Spring 2016 MAP
ELA and Mathematics Data Analysis Report
(Early Tables and Vertical Scaling)

DRAFT

DRAFT

Developed and published under contract with Missouri Department of Elementary and Secondary Education (DESE) by Data Recognition Corporation, 13490 Bass Lake Road, Maple Grove, MN 55311. Copyright © 2016 by Missouri DESE. 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/college-career-readiness/assessment/assessment-technical-support-materials. Any other use or reproduction of this document, in whole or in part, requires the prior written permission of Missouri DESE.

TABLE OF CONTENTS

CHAPTER 3: TEST DEVELOPMENT 6
CHAPTER 5: SCORING OF CONSTRUCTED-RESPONSE 11
5.1. Inter-Rater Reliability 11
CHAPTER 6: OPERATIONAL DATA ANALYSES 13
6.1. Classical Item Statistics 13
6.1.1. Test-Level Statistics 13
6.1.2. Item-Level Statistics 13
6.2. Reliability 15
6.2.1. Test Reliability 15
6.2.2. Standard Error of Measurement 16
6.3. Vertical Scaling Design 17
6.3.1 Evaluation of Student Performance on Linking Items 18
6.4. Item Response Theory 20
6.4.1. Calibration Sample 20
6.4.2. Data Calibration and Scaling 21
6.4.3. Model Fit 23
6.4.4. Vertical Scale Evaluation 24
6.5. Lowest and Highest Obtainable Scale Scores 28
6.6. Item-Pattern Scoring 29
CHAPTER 7: TEST RESULTS 73
CHAPTER 10: FAIRNESS 80
10.1. Evaluating Bias through Differential Item Functioning (DIF) Statistics 80
10.2. Evaluating Bias through Impact Analysis 83
10.2.1. Effect Size 83
10.2.2. Reliability 85
References 102
Appendix A 104
A.1. Separate calibrations and chain linking 104
A.2. ELA Scale Developed Using Below-Grade Level Linking Items 105
A.3. ELA Scale Developed Using Above-Grade Level Linking Items 107
A.4. Mathematics Scale Developed Using Below-Grade Level Linking Items 109
A.5. Mathematics Scale Developed Using Above-Grade Level Linking Items 110
A.6. Summary 111

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

LIST OF TABLES

Table 3.1: College- and Career-Readiness Item bank development activities 8
Table 3.2: CCR item types 9
Table 5.1: Inter-rater reliability: English Language Arts 12
Table 6.1: MAP Means and Standard Deviations for Raw Scores and *p*-values: English Language Arts 2016 30
Table 6.2: MAP Means and Standard Deviations for Raw Scores and *p*-values: Mathematics 2016 30
Table 6.3: Item Statistics English Language Arts Grade 3 31
Table 6.4: Item Statistics English Language Arts Grade 4 33
Table 6.5: Item Statistics English Language Arts Grade 5 35
Table 6.6: Item Statistics English Language Arts Grade 6 37
Table 6.7: Item Statistics English Language Arts Grade 7 39
Table 6.8: Item Statistics English Language Arts Grade 8 41
Table 6.9: Item Statistics Mathematics Grade 3 43
Table 6.10: Item Statistics Mathematics Grade 4 44
Table 6.11: Item Statistics Mathematics Grade 5 45
Table 6.12: Item Statistics Mathematics Grade 6 46
Table 6.13: Item Statistics Mathematics Grade 7 47
Table 6.14: Item Statistics Mathematics Grade 8 48
Table 6.15: Reliability in English Language Arts 49
Table 6.16: Reliability in Mathematics 49
Table 6.17: ELA Test Blueprint and Vertical Linking Set Content Coverage 50
Table 6.18: Mathematics Test Blueprint and Vertical Linking Set Content Coverage 51
Table 6.19: ELA Grade 3 vs. Grade 4 Vertical Linking Item Statistics 52
Table 6.20: ELA Grade 4 vs. Grade 5 Vertical Linking Item Statistics 53
Table 6.21: ELA Grade 5 vs. Grade 6 Vertical Linking Item Statistics 54
Table 6.22: ELA Grade 6 vs. Grade 7 Vertical Linking Item Statistics 55
Table 6.23: ELA Grade 7 vs. Grade 8 Vertical Linking Item Statistics 56
Table 6.24: Mathematics Grade 3 vs. Grade 4 Vertical Linking Item Statistics 57
Table 6.25: Mathematics Grade 4 vs. Grade 5 Vertical Linking Item Statistics 58
Table 6.26: Mathematics Grade 5 vs. Grade 6 Vertical Linking Item Statistics 59
Table 6.27: Mathematics Grade 6 vs. Grade 7 Vertical Linking Item Statistics 60
Table 6.28: Mathematics Grade 7 vs. Grade 8 Vertical Linking Item Statistics 61
Table 6.29: Summary of Calibration and Census Data: ELA 62
Table 6.30: Summary of Calibration and Census Data: Mathematics 64
Table 6.31: ELA population ability estimates across multiple groups on all items 66
Table 6.32: Math population ability estimates across multiple groups on all items 66
Table 6.33: Transformation constants for ELA and Mathematics base grades 66
Table 6.34: ELA items flagged for poor fit 66
Table 6.35: Mathematics items flagged for poor fit 67
Table 6.36: ELA scale score means and standard deviations 67
Table 6.37: ELA scale scores at different percentiles across grades 67
Table 6.38: Mathematics scale score means and standard deviations 67
Table 6.39: Mathematics scale scores at different percentiles across grades 68

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.40: Spring 2015 Grade 7 Mathematics scale score summary 68
Table 6.41: ELA and Mathematics Lowest and Highest Obtainable Scores 68
Table 7.1: State-Level Scale Score Statistics: English Language Arts 75
Table 7.2: State-Level Scale Score Statistics: Mathematics 75
Table 7.3: Comparison of Percentage of Students in Each Achievement Level, English Language Arts 2006 through 2016 Data 76
Table 7.4: Comparison of Percentage of Students in Each Achievement Level, Mathematics 2006 through 2016 Data 78
Table 10. 1: 2016 MAP DIF Statistics: Number of Flagged Items, English Language Arts 86
Table 10. 2: 2016 MAP DIF Statistics: Number of Flagged Items, Mathematics 87
Table 10.3: Impact Analysis, Grade 3 88
Table 10.4: Impact Analysis, Grade 4 89
Table 10.5: Impact Analysis, Grade 5 90
Table 10.6: Impact Analysis, Grade 6 91
Table 10.7: Impact Analysis, Grade 7 92
Table 10.8: Impact Analysis, Grade 8 93
Table 10.9: Grade 3 English Language Arts Reliability and SEM by Subgroup 94
Table 10.10: Grade 4 English Language Arts Reliability and SEM by Subgroup 94
Table 10.11: Grade 5 English Language Arts Reliability and SEM by Subgroup 95
Table 10.12: Grade 6 English Language Arts/Literacy Reliability and SEM by Subgroup 97
Table 10.13: Grade 7 English Language Arts/Literacy Reliability and SEM by Subgroup 97
Table 10.14: Grade 8 English Language Arts/Literacy Reliability and SEM by Subgroup 98
Table 10.15: Grade 3 Mathematics Reliability and SEM by Subgroup 99
Table 10.16: Grade 4 Mathematics Reliability and SEM by Subgroup 99
Table 10.17: Grade 5 Mathematics Reliability and SEM by Subgroup 100
Table 10.18: Grade 6 Mathematics Reliability and SEM by Subgroup 100
Table 10.19: Grade 7 Mathematics Reliability and SEM by Subgroup 101
Table 10.20: Grade 8 Mathematics Reliability and SEM by Subgroup 101

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

LIST OF FIGURES

Figure 3.1: Operational/Field Test Development 10
Figure 6.1: ELA Test Characteristic Curves 69
Figure 6.2: ELA Standard Error Curves 69
Figure 6.3: ELA Test Characteristic Curves for Grade 5 forms 70
Figure 6.4: ELA Test Characteristic Curves for Grade 5 forms 70
Figure 6.5: ELA Test Characteristic Curves for Grade 8 forms 71
Figure 6.6: ELA Test Characteristic Curves for Grade 8 forms 71
Figure 6.7: Mathematics Test Characteristic Curves 72
Figure 6.8: Mathematics Standard Error Curves 72

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

This document contains an overview of the development of the 2016 English Language Arts (ELA) and Mathematics Grades 3-8 summative assessments administered to Missouri students in Spring 2016 and highlights of the operational test data analysis. A description of item an bank from which the ELA and Mathematics test forms were selected and test from development process is outlined in Chapter 3. The early results include: inter-rater reliability statistics (Chapter 5), state-level impact data and longitudinal impact data (Chapter 7), differential item functioning analysis, scale score summaries by subgroups (gender, ethnicity, and using or not using testing accommodations), and test reliability coefficients and standard error of measurement for total population and by subgroups (Chapter 10). In addition, detailed description of methodology and results related to the development of ELA and Mathematics vertical scales are presented in Chapter 6. The chapter numbers in this document are not consecutive and they reflect the chapter numbers in a upcoming ELA and Mathematics Technical Report.

CHAPTER 3: TEST DEVELOPMENT

DRAFT

This section of the document provides a high-level overview of the development of the DRC College-and Career-Ready item bank and its use for Missouri Assessment Program English Language Arts and Mathematics assessments in Grades 3 through 8.

The ***College- and Career-Ready*** (CCR) item bank contains nationally standardized college- and career-ready items that support the next generation of standards and assessments. It is aligned to the college- and career-ready standards in mathematics and English language arts in grades 3–8 and is designed to support states like Missouri that are transitioning to more rigorous content standards, curricula, and assessments that better prepare students for college and careers.

Alignment to the College and Career Readiness standards, grade-level appropriateness, depth of knowledge (DOK), item/task level of complexity, estimated difficulty level, relevancy of context, rationale for distractors, style, accuracy, and correct terminology were major considerations in the item development process. Our item development processes for the College-and Career-Ready (CCR) item bank followed the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 2014). DRC's item development work was and continues to be designed to produce reliable and instructionally valid tests that reflect the complete range of performance articulated in the AERA, APA, NCME standards.

Furthermore, DRC's item development work for the CCR item bank adheres to the Principles of Universal Design (Thompson, Johnstone, & Thurlow, 2002) and reflects how items and tests must lend themselves to accessibility by diverse groups of students. Members of our leadership item development team have received direct training from the National Center on Educational Outcomes (NCEO). DRC employs the

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Principles of Universal Design throughout all stages of both the item development process and the test development process.

All items were reviewed for content and for fairness not only by DRC's internal experts but by a panel of external experts as well. The external reviewers have a broad range of experience in the educational field. All of the reviewers have bachelor-level, master-level, or doctoral-level degrees and teaching experience in their specific area of expertise. Table 3.1 shows the high level sequence of the activities that occurred in the development of the CCR item bank.

Various item types were developed in order to best assess students' understandings of the standards. Descriptions of each item type (in alphabetical order) used in the CCR item bank are included in Table 3.2.

In 2015, it was determined that Missouri would license ELA and MA items from DRC's College- and Career-Ready (CCR) item bank. These items would be used on the Spring 2016 Summative Assessments. Since MO had not participated in DRC's national field-test, the test design incorporated an operational/field test model. Figure 3.1 outlines the steps that were used to develop the summative assessments administered in Spring 2016.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 3.1: College- and Career-Readiness Item bank development activities

College- and Career-Readiness Item bank Development Activities
Establish item/passage development specifications and style guides, and prepare item writing training manuals.
Determine Item Development Plans
Train item writers and/or passage developers in the project requirements and specifications.
Develop passages and write items.
Review, edit, code, and track items and produce graphics.
Produce review forms for content and bias/fairness/sensitivity reviews by external reviewers.
Modify items based on external reviewers' recommendations.
Review and approve field test ready items and passages.
Develop field-test forms and administer field test.
Internal review of field-test item data.
Approve items to be included in the item bank.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

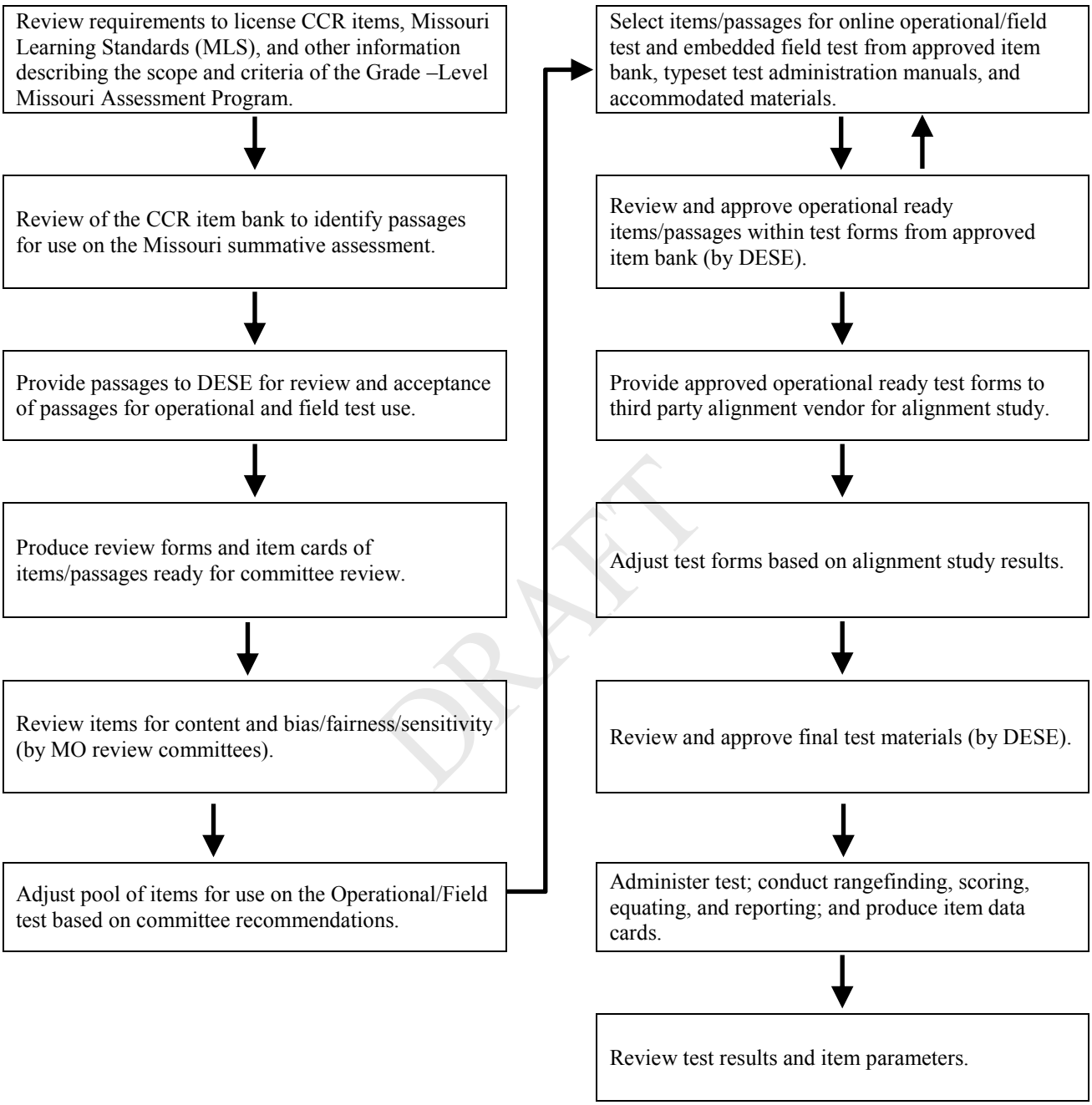
Table 3.2: CCR item types

tem Type	Name	Description
ESR	Evidence-based Selected Response	Each evidence-based selected-response item has two parts, and each two-part item is designed to elicit an evidence-based response from a student who has read a literature text passage, an informational text passage, or a writing concept. In part one, which is similar to a multiple-choice item, the student analyzes a passage or writing concept and chooses the best answer from four response options. In part two, the student elicits evidence from the passage or writing concept to select one or more answers based on the response to Part One. Each of these items is worth one point.
MC	Multiple Choice	Each multiple-choice item has four response options, only one of which is correct. Multiple-choice items are used to assess a variety of skill levels, from short-term recall of information to inference and problem solving. Each of these items is worth one point.
MS	Multiple Select	Each multiple-select item requires a student to evaluate information presented and respond by choosing two or more correct responses. Multiple-select items can be used to assess multiple skills and concepts in both mathematics and English language arts. Each of these items is worth one point.
SA	Short Answer	Each short answer item requires a student to enter a short numeric or algebraic response. These items are designed to assess a student's ability to formulate a solution to a pure or applied math problem without the assistance of response options. The short answer items are scored on a 0–1 point scale using item-specific autoscoring rules.
SCR	Short Constructed Response	Each short-constructed response item is designed to address writing through a short response as opposed to an essay. It assesses writing skills in ways a multiple-choice item cannot. The short-constructed response items are scored on a 0–2 point scale using item-specific scoring rubric.
TE	Technology Enhanced	Each technology-enhanced item is designed to elicit evidence of a broad range of student understanding. A student interacts with the enhanced features of these computer-delivered, auto-scoreable test items to show understanding of skills and concepts. Item types such as drag-and-drop, hot-spot, number line and coordinate graphing, data displays, matching interaction, and drop-down menus are just some of the technology-enhanced items presented to a student. The technology-enhanced items are scored on a 0–1 point scale using item-specific scoring rules.
WP	Writing Prompt	Each writing-prompt item is designed to assess the student's ability to write an on-demand essay response to a mode-specific task. A student demonstrates their writing skills by producing a narrative, informative/expository, or opinion/argumentative essay. The writing prompt items are scored using a three-trait holistic rubric for a total of 10 points. The rubric traits include: purpose/organization (1–4 points), development/elaboration (1–4 points), and conventions (1–2 points).

