.0	3	9.4	22	68.8	5	15.6	2	6.3	718.84
Male	Mathematics	05	2007	21	21	0	0.0	1	4.8	14	66.7	5	23.8	0	0.0	723.00
Male	Mathematics	05	2008	23	23	0	0.0	3	13.0	14	60.9	2	8.7	0	0.0	705.00
Male	Mathematics	06	2006	26	26	0	0.0	4	15.4	14	53.8	7	26.9	1	3.8	719.25
Male	Mathematics	06	2007	25	25	0	0.0	8	32.0	13	52.0	4	16.0	2	8.0	748.58
Male	Mathematics	06	2008	19	19	0	0.0	2	11.1	12	63.2	4	21.1	0	0.0	711.88

ACC=Accomplished, REP=Proficient, LMB=Level Not Determined, LMD=Level Determined, BB=Proficient, Adv=Advanced, Top2 = Percent of Students Scoring Advanced or Proficient
MAP Score = (pp BB %)+(pp Basic %)+(pp Proficient%)+(pp Advanced%)

Report as of: 8/27/2008

Content Standards/Strands Reports

- Show the average percentage of raw-score points earned by students on each strand
- Identify trends over the multiple years
 - Strengths
 - Weaknesses
- Determine possible reasons for improvement or lack of improvement

Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
Content Standards Report

District	Mathematics	CS-1 Number and Operations	CS-2 Algebraic Relationships	CS-3 Geometric and Spatial Relationships	CS-4 Measurement	CS-5 Data and Probability					
Mathematics 6000	Category/Type	pp	pp	pp	pp	pp					
Warrensburg 05 Total/Total	2008	72%	18	65%	14	35%	13	47%	12	75%	18
Warrensburg 05 Total/Total	2007	72%	18	59%	14	35%	13	42%	12	84%	16
Warrensburg 05 Total/Total	2006	74%	17	59%	13	35%	12	55%	13	50%	13
School Mathematics 4050	Category/Type	pp	pp	pp	pp	pp					
Warrensburg 05 Total/Total	2008	72%	18	59%	14	35%	13	45%	12	72%	18
Warrensburg 05 Total/Total	2007	87%	18	56%	14	31%	13	45%	13	81%	16
Warrensburg 05 Total/Total	2006	89%	17	59%	13	31%	12	55%	13	82%	13
School Mathematics 4080	Category/Type	pp	pp	pp	pp	pp					
Warrensburg 05 Total/Total	2008	72%	18	62%	14	34%	13	47%	12	75%	18
Warrensburg 05 Total/Total	2007	89%	18	74%	14	39%	13	50%	13	87%	16
Warrensburg 05 Total/Total	2006	71%	17	59%	13	35%	12	72%	12	65%	13

Deeper Analysis

Activity: Prioritize the order in which you will analyze the content standards/strands based on what you have discovered from the Achievement Level and Content Standards/Strands Reports.

Make sure you come to consensus as a group.

Item Benchmark Descriptors

- The IBD Report gives the mean score point for an item for the group selected for analysis.

Course Level Expectations

Geometric and Spatial Relationships

1. Analyze characteristics and properties of two- and three-dimensional geometric shapes and develop mathematical arguments about geometric relationships

	Grade 7	Grade 8	Algebra I	Geometry
A Describe and use geometric relationships	*identify the 2-dimensional cross-section of a 3-dimensional shape	*describe, classify and generalize relationships between and among types of a) 2-dimensional objects and b) 3-dimensional objects using their defining properties, including Pythagorean Theorem		use inductive and deductive reasoning to establish the validity of geometric relationships; prove theorems and critique arguments made by others
DOK	2	3		3
ST	Math 2 1.10	Math 2 1.6		Math 2 3.2
B Apply geometric relationships	describe relationships between corresponding sides, corresponding angles and corresponding perimeters of similar polygons.		*apply geometric properties such as similarity and angle relationship to solve multi-step problems in 2-dimensions	
DOK	2		2	
ST	Math 2 1.6		Math 2 3.6	

G1Ageo

Process Standards – an integral part of the process of interpreting GLEs and CLEs

Process Standards

Students will demonstrate within and integrate across all content areas the ability to:

1:6 discover and evaluate patterns and relationships in information, ideas and structures

1:8 organize data, information and ideas into useful forms (including charts, graphs, outlines) for analysis or presentation

1:10 apply acquired information, ideas and skills to different context as students, workers, citizens and consumers

- ### Analyzing the Items
- Use the Data Analysis Chart
 - Use several years of data
 - Define the items that are below 50% or whatever cut-point you decide
 - Identify the lowest items for a specific content strand

Form A **Crystal Reports Data Analysis**

Content Standard _____ Grade _____

Year	Item	Item Type	Benchmark Description	Process Standard	Notable Trends General Conclusions
	Session _____ Item _____	MC CR FE			
	Session _____ Item _____	MC CR FE			
	Session _____ Item _____	MC CR FE			
	Session _____ Item _____	MC CR FE			
	Session _____ Item _____	MC CR FE			
	Session _____ Item _____	MC CR FE			

Crystal Reports Data Analysis

Update with this year's info based on the local example being used.

Content Standard 3 Nonfiction _____

Year	Item	Item Type	Benchmark Description	Process Standard	Notable Trends General Conclusions
2000	Session _____ Item _____	MC CR FE	Analyze the problem-solving strategies of individual characters.	14	
2000	Session _____ Item _____	MC CR FE	Identify and state problem and solution clearly.	11	
2000	Session _____ Item _____	MC CR FE	Formulate and support information and ideas by making predictions and comparisons, drawing general and specific conclusions.	24	
2000	Session _____ Item _____	MC CR FE	Determine the reliability of information.	17	
2000	Session _____ Item _____	MC CR FE	Write about a variety of communication.	14	
2000	Session _____ Item _____	MC CR FE	Formulate and support information and ideas by making predictions and comparisons, drawing general and specific conclusions. Present possible results of argument solutions.	24 17	

Crystal Reports Data Analysis
Update with local info based on example being used

Content Standard ____3 Nonfiction ____

Year	Item	Item Type	Benchmark Description	Process Standards	Notable Trends, General Conclusions
2002	Session 1, Item 12	MC, DE, K, PE	Formulate and support information and ideas by testing predictions and comparisons, developing options, and drawing conclusions.	2.4	
2002	Session 1, Item 18a	MC, DE, K, PE	Summarize main ideas.	1.6	
2002	Session 1, Item 18b	MC, DE, K, PE	Formulate and support information and ideas by testing predictions and comparisons, developing options, and drawing conclusions.	2.4	
2002	Session 1, Item 18c	MC, DE, K, PE	Formulate and support information and ideas by testing predictions and comparisons, developing options, and drawing conclusions.	2.4	
2002	Session 1, Item 18d	MC, DE, K, PE	Draw conclusions and form options.	3.5	
2002	Session 1, Item 20a	MC, DE, K, PE	Summarize main ideas.	1.6	

Crystal Reports Data Analysis
Update with local information

Year	Item	Item Type	Benchmark Description	Process Standards	Notable Trends, General Conclusions
2002	Session 1, Item 12	MC, DE, K, PE	Determine key words to use in locating multiple resources.	1.1	After analyzing three years of this data, had all other data on these students available from the district - other reading levels, comprehension, etc.
2002	Session 1, Item 18a	MC, DE, K, PE	Identify and state problems and solutions clearly.	3.1	Use what factors influenced the scores on SAMP before the list of Questions to Ask About the state of where to look.
2002	Session 1, Item 18b	MC, DE, K, PE	Determine ideas to use in locating multiple resources.	1.1	Determine whether instruction in the domain of whether it is better. Factors influencing student success. Guard against changing instruction that is already used.
2002	Session 1, Item 18c	MC, DE, K, PE	Predict possible results of proposed solutions.	3.6	Use resources to help improve instruction where needed before to determine if help was.
2002	Session 1, Item 18d	MC, DE, K, PE	Determine the reliability of information.	1.7	
2002	Session 1, Item 20a	MC, DE, K, PE	Predict possible results of proposed solutions.	3.7	

Ponder – Write – Share

- What assumptions can you make about instruction related to your area of focus?
- How does your analysis differ from the evidence gathered locally about the group of students?
- How does knowledge of the test blue print impact the evidence from the reports and what you know locally?