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure 3.1: Operational/Field Test Development



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

CHAPTER 5: SCORING OF CONSTRUCTED-RESPONSE

In the section of the document, the inter-rated reliability statistics for ELA Grade 5 and 8 writing prompts are provided. The writing prompts were scored by human raters. The student writing performance was evaluated on three 'traits': Conventions (maximum of 2 points), Evidence/Elaboration (maximum of 4 points), and Organization/Purpose (maximum of 4 points).

5.1. Inter-Rater Reliability

Approximately 10% of the papers in ELA Grades 5 and 8 were scored independently by a second reader. The statistics for the inter-rater reliability were calculated for all items at all grades. To determine the reliability of scoring, the percentage of perfect agreement and adjacent agreement between the two readers was examined.

For each item, a quadratic weighted kappa statistic was calculated to reflect the level of improvement beyond the chance level in the consistency of scoring. These quadratic weighted kappa values are presented in Table 5.1. To aid in the interpretation of kappa statistic, the following cutoffs have been suggested (Landis & Koch, 1977; Altman, 1991):

Kappa Value	Strength of Agreement
0	None
<0.20	Poor
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Good
0.81 – 1.00	Very Good

A total of 7 items for Grade 5 and 4 items for Grade 8 were scored by human readers across all test forms. Each item was scored on three components: Conventions, Evidence/Elaboration, and Organization/Purpose. A total of 21 item components were scored for Grade 5 and a total of 12 items components were scored for Grade 8. As shown in Table 5.1, raters demonstrated at least 99% perfect and adjacent agreement for the item component scoring. The quadratic weighted kappa values indicate that there was moderate or good inter-rater agreement for all item components except for 1 item component in Grade 5 which showed a fair agreement.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 5.1: Inter-rater reliability: English Language Arts

Grade	Form	Trait	Item #	Score Range	% Perfect	% Adjacent	% Perfect & Adjacent*	Quadratic Weighted Kappa
5	1	Conventions		0–2	77.1%	22.6%	99.7%	0.54
	1	Evidence/Elaboration	55	0–4	68.2%	30.9%	99.1%	0.67
	1	Organization/Purpose		0–4	68.3%	30.6%	99.0%	0.61
	2	Conventions		0–2	83.6%	16.1%	99.7%	0.50
	2	Evidence/Elaboration	55	0–4	69.9%	29.0%	98.9%	0.66
	2	Organization/Purpose		0–4	68.6%	31.0%	99.6%	0.60
	3	Conventions		0–2	71.1%	28.1%	99.2%	0.56
	3	Evidence/Elaboration	55	0–4	70.2%	29.0%	99.1%	0.69
	3	Organization/Purpose		0–4	70.9%	28.3%	99.2%	0.68
	4	Conventions		0–2	80.3%	19.7%	100.0%	0.61
	4	Evidence/Elaboration	55	0–4	71.4%	28.2%	99.6%	0.72
	4	Organization/Purpose		0–4	74.3%	25.3%	99.6%	0.72
	5	Conventions		0–2	72.4%	26.8%	99.2%	0.39
	5	Evidence/Elaboration	55	0–4	61.8%	36.2%	98.0%	0.64
	5	Organization/Purpose		0–4	65.0%	33.5%	98.5%	0.65
	6	Conventions		0–2	83.5%	16.5%	100.0%	0.58
	6	Evidence/Elaboration	55	0–4	73.6%	26.4%	100.0%	0.75
	6	Organization/Purpose		0–4	74.4%	25.6%	100.0%	0.74
	7	Conventions		0–2	82.8%	17.2%	100.0%	0.65
	7	Evidence/Elaboration	55	0–4	70.3%	29.4%	99.7%	0.68
	7	Organization/Purpose		0–4	78.8%	21.0%	99.7%	0.73
8	1	Conventions		0–2	72.6%	25.9%	98.5%	0.56
	1	Evidence/Elaboration	55	0–4	69.0%	30.4%	99.4%	0.69
	1	Organization/Purpose		0–4	71.0%	28.4%	99.4%	0.68
	2	Conventions		0–2	85.0%	14.8%	99.8%	0.49
	2	Evidence/Elaboration	55	0–4	63.2%	35.7%	99.0%	0.68
	2	Organization/Purpose		0–4	66.9%	31.7%	98.6%	0.69
	3	Conventions		0–2	85.4%	14.4%	99.7%	0.42
	3	Evidence/Elaboration	55	0–4	66.4%	32.5%	98.9%	0.65
	3	Organization/Purpose		0–4	66.4%	32.7%	99.1%	0.63
	4	Conventions		0–2	86.7%	13.1%	99.9%	0.45
	4	Evidence/Elaboration	55	0–4	70.1%	29.5%	99.6%	0.74
	4	Organization/Purpose		0–4	69.7%	30.0%	99.7%	0.73

* The percent perfect & adjacent may not add up to 100 due to the percent discrepant (the cases where the assigned score varied by more than 1 point).

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

CHAPTER 6: OPERATIONAL DATA ANALYSES

This section describes the analyses conducted on the Spring 2016 English Language Arts (ELA) and Mathematics Missouri Assessment Program (MAP) operational data. These analyses include a classical item analysis, examination of test reliability, and new scale development using Item Response Theory methodology. The analyses were conducted on test data available as of June 2, 2016.

6.1. Classical Item Statistics

In this section, we present summary test statistics for ELA and Mathematics. This is followed by item-level statistics for each grade/content area of MAP. These statistics were produced using sample data.

6.1.1. Test-Level Statistics

Tables 6.1 and 6.2 present the number of items and score points on each test, as well as the mean and standard deviation of the raw scores and *p*-values for each test form at each grade level of ELA and Mathematics, respectively. The mean *p*-value is the average of all item *p*-values of a specific grade/content area and it is explained in the next section.

6.1.2. Item-Level Statistics

Tables 6.3 through 6.8 present the item statistics for each item by grade for ELA. The data for Grades 5 and 8 ELA writing prompts are shown for the three components that were scored separately. Tables 6.9 through 6.14 show the item statistics for each item by grade for Mathematics. The tables include item ID, item type, item number (in data analysis), *p*-value, item-total correlation (R_{it}), omit rates and adjusted N count for each item by grade and content area. Note that technology-enhanced items and evidence-based selected response items are treated in the data analysis as constructed-response items; therefore the item type in Tables 6.3 through 6.14 is either multiple-choice (MC) or constructed-response (CR).

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 responded to an item divided by the total number of students who attempted 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 shown in Tables 6.1 and 6.2, the average *p*-values associated with the ELA forms range from 0.51 (Grade 3) to 0.63 (Grade 5 Form 1). The average *p*-values associated with the Mathematics forms range from 0.36 (Grade 8) to 0.56 (Grade 3). A trend of

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

higher mean *p*-values for lower grade levels and lower mean *p*-values for higher grade levels was observed for Mathematics.

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 the performance of students in the center of the distribution. Having a range of *p*-values also helps to prevent floor and ceiling effects so that the test does not have large numbers of students at the minimum or maximum possible scores. The ELA forms have items with *p*-values ranging from 0.06 to 0.93 (see Tables 6.3 through 6.8) across all grade levels. The *p*-values on the Mathematics forms range from 0.03 to 0.92 (see Tables 6.9 through 6.14). Such a broad range of *p*-values, which indicates the items measure well throughout the range of skills at a given grade, 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 excluded from the total score. It indicates how well an item differentiates between low- and high-achieving students. In general, items with correlations below 0.20 are said to be poorly discriminating. Over 95% of ELA items and 90% of Mathematics item in the MAP had item-test correlations above this threshold. Any item with an item-total correlation below the 0.20 threshold was further analyzed to ensure that the item was correctly keyed.

Omit Rates: The omit rate for each item indicates 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.

This examination of omit rates complies with Standard 4.14 of the AERA, APA, & NCME (2014) *Standards for educational and psychological testing*. This standard is concerned with speededness of a test:

> For a test that has a time limit, test development research should examine the degree to which scores include a speed component and should evaluate the appropriateness of that component, given the domain the test is designed to measure. (90)

The results presented in Tables 6.3 through 6.14 show that omit rates were under 2.7% for ELA and under 1.5% for Mathematics items.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

6.2. Reliability

Reliability refers to the consistency of 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 total score reliabilities are discussed in Section 6.2.1 of this document. The standard error of measurement (SEM) of the total score is discussed in Section 6.2.2.

6.2.1. Test Reliability

The reliability of raw scores by test form 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 the variance of the total 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 are, 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

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_X^2}\right], \tag{1.1}$$

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 SEM, 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 6.15 and 6.16 for ELA and Mathematics, respectively. These reliability coefficients were computed using available Missouri student sample data. The reliability statistics were 0.90 or higher for all ELA

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

forms. For Mathematics, the reliabilities ranged from 0.84 (Grade 8) to 0.91 (Grades 3 and 4). These results indicate acceptable reliability coefficients for MAP tests.

6.2.2. Standard Error of Measurement

The reliability of reported test scores can be characterized by the standard errors associated with the scores. The 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 SEM is an index of the random variability in test scores and is defined as follows:

$$\text{SEM} = SD\sqrt{1 - R_{xx'}}\ , \tag{1.2}$$

where SD represents standard deviation of the raw score distribution, and $R_{xx'}$ is estimated by $\hat{\alpha}$, as expressed in Equation 1.1.

The SEM at the test form level was computed in raw score metric and is also presented in Tables 6.15 and 6.16 for ELA and Mathematics, respectively.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

6.3. Vertical Scaling Design

A common item-linking design was implemented to facilitate Missouri vertical scale development. In this design, samples of students were administered test forms with embedded test items from adjacent grades. These off-grade level items were used for linking adjacent grades but did not contribute to the test score. Using off-grade level items for linking adjacent grades is possible because of normal overlap in content and difficulty across adjacent grades. The content of the off-grade level items conformed to the Missouri Learning Standards for each grade. The linking items were selected to ensure that the tests for all grades were anchored and continuous, and conformed to the learning standards assessed in grades 3-8 ELA and Mathematics tests.

For ELA, either 10 or 11 items below and above grade level were administered to student samples in grades 4 through 7. Grade 3 students were administered above grade level items and Grade 8 students were administered below grade level items only. The off-grade level items were administered in six different forms in grades 4-7 and in 3 different forms in grades 3-7.

For Mathematics, 10 items below and above grade level were administered to student samples in grades 4 through 7. The exception was Grade 6 in which students were administered 8 items from Grade 7. Similarly to the ELA design, Grade 3 students were administered above grade level items and Grade 8 students were administered below grade level items only. The off-grade level items were administered in four different forms in grades 4-7 and in two different forms in grades 3-7. All test forms for ELA and Mathematics were administered in a spiraled manner.

The off-grade level items that were to be administered in each grade were carefully selected and their content was matched to the on-grade operational test blueprint as close as possible while at the same time being appropriate for grades above or below.

For ELA, vertical linking items were selected from three content categories: Reading, Research, and Writing. Due to the test structure and limitations in regard to the total number of items that could be administered to students, no items were selected from the Listening part of the test.

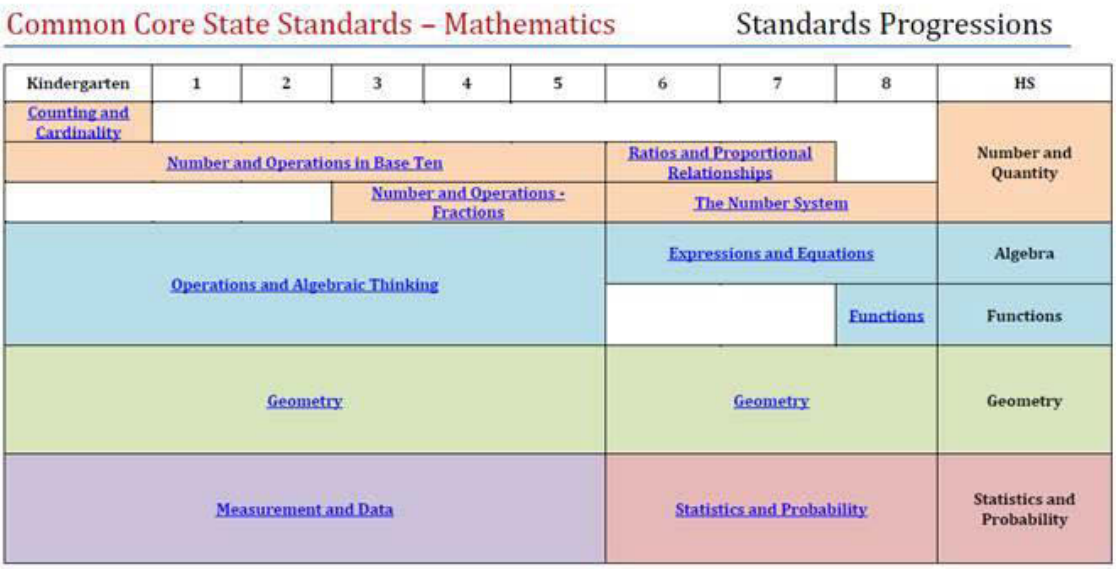
For Mathematics, with a few exceptions, the items were selected from all content categories. The exceptions were Geometry, and Statistics and Probability categories in Grade 6, and Geometry category in Grade 7. No Geometry or Statistics and Probability items from Grade 6 were selected for administration in Grade 5 because these items required students to use a calculator and calculators were not allowed in the Grade 5 test. No Geometry items from Grade 7 were selected for administration in Grade 6 because it was determined by the content experts that none of the Grade 7 Geometry items were appropriate for Grade 6.

Tables 6.17 and 6.18 show content alignment of operational on-grade level tests with off-grade level linking items across ELA and Mathematics tests, respectively. The percentage

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

of points obtainable across ELA and Mathematics strands (content categories) in the operational assessments as well as the percentage of points obtainable in the vertical linking sets are presented. It should be noted for Mathematics, that while the domain names change between Grade 5 and Grade 6, there is continuity of the construct being measured by the Mathematics assessment between all grades. The diagram below shows the progression of the mathematics concepts in the Common Core Standards environment and the continuity of the domains in mathematics. The diagram was constructed based on the work of University of Arizona, Institute for Mathematics and Education (http://math.arizona.edu/~ime/progressions/).



6.3.1 Evaluation of Student Performance on Linking Items

Classical item analysis was performed on the data used for vertical scale development. Tables 6.19 to 6.23 present the item analysis results for on-grade level operational items and the same items administered off-grade level for ELA and Tables 6.24 to 6.28 show similar item analysis results for Mathematics. The following information is provided in Tables 6.19 to 6.28: item type, item classification by test strand (or content category), item difficulty (p-value) on- and off-grade level, item-total test correlation on- and off-grade level, omit rates on- and off-grade level, and the number of students who took each item on-and off-grade level. The table headers are labeled as follows: PvalGx is the item p-value, RttGx is the item-total test correlation, OmitGx is proportion of students who omitted the item, and NobsGx is the total number of students who took the items (x is a Grade level 3, 4, 5, 6, 7, or 8 in which the item was administered).

As demonstrated by average p-values of the ELA linking sets in Tables 6.19 to 6.23, when items from adjacent grades were administered to students in a given grade level the students performed, on average, better on the items from the lower grade level than on

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

the items coming from the higher grade level. The exception to this pattern was Grade 5 students performing, on average, slightly better on Grade 6 items than Grade 6 students. When looking at the average mean item-total test correlations, the items displayed, on average, higher discrimination when administered on-grade level compared to the administration of the same items in adjacent grades. The exception was the average item-total test correlation of Grade 6 items in the vertical linking set administered to Grade 7 students which was the same in both grades.

Similar pattern was observed for Mathematics vertical linking sets (Tables 6.24 to 6.28). Students in a given grade level tended to perform better, on average, on the below-grade level Mathematics items compared to the above-grade level items. The exception was the linking set of Grade 8 items administered to Grade 7 students on which Grade 7 students performed, on average, better than Grade 8 students. Evaluation of the average item-total test correlations of the linking sets revealed that the items were more discriminating when administered on grade level compared to being administered off-grade level. The exception was the average item-total test correlation of set of Grade 3 items administered to Grade 4 students which was slightly higher for Grade 4 students.

The proportions of students who omitted linking items were very small and comparable in the on- and off-grade level administration for both ELA and Mathematics.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

6.4. Item Response Theory

Item parameter for items contained in ELA and Mathematics tests were estimated using a marginal maximum-likelihood procedure to simultaneously estimate the item parameters for MC and CR items using the 3-parameter logistic (3PL) model and 2-parameter partial credit (2PPC) IRT model (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

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

In the equation, 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-ability 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

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

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

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, 2002). 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.

Using the 3PL/2PPC model for estimation of ELA and Mathematics item parameters was consistent with the past methodology (except for administration year 2014–15) implemented for these content areas. Item parameters estimated after the 2015–16 ELA and Mathematics test administration were used to score Missouri students who took these tests.

6.4.1. Calibration Sample

In this section we describe the calibration sample in adherence to Standard 1.8 of the AERA, APA, & NCME (2014) *Standards*:

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

> The composition of any sample of test takers from which validity evidence is obtained should be described in as much detail as is practical and permissible, including major relevant socio-demographic and developmental characteristics. (25)

ELA and Mathematics test data were analyzed using calibration samples acquired after the testing window ended. The calibration samples contained close to 100% of the student data. Only a very small number of students, for which the completed test data were not available at the time of the data analysis, were excluded from the data analysis. Tables 6.29 and 6.30 show the characteristics of the calibration samples compared to the Spring 2015 census data for ELA and Mathematics, respectively. The characteristics of the Spring 2016 calibration sample were very similar to that of the Spring 2015 Missouri population.

6.4.2. Data Calibration and Scaling

The purpose of scaling a test is to enhance its validity by increasing the comparability of test takers' scores. In this section, we explicate the way in which the MAP scales are produced to comply with Standard 5.2 of the AERA, APA, & NCME (2014) *Standards*, which states the following:

> The procedures for constructing scales used for reporting scores and the rationale for these procedures should be described clearly. (102)

The MAP scores are produced using the three-parameter logistic, two-parameter partial credit (3PL/2PPC) IRT model (explained previously) 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).

Calibrating and scaling ELA and Mathematics data were performed using PARDUX software (Burket, 2002). PARDUX is designed to produce a single scale by jointly analyzing data resulting from students' responses to both MC items and CR items. In PARDUX, items are calibrated based on IRT, using the 3PL model (Lord & Novick, 1968) for MC items and the 2PPC model (Yen, 1993) for CR items.

In a process of item calibration, the number of estimation cycles was set to 200 with the convergence criterion of 0.001 for all content areas. The maximum value of *a*-parameter was set to 5.0, and the range for *b*-parameter was set between –7.5 and 7.5. For all items, the estimated *a*- and *b*-parameters were within the prescribed parameter ranges. It should be noted that there was a small number of items with the default value for the *c*-parameter on the ELA and Mathematics tests. When the PARDUX (Burket, 2002) program, used to calibrate the items, encounters difficulty estimating the *c*-parameter, it assigns a default *c*-parameter value of 0.20.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

New scales were established for ELA and Mathematics after the 2015–16 test administration. The test forms in adjacent grade levels of each content area shared common items and were calibrated concurrently at that grade level.

Concurrent calibration is a method that allows for establishing the common scale in a single step—the calibration phase—by simultaneously estimating parameters for all items at all grades. The estimated parameters in the theta metric are on the same scale. In addition, population ability estimates are obtained for multiple groups. The population mean and standard deviation for the base grade are then used to compute the M1 and M2 transformation parameters to convert the parameter estimates of the other grades onto the common scale score metric. Tables 6.31 and 6.32 present the sample mean and standard deviation ability estimates for multiple groups, as obtained from the concurrent calibration for ELA and Mathematics, respectively.

After placing item parameters on common scales for ELA and Mathematics, the Grade 5 theta means were re-estimated using only item parameters for on-grade level items. These estimates were then used to identify transformation constants that would allow transformation of item parameter estimates in a theta metric into a scale score metric and produce a scale with a mean of 500 and the standard deviation of 50 for Grade 5 of both ELA and Mathematics assessments.

DRAFT

The following formulae were used to compute transformation constants for the transformation of the base grade item parameter estimates from the theta metric to the scale score metric:

$$M1 = \frac{SD_{ss,5}}{SD_{\theta,5}}, \text{ and}$$

$$M2 = \overline{X}_5 - (\overline{\theta}_5 * M1)$$

where:
M1 and *M2* are the transformation constants
$SD_{ss,\,5}$ is the desired standard deviation in scale score metric for the base grade
$SD_{\theta,\,5}$ is the estimated standard deviation in the theta metric for the base grade
$\overline{\theta}_5$ is the estimated population mean in theta metric for the base grade
$\overline{X}_5$ is the desired mean in scale score metric for the base grade

The *M1* and *M2* transformation constants were then applied to item parameter estimates in the theta metric to transform them into scale score metric using the following formulas:

$$A_{ss} = a_\theta / M1$$
$$B_{ss} = M1 * b_\theta + M2$$
$$F_{ss} = f_\theta / M1$$
$$G_{ss} = g_\theta + (f_\theta / M1) * M2$$

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

$$C_{ss} = c_\theta$$

where:
A_{ss} is a discrimination parameter in scale score metric for MC items
B_{ss} is a difficulty parameter in scale score metric for MC items
F_{ss} is a discrimination parameter in scale score metric for CR items
G_{ss} is a difficulty for category m_j in scale score metric for CR items
a_θ is a discrimination parameter in the original theta metric for MC items
b_θ is a difficulty parameter in the original theta metric for MC items
f_θ is a discrimination parameter in the original theta metric for CR items
g_θ is a difficulty level for category m_j in the original theta metric for CR items
C_{ss} and c_θ is a guessing parameter in the original theta metric

Table 6.33 presents the initial population mean and standard deviation estimates and the transformation constants used for scale transformation of the base grade (5) for ELA and Mathematics.

Because the parameter estimates in theta metric were estimated for all grades (within each content area) and they were already on the same scale the same M1 and M2 transformation parameter constants were applied to all (grades 3 through 8) item parameter estimates.

DRAFT

6.4.3. 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. The fit index for item i is

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

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

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

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 x 4) for a given test, where N is the sample size.

Of the 337 ELA items included in the concurrent calibration, nine items were flagged for poor fit. Of the nine flagged items two items were linking items. Of the 264 Mathematics items included in the concurrent calibrations, four items were flagged for poor fit. Of these four items, one was a linking item.

Tables 6.34 and 6.35 show the chi-square statistic and the Z-statistic for each flagged item for ELA and Mathematics, respectively. The average percentage correct 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 nonzero 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 of these regions of the underlying ability distribution where there are relatively few students. The model tends to show good model-data fit for the flagged items in the middle of the theta distribution where the majority of students perform. All items flagged for poor fit were retained and contributed to student scores. Linking items flagged for poor fit were retained as part of the linking sets after determining that off-grade level item administration was not a cause of poor fit. On-grade level data calibrations were conducted and the results indicated that the flagged linking items displayed poor fit when calibrated on-grade level.

DRAFT

It is important to notice that while items may be flagged for misfit, these flags may not be of practical importance. Misfitting items that have content validity are often retained for use in one 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.

6.4.4. Vertical Scale Evaluation

In this section, the results of vertical scaling of ELA and Mathematics are described and evaluated. The scale evaluation includes examination of the pattern of grade-to-grade growth (means), grade-to-grade variability (standard deviations), and separation of scale score distributions across grades; as well as the test characteristic curves (TCC's) and standard error (SE) curves. Only on-grade level operational test items were used in computation of statistics used in scale evaluation.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

ELA Scale

Table 6.36 shows the scale score means, standard deviations, and change in mean from previous grade for ELA. As seen in Table 6.36, the ELA scale score means increase as grade level increases. The standard deviations range from 44.4 for Grade 6 to 53.5 for Grade 7 and do not show a consistent pattern across grades. The mean difference between grades is not uniform across grade levels. Most growth across grades is observed between Grades 3 and 4, followed by growth between Grades 4 and 5, and between Grades 7 and 8. Less growth is observed between Grades 5 and 6, and between Grades 6 and 7.

In addition to the evaluation of grade to grade growth using scale score mean changes across grades, the pattern of scale scores at the 10th, 25th, 50th, 75th, and 90th percentiles was examined across grades. Ideally, the scale score associated with each percentile will increase from grade to grade. Table 6.37 summarizes this information for ELA. The data in Table 6.37 show that the scale scores increase as the percentile and grade level increase showing continuous progress upward from Grades 3 through 8 at all selected percentiles except for the 10th percentile for Grades 6 and 7. Higher scale scores for Grade 6 at the lower ability end indicate that lower ability Grade 6 students may perform better on the ELA assessment compared to lower ability Grade 7 students.

Figures 6.1 and 6.2 show the TCCs and SE curves for ELA tests. In these Figures, in order to maintain the graph clarity, only one Grade 5 and Grade 8 test form (Form 1) in presented. As shown in Figure 6.1, the ELA test TCCs, with the exception of TCCs for Grades 3 and 4, are generally ordinal across grades indicating that the test difficulty increases as the grade level increases. Grade 3 and 4 TCCs are overlapping or crossing at the upper end of ability scale indicating comparable difficulty or Grade 3 test being more difficult compared to the Grade 4 test for the highest ability students. Grade 6, 7, and 8 TCCs are either overlapping or crossing at the lower end of ability scale indicating that these assessments are of comparable difficulty for the lowest ability students. However, it should be noted that even if the adjacent grade assessments are of comparable difficulty for some students, the higher grade students are of higher ability as demonstrated by the increasing scale score means across grade in Table 6.36 and increasing scale scores associated with selected percentiles (refer to Table 6.37).

The standard error curves presented in Figure 6.2 are U-shaped (as expected) indicating smaller errors around ability estimates roughly in the middle of the scale score distribution. The SE is expected to be higher at the top and bottom ends of the ability scale where fewer items measuring very high and very low achieving students are found. Overall, the standard errors around the scale score were found to be reasonable for ELA assessments.

Seven forms, each containing a different writing prompt, were administered to Grade 5 students. The Grade 5 TCCs for the seven test forms, presented in Figure 6.3, demonstrate good alignment of the form difficulty. The standard error curves for the Grade 5 test forms, presented in Figure 6.4, are also aligned.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Four forms, each containing a different writing prompt, were administered to Grade 8 students. The Grade 8 TCCs for the four test forms, presented in Figure 6.5, show good alignment of the form difficulty. The standard error curves for the Grade 8 test forms, presented in Figure 6.6, are also aligned.

Mathematics Scale

A growth pattern similar to the one described in ELA is observed for Mathematics from Grade 3 to Grade 7. The scale score means increase as the grade increases and most growth is observed between Grades 3 and 4, followed by growth between Grades 4 and 5. Less growth is observed between Grades 5 and 6, and between Grades 6 and 7. Unlike for ELA, almost no growth (as demonstrated by Grade 7 and 8 scale score means in Table 6.38) was observed between Grades 7 and Grade 8. The standard deviations ranged from 48.0 for Grade 6 to 55.9 for Grade 4.

As shown in Table 6.39, there is an upward progression of scale scores across Grades 3 through 7 and all percentiles. The evaluation of growth between Grades 7 and 8 revealed a different pattern. While grade-to-grade growth between two highest grades was observed at the 25th and the 50th percentiles, higher scale scores for Grade 7 were found at and below the 10th percentile and at and above the 75th percentiles than for Grade 8. This scale score pattern between Grades 7 and 8 indicates that while lower-to-middle ability Grade 8 students performed better on the Mathematics assessment than lower-to-middle ability Grade 7 students, the opposite may be true for the higher ability students. While this is not an expected growth pattern, a contributing factor may be the fact that the population of Grade 8 students who are administered the Mathematics tests is not fully comparable to the Grade 7 population from the previous administration year. Rather, the Grade 8 Mathematics test takers are sub-population of the Grade 7 students who took the Mathematics test in the previous administration year (Spring 2015).

Missouri Grade 8 students have an option of participating in the Algebra 1 assessment instead of the Mathematics assessment. Annually, approximately 20% of Grade 8 students participate in the Algebra 1 assessment. It was hypothesized that a sub-population of students taking the Algebra 1 test is of higher ability than the total student population for Grade 8. This hypothesis was tested through evaluation of the Grade 8 students' performance on the Grade 7 Mathematics test in the previous administration. Grade 8 student records from the Spring 2016 administration of Mathematics assessment were matched to the Grade 7 Mathematics data from the Spring 2015 administration using unique Missouri Student Information System (MOSIS) IDs which students retain as they progress from grade to grade. Following the data matching, Grade 7 Mathematics mean scale scores and scale score standard deviations were computed for the following groups of students: total population of students who took the Grade 7 Mathematics test in Spring 2015, a sub-group of students who took the Grade 7 Mathematics test in Spring 2015 and the Grade 8 Mathematics test in Spring 2016 (Matched Students), and a subgroup of students who took the Grade 7 Mathematics test in Spring 2015 but were not found in the Spring 2016 Grade 8 Mathematics data set (Unmatched Students).