Now What?

Get Ready to Act!

Act as a Team!

An awful lot of our teachers—even brand new ones—are left to figure out on their own what to teach and what constitutes “good enough” work.

What does this do?

Leaves teachers entirely on their own to figure out what to teach, what order to teach it in, HOW to teach it...and to what level.

Students can do no better than the assignments they are given...

Time to work on your Data Analysis Action Plan




**Leave
Blame
At the
Door!**

Data Analysis Action Plan

- **Develop hypotheses about causes of low performance**
- **Identify appropriate solutions**
- **Specify the actions needed**
 - **Be very specific**
 - **Include dates**
- **Plan for follow-through**

Form B
Crystal Reports Data Analysis Action Plan

Cause	Solution	Action Steps/Timeline	Resources Needed	Follow Up/Timeline

Crystal Reports Data Analysis Action Plan
Update to connect to local information

Cause and Instructional Strategy used to teach	Solution and new Instructional Strategy to try	Action Steps/Timeline	Learning Tools/Professional Development Needed	Follow Up/Timeline
Trades writing quality answers to CR questions	Have students practice with CRs in all content areas	From the beginning of the year, model and have students answer CRs in instruction using the question type in appropriate content areas. Use CRs in classroom with individual programs. Use CRs in instruction to practice for the year and is practiced in each quarterly test. Have students collect CRs from all content areas every two weeks to share with colleagues for shared CRs and CRs needed for improvement.	Local Strategy: Anrich CRs gives a good model for answers. Instructional time and CR collection from MAP are also useful.	Have a certain number of CRs on each quarterly test. Collect CRs from the test score. Agree, collect student answers and have them with colleagues to continue to track improvements by students.
Trades knowing how to draw conclusions	Have students practice drawing conclusions in all content areas	From the beginning of the year, give a model of how to draw conclusions. Use CRs in instruction and from information that is a good model. Use CRs in instruction and from information that is a good model. Have students collect CRs from all content areas every two weeks to share with colleagues for shared CRs and CRs needed for improvement.	Communication Arts Standards Interpretations gives questions under 1 and 2 that will be helpful in completing CRs.	Have test specific questions on quarterly tests. Have students collect CRs from the test score. Agree, collect student answers and have them with colleagues to continue to track improvements by students.
Trades with problem solving - identification, making conclusions, problem solving, and solutions are effective.	Have students practice problem solving in all content areas. This might be done by a problem solving session in content other than CRs.	From the beginning of the year, give a model of how to solve problems. Use CRs in instruction and from information that is a good model. Use CRs in instruction and from information that is a good model. Have students collect CRs from all content areas every two weeks to share with colleagues for shared CRs and CRs needed for improvement.	Communication Arts Standards Interpretations gives questions under 1 and 2 that will be helpful in completing CRs.	Agree, have test specific questions on quarterly tests. Have students collect CRs from the test score. Agree, collect student answers and have them with colleagues to continue to track improvements by students.

Other Ways of Making Sense of the Data

By exporting the IBD report from Crystal to an Excel document you can:

• Sort by **Strand (Algebra)** in order from least to greatest.

Or

• Sort by **GLE** in order from least to greatest.

Or

• Sort by **Process Standard** from least to greatest.

Sorted by Strand (Algebra) in order from least to greatest.

YR	Str	GLE	Code	GLE Desc.	DOM	SA	Q1	IP	%		
2006	06	A	1	B	represent and describe patterns with tables, graphs, pictures, symbolic rules or words	2	1	14	CR	2	34.00
2006	06	A	1	A	compare situations with constant or varying rates of change	2	1	14	MC	1	36.00
2006	06	A	1	D	identify functions as linear or nonlinear from a table or graph	2	1	11	MC	1	36.00
2006	06	A	1	A	compare situations with constant or varying rates of change	2	1	11	CR	2	40.00
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	1	13	MC	1	42.00
2007	06	A	1	A	compare situations with constant or varying rates of change	1	1	13	MC	1	41.00
2007	06	A	3	A	model and solve problems, using multiple representations such as graphs, tables, expressions and equations	1	1	13	MC	1	44.00
2007	06	A	1	D	identify functions as linear or nonlinear from a table or graph	2	1	13	MC	1	48.00
2004	06	A	1	B	describe equivalent forms for simple algebraic expressions including associative and distributive properties	1	1	13	MC	1	48.00
2007	06	A	3	A	model and solve problems, using multiple representations such as graphs, tables, expressions and equations	1	1	14	MC	1	48.00
2007	06	A	1	B	represent and describe patterns with tables, graphs, pictures, symbolic rules or words	1	1	13	MC	1	54.00
2006	06	A	1	A	compare situations with constant or varying rates of change	1	1	11	MC	1	57.00
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	2	23	MC	1	57.00
2004	06	A	2	A	compare situations with constant or varying rates of change	2	1	13	MC	1	60.00
2004	06	A	2	A	use variables to represent unknown quantities in expressions	1	1	13	MC	1	62.00
2006	06	A	1	B	represent and describe patterns with tables, graphs, pictures, symbolic rules or words	2	2	13	MC	1	62.00
2007	06	A	1	C	compare various forms of representations to identify a pattern	1	1	13	MC	1	62.00
2004	06	A	4	A	compare situations with constant or varying rates of change	1	1	13	MC	1	66.00
2006	06	A	1	A	use variables to represent unknown quantities in expressions	1	1	13	MC	1	66.00
2004	06	A	3	A	model and solve problems, using multiple representations such as graphs, tables, expressions and equations	1	1	13	MC	1	68.00
2007	06	A	1	B	represent and describe patterns with tables, graphs, pictures, symbolic rules or words	1	1	13	MC	1	68.00
2006	06	A	1	C	compare various forms of representations to identify a pattern	1	1	13	MC	1	68.00
2004	06	A	4	A	use variables to represent unknown quantities in expressions	1	1	13	MC	1	68.00
2004	06	A	1	B	represent and describe patterns with tables, graphs, pictures, symbolic rules or words	1	1	13	MC	1	70.00
2004	06	A	1	B	represent and describe patterns with tables, graphs, pictures, symbolic rules or words	1	2	13	MC	1	70.00

Sorted by GLE in order from least to greatest.

YR	Str	S	C	GLE	Desc.	DOM	SA	Q1	IP	%	
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	2	31	CR	2	56.00
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	2	21	CR	2	74.00
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	2	21	MC	4	88.00
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	2	21	MC	1	83.00
2007	06	A	1	A	make and describe generalizations about geometric and numeric patterns	2	1	13	MC	1	54.00
2006	06	A	1	A	make and describe generalizations about geometric and numeric patterns	1	2	16	MC	1	36.00
2006	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	2	1	18	MC	1	36.00
2007	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	3	1	12	CR	2	48.00
2006	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	3	1	14	CR	2	44.00
2007	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	2	1	23	MC	1	48.00
2006	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	2	1	31	MC	4	57.00
2006	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	2	1	18	MC	1	63.00
2006	06	A	1	B	generate patterns represented graphically or numerically using words or symbolic rules, including recursive notation	2	1	19	MC	1	63.00
2006	06	A	1	C	compare and contrast various forms of representations of patterns	2	2	26	MC	1	36.00
2006	06	A	1	C	compare and contrast various forms of representations of patterns	2	2	26	MC	1	36.00
2006	06	A	1	C	compare and contrast various forms of representations of patterns	2	1	23	MC	1	73.00
2006	06	A	1	C	compare and contrast various forms of representations of patterns	2	2	30	MC	1	36.00
2006	06	A	1	C	compare and contrast various forms of representations of patterns	2	2	30	MC	1	37.00
2006	06	A	1	D	compare properties of linear functions between or among tables, graphs and equations	2	3	3	CR	2	17.00
2007	06	A	1	C	compare properties of linear functions between or among tables, graphs and equations	2	1	26	MC	1	88.00
2006	06	A	1	D	compare properties of linear functions between or among tables, graphs and equations	2	1	19	MC	1	82.00

Other Ways of Making Sense of the Data

A longitudinal item analysis:

- across grades
- examining one grade at a time
- by building or
- by district

Analyzing performance against total number of items per GLE to get percentages

Things to Remember

- MAP results should be reviewed and analyzed with all appropriate grade level or grade span teachers
- Multiple years of data must be used to ensure the most meaningful analysis
- Incorrectly identifying causes for poor performance will result in wasted effort



Things to Remember (con't.)