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

The Grade 7 Mathematics scale score summary statistics for these three groups of students are presented in Table 6.40.

Certain limitations of matching students solely on the MOSIS ID and interpreting the data in Table 6.40 should be noted. First, it is possible that not all 'Unmatched Students' took the Algebra test in the 2015-2016 school year. Some of them might have moved out of the state or transferred to a non-public institution and, as such, they would not be in the Spring 2016 Grade 8 Mathematics test data. Second, while the Spring 2016 data calibration sample was close to 100% of the student population, the calibration sample did not contain a full 100% of the student records. A slightly higher match sample may be obtained after all Spring 2016 student records are considered.

It should also be noted that there were 3,268 unique MOSIS IDs in the Spring 2016 Grade 8 Mathematics student data that were not found in the Spring 2015 Grade 7 Mathematics student data, possibly indicating new students in Grade 8 who were either not in the Missouri public school system in Spring 2015 or did not take Grade 7 Mathematics in Spring 2015. The previous performance on the Mathematics test is not available for these students. Also, the 2015-16 Algebra 1 test data were not available to DRC to confirm which students took the Algebra test in the 2015-16 school year.

However, even given these limitations, it was observed that students who took the Grade 7 Mathematics test in Spring 2015 and the Grade 8 Mathematics test in Spring 2016 ('Matched Students') tended to be of a lower ability compared to the total population of students who took the Grade 7 Mathematics test in Spring 2015 as measured by the Spring 2015 Grade 7 Mathematics test. The mean scale score for the Spring 2016 Grade 8 students ('Matched Students') on the Spring 2015 Grade 7 test was 2509.5 scale score points while the corresponding mean for the total population of Grade 7 students was 2529.8 scale score points. The mean scale score of 'Unmatched Student's (most of whom were assumed to be taking the Algebra test in 2015-16 school year) was 2594.12 which was close to 1 standard deviation above the scale score mean of 'Matched students'.

This finding appears to support the hypothesis of attrition of higher ability Grade 8 students from MAP Mathematics. Removing higher ability students from the population of Grade 8 students taking the Mathematics test, may subsequently contribute to little or no growth between Grades 7 and 8 which was observed at the upper end of ability scale.

The Spring 2016 Grade 8 Mathematics results are also, to some degree, supported by the Spring 2015 results. No growth was observed at any of the selected percentiles between Grades 7 and 8 after the Spring 2015 Mathematics test administration; the mean scale score for these grades at all selected percentiles did not differ by more than 1 scale score point in either direction (see Table 7.12, page 162 of the Spring 2015 MAP technical report posted at https://dese.mo.gov/sites/default/files/asmt-gl-2015-tech-report.pdf).

Figures 6.7 and 6.8 show the TCCs and SE curves for Mathematics tests assessments. As observed in Figure 6.7, the TCCs for Mathematics, with the exception of Grade 5 and 6 TCCs, are ordinal indicating increasing difficulty of the assessment as the grade level

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

increases. The crossing of Grade 5 and 6 TCCs indicates that the Grade 6 assessment is still more difficult for the middle-to-higher ability students compared to the Grade 5 assessment but appears to be easier than the Grade 5 test for the lower ability students. However, it should be noted that despite Grade 5 and 6 tests being of similar difficulty, the Grade 6 students are of higher ability than the Grade 5 students (as demonstrated by the higher scale scores for Grade 6 compared to scale scores for Grade 5 at selected percentiles; see Table 6.38).

The standard error curves presented in Figure 6.8 are U-shaped (as expected) indicating smaller errors around ability estimates roughly in the middle of the scale score distribution. The SE is expected to be higher at the top and bottom ends of the ability scale where fewer items measuring these student are found. Overall, the standard errors around the scale score were found to be reasonable for Mathematics assessments.

Summary of vertical linking analysis and results

The methodology of vertical linking using concurrent calibration and all available linking items were deemed to result in most reasonable outcome for Missouri ELA and Mathematics assessments. Concurrent calibration is an efficient way of scaling multiple-group data and results in a smaller linking error compared to on-grade level separate calibrations and chain linking. Full linking set, including items from below- and above-grade level provided the students with an opportunity to demonstrate their ability on a wider range of Missouri Learning Standards compared to linking sets from only below-grade or only above-grade level items.

In summary, with a few exceptions, the increasing scale score means as the grade level increases, the upward progress of scale scores at selected percentiles, and increasing form difficulty across grade levels provides evidence of the validity of the new MAP ELA and Mathematics vertical scales.

As recommended by Missouri TAC, alternative options of ELA and Mathematics vertical linking were also attempted and their results were ultimately deemed to be overall not as good as the concurrent calibration with all linking items included. The results of the alternative ELA and Mathematics scaling approaches as well as a rationale for implementing the concurrent calibration results for MAP operational use are provided in Appendix A of this document.

6.5. 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. In addition, 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

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

called the lowest obtainable scale score (LOSS) and the highest obtainable scale score (HOSS). The LOSS and HOSS for ELA and Mathematics were set to increase as the grade level increases while minimizing the standard error around them. Table 6.41 shows the LOSS and HOSS values used for each grade of the ELA and Mathematics MAP tests.

6.6. Item-Pattern Scoring

The MAP scale scores are derived using item-pattern scoring; thus, these scale scores are based on the student's responses to all items on a given test, and scale scores account for the characteristics of the items that are in the test (such as item difficulty). A scale score can be interpreted as a highly probable estimate of a student's ability in a given content area.

Using item-pattern scoring, a student's scale score is based on the student's responses to each item (his/her item-response vector). Each item uses optimal item weights in terms of item information, meaning that items do not contribute equally to the overall scale score. Students with the same raw score may be assigned to different scale scores, depending on which items they answered correctly.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.1: MAP Means and Standard Deviations for Raw Scores and *p*-values: English Language Arts 2016

Grade	Total Items	Total Points	Mean Raw Score	Raw Score SD	Mean p-value	P-value SD	Mean R_{itt}	R_{itt} SD
3	52	52	27.12	9.98	0.53	0.19	0.37	0.09
4	52	52	31.67	10.19	0.61	0.17	0.39	0.09
5 F1	49	58	36.82	10.09	0.63	0.16	0.38	0.12
5 F2	49	58	36.67	9.88	0.63	0.16	0.37	0.12
5 F3	49	58	32.19	10.78	0.55	0.16	0.39	0.11
5 F4	49	58	37.00	10.15	0.63	0.16	0.38	0.12
5 F5	49	58	36.71	9.99	0.63	0.16	0.37	0.12
5 F6	49	58	36.45	10.12	0.62	0.16	0.38	0.12
5 F7	49	58	36.96	10.13	0.63	0.16	0.38	0.12
5 Average			35.00		0.60		0.38	
6	52	52	28.83	9.61	0.55	0.18	0.36	0.11
7	52	52	28.03	10.51	0.54	0.14	0.38	0.09
8 F1	49	58	31.71	10.79	0.54	0.19	0.40	0.10
8 F2	49	58	34.93	10.31	0.58	0.20	0.39	0.10
8 F3	49	58	34.45	10.14	0.58	0.20	0.39	0.09
8 F4	49	58	34.77	10.31	0.58	0.20	0.39	0.09
8 Average			33.47		0.56		0.39	

Note that Form 3 (F3) in Grade 5 and From 1 (F1) in Grade 8, in addition to being spiraled with other forms, were administered to students using testing accommodations. Students using accommodations tend to perform less well on the test compared to students not using accommodations resulting in lower mean raw score and mean p-value for these forms. Weighted mean raw scores, weighted mean p-values, and average item total test correlations across all Grade 5 and Grade 8 forms were computed to provide test summary statistics at the grade level for these two grades.

DRAFT

Table 6.2: MAP Means and Standard Deviations for Raw Scores and *p*-values: Mathematics 2016

Grade	Total Items	Total Points	Mean Raw Score	Raw Score SD	Mean p-value	P-value SD	Mean R_{itt}	R_{itt} SD
3	42	42	23.63	8.74	0.56	0.20	0.41	0.11
4	42	42	21.99	9.29	0.52	0.17	0.43	0.11
5	42	42	18.15	8.20	0.43	0.19	0.38	0.13
6	46	46	21.05	8.60	0.46	0.24	0.39	0.11
7	46	46	17.55	8.82	0.38	0.17	0.38	0.13
8	46	46	16.72	7.28	0.36	0.17	0.30	0.13

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.3: Item Statistics English Language Arts Grade 3

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
699752	CR	1	0.53	0.30	1.49%	68162
699751	CR	2	0.33	0.37	0.08%	69136
699754	MC	3	0.76	0.44	0.07%	69142
699750	MC	4	0.70	0.33	0.08%	69137
699749	MC	5	0.55	0.47	0.08%	69135
699976	MC	6	0.47	0.37	0.17%	69073
699984	MC	7	0.62	0.48	0.10%	69123
699980	MC	8	0.71	0.44	0.13%	69097
699981	CR	9	0.25	0.36	0.05%	69156
699977	MC	10	0.49	0.11	0.11%	69113
697843	MC	11	0.57	0.30	0.15%	69088
697836	MC	12	0.52	0.42	0.11%	69114
697835	CR	13	0.29	0.35	1.28%	68304
697834	CR	14	0.70	0.50	0.24%	69025
697840	CR	15	0.29	0.44	0.09%	69129
697477	MC	16	0.39	0.24	0.17%	69075
697483	CR	17	0.29	0.42	0.10%	69122
697481	CR	18	0.83	0.22	0.78%	68652
697484	MC	19	0.51	0.41	0.16%	69076
705683	MC	20	0.52	0.35	0.18%	69064
696655	CR	21	0.39	0.34	0.45%	68879
696833	MC	22	0.64	0.33	0.06%	69148
696656	CR	23	0.24	0.31	0.39%	68918
693011	MC	24	0.64	0.31	0.06%	69148
696794	CR	25	0.27	0.38	0.07%	69145
696580	MC	26	0.61	0.41	0.08%	69137
693016	CR	27	0.35	0.43	0.08%	69132
696692	MC	28	0.81	0.45	0.11%	69117
755160	CR	29	0.56	0.51	0.09%	69126
692599	MC	30	0.55	0.29	0.10%	69122
692629	CR	31	0.56	0.42	0.20%	69051
692589	MC	32	0.49	0.33	0.09%	69131
755162	CR	33	0.78	0.50	0.18%	69062
692628	CR	34	0.86	0.33	0.10%	69118
692626	CR	35	0.45	0.22	0.59%	68781

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.3: Item Statistics English Language Arts Grade 3 (cont.)

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
692585	MC	36	0.66	0.35	0.07%	69140
755171	CR	37	0.89	0.36	0.10%	69122
692022	CR	38	0.16	0.30	0.21%	69046
689528	MC	39	0.30	0.15	0.11%	69114
755159	CR	40	0.19	0.41	0.36%	68942
684311	MC	41	0.46	0.30	0.10%	69118
692021	CR	42	0.48	0.48	0.10%	69120
692016	MC	43	0.69	0.54	0.12%	69104
692017	MC	44	0.51	0.32	0.13%	69097
755361	CR	45	0.22	0.39	0.04%	69161
755360	MC	46	0.61	0.32	0.07%	69145
755363	CR	47	0.39	0.40	0.04%	69160
755362	MC	48	0.63	0.32	0.07%	69140
755468	MC	49	0.69	0.50	0.09%	69127
755469	MC	50	0.68	0.43	0.09%	69129
755411	CR	51	0.53	0.47	0.08%	69134
755397	CR	52	0.56	0.49	0.06%	69146

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.4: Item Statistics English Language Arts Grade 4

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
698250	MC	1	0.83	0.46	0.02%	67658
698255	CR	2	0.55	0.56	0.10%	67607
698251	MC	3	0.83	0.46	0.04%	67647
698256	CR	4	0.31	0.47	0.39%	67412
698259	MC	5	0.54	0.40	0.04%	67643
697462	MC	6	0.67	0.49	0.10%	67604
697465	MC	7	0.76	0.37	0.05%	67641
742995	MC	8	0.66	0.28	0.07%	67629
697458	CR	9	0.77	0.39	0.36%	67427
697464	MC	10	0.75	0.42	0.06%	67633
698164	MC	11	0.48	0.31	0.10%	67604
698159	CR	12	0.51	0.55	0.07%	67628
698156	MC	13	0.66	0.47	0.07%	67624
698163	MC	14	0.71	0.38	0.06%	67635
698157	MC	15	0.44	0.27	0.07%	67629
697508	MC	16	0.78	0.42	0.08%	67619
697509	MC	17	0.79	0.49	0.08%	67620
697515	MC	18	0.56	0.43	0.08%	67618
706002	CR	19	0.28	0.44	0.05%	67639
697510	MC	20	0.79	0.48	0.09%	67612
693018	MC	21	0.93	0.29	0.01%	67666
696686	CR	22	0.28	0.32	0.03%	67651
692969	MC	23	0.45	0.25	0.03%	67651
692944	CR	24	0.93	0.31	0.03%	67654
693020	MC	25	0.44	0.32	0.04%	67649
692310	CR	26	0.93	0.39	0.04%	67647
697141	CR	27	0.40	0.53	0.14%	67575
696687	MC	28	0.76	0.45	0.09%	67614
692946	MC	29	0.52	0.33	0.09%	67614
692957	MC	30	0.51	0.43	0.05%	67636
755244	CR	31	0.62	0.46	0.08%	67619
692959	MC	32	0.70	0.32	0.06%	67635
692940	MC	33	0.64	0.35	0.09%	67612
692938	MC	34	0.80	0.42	0.05%	67637
755243	CR	35	0.33	0.28	0.09%	67610

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.4: Item Statistics English Language Arts Grade 4 (cont.)

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
692954	MC	36	0.62	0.28	0.05%	67638
684817	MC	37	0.28	0.20	0.07%	67627
692934	CR	38	0.47	0.36	0.06%	67633
692309	CR	39	0.64	0.57	0.08%	67620
638903	MC	40	0.77	0.49	0.08%	67617
692973	CR	41	0.50	0.40	0.34%	67440
755247	CR	42	0.59	0.28	0.08%	67618
755242	CR	43	0.64	0.34	0.08%	67620
755248	CR	44	0.86	0.39	0.05%	67637
755466	MC	45	0.52	0.26	0.04%	67643
755421	CR	46	0.41	0.48	0.04%	67646
755422	CR	47	0.58	0.51	0.04%	67649
755475	MC	48	0.63	0.37	0.07%	67628
755399	CR	49	0.52	0.47	0.06%	67631
755385	CR	50	0.51	0.35	0.04%	67645
755476	MC	51	0.60	0.35	0.09%	67610
755477	MC	52	0.66	0.42	0.14%	67580

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.5: Item Statistics English Language Arts Grade 5

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
701082	MC	1	0.54	0.39	0.03%	66614
701089	CR	2	0.77	0.42	0.02%	66623
701084	MC	3	0.61	0.44	0.08%	66584
701085	CR	4	0.17	0.32	0.04%	66609
701088	MC	5	0.76	0.28	0.06%	66596
698286	MC	6	0.49	0.19	0.07%	66591
698289	MC	7	0.38	0.13	0.08%	66585
698285	CR	8	0.57	0.47	0.02%	66621
698287	MC	9	0.82	0.44	0.08%	66579
698280	MC	10	0.71	0.38	0.06%	66596
697380	MC	11	0.73	0.25	0.09%	66577
697379	MC	12	0.43	0.45	0.05%	66600
697377	CR	13	0.48	0.40	0.29%	66442
697382	MC	14	0.52	0.51	0.07%	66590
697384	MC	15	0.65	0.51	0.07%	66591
697726	MC	16	0.32	0.24	0.07%	66587
697733	CR	17	0.61	0.62	0.30%	66433
697725	MC	18	0.62	0.43	0.11%	66562
697729	MC	19	0.65	0.54	0.09%	66572
697728	MC	20	0.58	0.37	0.12%	66557
692568	MC	21	0.50	0.04	0.03%	66612
693014	MC	22	0.69	0.27	0.06%	66596
696587	CR	23	0.51	0.38	0.16%	66531
693023	MC	24	0.64	0.39	0.06%	66598
696615	CR	25	0.45	0.57	0.15%	66535
693027	MC	26	0.86	0.38	0.05%	66599
696652	MC	27	0.74	0.42	0.06%	66593
696658	MC	28	0.74	0.46	0.06%	66597
693123	MC	29	0.69	0.27	0.05%	66599
755174	CR	30	0.82	0.43	0.12%	66557
755178	CR	31	0.28	0.21	0.16%	66529
693131	MC	32	0.81	0.45	0.06%	66594
755175	CR	33	0.47	0.44	0.11%	66562
693132	MC	34	0.75	0.45	0.06%	66593
693154	CR	35	0.41	0.36	0.08%	66580

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.5: Item Statistics English Language Arts Grade 5 (cont.)

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
638617	MC	36	0.49	0.31	0.07%	66586
684651	MC	37	0.71	0.46	0.10%	66571
693114	CR	38	0.51	0.47	0.11%	66561
692572	CR	39	0.56	0.45	1.06%	65930
687266	MC	40	0.60	0.37	0.07%	66587
755498	MC	41	0.75	0.34	0.04%	66606
755403	CR	42	0.59	0.53	0.02%	66619
755404	CR	43	0.39	0.40	0.04%	66607
755499	MC	44	0.83	0.46	0.06%	66597
755497	MC	45	0.71	0.21	0.07%	66590
755482	MC	46	0.41	0.38	0.05%	66605
755387	CR	47	0.45	0.43	0.03%	66613
755496	MC	48	0.62	0.35	0.21%	66496
755137A	CR	49	0.64	0.53	n/a	7501
755137B	CR	50	0.64	0.52	n/a	7501
755137C	CR	51	0.85	0.49	n/a	7501
755210A	CR	52	0.60	0.46	n/a	7506
755210B	CR	53	0.62	0.46	n/a	7506
755210C	CR	54	0.91	0.42	n/a	7506
755212A	CR	55	0.59	0.56	n/a	25557
755212B	CR	56	0.59	0.55	n/a	25557
755212C	CR	57	0.77	0.49	n/a	25557
755140A	CR	58	0.65	0.56	n/a	7430
755140B	CR	59	0.64	0.57	n/a	7430
755140C	CR	60	0.84	0.47	n/a	7430
755139A	CR	61	0.64	0.52	n/a	7446
755139B	CR	62	0.64	0.51	n/a	7446
755139C	CR	63	0.86	0.44	n/a	7446
755141A	CR	64	0.62	0.52	n/a	7469
755141B	CR	65	0.63	0.53	n/a	7469
755141C	CR	66	0.90	0.45	n/a	7469
755138A	CR	67	0.66	0.54	n/a	3726
755138B	CR	68	0.67	0.52	n/a	3726
755138C	CR	69	0.83	0.51	n/a	3726

Note: WRT prompt statistics are presented separately for the three traits: A, B, and C. The omit rate for WRT item traits is not available because condition codes are given at the item level. Condition codes at the item level are converted to a score of 0 at the trait level. The omit rates for WRT prompt at the item level were inspected and they were found to be 0.25% or less in Grade 5.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.6: Item Statistics English Language Arts Grade 6

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
699161	CR	1	0.75	0.49	0.31%	63690
699160	CR	2	0.65	0.31	0.09%	63834
699162	MC	3	0.84	0.35	0.03%	63870
699157	MC	4	0.71	0.33	0.06%	63851
699163	MC	5	0.58	0.33	0.04%	63861
699175	CR	6	0.52	0.44	0.02%	63878
698192	MC	7	0.63	0.36	0.08%	63839
698190	CR	8	0.49	0.47	0.03%	63868
698191	MC	9	0.67	0.49	0.05%	63854
698196	MC	10	0.52	0.18	0.10%	63828
699811	MC	11	0.61	0.49	0.05%	63857
699808	MC	12	0.56	0.39	0.09%	63833
699812	CR	13	0.15	0.35	0.05%	63858
699814	CR	14	0.21	0.37	0.28%	63710
699807	CR	15	0.12	0.23	0.04%	63862
699809	MC	16	0.51	0.37	0.07%	63842
697793	MC	17	0.71	0.34	0.08%	63840
697787	CR	18	0.68	0.55	0.04%	63862
697788	MC	19	0.52	0.36	0.11%	63820
697786	MC	20	0.65	0.43	0.11%	63817
692715	MC	21	0.67	0.43	0.04%	63866
692713	MC	22	0.79	0.39	0.09%	63830
693031	CR	23	0.42	0.38	0.04%	63864
693029	MC	24	0.66	0.26	0.08%	63837
696919	MC	25	0.37	0.11	0.09%	63834
696799	CR	26	0.65	0.14	0.15%	63795
696677	MC	27	0.66	0.42	0.07%	63846
696834	CR	28	0.50	0.35	0.14%	63798
693911	MC	29	0.59	0.33	0.12%	63810
693930	MC	30	0.06	0.08	0.08%	63839
693928	MC	31	0.67	0.34	0.10%	63827
755180	CR	32	0.45	0.23	0.08%	63841
693909	MC	33	0.42	0.22	0.09%	63834
693923	MC	34	0.61	0.45	0.07%	63845
755188	CR	35	0.65	0.40	0.21%	63753

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.6: Item Statistics English Language Arts Grade 6 (cont.)

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
693913	MC	36	0.42	0.11	0.07%	63844
755179	CR	37	0.51	0.29	0.15%	63793
692841	MC	38	0.73	0.45	0.10%	63826
692722	CR	39	0.58	0.55	0.07%	63843
693905	CR	40	0.53	0.47	0.12%	63812
755165	CR	41	0.41	0.34	0.10%	63827
640233	MC	42	0.77	0.40	0.13%	63807
705478	CR	43	0.34	0.38	0.18%	63776
693901	CR	44	0.31	0.24	0.15%	63796
755438	CR	45	0.31	0.35	0.04%	63862
755437	CR	46	0.72	0.35	0.03%	63872
755488	MC	47	0.82	0.42	0.06%	63850
755512	MC	48	0.83	0.38	0.05%	63855
755490	MC	49	0.66	0.51	0.08%	63837
755440	CR	50	0.28	0.47	0.05%	63858
755439	CR	51	0.61	0.31	0.04%	63865
755492	MC	52	0.76	0.47	0.25%	63729

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.7: Item Statistics English Language Arts Grade 7

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
697302	CR	1	0.63	0.39	2.67%	61776
697297	MC	2	0.69	0.37	0.08%	63420
697299	MC	3	0.63	0.39	0.06%	63431
697301	MC	4	0.64	0.26	0.09%	63415
697304	CR	5	0.26	0.12	0.07%	63428
698235	MC	6	0.80	0.41	0.11%	63401
698231	MC	7	0.65	0.28	0.08%	63417
698238	CR	8	0.58	0.50	0.04%	63445
698239	CR	9	0.72	0.43	2.45%	61917
697329	MC	10	0.58	0.33	0.16%	63369
697327	MC	11	0.71	0.42	0.08%	63422
697332	CR	12	0.29	0.38	0.09%	63412
697331	MC	13	0.45	0.32	0.09%	63411
697324	CR	14	0.48	0.44	0.07%	63427
697325	MC	15	0.57	0.45	0.12%	63394
700936	MC	16	0.56	0.36	0.13%	63390
700940	CR	17	0.40	0.48	0.13%	63390
700932	CR	18	0.71	0.48	0.23%	63323
700939	MC	19	0.44	0.32	0.15%	63376
700935	MC	20	0.48	0.30	0.16%	63369
694014	CR	21	0.20	0.20	0.17%	63362
693053	MC	22	0.67	0.38	0.07%	63428
693022	CR	23	0.43	0.59	0.15%	63373
696688	CR	24	0.65	0.52	0.06%	63434
696659	MC	25	0.57	0.22	0.10%	63409
696803	CR	26	0.73	0.43	0.34%	63253
696660	MC	27	0.63	0.42	0.09%	63413
693049	CR	28	0.47	0.46	0.05%	63438
755285	CR	29	0.46	0.44	0.12%	63397
694019	MC	30	0.48	0.33	0.13%	63388
694024	MC	31	0.63	0.32	0.12%	63394
694049	CR	32	0.51	0.50	0.09%	63411
694035	MC	33	0.51	0.33	0.13%	63388
755288	CR	34	0.31	0.33	0.11%	63402
694027	MC	35	0.57	0.32	0.10%	63407

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.7: Item Statistics English Language Arts Grade 7 (cont.)

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
755290	CR	36	0.65	0.34	0.24%	63318
694055	CR	37	0.43	0.34	0.23%	63325
693160	MC	38	0.76	0.48	0.14%	63379
638620	MC	39	0.79	0.38	0.14%	63384
692845	MC	40	0.45	0.26	0.12%	63393
694010	CR	41	0.30	0.38	0.16%	63367
694054	MC	42	0.47	0.40	0.15%	63375
692844	MC	43	0.55	0.33	0.14%	63378
693159	MC	44	0.51	0.27	0.15%	63374
755444	CR	45	0.39	0.48	0.07%	63428
755445	CR	46	0.48	0.39	0.05%	63439
755446	CR	47	0.55	0.57	0.09%	63414
755531	MC	48	0.44	0.41	0.09%	63414
755378	CR	49	0.66	0.31	0.16%	63370
755430	CR	50	0.49	0.50	0.06%	63429
755456	CR	51	0.40	0.48	0.08%	63419
755518	MC	52	0.69	0.31	0.13%	63388

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.8: Item Statistics English Language Arts Grade 8

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
698986	MC	1	0.50	0.18	0.04%	63442
698987	MC	2	0.74	0.39	0.07%	63425
698992	CR	3	0.66	0.50	0.03%	63448
698995	MC	4	0.74	0.48	0.07%	63425
699080	MC	5	0.85	0.37	0.11%	63398
699087	CR	6	0.44	0.29	0.04%	63442
699084	CR	7	0.37	0.34	0.08%	63419
699081	MC	8	0.58	0.36	0.07%	63426
700925	MC	9	0.43	0.46	0.15%	63374
700929	CR	10	0.45	0.48	0.08%	63417
700930	MC	11	0.61	0.46	0.11%	63400
700927	MC	12	0.57	0.53	0.08%	63420
700924	MC	13	0.65	0.45	0.11%	63396
700922	CR	14	0.51	0.39	0.34%	63251
698855	MC	15	0.65	0.50	0.14%	63376
698857	MC	16	0.39	0.23	0.12%	63392
698862	CR	17	0.57	0.39	0.41%	63209
698860	CR	18	0.23	0.43	0.11%	63397
698856	MC	19	0.49	0.56	0.10%	63402
698853	CR	20	0.59	0.61	0.21%	63333
695629	CR	21	0.17	0.33	0.10%	63407
693318	CR	22	0.76	0.43	0.11%	63399
696830	CR	23	0.47	0.32	0.29%	63281
696828	CR	24	0.35	0.31	0.05%	63435
693311	MC	25	0.64	0.41	0.06%	63428
693312	MC	26	0.90	0.38	0.08%	63417
693316	CR	27	0.93	0.35	0.09%	63410
693061	MC	28	0.57	0.31	0.10%	63405
692970	MC	29	0.71	0.37	0.09%	63409
693313	CR	30	0.10	0.35	0.33%	63257
695634	MC	31	0.45	0.42	0.09%	63410
755270	CR	32	0.27	0.29	0.25%	63307
639837	MC	33	0.20	0.14	0.12%	63391
695655	MC	34	0.61	0.38	0.14%	63382
755267	CR	35	0.38	0.43	0.14%	63377

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.8: Item Statistics English Language Arts Grade 8 (cont.)

English Language Arts						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
755268	CR	36	0.57	0.43	0.16%	63368
695626	CR	37	0.39	0.51	0.29%	63284
755262	CR	38	0.42	0.39	0.25%	63311
695659	MC	39	0.71	0.46	0.15%	63370
695658	MC	40	0.74	0.40	0.14%	63376
755548	MC	41	0.60	0.24	0.05%	63437
755546	MC	42	0.88	0.38	0.07%	63421
787749	MC	43	0.66	0.25	0.06%	63431
755453	CR	44	0.50	0.42	0.05%	63436
755460	CR	45	0.39	0.38	0.04%	63443
755551	MC	46	0.80	0.39	0.07%	63421
755552	MC	47	0.70	0.40	0.10%	63406
755553	MC	48	0.71	0.39	0.14%	63381
755148A	CR	49	0.58	0.58	n/a	26143
755148B	CR	50	0.58	0.57	n/a	26143
755148C	CR	51	0.81	0.50	n/a	26143
755209A	CR	52	0.71	0.55	n/a	14937
755209B	CR	53	0.72	0.56	n/a	14937
755209C	CR	54	0.92	0.43	n/a	14937
755145A	CR	55	0.65	0.53	n/a	14924
755145B	CR	56	0.66	0.53	n/a	14924
755145C	CR	57	0.93	0.36	n/a	14924
755208A	CR	58	0.70	0.55	n/a	7464
755208B	CR	59	0.71	0.55	n/a	7464
755208C	CR	60	0.93	0.41	n/a	7464

Note: WRT prompt statistics are presented separately for the three traits: A, B, and C. The omit rate for WRT item traits is not available because condition codes are given at the item level. Condition codes at the item level are converted to a score of 0 at the trait level. The omit rates for WRT prompt at the item level were inspected and they were found to be smaller than 0.5% in Grade 8.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.9: Item Statistics Mathematics Grade 3

Mathematics						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
685065	MC	1	0.63	0.45	0.09%	69251
682755	MC	2	0.77	0.43	0.04%	69286
685432	MC	3	0.76	0.46	0.09%	69252
682765	MC	4	0.73	0.45	0.09%	69251
685422	MC	5	0.53	0.43	0.07%	69266
687927	CR	6	0.84	0.40	0.23%	69154
685073	MC	7	0.45	0.40	0.09%	69255
690824	MC	8	0.83	0.43	0.07%	69263
689588	CR	9	0.53	0.37	0.11%	69236
685457	MC	10	0.55	0.44	0.07%	69263
685944	MC	11	0.70	0.39	0.14%	69216
687915	CR	12	0.61	0.58	0.08%	69262
686642	MC	13	0.42	0.37	0.11%	69235
690830	MC	14	0.46	0.24	0.07%	69265
682763	MC	15	0.45	0.35	0.23%	69154
682823	MC	16	0.18	0.28	0.09%	69254
686203	MC	17	0.40	0.45	0.11%	69235
690361	CR	18	0.30	0.40	0.08%	69259
687078	MC	19	0.72	0.50	0.13%	69227
686689	MC	20	0.47	0.49	0.11%	69235
686629	MC	21	0.46	0.39	0.11%	69238
689587	CR	22	0.45	0.61	0.04%	69283
685959	MC	23	0.89	0.33	0.08%	69256
688691	MC	24	0.63	0.43	0.11%	69241
690836	MC	25	0.57	0.47	0.10%	69246
685460	MC	26	0.62	0.36	0.07%	69267
689372	CR	27	0.59	0.59	0.08%	69261
682729	MC	28	0.80	0.43	0.07%	69268
685436	MC	29	0.56	0.35	0.13%	69225
687910	MC	30	0.39	0.19	0.08%	69261
685445	MC	31	0.66	0.47	0.12%	69234
685455	MC	32	0.44	0.42	0.10%	69244
685956	MC	33	0.69	0.49	0.14%	69217
687097	MC	34	0.19	0.26	0.12%	69232
686197	MC	35	0.68	0.50	0.10%	69246
686657	MC	36	0.81	0.35	0.11%	69239
686215	MC	37	0.81	0.46	0.13%	69224
690838	MC	38	0.10	0.05	0.10%	69246
689596	CR	39	0.66	0.56	0.10%	69247
688679	MC	40	0.49	0.49	0.11%	69235
684655	CR	41	0.10	0.33	0.16%	69205
687090	MC	42	0.71	0.48	0.09%	69251