- Acknowledge strengths in order to maintain the performance
- Action plans should have timelines and follow-up
- Teachers need time and encouragement to consider the implications for change in curriculum, instruction, and classroom assessment

Examples of Score Reports

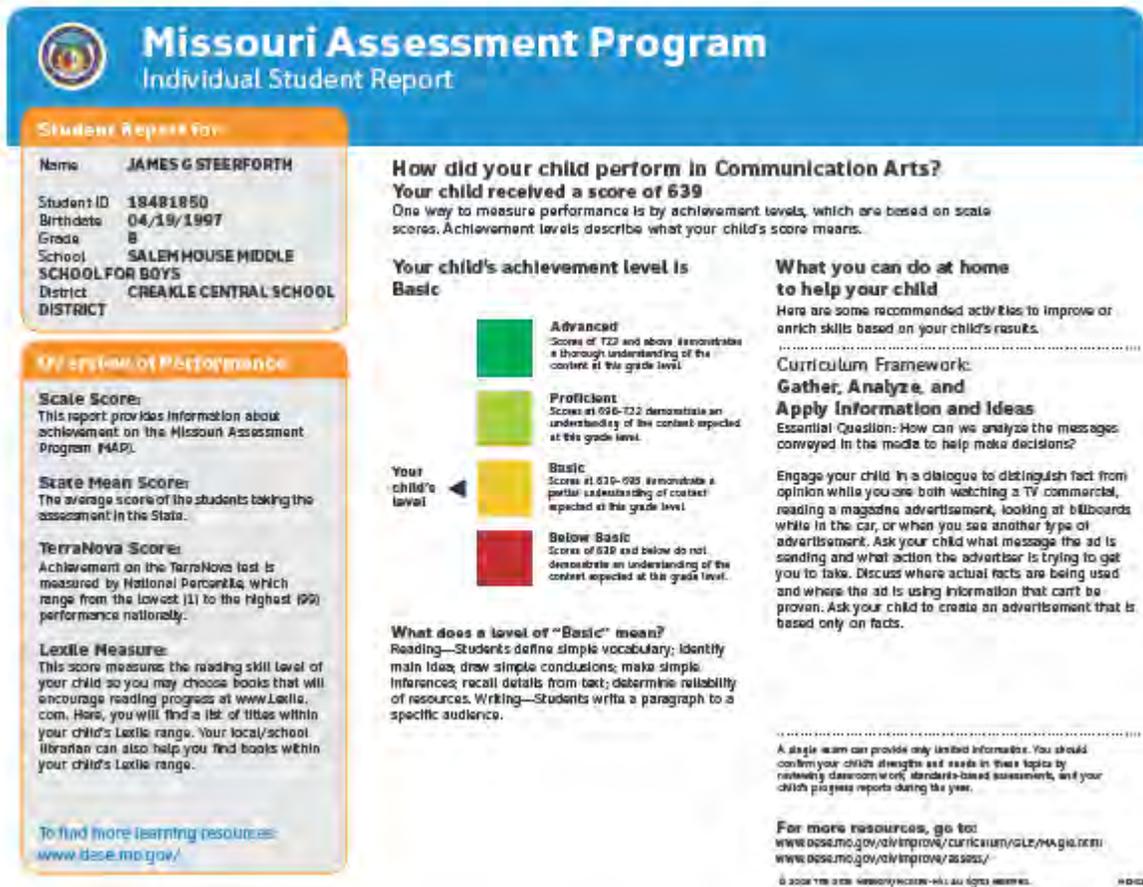


Figure C. 1 Example of Missouri Assessment Program Individual Student Report

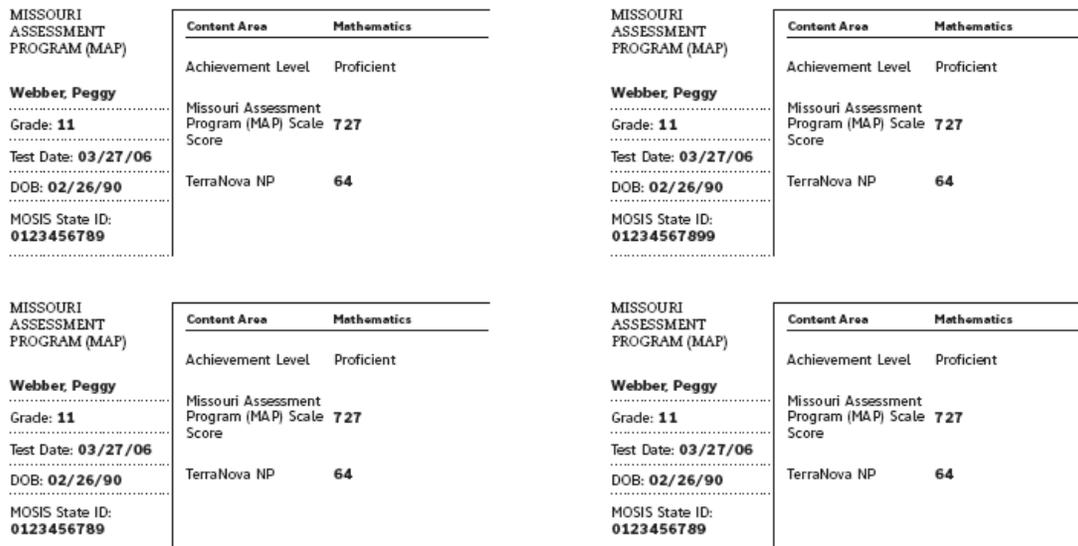


Figure C. 2 Example of Missouri Assessment Program Student Label

**Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
MAP Summary Report**

Communication Arts
03
2008
ELEM.

MOSIS ID	Year	CONTENT AREA	GRADE	MAP SCALE SCORE	Achievement Level	Terranova National Percentile
	2008	Communication Arts	03	650	Proficient	84
	2008	Communication Arts	03	608	Basic	30
	2008	Communication Arts	03	665	Proficient	85
	2008	Communication Arts	03	599	Basic	18
	2008	Communication Arts	03	679	Advanced	84
	2008	Communication Arts	03	644	Basic	72
	2008	Communication Arts	03	690	Advanced	88
	2008	Communication Arts	03	675	Advanced	88
	2008	Communication Arts	03	643	Basic	59
	2008	Communication Arts	03	599	Basic	20
	2008	Communication Arts	03	558	Below Basic	9
	2008	Communication Arts	03	602	Basic	36
	2008	Communication Arts	03	578	Below Basic	13
	2008	Communication Arts	03	626	Basic	54
	2008	Communication Arts	03	608	Basic	24
	2008	Communication Arts	03	641	Basic	63
	2008	Communication Arts	03	667	Proficient	80
	2008	Communication Arts	03	662	Proficient	77

04
2008
ELEM.