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.10: Item Statistics Mathematics Grade 4

Mathematics						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
688755	MC	1	0.32	0.47	0.06%	67830
685187	MC	2	0.45	0.29	0.06%	67826
692060	CR	3	0.57	0.34	0.07%	67821
685968	MC	4	0.67	0.53	0.07%	67819
685191	MC	5	0.73	0.44	0.05%	67836
690498	MC	6	0.62	0.51	0.05%	67835
682912	MC	7	0.75	0.39	0.04%	67839
688696	MC	8	0.38	0.32	0.07%	67822
685470	MC	9	0.61	0.40	0.08%	67816
685967	MC	10	0.47	0.60	0.06%	67830
685724	MC	11	0.41	0.30	0.09%	67808
685705	MC	12	0.42	0.36	0.09%	67807
688700	CR	13	0.32	0.47	0.05%	67833
689001	MC	14	0.50	0.61	0.08%	67818
687992	CR	15	0.51	0.56	0.07%	67821
686728	MC	16	0.41	0.45	0.11%	67796
685701	MC	17	0.60	0.47	0.14%	67773
690505	MC	18	0.63	0.32	0.11%	67793
690880	MC	19	0.69	0.40	0.13%	67783
690864	MC	20	0.66	0.40	0.07%	67824
690578	MC	21	0.44	0.47	0.09%	67811
685188	MC	22	0.88	0.32	0.09%	67810
692059	MC	23	0.65	0.42	0.04%	67842
685731	MC	24	0.47	0.50	0.07%	67823
689037	MC	25	0.44	0.60	0.10%	67798
683362	CR	26	0.60	0.48	0.06%	67829
686731	MC	27	0.69	0.40	0.06%	67831
686224	MC	28	0.31	0.36	0.05%	67832
687112	MC	29	0.40	0.29	0.11%	67797
685998	MC	30	0.52	0.58	0.07%	67820
689234	CR	31	0.29	0.47	0.06%	67828
685972	MC	32	0.74	0.43	0.06%	67830
690375	CR	33	0.07	0.34	0.72%	67383
689004	CR	34	0.43	0.49	0.07%	67822
682917	MC	35	0.72	0.44	0.06%	67830
685996	MC	36	0.35	0.53	0.07%	67821
689007	CR	37	0.49	0.62	0.10%	67801
690873	MC	38	0.26	0.21	0.06%	67827
690372	CR	39	0.57	0.52	0.09%	67806
686225	MC	40	0.61	0.42	0.09%	67811
685493	MC	41	0.54	0.09	0.10%	67803
692063	MC	42	0.82	0.36	0.07%	67821

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.11: Item Statistics Mathematics Grade 5

Mathematics						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
687156	MC	1	0.83	0.29	0.07%	66798
689654	CR	2	0.15	0.52	0.06%	66808
683206	MC	3	0.56	0.19	0.08%	66790
685513	MC	4	0.38	0.38	0.08%	66791
687140	MC	5	0.47	0.59	0.06%	66804
686234	MC	6	0.64	0.39	0.05%	66813
684498	CR	7	0.33	0.56	0.07%	66800
689395	CR	8	0.18	0.54	0.09%	66783
685746	MC	9	0.39	0.32	0.09%	66785
687133	MC	10	0.73	0.41	0.05%	66812
684648	CR	11	0.32	0.62	0.14%	66750
685194	MC	12	0.74	0.43	0.07%	66802
686764	MC	13	0.64	0.11	0.07%	66802
689107	CR	14	0.48	0.54	0.05%	66814
689405	CR	15	0.45	0.48	0.17%	66733
686008	MC	16	0.46	0.56	0.08%	66790
686253	MC	17	0.56	0.42	0.12%	66763
692651	CR	18	0.07	0.23	0.12%	66766
686283	MC	19	0.42	0.28	0.13%	66758
685740	MC	20	0.46	0.27	0.09%	66789
689634	CR	21	0.46	0.53	0.10%	66776
686256	MC	22	0.55	0.35	0.07%	66797
686003	MC	23	0.38	0.51	0.05%	66810
685755	MC	24	0.26	0.15	0.07%	66801
688787	CR	25	0.15	0.41	0.05%	66811
688030	CR	26	0.15	0.46	0.26%	66673
688799	CR	27	0.40	0.50	0.22%	66702
688776	MC	28	0.48	0.37	0.06%	66803
686005	MC	29	0.55	0.27	0.09%	66783
686011	MC	30	0.51	0.45	0.07%	66797
689656	CR	31	0.73	0.19	0.07%	66801
690383	CR	32	0.22	0.50	0.83%	66293
686257	MC	33	0.72	0.33	0.07%	66796
683186	MC	34	0.47	0.47	0.07%	66797
692707	CR	35	0.12	0.28	0.07%	66799
683011	MC	36	0.52	0.31	0.07%	66796
689401	CR	37	0.40	0.18	0.16%	66738
688813	MC	38	0.65	0.42	0.09%	66783
686766	MC	39	0.28	0.44	0.12%	66768
689400	CR	40	0.04	0.29	0.13%	66761
690532	MC	41	0.47	0.33	0.09%	66784
687153	MC	42	0.42	0.21	0.11%	66775

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.12: Item Statistics Mathematics Grade 6

Mathematics						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
688822	MC	1	0.54	0.43	0.09%	66297
688164	MC	2	0.30	0.15	0.05%	66321
686030	MC	3	0.15	0.45	0.09%	66295
685108	MC	4	0.86	0.35	0.06%	66312
688201	CR	5	0.32	0.45	0.05%	66321
685106	MC	6	0.61	0.44	0.06%	66313
686051	MC	7	0.50	0.32	0.09%	66295
692227	CR	8	0.12	0.39	0.07%	66306
688171	CR	9	0.32	0.56	0.18%	66238
685222	MC	10	0.56	0.33	0.08%	66301
688133	MC	11	0.91	0.37	0.07%	66307
687327	MC	12	0.92	0.32	0.08%	66304
686881	MC	13	0.30	0.14	0.10%	66290
692377	CR	14	0.39	0.50	0.12%	66275
683354	MC	15	0.35	0.24	0.11%	66282
684515	CR	16	0.55	0.52	0.38%	66105
687330	CR	17	0.20	0.49	0.11%	66285
685218	MC	18	0.63	0.33	0.11%	66285
687328	MC	19	0.64	0.42	0.09%	66292
685825	MC	20	0.68	0.38	0.05%	66322
683446	MC	21	0.39	0.29	0.08%	66301
687187	MC	22	0.27	0.44	0.10%	66291
685807	MC	23	0.76	0.29	0.09%	66293
689665	CR	24	0.14	0.37	0.06%	66317
683347	MC	25	0.56	0.41	0.11%	66279
685552	MC	26	0.73	0.36	0.08%	66302
685785	MC	27	0.92	0.30	0.14%	66265
688430	MC	28	0.37	0.46	0.08%	66303
688103	CR	29	0.60	0.51	0.12%	66274
687180	MC	30	0.33	0.32	0.08%	66301
689440	CR	31	0.06	0.18	0.11%	66279
689435	CR	32	0.43	0.38	0.06%	66315
688179	MC	33	0.72	0.47	0.11%	66282
687338	MC	34	0.65	0.45	0.09%	66298
686047	MC	35	0.21	0.15	0.10%	66288
686870	MC	36	0.64	0.44	0.10%	66287
688140	CR	37	0.09	0.45	0.25%	66186
687175	MC	38	0.48	0.42	0.10%	66291
688185	CR	39	0.65	0.54	0.17%	66241
687339	MC	40	0.37	0.54	0.18%	66234
688218	CR	41	0.26	0.53	0.31%	66149
685562	MC	42	0.67	0.39	0.17%	66242
688246	CR	43	0.07	0.17	0.10%	66291
689267	CR	44	0.15	0.45	0.14%	66261
690544	MC	45	0.51	0.34	0.18%	66238
688232	CR	46	0.17	0.56	0.13%	66266

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.13: Item Statistics Mathematics Grade 7

Mathematics						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
685261	MC	1	0.43	0.38	0.04%	65120
686972	MC	2	0.16	0.31	0.04%	65121
686331	MC	3	0.37	0.51	0.05%	65114
687365	CR	4	0.53	0.55	0.12%	65070
685848	MC	5	0.42	0.38	0.05%	65117
689339	MC	6	0.18	0.14	0.08%	65097
687356	CR	7	0.12	0.42	0.37%	64904
688333	CR	8	0.15	0.43	0.04%	65119
685258	MC	9	0.34	0.38	0.05%	65117
687000	MC	10	0.51	0.28	0.08%	65096
686341	MC	11	0.41	0.35	0.13%	65061
686988	MC	12	0.27	0.14	0.06%	65106
688311	CR	13	0.20	0.41	0.05%	65115
686961	MC	14	0.43	0.39	0.09%	65091
686085	MC	15	0.46	0.60	0.08%	65098
687007	MC	16	0.32	0.40	0.14%	65060
683545	MC	17	0.61	0.45	0.16%	65047
689143	MC	18	0.22	0.45	0.13%	65065
685573	MC	19	0.42	0.18	0.17%	65039
688324	CR	20	0.19	0.39	0.14%	65059
686311	MC	21	0.29	0.23	0.16%	65045
686997	MC	22	0.49	0.43	0.14%	65056
686078	MC	23	0.78	0.40	0.18%	65030
689155	MC	24	0.33	0.31	0.10%	65083
689271	CR	25	0.67	0.23	0.50%	64825
686069	MC	26	0.59	0.38	0.16%	65047
687231	MC	27	0.51	0.43	0.15%	65052
688335	CR	28	0.54	0.53	0.25%	64983
688423	CR	29	0.12	0.33	0.16%	65044
686945	MC	30	0.59	0.53	0.13%	65061
688304	CR	31	0.23	0.56	0.29%	64957
689158	MC	32	0.67	0.43	0.17%	65039
686910	MC	33	0.22	-0.04	0.13%	65062
686991	MC	34	0.43	0.48	0.15%	65051
687226	MC	35	0.53	0.30	0.17%	65040
686998	MC	36	0.29	0.44	0.24%	64993
685834	MC	37	0.55	0.23	0.21%	65008
687001	MC	38	0.67	0.26	0.23%	64997
688389	CR	39	0.32	0.52	0.24%	64990
689491	CR	40	0.08	0.41	0.27%	64972
688308	CR	41	0.47	0.22	0.51%	64819
686999	MC	42	0.31	0.40	0.22%	65003
688336	CR	43	0.18	0.56	0.59%	64762
683630	MC	44	0.50	0.50	0.20%	65016
688411	CR	45	0.25	0.62	0.51%	64816
687381	MC	46	0.25	0.29	0.18%	65028

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.14: Item Statistics Mathematics Grade 8

Mathematics						
Item ID	**Item Type**	**Item**	***p*-value**	**R_{it}**	**Omit Rate**	**Adj. N**
685609	MC	1	0.21	0.04	0.12%	52628
683744	MC	2	0.45	0.30	0.07%	52654
686166	MC	3	0.42	0.34	0.09%	52646
685138	MC	4	0.42	0.36	0.09%	52644
687396	CR	5	0.05	0.33	0.43%	52463
685124	MC	6	0.36	0.19	0.08%	52650
685295	MC	7	0.35	0.36	0.08%	52652
688592	CR	8	0.28	0.53	0.65%	52352
687028	MC	9	0.42	0.30	0.12%	52631
688506	CR	10	0.03	0.24	0.59%	52381
685610	MC	11	0.34	0.39	0.10%	52638
685267	MC	12	0.59	0.23	0.13%	52624
686117	MC	13	0.54	0.37	0.14%	52617
690390	CR	14	0.07	0.37	0.38%	52491
688567	CR	15	0.09	0.25	0.09%	52645
687233	MC	16	0.56	0.33	0.17%	52601
692725	CR	17	0.16	0.43	1.02%	52154
694848	MC	18	0.17	0.19	0.18%	52597
685273	MC	19	0.46	0.22	0.17%	52601
688537	MC	20	0.36	0.25	0.17%	52604
685881	MC	21	0.62	0.29	0.28%	52546
690598	MC	22	0.21	0.15	0.17%	52601
688620	MC	23	0.43	0.47	0.16%	52610
685643	MC	24	0.32	0.01	0.14%	52616
684669	CR	25	0.40	0.45	0.82%	52261
686139	MC	26	0.45	0.11	0.19%	52594
689682	CR	27	0.29	0.28	0.15%	52612
689677	CR	28	0.20	0.48	1.02%	52155
685635	MC	29	0.45	0.23	0.16%	52609
685896	MC	30	0.57	0.24	0.20%	52585
686159	MC	31	0.70	0.40	0.18%	52599
685626	MC	32	0.32	0.11	0.25%	52562
685296	MC	33	0.48	0.39	0.23%	52570
688499	MC	34	0.72	0.37	0.24%	52567
683999	MC	35	0.44	0.31	0.30%	52534
686375	MC	36	0.51	0.39	0.25%	52562
688606	CR	37	0.20	0.48	0.25%	52562
686113	MC	38	0.35	0.23	0.28%	52547
690911	MC	39	0.21	0.18	0.24%	52567
688647	MC	40	0.27	0.05	0.23%	52572
685297	MC	41	0.37	0.29	0.24%	52565
701237	CR	42	0.27	0.39	1.08%	52124
685898	MC	43	0.62	0.47	0.25%	52562
688454	CR	44	0.15	0.49	1.28%	52017
685886	MC	45	0.65	0.32	0.22%	52575
685291	MC	46	0.23	0.28	0.35%	52507

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.15: Reliability in English Language Arts

Grade/ Form	Number of Items	Number of Score Points	Cronbach's Alpha	SEM
3	52	52	0.90	3.09
4	52	52	0.91	3.01
5 F1	49	58	0.91	3.10
5 F2	49	58	0.90	3.08
5 F3	49	58	0.91	3.19
5 F4	49	58	0.91	3.10
5 F5	49	58	0.90	3.15
5 F6	49	58	0.91	3.09
5 F7	49	58	0.91	3.08
6	52	52	0.90	3.09
7	52	52	0.91	3.17
8 F1	49	58	0.92	3.10
8 F2	49	58	0.91	3.08
8 F3	49	58	0.91	3.05
8 F4	49	58	0.91	3.09

Note: Each of the seven Grade 5 forms and each of the four Grade 8 forms contain a different Writing prompt.

Table 6.16: Reliability in Mathematics

Grade	Number of Items	Number of Score Points	Cronbach's Alpha	SEM
3	42	42	0.91	2.66
4	42	42	0.91	2.73
5	42	42	0.89	2.71
6	46	46	0.90	2.70
7	46	46	0.90	2.82
8	46	46	0.84	2.88

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.17: ELA Test Blueprint and Vertical Linking Set Content Coverage

Grade	Item Set	Content Categories (Strands)				Total
		Reading	Research	Writing	Listening	
3	OP	38%	15%	31%	15%	100%
3	VS set in G4	45%	18%	36%	0	100%
4	OP	38%	15%	31%	15%	100%
4	VS set in G3	45%	18%	36%	0	100%
4	VS set in G5	40%	20%	40%	0	100%
5	OP	34%	14%	38%	14%	100%
5	VS set in G4	40%	20%	40%	0	100%
5	VS set in G6	45%	18%	36%	0	100%
6	OP	38%	15%	31%	15%	100%
6	VS set in G5	40%	20%	40%	0	100%
6	VS set in G7	43%	19%	38%	0	100%
7	OP	38%	15%	31%	15%	100%
7	VS set in G6	40%	20%	40%	0	100%
7	VS set in G8	40%	20%	40%	0	100%
8	OP	34%	14%	38%	14%	100%
8	VS set in G7	40%	20%	40%	0	100%

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.18: Mathematics Test Blueprint and Vertical Linking Set Content Coverage

Grade	Item Set	Content Categories										Total
		OA	NBT	NF	MD	GE	RP	NS	EE	SP	FN	
3	OP	36%	14%	17%	24%	10%						100%
3	VS set in G4	30%	10%	20%	30%	10%						100%
4	OP	24%	19%	29%	19%	10%						100%
4	VS set in G3	20%	20%	30%	20%	10%						100%
4	VS set in G5	20%	20%	30%	20%	10%						100%
5	OP	14%	19%	38%	17%	12%						100%
5	VS set in G4	20%	20%	30%	20%	10%						100%
5	VS set in G6	20%	20%	30%	20%	10%						100%
6	OP					13%	13%	28%	33%	13%		100%
6	VS set in G5						20%	40%	40%			100%
6	VS set in G7					10%	20%	20%	40%	10%		100%
7	OP					15%	22%	17%	28%	17%		100%
7	VS set in G6						13%	37%	37%	13%		100%
7	VS set in G8					20%	20%	20%	20%	20%		100%
8	OP					24%		9%	33%	13%	22%	100%
8	VS set in G7					30%		10%	30%	10%	20%	100%

Note: Content categories are as follows: OA=Operations and Algebraic Thinking; NBT= Numbers and Operations in Base Ten; NF= Numbers and Operations - Fractions; MD=Measurement and Data; GE=Geometry; RP= Ratios and Proportional Relationships; NS=The Number System; EE=Expressions and Equations; SP=Statistics and Probability; and F=Functions.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.19: ELA Grade 3 vs. Grade 4 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	Item Statistics in Administration Grade							
			PvalG3	PvalG4	RttG3	RttG4	OmitG3	OmitG4	NobsG3	NobsG4
3	TE	Writing	0.48	0.70	0.48	0.36	0.00	0.00	69120	7307
3	MC	Writing	0.49	0.58	0.33	0.34	0.00	0.00	69131	24810
3	TE	Writing	0.56	0.71	0.42	0.35	0.00	0.00	69051	7341
3	MC	Research	0.81	0.89	0.45	0.38	0.00	0.00	69117	24808
3	MC	Research	0.64	0.78	0.33	0.31	0.00	0.00	69148	7309
3	MC	Reading	0.47	0.52	0.37	0.33	0.00	0.00	69073	32096
3	MC	Reading	0.49	0.50	0.11	0.08	0.00	0.00	69113	24796
3	MC	Reading	0.71	0.76	0.44	0.43	0.00	0.00	69097	32099
3	ESR	Reading	0.25	0.39	0.36	0.39	0.00	0.00	69156	7307
3	MC	Reading	0.62	0.61	0.48	0.47	0.00	0.00	69123	32092
3	TE	Writing	0.19	0.40	0.42	0.44	0.00	0.00	68942	7340
3		AVERAGE	0.52	0.62	0.38	0.35	0.00	0.00		
4	MC	Writing	0.67	0.77	0.44	0.49	0.00	0.00	13571	67617
4	MC	Writing	0.72	0.80	0.39	0.42	0.00	0.00	13496	67637
4	TE	Research	0.88	0.93	0.33	0.31	0.00	0.00	13570	67654
4	MC	Research	0.84	0.93	0.34	0.29	0.00	0.00	32493	67666
4	MC	Reading	0.45	0.66	0.33	0.47	0.00	0.00	32451	67624
4	MC	Reading	0.40	0.44	0.21	0.27	0.00	0.00	13557	67629
4	TE	Reading	0.29	0.51	0.47	0.55	0.00	0.00	46010	67628
4	MC	Reading	0.54	0.71	0.32	0.38	0.00	0.00	46000	67635
4	MC	Reading	0.32	0.48	0.21	0.31	0.00	0.00	45988	67604
4	TE	Writing	0.65	0.64	0.25	0.34	0.00	0.00	13487	67620
4	TE	Writing	0.50	0.62	0.36	0.46	0.00	0.00	32438	67619
4		AVERAGE	0.57	0.68	0.33	0.39	0.00	0.00		

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.20: ELA Grade 4 vs. Grade 5 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG4	PvalG5	RttG4	RttG5	OmitG4	OmitG5	NobsG4	NobsG5
4	TE	Writing	0.64	0.80	0.57	0.47	0.00	0.00	67620	7421
4	MC	Writing	0.51	0.65	0.43	0.44	0.00	0.00	67636	7467
4	TE	Writing	0.51	0.64	0.40	0.31	0.00	0.00	67440	7454
4	MC	Research	0.44	0.45	0.32	0.28	0.00	0.00	67649	23210
4	MC	Research	0.77	0.84	0.45	0.42	0.00	0.00	67614	7426
4	MC	Reading	0.67	0.64	0.49	0.46	0.00	0.00	67604	30621
4	MC	Reading	0.75	0.73	0.42	0.43	0.00	0.00	67633	30621
4	MC	Reading	0.76	0.73	0.37	0.38	0.00	0.00	67641	30620
4	MC	Reading	0.66	0.66	0.28	0.32	0.00	0.00	67629	30628
4	TE	Writing	0.33	0.33	0.28	0.26	0.00	0.00	67610	23185
4		AVERAGE	0.60	0.65	0.40	0.38	0.00	0.00		
5	MC	Writing	0.45	0.49	0.19	0.31	0.00	0.00	7326	66586
5	MC	Research	0.64	0.69	0.21	0.27	0.00	0.00	7236	66596
5	MC	Research	0.56	0.64	0.34	0.39	0.00	0.00	7301	66598
5	MC	Research	0.73	0.81	0.41	0.45	0.00	0.00	7328	66594
5	TE	Writing	0.37	0.41	0.22	0.36	0.00	0.00	7300	66580
5	MC	Reading	0.55	0.62	0.37	0.43	0.00	0.00	14537	66562
5	MC	Reading	0.28	0.32	0.16	0.24	0.00	0.00	14538	66587
5	MC	Reading	0.54	0.58	0.28	0.37	0.00	0.00	7300	66557
5	MC	Reading	0.60	0.66	0.48	0.54	0.00	0.00	7236	66572
5	TE	Reading	0.45	0.61	0.55	0.62	0.00	0.00	14499	66433
5	TE	Writing	0.81	0.82	0.38	0.43	0.00	0.00	7238	66557
5		AVERAGE	0.54	0.60	0.32	0.40	0.00	0.00		

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.21: ELA Grade 5 vs. Grade 6 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG5	PvalG6	RttG5	RttG6	OmitG5	OmitG6	NobsG5	NobsG6
5	MC	Writing	0.71	0.86	0.46	0.32	0.00	0.00	66571	7866
5	TE	Writing	0.51	0.66	0.47	0.42	0.00	0.00	66561	7865
5	MC	Writing	0.75	0.84	0.45	0.40	0.00	0.00	66593	18208
5	TE	Research	0.51	0.48	0.38	0.29	0.00	0.00	66531	18200
5	MC	Research	0.74	0.74	0.42	0.30	0.00	0.00	66593	7816
5	TE	Reading	0.48	0.48	0.41	0.40	0.00	0.00	66442	25961
5	MC	Reading	0.43	0.41	0.45	0.43	0.00	0.00	66600	18197
5	MC	Reading	0.73	0.75	0.25	0.24	0.00	0.00	66577	26013
5	MC	Reading	0.52	0.52	0.51	0.48	0.00	0.00	66590	26012
5	MC	Reading	0.65	0.66	0.51	0.49	0.00	0.00	66591	7813
5	TE	Writing	0.47	0.58	0.44	0.33	0.00	0.00	66562	7818
5		AVERAGE	0.59	0.64	0.43	0.37	0.00	0.00		
6	MC	Research	0.76	0.79	0.46	0.39	0.00	0.00	7501	63830
6	ESR	Writing	0.53	0.58	0.44	0.55	0.00	0.00	7443	63843
6	TE	Writing	0.62	0.53	0.46	0.48	0.00	0.00	7496	63812
6	MC	Writing	0.45	0.42	0.07	0.11	0.00	0.00	7503	63844
6	MC	Writing	0.05	0.06	0.04	0.08	0.00	0.00	7442	63839
6	MC	Research	0.76	0.67	0.46	0.42	0.00	0.00	7495	63846
6	MC	Reading	0.60	0.65	0.41	0.43	0.00	0.00	14991	63817
6	ESR	Reading	0.69	0.68	0.53	0.55	0.00	0.00	15001	63862
6	MC	Reading	0.50	0.52	0.30	0.36	0.00	0.00	14994	63820
6	MC	Reading	0.74	0.71	0.38	0.34	0.00	0.00	14994	63840
6		AVERAGE	0.57	0.56	0.35	0.37	0.00	0.00		

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.22: ELA Grade 6 vs. Grade 7 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG6	PvalG7	RttG6	RttG7	OmitG6	OmitG7	NobsG6	NobsG7
6	MC	Writing	0.77	0.85	0.40	0.34	0.00	0.00	63807	7953
6	MC	Research	0.67	0.66	0.43	0.41	0.00	0.00	63866	17675
6	TE	Writing	0.31	0.42	0.24	0.23	0.00	0.00	63796	7973
6	MC	Writing	0.61	0.77	0.45	0.42	0.00	0.00	63845	7987
6	TE	Research	0.50	0.52	0.35	0.30	0.00	0.00	63798	7937
6	MC	Reading	0.71	0.61	0.33	0.38	0.00	0.00	63851	25608
6	TE	Reading	0.65	0.57	0.31	0.36	0.00	0.00	63834	17647
6	TE	Reading	0.75	0.71	0.49	0.50	0.00	0.02	63690	25101
6	MC	Reading	0.84	0.77	0.35	0.38	0.00	0.00	63870	25612
6	MC	Reading	0.58	0.56	0.33	0.33	0.00	0.00	63861	7949
6	TE	Writing	0.45	0.48	0.23	0.19	0.00	0.00	63841	17674
6		AVERAGE	0.62	0.63	0.35	0.35	0.00	0.00		
7	MC	Writing	0.84	0.79	0.31	0.38	0.00	0.00	7836	63384
7	TE	Writing	0.31	0.30	0.30	0.38	0.00	0.00	7852	63367
7	MC	Writing	0.58	0.63	0.19	0.32	0.00	0.00	7853	63394
7	MC	Research	0.61	0.57	0.21	0.22	0.00	0.00	7829	63409
7	TE	Research	0.66	0.73	0.46	0.43	0.00	0.00	7836	63253
7	MC	Reading	0.50	0.65	0.27	0.28	0.00	0.00	15655	63417
7	MC	Reading	0.69	0.80	0.42	0.41	0.00	0.00	15663	63401
7	ESR	Reading	0.45	0.58	0.50	0.50	0.00	0.00	15667	63445
7	TE	Reading	0.68	0.72	0.40	0.44	0.01	0.02	15601	61917
7	TE	Writing	0.62	0.65	0.26	0.34	0.00	0.00	7823	63318
7		AVERAGE	0.59	0.64	0.33	0.37	0.00	0.00		

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.23: ELA Grade 7 vs. Grade 8 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG7	PvalG8	RttG7	RttG8	OmitG7	OmitG8	NobsG7	NobsG8
7	MC	Writing	0.45	0.54	0.26	0.21	0.00	0.00	63393	14921
7	TE	Research	0.20	0.22	0.20	0.24	0.00	0.00	63362	14918
7	MC	Writing	0.48	0.51	0.34	0.34	0.00	0.00	63388	14916
7	MC	Writing	0.51	0.56	0.33	0.22	0.00	0.00	63388	24211
7	MC	Writing	0.47	0.59	0.40	0.39	0.00	0.00	63375	14900
7	MC	Research	0.63	0.70	0.42	0.41	0.00	0.00	63413	24236
7	MC	Reading	0.57	0.59	0.45	0.35	0.00	0.00	63394	39138
7	MC	Reading	0.71	0.71	0.42	0.38	0.00	0.00	63422	39140
7	MC	Reading	0.58	0.58	0.33	0.26	0.00	0.00	63369	39146
7	TE	Reading	0.29	0.31	0.38	0.37	0.00	0.00	63412	39151
7		AVERAGE	0.49	0.53	0.35	0.32	0.00	0.00		
8	MC	Writing	0.67	0.71	0.29	0.37	0.00	0.00	7957	63409
8	MC	Research	0.56	0.64	0.36	0.41	0.00	0.00	7948	63428
8	TE	Research	0.93	0.93	0.31	0.35	0.00	0.00	7955	63410
8	TE	Writing	0.38	0.39	0.43	0.51	0.00	0.00	7939	63284
8	MC	Writing	0.60	0.61	0.33	0.38	0.00	0.00	7916	63382
8	MC	Writing	0.77	0.74	0.33	0.40	0.00	0.00	7914	63376
8	MC	Reading	0.40	0.51	0.13	0.18	0.00	0.00	15895	63442
8	MC	Reading	0.64	0.74	0.42	0.40	0.00	0.00	15888	63425
8	ESR	Reading	0.53	0.66	0.52	0.50	0.00	0.00	15900	63448
8	MC	Reading	0.63	0.75	0.49	0.48	0.00	0.00	15874	63425
8		AVERAGE	0.61	0.67	0.36	0.40	0.00	0.00		

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.24: Mathematics Grade 3 vs. Grade 4 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG3	PvalG4	RttG3	RttG4	OmitG3	OmitG4	NobsG3	NobsG4
3	MC	NF	0.45	0.67	0.35	0.44	0.00	0.00	69154	10010
3	MC	MD	0.18	0.28	0.28	0.40	0.00	0.00	69254	10005
3	MC	NBT	0.45	0.42	0.40	0.39	0.00	0.00	69255	29274
3	MC	GE	0.53	0.58	0.43	0.39	0.00	0.00	69266	29270
3	MC	NF	0.89	0.91	0.33	0.24	0.00	0.00	69256	29255
3	MC	OA	0.40	0.59	0.45	0.39	0.00	0.00	69235	10011
3	MC	MD	0.47	0.52	0.39	0.39	0.00	0.00	69238	29285
3	MC	OA	0.42	0.53	0.37	0.40	0.00	0.00	69235	10009
3	MC	MD	0.19	0.27	0.26	0.39	0.00	0.00	69232	10011
3	MC	OA	0.63	0.72	0.43	0.39	0.00	0.00	69241	29275
3		AVERAGE	0.46	0.55	0.37	0.38	0.00	0.00		
4	MC	OA	0.59	0.72	0.27	0.44	0.00	0.00	18131	67830
4	MC	OA	0.40	0.45	0.16	0.30	0.00	0.00	38557	67826
4	MC	NBT	0.79	0.73	0.42	0.44	0.00	0.00	18131	67836
4	MC	NF	0.36	0.67	0.41	0.53	0.00	0.00	38523	67819
4	MC	NF	0.36	0.74	0.11	0.43	0.00	0.00	18131	67830
4	MC	NF	0.40	0.52	0.43	0.58	0.00	0.00	18129	67820
4	MC	MD	0.24	0.31	0.16	0.36	0.00	0.00	18130	67832
4	MC	NBT	0.14	0.32	0.25	0.47	0.00	0.00	38545	67830
4	MC	MD	0.51	0.62	0.49	0.51	0.00	0.00	38559	67835
4	MC	GE	0.28	0.65	0.13	0.42	0.00	0.00	38543	67842
4		AVERAGE	0.41	0.57	0.28	0.45	0.00	0.00		