Missing data indicates student took MAP-Alternate

Report as of: 10/23/20

1

Figure C. 3 Example of Missouri Assessment Program Crystal Report, MAP Scale Score Summary Report

**Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
MAP Student Demographics Report**

Disability Diag

Year GR	Last Name	FName	DOB	Content	CTB#	MOSIS ID	DIST ID	DIST < 1YR	BUILD < 1YR	LEP	RACE	SES	IEP	ELL/LEP <1 Yr & <3 Yr	LEP ELL Title 3	LEP ELL Mth US	DIS Diag	MAP-A	TITLE I
2008	03			CA							5	Y	Y		00		Speech Impair		Y
2008	03			CA							5		Y		00		Other Health Ii		Y
2008	03			CA							5	Y			00				Y
2008	03			CA							5	Y	Y		00		Speech Impair		Y
2008	03			CA							5				00				Y
2008	03			CA							5				00				Y
2008	03			CA							5	Y			00				Y
2008	03			CA							5	Y	Y		00		Other Health Ii		Y
2008	03			CA							5		Y		00		Other Health Ii		Y
2008	03			CA							5	Y			00				Y
2008	03			CA							5	Y			00				Y
2008	03			CA							5	Y			00				Y
2008	03			CA							5	Y	Y		00		Speech Impair		Y
2008	03			CA							5		Y		00		Speech Impair		Y
2008	03			CA							5	Y			00				Y

Figure C. 4 Example of Missouri Assessment Program Crystal Report, Missouri Student Demographic Report

**Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
Student Achievement Level**

School Name	YEAR	Content Area	GRADE	Full Name	Achievement Level	MOSIS ID
2008	2008	Mathematics	10		Advanced	6208295203
	2008	Mathematics	08		Advanced	9946549972
	2008	Communication Arts	07		Advanced	4798940127
	2008	Communication Arts	08		Advanced	5889800272
	2008	Communication Arts	08		Advanced	7083159548
	2008	Mathematics	08		Advanced	7083159548
	2008	Science	08		Advanced	7083159548
	2008	Communication Arts	07		Advanced	7152450994
	2008	Mathematics	07		Advanced	7152450994
	2008	Mathematics	08		Advanced	8654154047
	2008	Communication Arts	08		Advanced	7878534535
	2008	Mathematics	08		Advanced	7878534535
	2008	Mathematics	08		Advanced	5609973635
	2008	Communication Arts	07		Advanced	5421542475
	2008	Mathematics	07		Advanced	5421542475
	2008	Communication Arts	11		Advanced	8548287977
	2008	Science	11		Advanced	6311573466
	2008	Communication Arts	06		Advanced	9775283205
	2008	Communication Arts	03		Advanced	8147981244
	2008	Mathematics	03		Advanced	8147981244
	2008	Communication Arts	04		Advanced	9725894227
	2008	Communication Arts	04		Advanced	9552970911
	2008	Communication Arts	03		Advanced	5480522442
	2008	Communication Arts	06		Advanced	5117801954
	2008	Communication Arts	03		Advanced	1660619017

Report as Of: 10/23/2008

1

Figure C. 5 Example of Missouri Assessment Program Crystal Report, Student Achievement Level Report

**Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
Student Report**

2008

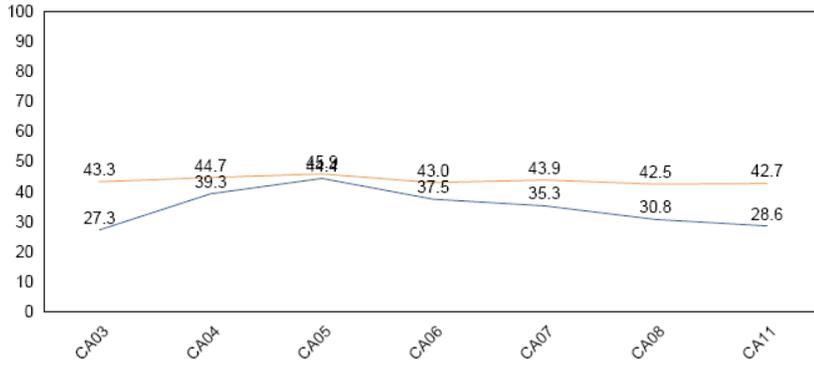
District:
County District Code:
School Name
School Code

DOB: 04-17-91	DISTRICT STUDENT #:	CTB STUDENT #:		
MOSIS ID #: CONTENT AREA: Mathematics	GRADE:10	ACHIEVEMENT LEVEL: Proficient	MAP SCALE: 741	TERRA NOVA: 86
DOB: 03-27-91	DISTRICT STUDENT #:	CTB STUDENT #:		
MOSIS ID #: CONTENT AREA: Communication Arts CONTENT AREA: Science	GRADE:11 GRADE:11	ACHIEVEMENT LEVEL: Proficient ACHIEVEMENT LEVEL: Proficient	MAP SCALE: 732 MAP SCALE: 731	TERRA NOVA: 74 TERRA NOVA: 51
DOB: 12-12-92	DISTRICT STUDENT #:	CTB STUDENT #:		
MOSIS ID #: CONTENT AREA: Communication Arts CONTENT AREA: Mathematics CONTENT AREA: Science	GRADE:08 GRADE:08 GRADE:08	ACHIEVEMENT LEVEL: Basic ACHIEVEMENT LEVEL: Below Basic ACHIEVEMENT LEVEL: Below Basic	MAP SCALE: 641 MAP SCALE: 638 MAP SCALE: 647	TERRA NOVA: 12 TERRA NOVA: 13 TERRA NOVA: 13
DOB: 06-07-91	DISTRICT STUDENT #:	CTB STUDENT #:		
MOSIS ID #: CONTENT AREA: Communication Arts CONTENT AREA: Science	GRADE:11 GRADE:11	ACHIEVEMENT LEVEL: Proficient ACHIEVEMENT LEVEL: Basic	MAP SCALE: 733 MAP SCALE: 723	TERRA NOVA: 90 TERRA NOVA: 53
DOB: 07-02-93	DISTRICT STUDENT #:	CTB STUDENT #:		
MOSIS ID #: CONTENT AREA: Communication Arts CONTENT AREA: Mathematics CONTENT AREA: Science	GRADE:08 GRADE:08 GRADE:08	ACHIEVEMENT LEVEL: Basic ACHIEVEMENT LEVEL: Below Basic ACHIEVEMENT LEVEL: Basic	MAP SCALE: 682 MAP SCALE: 663 MAP SCALE: 678	TERRA NOVA: 45 TERRA NOVA: 17 TERRA NOVA: 34
DOB: 04-09-91	DISTRICT STUDENT #:	CTB STUDENT #:		
MOSIS ID #: CONTENT AREA: Communication Arts	GRADE:11	ACHIEVEMENT LEVEL: Basic	MAP SCALE: 688	TERRA NOVA: 37

1

Figure C. 6 Example of Missouri Assessment Program Crystal Report, Student Report

Missouri Assessment Program
MAP Achievement 4 Levels
Percent of of Students Scoring Proficient or Advanced
Year: 2006



Total/Total

Data type = Total

Top 2 = Percent of Students Scoring Advanced or Proficient

Report as of: 10/23/2008

1

Figure C. 7 Example of Missouri Assessment Program Crystal Report, Achievement Level 4 Chart

Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
MAP Achievement Level 4 Report

District Totals

Total	Content Area	Grade	Year	ACC	REP	LND	LND%	BB	% BB	Basic	% Basic	Prof	%Prof	Adv	%Adv	MAP Index*	Mean Scale	Median Testscore
Total	Com. Arts	03	2008	18	18	0	0.0	2	11.1	9	50.0	4	22.2	3	16.7	744.40	633.00	61.00
Total	Com. Arts	04	2008	17	17	0	0.0	0	0.0	8	47.1	4	23.5	5	29.4	782.40	665.00	72.50
Total	Com. Arts	05	2008	23	23	0	0.0	3	13.0	12	52.2	7	30.4	1	4.3	726.10	664.50	55.00
Total	Com. Arts	06	2008	30	30	0	0.0	3	10.0	12	40.0	11	36.7	4	13.3	753.30	665.40	66.50
Total	Com. Arts	07	2008	17	17	0	0.0	3	17.6	5	29.4	6	35.3	3	17.6	752.90	670.70	70.50
Total	Com. Arts	08	2008	26	26	0	0.0	2	7.7	11	42.3	10	38.5	3	11.5	753.80	693.80	61.00
Total	Com. Arts	11	2008	22	22	0	0.0	1	4.5	16	72.7	4	18.2	1	4.5	722.70	706.70	51.00

ACC=Accountable; REP=Reportable; LND=Level Not Determined; Prof=Proficient; Adv=Advanced;
Top2 = Percent of Students Scoring Advanced or Proficient
MAP Index* = (pct BB*6)+(pct Basic*7)+(pct Proficient*8)+(pct Advanced*9)

Report as of: 10/23/200

2

Figure C. 8 Example of Missouri Assessment Program Crystal Report, Achievement Level 4 Report

Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
Content Standards Report

District Mathematics 0000	Category/Type	2008	CS-1 Number and Operations		CS-2 Algebraic Relationshi ps		CS-3 Geometric and Spatial Relationsh		CS-4 Measurement		CS-5 Data and Probability	
			PP	PP	PP	PP	PP	PP				
Mathematics	03 Total/Total	2008	71%	24	70%	13	68%	13	64%	10	68%	7
Mathematics	04 Total/Total	2008	86%	21	65%	15	84%	13	71%	15	60%	13
Mathematics	05 Total/Total	2008	75%	16	55%	13	57%	13	62%	14	69%	13
Mathematics	06 Total/Total	2008	81%	17	61%	13	54%	12	66%	13	57%	13
Mathematics	07 Total/Total	2008	61%	16	55%	13	47%	14	44%	13	52%	13
Mathematics	08 Total/Total	2008	71%	14	50%	21	46%	15	52%	12	59%	14
Mathematics	10 Total/Total	2008	62%	11	43%	20	44%	14	54%	11	60%	14

Average percentage of raw-score points earned by students on each standard.
 PP = Points Possible.

Report as of: 10/23/2008

1

Figure C. 9 Example of Missouri Assessment Program Crystal Report, Content Standards Report

**Missouri Assessment Program
Content Standards By Student
Missouri Department of Elementary and Secondary Education**

Mathematics

			CS-1 Number and Operations	CS-2 Algebraic Relationships	CS-3 Geometric and Spatial Relationships	CS-4 Measurement	CS-5 Data and Probability
Examiner Name:	Grade Level	Year	Points Possible-24	Points Possible-13	Points Possible-13	Points Possible-10	Points Possible-7
	03	2008	63%	85%	46%	30%	71%
	03	2008	92%	92%	92%	90%	100%
	03	2008	83%	92%	69%	90%	86%
	03	2008	63%	54%	46%	60%	57%
	03	2008	58%	54%	39%	60%	29%
	03	2008	92%	92%	92%	100%	86%
	03	2008	96%	100%	77%	90%	71%
	03	2008	50%	46%	31%	50%	43%
	03	2008	50%	39%	92%	40%	43%
	03	2008	38%	54%	23%	40%	43%
	03	2008	38%	54%	31%	30%	71%
	03	2008	92%	62%	77%	60%	57%
	03	2008	96%	100%	85%	100%	71%
	03	2008	54%	62%	62%	70%	57%
	03	2008	63%	62%	100%	40%	86%
	03	2008	96%	69%	77%	90%	86%
	03	2008	58%	39%	77%	20%	86%
	03	2008	92%	100%	100%	90%	86%

			CS-1 Number and Operations	CS-2 Algebraic Relationships	CS-3 Geometric and Spatial Relationships	CS-4 Measurement	CS-5 Data and Probability
Examiner Name:	Grade Level	Year	Points Possible-21	Points Possible-15	Points Possible-13	Points Possible-15	Points Possible-13
	04	2008	100%	80%	85%	80%	77%
	04	2008	95%	53%	77%	73%	54%
	04	2008	86%	60%	85%	67%	69%
	04	2008	91%	53%	77%	53%	77%
	04	2008	100%	100%	85%	100%	85%

Average percentage of raw-score points earned by students on each standard.

Report as of: 10/23/200

1

Figure C. 10 Example of Missouri Assessment Program Crystal Report, Content Standards Detail Report

Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
Average Points Earned with Item Benchmark Descriptions
Content Standard IBD Extended

District		School		Communication Arts		Session/Item		Pts Avg %	
2008		0000							
SC	GR	Standard Desc.	GLE Code	GLE Description	D_O_K	QT	Pts Poss	Avg Pts	% Earn
0000	D3	1 speaking/writing standard English	W 2 E	In writing, use <input type="checkbox"/> correct spelling of simple compounds, homophones, contractions and words with affixes <input type="checkbox"/> standard spelling <input type="checkbox"/> classroom resources and dictionary to verify correct spelling	2	Skill/Concept 1 / 10	MC	1	0.56 56.00
0000	D3	1 speaking/writing standard English	W 2 E	In writing, use <input type="checkbox"/> correct spelling of simple compounds, homophones, contractions and words with affixes <input type="checkbox"/> standard spelling <input type="checkbox"/> classroom resources and dictionary to verify correct spelling	2	Skill/Concept 1 / 11	MC	1	0.67 67.00
0000	D3	1 speaking/writing standard English	W 2 E	In writing, use <input type="checkbox"/> correct spelling of simple compounds, homophones, contractions and words with affixes <input type="checkbox"/> standard spelling <input type="checkbox"/> classroom resources and dictionary to verify correct spelling	2	Skill/Concept 1 / 12	MC	1	0.44 44.00
0000	D3	1 speaking/writing standard English	W 2 C	In composing text, use <input type="checkbox"/> correct ending punctuation in Imperative and exclamatory sentences <input type="checkbox"/> comma in the greeting and closing of a letter	2	Skill/Concept 1 / 7	MC	1	0.61 61.00
0000	D3	1 speaking/writing standard English	W 2 C	In composing text, use <input type="checkbox"/> correct ending punctuation in Imperative and exclamatory sentences <input type="checkbox"/> comma in the greeting and closing of a letter	2	Skill/Concept 1 / 8	MC	1	0.72 72.00
0000	D3	1 speaking/writing standard English	W 2 C	In composing text, use <input type="checkbox"/> correct ending punctuation in Imperative and exclamatory sentences <input type="checkbox"/> comma in the greeting and closing of a letter	2	Skill/Concept 1 / 9	MC	1	0.67 67.00
0000	D3	1 speaking/writing standard English	W 2 D	Use parts of speech correctly in written text <input type="checkbox"/> verbs that agree with the subject <input type="checkbox"/> words that answer when, where, why and how questions (adverbs) <input type="checkbox"/> words to compare (adverbs)	2	Skill/Concept 3 / 10	MC	1	0.89 89.00
0000	D3	1 speaking/writing standard English	W 2 D	Use parts of speech correctly in written text <input type="checkbox"/> verbs that agree with the subject <input type="checkbox"/> words that answer when, where, why and how questions (adverbs) <input type="checkbox"/> words to compare (adverbs)	2	Skill/Concept 3 / 11	MC	1	0.89 89.00
0000	D3	1 speaking/writing standard English	R 1 E	Develop vocabulary through text, using <input type="checkbox"/> base words <input type="checkbox"/> synonyms and antonyms <input type="checkbox"/> context clues <input type="checkbox"/> glossary <input type="checkbox"/> dictionary, with assistance	2	Skill/Concept 3 / 18	MC	1	0.67 67.00
0000	D3	1 speaking/writing standard English	W 2 D	Use parts of speech correctly in written text <input type="checkbox"/> verbs that agree with the subject <input type="checkbox"/> words that answer when, where, why and how questions (adverbs) <input type="checkbox"/> words to compare (adverbs)	2	Skill/Concept 3 / 19	MC	1	0.61 61.00
0000	D3	1 speaking/writing standard English	W 2 D	Use parts of speech correctly in written text <input type="checkbox"/> verbs that agree with the subject <input type="checkbox"/> words that answer when, where, why and how questions (adverbs) <input type="checkbox"/> words to compare (adverbs)	2	Skill/Concept 3 / 20	MC	1	0.78 78.00
0000	D3	1 speaking/writing standard English	W 1 A	Follow a writing process to <input type="checkbox"/> independently use a simple graphic organizer in pre-writing <input type="checkbox"/> generate a draft <input type="checkbox"/> routinely reread and revise work <input type="checkbox"/> routinely edit and proofread for capitalization and ending punctuation <input type="checkbox"/> independently publish writing	2	Skill/Concept 3 / 21	MC	1	0.83 83.00
0000	D3	1 speaking/writing standard English	W 1 A	Follow a writing process to <input type="checkbox"/> independently use a simple graphic organizer in pre-writing <input type="checkbox"/> generate a draft <input type="checkbox"/> routinely reread and revise work <input type="checkbox"/> routinely edit and proofread for capitalization and ending punctuation <input type="checkbox"/> independently publish writing	2	Skill/Concept 3 / 27	MC	1	0.39 39.00
0000	D3	1 speaking/writing standard English	W 1 A	Follow a writing process to <input type="checkbox"/> independently use a simple graphic organizer in pre-writing <input type="checkbox"/> generate a draft <input type="checkbox"/> routinely reread and revise work <input type="checkbox"/> routinely edit and proofread for capitalization and ending punctuation <input type="checkbox"/> independently publish writing	2	Skill/Concept 3 / 28	MC	1	0.44 44.00
0000	D3	1 speaking/writing standard English	W 2 B	Use conventions of capitalization in written text <input type="checkbox"/> months of year <input type="checkbox"/> titles of individuals <input type="checkbox"/> greeting and closing of letter	2	Skill/Concept 3 / 38	MC	1	0.44 44.00
0000	D3	1 speaking/writing standard English	W 2 B	Use conventions of capitalization in written text <input type="checkbox"/> months of year <input type="checkbox"/> titles of individuals <input type="checkbox"/> greeting and closing of letter	2	Skill/Concept 3 / 39	MC	1	0.33 33.00

Report as of: 10/23/2008

Figure C. 11 Example of Missouri Assessment Program Crystal Report, Content Standard IBD EX Report

Missouri Department of Elementary and Secondary Education
Missouri Assessment Program
Average Points Earned with Item Benchmark Descriptions
Goal Process Standard IBD Expanded

District		School		0000								
2008						Communication Arts						
SC	GR	Goal	Standard Desc.	GLE Code		D.O.K	Session/ Item	QT	Pts. Pos.	Avg. Pts.	% Earn	
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 1	MC	1	1.00	100.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 17	MC	1	0.72	72.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 2	MC	1	0.94	94.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 22	MC	1	0.94	94.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 23	MC	1	0.78	78.00
0000	03	1.5	comprehend/evaluate resources	R 1 E	Develop vocabulary through text, using □ base words □ synonyms and antonyms □ context clues □ glossary □ dictionary, with assistance	1	Recall	3 / 24	MC	1	0.61	61.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 29	MC	1	0.67	67.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 3	MC	1	0.61	61.00
0000	03	1.5	comprehend/evaluate resources	R 2 C	Use details from text to □ make inferences about setting, character traits and problem and solution □ make predictions □ draw conclusions □ compare and contrast characters and changes in problems and settings □ identify the narrator □ identify cause and effect □ identify events from the beginning, middle and end □ identify author's purpose	1	Recall	3 / 30	MC	1	0.50	50.00
0000	03	1.5	comprehend/evaluate resources	R 1 H	Apply post-reading skills to identify and explain the relationship between the main idea and supporting details □ question to clarify □ reflect □ analyze □ draw conclusions □ summarize □ paraphrase	1	Recall	3 / 33	MC	1	0.61	61.00
0000	03	1.5	comprehend/evaluate resources	R 1 E	Develop vocabulary through text, using □ base words □ synonyms and antonyms □ context clues □ glossary □ dictionary, with assistance	1	Recall	3 / 35	MC	1	1.00	100.00

Report as of: 10/23/200

1

Figure C. 12 Example of Missouri Assessment Program Crystal Report, Goal Process Standard IBD EX Report

**Missouri
Assessment
Program
(MAP)**

Summary Report

School: PINE VALLEY

Grade: 7

Simulated Data

Purpose

This report shows the number and percent of students locally in each of the four achievement levels. Instructional priorities can be established using this information along with other sources.



Test Date: 03/31/08

CODES: 048-079-2509
District: BIG CREEK
State: MISSOURI

City/State: ANYWHERE, MO

Communication Arts

Achievement Levels	Reportable	Accountable	Descriptions
ADVANCED	1% 3 Students	1% 3 Students	Reading-Students interpret complex figurative language and vocabulary; support a position; make predictions; summarize, analyze, and synthesize information and techniques; paraphrase ideas. Writing-Students consistently use the rules and conventions of Standard English; use logical order, cohesive devices, clear and varied sentences, writing techniques; target specific audience and purpose. MAP score range: 712-865.
PROFICIENT	10% 33 Students	9% 33 Students	Reading-Students make inferences; summarize; make comparisons and predictions using complex text; analyze characters; determine word meaning, point of view, supporting information; locate resources. Writing-Students stay on topic; write for a specific audience and purpose; demonstrate consistent use of a controlling idea; use rules and conventions of Standard English; use complex sentences, cohesive devices, clear and varied sentences. MAP score range: 680-711.
BASIC	29% 96 Students	27% 96 Students	Reading-Students identify text-based details; identify main idea; make simple summaries; identify the meaning of figurative language; draw simple conclusions; make simple inferences. Writing-Students use a writing process; edit for appropriate support; revise for a controlling idea; generally use the rules of Standard English. MAP score range: 634-679.
BELOW BASIC	38% 125 Students	37% 125 Students	Reading-Students locate and apply information in text; identify figurative language, text elements, and problems/solutions, character traits; make obvious predictions. Writing-Students organize information; use some components of letter writing format; generally stay on topic; show awareness of audience and purpose; minimally use rules and conventions of Standard English. MAP score range: 515-633.
LEVEL NOT DETERMINED		6% 20 Students	Students in this category are absent or did not have a valid attempt on one or more test sessions. A valid attempt on any item of the MAP test is necessary in order to receive a MAP score. The valid attempt rules for a TerraNova score are as follows: Attempt any five items or get one correct in the TerraNova item group.
Total Number of Students		330	350
TerraNova National Percentile		The number of students reported in each of the 4 achievement levels is the same for "Reportable" and "Accountable" because only students with a valid attempt are assigned an achievement level. The percentage of students in the Reportable column is based on the sum of all 4 achievement levels. The Accountable column adds a "Level Not Determined" that includes all students who do not have an assigned achievement level. The percentage of students is based on the sum of all 4 levels plus Level Not Determined.	
NP of Mean NCE*:		55	
Median NP:		54.0	
No. Students with TerraNova scores:		330	

*National Percentile of the Mean Normal Curve Equivalent

8/14/08

Figure C. 13 Example of Missouri Assessment Program Summary Report

Layout of General Research File

MAP'08 GRT				
Start	End	Length	Field	Values
Hierarchical Data				
				1 = State 2 = District 3 = School 4 = Class
1	1	1	Mode level	
2	11	10	Organization ID	A-Z, 0-9
12	41	30	Element Name - District	Any character, blank
42	43	2	Element Structure Level Number - District	02
44	50	7	Element Number - District	0-9
51	53	3	District/Element Special Codes A-C (Region Code)	Any character, ''
54	56	3	District/Element Special Codes D-F (District Code)	Any character, ''
57	76	20	District Special Codes G-Z	Any character, ''
77	78	2	Grade	03-08, 10, 11
79	108	30	City	Any character, blank
109	110	2	State	MO
111	140	30	Element Name- School	Any character, blank
141	142	2	Element Structure Level Number - School	03
143	149	7	Element Number - School	0-9
150	152	3	School/Element Special Codes A-C (Region Code)	Any character, ''
153	155	3	School/Element Special Codes D-F (District Code)	Any character, ''
156	159	4	School/Element Special Codes G-J (School Code)	Any character, ''
160	175	16	School/Element Special Codes K-Z	Any character, ''
176	205	30	Element Name- Class	Any character, blank
206	207	2	Element Structure Level Number- Class	04
208	214	7	Element Number-Class	0-9
215	217	3	Class/Element Special Codes A-C (Region Code)	Any character, ''
218	220	3	Class/Element Special Codes D-F (District Code)	Any character, ''
221	224	4	Class/Element Special Codes G-J (School Code)	Any character, ''
225	240	16	Class/Element Special Codes K-Z	Any character, ''
241	247	7	Student Element Number	0-9
248	275	28	Test Name	"Missouri Assessment Program"
276	278	3	<i>TerraNova Form/Level: Communication Arts</i>	D13 = Gr. 3, D14 = Gr. 4, D15 = Gr. 5, D16 = Gr. 6, D17 = Gr. 7 H18 = Gr. 8, A21 = Gr. 11 blank
279	281	3	<i>TerraNova Form/Level: Mathematics</i>	D13 = Gr. 3, D14 = Gr. 4, D15 = Gr. 5, D16 = Gr. 6, D17 = Gr. 7 D18 = Gr. 8, D20 = Gr. 10 blank
282	284	3	<i>TerraNova Form/Level: Science</i>	D15 = Gr. 5, D18 = Gr. 8, C21 = Gr. 11 blank
285	290	6	Test Date (MMDDYY)	
Special codes (Length 26)				
291	300	10	MOSIS State ID	See definition
301	310	10	CTB Use	0-9
311	311	1	Race/ Ethnicity	0 = Native American or Alaska Native 1 = Asian/Pacific Islander 3 = Black (not Hispanic) 4 = Hispanic 5 = White (not Hispanic) '-' = multi-mark ' ' = blank
312	314	3	CTB Use	''

Start	End	Length	Field	Values
315	315	1	Flag for Grade 11 Science book	1=Grade 11 science or '.'
316	316	1	CTB Use	'.'
			User Defined Data	
			Accommodation - CA	
317	317	1	01 Braille edition	Blank=Not Marked 0=Marked (Communication Arts)
318	318	1	02 Large Print edition	Blank=Not Marked 0=Marked (Communication Arts)
319	319	1	04 Oral reading – invalidates CA	Blank=Not Marked 0=Marked (Communication Arts)
320	320	1	04 Oral reading – (Blind/Partial Sight)	Blank=Not Marked 0=Marked (Communication Arts)
321	321	1	05 Signing of assessment – invalidates CA	Blank=Not Marked 0=Marked (Communication Arts)
322	322	1	06 Paraphrasing – invalidates all tests	Blank=Not Marked 0=Marked (Communication Arts)
323	323	1	10 Other Administration	Blank=Not Marked 0=Marked (Communication Arts)
324	324	1	11 Oral reading in native language – invalidates CA	Blank=Not Marked 0=Marked (Communication Arts)
325	325	1	20 Extend time– <i>TerraNova</i> session	Blank=Not Marked 0=Marked (Communication Arts)
326	326	1	21 Administer using > allotted periods	Blank=Not Marked 0=Marked (Communication Arts)
327	327	1	22 Other timing	Blank=Not Marked 0=Marked (Communication Arts)
328	328	1	35 Use of scribe	Blank=Not Marked 0=Marked (Communication Arts)
329	329	1	39 Use of calculator, math table, etc.	Blank=Not Marked 0=Marked (Communication Arts)
330	330	1	43 Use of bilingual dictionary	Blank=Not Marked 0=Marked (Communication Arts)
331	331	1	44 Other response	Blank=Not Marked 0=Marked (Communication Arts)
332	332	1	50 Testing individually	Blank=Not Marked 0=Marked (Communication Arts)
333	333	1	51 Testing in small group	Blank=Not Marked 0=Marked (Communication Arts)

Start	End	Length	Field	Values
				Blank=Not Marked 0=Marked (Communication Arts)
334	334	1	53 Other setting	
335	338	4	Blank for Future Use(4)	
			Accommodation - MA	
				Blank=Not Marked 0=Marked Mathematics
339	339	1	01 Braille edition	
				Blank=Not Marked 0=Marked Mathematics
340	340	1	02 Large Print edition	
				Blank=Not Marked 0=Marked Mathematics
341	341	1	04 Oral reading	
				Blank=Not Marked 0=Marked Mathematics
342	342	1	05 Signing of assessment	
				Blank=Not Marked 0=Marked Mathematics
343	343	1	06 Paraphrasing – invalidates all tests	
				Blank=Not Marked 0=Marked Mathematics
344	344	1	10 Other Administration	
				Blank=Not Marked 0=Marked Mathematics
345	345	1	11 Oral reading in native language	
				Blank=Not Marked 0=Marked Mathematics
346	346	1	20 Extend time– <i>TerraNova</i> session	
				Blank=Not Marked 0=Marked Mathematics
347	347	1	21 Administer using > allotted periods	
				Blank=Not Marked 0=Marked Mathematics
348	348	1	22 Other timing	
				Blank=Not Marked 0=Marked Mathematics
349	349	1	35 Use of scribe	
				Blank=Not Marked 0=Marked Mathematics
350	350	1	39 Use of calculator, math table, etc.	
				Blank=Not Marked 0=Marked Mathematics
351	351	1	43 Use of bilingual dictionary	
				Blank=Not Marked 0=Marked Mathematics
352	352	1	44 Other response	
				Blank=Not Marked 0=Marked Mathematics
353	353	1	50 Testing individually	
				Blank=Not Marked 0=Marked Mathematics
354	354	1	51 Testing in small group	
				Blank=Not Marked 0=Marked Mathematics
355	355	1	53 Other setting	
356	360	5	Blank for Future Use(5)	

Start	End	Length	Field	Values
			Accommodation - SC	
361	361	1	01 Braille edition	Blank=Not Marked 0=Marked Science
362	362	1	02 Large Print edition	Blank=Not Marked 0=Marked Science
363	363	1	04 Oral reading	Blank=Not Marked 0=Marked Science
364	364	1	05 Signing of assessment	Blank=Not Marked 0=Marked Science
365	365	1	06 Paraphrasing – invalidates all tests	Blank=Not Marked 0=Marked Science
366	366	1	10 Other Administration	Blank=Not Marked 0=Marked Science
367	367	1	11 Oral reading in native language	Blank=Not Marked 0=Marked Science
368	368	1	20 Extend time– <i>TerraNova</i> session	Blank=Not Marked 0=Marked Science
369	369	1	21 Administer using > allotted periods	Blank=Not Marked 0=Marked Science
370	370	1	22 Other timing	Blank=Not Marked 0=Marked Science
371	371	1	35 Use of scribe	Blank=Not Marked 0=Marked Science
372	372	1	39 Use of calculator, math table, etc.	Blank=Not Marked 0=Marked Science
373	373	1	43 Use of bilingual dictionary	Blank=Not Marked 0=Marked Science
374	374	1	44 Other response	Blank=Not Marked 0=Marked Science
375	375	1	50 Testing individually	Blank=Not Marked 0=Marked Science
376	376	1	51 Testing in small group	Blank=Not Marked 0=Marked Science
377	377	1	53 Other setting	Blank=Not Marked 0=Marked Science
378	392	15	Blank for Future Use	
			Teacher Invalidations	
393	393	1	Teacher Invalidation_CommArts_Session 1	Blank= No Invalidation marked 0 = Invalidated this session

Start	End	Length	Field	Values
394	394	1	Teacher Invalidation_CommArts_Session 2	Blank= No Invalidation marked 0 = Invalidated this session
395	395	1	Teacher Invalidation_CommArts_Session 3	Blank= No Invalidation marked 0 = Invalidated this session
396	396	1	Teacher Invalidation_CommArts_Session 4	Blank= No Invalidation marked 0 = Invalidated this session
397	397	1	Teacher Invalidation_Mathematics_Session 1	Blank= No Invalidation marked 0 = Invalidated this session
398	398	1	Teacher Invalidation_Mathematics_Session 2	Blank= No Invalidation marked 0 = Invalidated this session
399	399	1	Teacher Invalidation_Mathematics_Session 3	Blank= No Invalidation marked 0 = Invalidated this session
400	400	1	Teacher Invalidation_Science_Session 1	Blank= No Invalidation marked 0 = Invalidated this session
401	401	1	Teacher Invalidation_Science_Session 2	Blank= No Invalidation marked 0 = Invalidated this session
402	402	1	Teacher Invalidation_Science_Session 3	Blank= No Invalidation marked 0 = Invalidated this session
403	407	5	Blank for Future Use(5)	
			Absent in Session	
408	408	1	CA Absent Session 1	Blank= No Absent marked 0 = Absent this session
409	409	1	CA Absent Session 2	Blank= No Absent marked 0 = Absent this session
410	410	1	CA Absent Session 3	Blank= No Absent marked 0 = Absent this session
411	411	1	CA Absent Session 4	Blank= No Absent marked 0 = Absent this session
412	412	1	MA Absent Session 1	Blank= No Absent marked 0 = Absent this session
413	413	1	MA Absent Session 2	Blank= No Absent marked 0 = Absent this session
414	414	1	MA Absent Session 3	Blank= No Absent marked 0 = Absent this session
415	415	1	SC Absent Session 1	Blank= No Absent marked 0 = Absent this session
416	416	1	SC Absent Session 2	Blank= No Absent marked 0 = Absent this session
417	417	1	SC Absent Session 3	Blank= No Absent marked 0 = Absent this session
418	418	1	Filler	Blank
419	423	5	Blank for Future Use(5)	Blank
			Student BIO Information	
424	431	8	Student Barcode	0-9, blank
432	439	8	Book Lithocode	0-9, blank
440	447	8	Book Security Barcode	A-Z, 0-9, blank
448	453	6	Birth Date (MMDDYY)	MM=01-12, DD=0-3,&0-9, YY=0,8,9, & 0-9 blank, '-' = multi-mark
454	454	1	Scoring: 'I' = IRT (Communication arts, Math, Science)	'I', blank
455	456	2	Quarter-month - '30' (for TerraNova - Sessions)	30, blank
457	471	15	Last name	A - Z, a - z, blank
472	491	20	First Name	A - Z, a - z, blank
492	492	1	Middle Initial	A - Z, a - z, blank

Start	End	Length	Field	Values
493	496	4	Chronological Age (in months) - right aligned	0-9, blank
497	497	1	Gender	blank=None marked, 'F'=Female 'M'=Male, '-' = both marked
498	527	30	Blank for Future Use(30)	blank
528	545	18	Content Area Title - Communication Arts	"Communication Arts", blank
546	548	3	Content Standard # 1	Content Standard Scores - % of points earned - 000 through 100. If a student does not have a MAP score (a session not taken or invalidated), all content standards will be reported as blanks.
549	551	3	Content Standard # 2	
552	554	3	Content Standard # 3	
555	557	3	Content Standard # 4	
558	560	3	Content Standard # 5	
561	563	3	Content Standard # 6	
564	566	3	Content Standard # 7	
567	569	3	Content Standard # 8	
570	572	3	Content Standard # 9	
573	575	3	Content Standard # 10	
576	590	15	Reserved - Filler	Content Standard scores will carry leading zeros.
591	596	6	1st Process Standard reported	Process Standard scores will carry leading zeros.
597	602	6	2nd Process Standard reported	
603	608	6	3rd Process Standard reported	
609	614	6	4th Process Standard reported	
615	620	6	5th Process Standard reported	
621	626	6	6th Process Standard reported	
627	632	6	7th Process Standard reported	
633	638	6	8th Process Standard reported	
639	644	6	9th Process Standard reported	
645	650	6	10th Process Standard reported	
651	656	6	11th Process Standard reported	
657	662	6	12th Process Standard reported	
663	668	6	13th Process Standard reported	
669	674	6	14th Process Standard reported	
675	680	6	15th Process Standard reported	
681	681	1	Achievement Level for Comm. Arts	Position 1 = Goal number Positions 2 to 3 = Standard number Positions 4 to 6 = % of pts. earned
682	684	3	MAP Scale Score	0, 2, 3, 4, 5, blank. See definitions
685	687	3	CTB use - TerraNova Scale score	000-999, blank
688	691	4	CTB use - TerraNova Norm Year	2005, blank
692	693	2	TerraNova NP score	00-99, blank
694	697	4	TerraNova Lexile Scores (Comm. Arts only)	0000-9999, blank
698	698	1	MAP Test Status - Communication Arts	blank=valid test. See definitions
699	728	30	Blank for Future Use(30)	
729	746	18	Content Area Title - Mathematics	"Mathematics", blank
747	749	3	Content Standard # 1	Content Standard Scores - % of points earned - 000 through 100. If a student does not have a MAP score (a session not taken or invalidated), all content standards will be reported as blanks.
750	752	3	Content Standard # 2	
753	755	3	Content Standard # 3	
756	758	3	Content Standard # 4	
759	761	3	Content Standard # 5	
762	764	3	Content Standard # 6	
765	767	3	Content Standard # 7	
768	770	3	Content Standard # 8	
771	773	3	Content Standard # 9	
774	776	3	Content Standard # 10	
777	791	15	Content Standard # 11	Content Standard scores will carry leading zeros.
792	797	6	1st Process Standard reported	Process Standard Scores - % of pts. earned - 000 through 100;
798	803	6	2nd Process Standard reported	
804	809	6	3rd Process Standard reported	
810	815	6	4th Process Standard reported	
816	821	6	5th Process Standard reported	
				If a student does not have a MAP score (a session not taken or invalidated), all

Start	End	Length	Field	Values
822	827	6	6th Process Standard reported	process standards will be reported as
828	833	6	7th Process Standard reported	blanks including Goal and Standard
834	839	6	8th Process Standard reported	Number
840	845	6	9th Process Standard reported	
846	851	6	10th Process Standard reported	Process Standard scores will carry
852	857	6	11th Process Standard reported	leading zeros.
858	863	6	12th Process Standard reported	
864	869	6	13th Process Standard reported	Position 1 = Goal number
870	875	6	14th Process Standard reported	Positions 2 to 3 = Standard number
876	881	6	15th Process Standard reported	Positions 4 to 6 = % of pts. earned
882	882	1	Achievement Level for Math	0, 2, 3, 4, 5, blank. See definitions
883	885	3	MAP Scale Score	000-999, blank
886	888	3	CTB use - TerraNova Scale score	000-999, blank
889	892	4	CTB use - TerraNova Norm Year	2005, blank
893	894	2	TerraNova NP score	00-99, blank
895	895	1	MAP Test Status - Mathematics	blank=valid test. See definitions
896	925	30	Blank for Future Use(30)	
926	943	18	Content Area Title - Science	"Science", blank
944	946	3	Content Standard # 1	Content Standard Scores - % of points
947	949	3	Content Standard # 2	earned -
950	952	3	Content Standard # 3	000 through 100. If a student does not
953	955	3	Content Standard # 4	have a MAP score
956	958	3	Content Standard # 5	(a session not taken or invalidated), all
959	961	3	Content Standard # 6	content
962	964	3	Content Standard # 7	standards will be reported as blanks.
965	967	3	Content Standard # 8	
968	970	3	Content Standard # 9	Content Standard scores will carry
971	973	3	Content Standard # 10	leading zeros.
974	988	15	Reserved - Filler	blank
989	994	6	1st Process Standard reported	Process Standards Scores - % of pts.
995	1000	6	2nd Process Standard reported	earned - 000 through 100;
1001	1006	6	3rd Process Standard reported	
1007	1012	6	4th Process Standard reported	If a student does not have a MAP score
1013	1018	6	5th Process Standard reported	(a session not taken or invalidated), all
1019	1024	6	6th Process Standard reported	process standards will be reported as
1025	1030	6	7th Process Standard reported	blanks including Goal and Standard
1031	1036	6	8th Process Standard reported	Number
1037	1042	6	9th Process Standard reported	
1043	1048	6	10th Process Standard reported	Process Standard scores will carry
1049	1054	6	11th Process Standard reported	leading zeros.
1055	1060	6	12th Process Standard reported	
1061	1066	6	13th Process Standard reported	Position 1 = Goal number
1067	1072	6	14th Process Standard reported	Positions 2 to 3 = Standard number
1073	1078	6	15th Process Standard reported	Positions 4 to 6 = % of pts. earned
1079	1079	1	Achievement Level for Science	0, 2, 3, 4, 5, blank. See definitions
1080	1082	3	MAP Scale Score	000-999, blank
1083	1085	3	CTB use - TerraNova Scale score	000-999, blank
1086	1089	4	CTB use - TerraNova Norm Year	2005, blank
1090	1091	2	TerraNova NP score	00-99, blank
1092	1092	1	MAP Test Status - Science	blank=valid test. See definitions
1093	1122	30	Blank for Future Use(30)	blank
			Item Response	See definition
1123	1152	30	Communication Arts Session 1	
1153	1212	60	Communication Arts Session 2	
1213	1272	60	Communication Arts Session 3	
1273	1292	20	Communication Arts Session 4	
1293	1352	60	Mathematics Session 1	

Start	End	Length	Field	Values
1353	1412	60	Mathematics Session 2	
1413	1442	30	Mathematics Session 3	
1443	1482	40	Science Session 1	
1483	1542	60	Science Session 2	
1543	1602	60	Science Session 3	