Note: Content categories are as follows: OA=Operations and Algebraic Thinking; NBT= Numbers and Operations in Base Ten; NF= Numbers and Operations - Fractions; MD=Measurement and Data; GE=Geometry; RP= Ratios and Proportional Relationships; NS=The Number System; EE=Expressions and Equations; SP=Statistics and Probability; and F=Functions.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.25: Mathematics Grade 4 vs. Grade 5 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG4	PvalG5	RttG4	RttG5	OmitG4	OmitG5	NobsG4	NobsG5
4	MC	NBT	0.60	0.71	0.47	0.40	0.00	0.00	67773	27924
4	MC	NBT	0.42	0.53	0.36	0.34	0.00	0.00	67807	10187
4	MC	MD	0.47	0.52	0.50	0.44	0.00	0.00	67823	27933
4	MC	NF	0.47	0.60	0.60	0.51	0.00	0.00	67830	10188
4	MC	NF	0.35	0.40	0.53	0.56	0.00	0.00	67821	27903
4	MC	MD	0.61	0.63	0.42	0.24	0.00	0.00	67811	10193
4	MC	OA	0.41	0.41	0.29	0.38	0.00	0.00	67797	27930
4	MC	GE	0.38	0.39	0.32	0.31	0.00	0.00	67822	27922
4	MC	NF	0.44	0.57	0.60	0.49	0.00	0.00	67798	10185
4	MC	OA	0.26	0.32	0.21	0.23	0.00	0.00	67827	10194
4		AVERAGE	0.44	0.51	0.43	0.39	0.00	0.00		
5	MC	OA	0.47	0.52	0.25	0.31	0.00	0.00	10122	66796
5	MC	NF	0.33	0.47	0.19	0.47	0.00	0.00	10119	66797
5	MC	MD	0.59	0.56	0.06	0.19	0.00	0.00	9979	66790
5	MC	GE	0.49	0.46	0.18	0.27		0.00	9982	66789
5	MC	MD	0.33	0.39	0.16	0.32	0.00	0.00	10118	66785
5	MC	NF	0.55	0.55	0.13	0.27	0.00	0.00	9974	66783
5	MC	NF	0.35	0.47	0.49	0.56	0.00	0.00	10120	66790
5	MC	OA	0.39	0.51	0.31	0.45	0.00	0.00	9976	66797
5	MC	NBT	0.54	0.64	0.33	0.39	0.00	0.00	9979	66813
5	MC	NBT	0.82	0.83	0.30	0.29	0.00	0.00	10117	66798
5		AVERAGE	0.49	0.54	0.24	0.35	0.00	0.00		

Note: Content categories are as follows: OA=Operations and Algebraic Thinking; NBT= Numbers and Operations in Base Ten; NF= Numbers and Operations - Fractions; MD=Measurement and Data; GE=Geometry; RP= Ratios and Proportional Relationships; NS=The Number System; EE=Expressions and Equations; SP=Statistics and Probability; and F=Functions.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.26: Mathematics Grade 5 vs. Grade 6 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG5	PvalG6	RttG5	RttG6	OmitG5	OmitG6	NobsG5	NobsG6
5	MS	OA	0.33	0.31	0.56	0.53	0.00	0.00	66800	24592
5	TE	NF	0.32	0.27	0.62	0.50	0.00	0.00	66750	11261
5	MC	OA	0.74	0.81	0.43	0.38	0.00	0.00	66802	11260
5	MC	MD	0.28	0.27	0.44	0.40	0.00	0.00	66768	24564
5	MC	NF	0.42	0.48	0.21	0.29	0.00	0.00	66775	11261
5	MC	NBT	0.65	0.72	0.42	0.39	0.00	0.00	66783	24579
5	SA	NBT	0.45	0.52	0.48	0.39	0.00	0.00	66733	24565
5	MS	NF	0.15	0.16	0.52	0.46	0.00	0.00	66808	11266
5	TE	MD	0.22	0.18	0.50	0.40	0.01	0.02	66293	24187
5	MS	GE	0.12	0.12	0.28	0.23	0.00	0.00	66799	11270
5		AVERAGE	0.37	0.38	0.45	0.40	0.00	0.00		
6	MC	EE	0.23	0.35	-0.04	0.24	0.00	0.00	10199	66282
6	MC	NS	0.54	0.61	0.36	0.45	0.00	0.00	10279	66313
6	MC	NS	0.87	0.86	0.31	0.35	0.00	0.00	10281	66312
6	MC	EE	0.69	0.63	0.16	0.33	0.00	0.00	10195	66285
6	MC	RP	0.43	0.50	0.21	0.32	0.00	0.00	10201	66295
6	MC	NS	0.96	0.93	0.20	0.32	0.00	0.00	10203	66304
6	MC	NS	0.67	0.64	0.33	0.42	0.00	0.00	10201	66292
6	MC	RP	0.93	0.91	0.27	0.37	0.00	0.00	10281	66307
6	MS	EE	0.26	0.32	0.38	0.45	0.00	0.00	10281	66321
6	MC	EE	0.35	0.54	0.24	0.43	0.00	0.00	10282	66297
6		AVERAGE	0.59	0.63	0.24	0.37	0.00	0.00		

Note: Content categories are as follows: OA=Operations and Algebraic Thinking; NBT= Numbers and Operations in Base Ten; NF= Numbers and Operations - Fractions; MD=Measurement and Data; GE=Geometry; RP= Ratios and Proportional Relationships; NS=The Number System; EE=Expressions and Equations; SP=Statistics and Probability; and F=Functions.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.27: Mathematics Grade 6 vs. Grade 7 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG6	PvalG7	RttG6	RttG7	OmitG6	OmitG7	NobsG6	NobsG7
6	MC	NS	0.56	0.51	0.41	0.34	0.00	0.00	66279	23898
6	MC	EE	0.39	0.35	0.29	0.25	0.00	0.00	66301	23890
6	MC	RP	0.15	0.12	0.45	0.40	0.00	0.00	66295	23922
6	MC	RP	0.31	0.30	0.14	0.05	0.00	0.00	66290	11246
6	MC	EE	0.33	0.45	0.32	0.34	0.00	0.00	66301	11240
6	MC	NS	0.37	0.41	0.54	0.50	0.00	0.00	66234	11236
6	SA	GE	0.09	0.11	0.45	0.51	0.00	0.01	66186	23776
6	SA	EE	0.26	0.42	0.53	0.46	0.00	0.01	66149	11188
6	MS	SP	0.06	0.07	0.18	0.20	0.00	0.00	66279	11236
6	MS	EE	0.14	0.17	0.37	0.42	0.00	0.00	66317	23917
6		AVERAGE	0.27	0.29	0.37	0.35	0.00	0.00		
7	MC	RP	0.59	0.61	0.37	0.46	0.00	0.00	22416	65047
7	MC	NS	0.20	0.34	0.18	0.38	0.00	0.00	11212	65117
7	MC	NS	0.47	0.43	0.35	0.38	0.00	0.00	11199	65120
7	MC	SP	0.62	0.55	0.19	0.23	0.00	0.00	22414	65008
7	MC	EE	0.36	0.27	0.10	0.14	0.00	0.00	11213	65106
7	MC	EE	0.46	0.51	0.12	0.28	0.00	0.00	11203	65096
7	MS	EE	0.08	0.15	0.33	0.43	0.00	0.00	11199	65119
7	MC	NS	0.21	0.18	0.09	0.14	0.00	0.00	11213	65097
7		AVERAGE	0.37	0.38	0.22	0.30	0.00	0.00		

Note: Content categories are as follows: OA=Operations and Algebraic Thinking; NBT= Numbers and Operations in Base Ten; NF= Numbers and Operations - Fractions; MD=Measurement and Data; GE=Geometry; RP= Ratios and Proportional Relationships; NS=The Number System; EE=Expressions and Equations; SP=Statistics and Probability; and F=Functions.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.28: Mathematics Grade 7 vs. Grade 8 Vertical Linking Item Statistics

Item Grade	Item Type	Strand	PvalG7	PvalG8	RttG7	RttG8	OmitG7	OmitG8	NobsG7	NobsG8
7	MC	SP	0.50	0.57	0.50	0.37	0.00	0.00	65016	16144
7	MC	RP	0.42	0.40	0.18	0.10	0.00	0.00	65039	26676
7	MC	EE	0.31	0.34	0.40	0.30	0.00	0.00	65003	16156
7	MC	GE	0.32	0.39	0.40	0.38	0.00	0.00	65060	16146
7	MC	GE	0.53	0.51	0.30	0.14	0.00	0.00	65040	26668
7	SA	NS	0.12	0.08	0.42	0.26	0.00	0.01	64904	26540
7	MS	NS	0.20	0.16	0.42	0.24	0.00	0.00	65115	16165
7	SA	RP	0.25	0.27	0.62	0.47	0.01	0.01	64816	16063
7	MS	EE	0.12	0.10	0.33	0.22	0.00	0.00	65044	26682
7	MS	SP	0.08	0.07	0.41	0.34	0.00	0.00	64972	26668
7		AVERAGE	0.28	0.29	0.40	0.28	0.00	0.00		
8	MC	NS	0.40	0.36	0.21	0.19	0.00	0.00	11240	52650
8	MC	EE	0.58	0.59	0.20	0.23	0.00	0.00	11236	52624
8	MC	EE	0.58	0.46	0.18	0.22	0.00	0.00	11243	52601
8	MC	EE	0.33	0.36	0.30	0.36	0.00	0.00	11249	52652
8	MC	GE	0.45	0.45	0.16	0.23	0.00	0.00	11228	52609
8	MC	GE	0.37	0.32	-0.06	0.01	0.00	0.00	11248	52616
8	MC	SP	0.48	0.54	0.33	0.37	0.00	0.00	11223	52617
8	MC	F	0.43	0.45	0.01	0.11	0.00	0.00	11232	52594
8	MC	F	0.77	0.70	0.40	0.40	0.00	0.00	11233	52599
8	MC	GE	0.27	0.27	-0.02	0.05	0.00	0.00	11234	52572
8		AVERAGE	0.47	0.45	0.17	0.22	0.00	0.00		

Note: Content categories are as follows: OA=Operations and Algebraic Thinking; NBT= Numbers and Operations in Base Ten; NF= Numbers and Operations - Fractions; MD=Measurement and Data; GE=Geometry; RP= Ratios and Proportional Relationships; NS=The Number System; EE=Expressions and Equations; SP=Statistics and Probability; and F=Functions.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.29: Summary of Calibration and Census Data: ELA

	Calibration Sample		Census Data (2015)		Difference
	N	%	N	%	(Calib. % − Census %)
ELA, Grade 3					
All Students	69190	100.0%	68003	100.0%	
Gender					
Male	35206	50.9%	34889	51.3%	-0.4%
Female	33984	49.1%	33114	48.7%	0.4%
Race/Ethnicity					
White	48830	70.6%	48338	71.1%	-0.5%
Black	11576	16.7%	11248	16.5%	0.2%
Hispanic	4297	6.2%	4268	6.3%	-0.1%
Asian/Pacific Islander	1429	2.1%	1482	2.2%	-0.1%
American Indian/Alaska Native	274	0.4%	240	0.4%	0.0%
Other	2784	4.0%	2427	3.6%	0.4%
ELA, Grade 4					
All Students	67673	100.0%	67015	100.0%	
Gender					
Male	34659	51.2%	34251	51.1%	0.1%
Female	33014	48.8%	32764	48.9%	-0.1%
Race/Ethnicity					
White	48327	71.4%	48302	72.1%	-0.7%
Black	11077	16.4%	10774	16.1%	0.3%
Hispanic	4209	6.2%	4080	6.1%	0.1%
Asian/Pacific Islander	1455	2.2%	1486	2.2%	0.0%
American Indian/Alaska Native	239	0.4%	275	0.4%	0.0%
Other	2366	3.5%	2098	3.1%	0.4%
ELA, Grade 5					
All Students	66635	100.0%	66419	100.0%	
Gender					
Male	34021	51.1%	33971	51.1%	0.0%
Female	32614	48.9%	32448	48.9%	0.0%
Race/Ethnicity					
White	48170	72.3%	48340	72.8%	-0.5%
Black	10497	15.8%	10748	16.2%	-0.4%
Hispanic	4089	6.1%	3806	5.7%	0.4%
Asian/Pacific Islander	1456	2.2%	1483	2.2%	0.0%
American Indian/Alaska Native	275	0.4%	293	0.4%	0.0%
Other	2148	3.2%	1749	2.6%	0.6%

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.29: Summary of Calibration and Census Data: ELA (cont.)

	Calibration Sample		Census Data (2015)		Difference
	N	%	N	%	(Calib. % − Census %)
ELA, Grade 6					
All Students	63889	100.0%	66062	100.0%	
Gender					
Male	32314	50.6%	33821	51.2%	-0.6%
Female	31575	49.4%	32241	48.8%	0.6%
Race/Ethnicity					
White	46411	72.6%	48304	73.1%	-0.5%
Black	10267	16.1%	10646	16.1%	0.0%
Hispanic	3676	5.8%	3683	5.6%	0.2%
Asian/Pacific Islander	1467	2.3%	1463	2.2%	0.1%
American Indian/Alaska Native	270	0.4%	273	0.4%	0.0%
Other	1798	2.8%	1693	2.6%	0.2%
ELA, Grade 7					
All Students	63470	100.0%	66000	100.0%	
Gender					
Male	32056	50.5%	33930	51.4%	-0.9%
Female	31414	49.5%	32070	48.6%	0.9%
Race/Ethnicity					
White	46435	73.2%	48236	73.1%	0.1%
Black	10080	15.9%	10721	16.2%	-0.3%
Hispanic	3570	5.6%	3670	5.6%	0.0%
Asian/Pacific Islander	1439	2.3%	1455	2.2%	0.1%
American Indian/Alaska Native	244	0.4%	303	0.5%	-0.1%
Other	1702	2.7%	1615	2.4%	0.3%
ELA, Grade 8					
All Students	63468	100.0%	66528	100.0%	
Gender					
Male	32250	50.8%	33733	50.7%	0.1%
Female	31218	49.2%	32795	49.3%	-0.1%
Race/Ethnicity					
White	46338	73.0%	49036	73.7%	-0.7%
Black	10134	16.0%	10724	16.1%	-0.1%
Hispanic	3587	5.7%	3572	5.4%	0.3%
Asian/Pacific Islander	1454	2.3%	1414	2.1%	0.2%
American Indian/Alaska Native	303	0.5%	289	0.4%	0.1%
Other	1652	2.6%	1493	2.2%	0.4%

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.30: Summary of Calibration and Census Data: Mathematics

	Calibration Sample		Census Data (2015)		Difference
	N	%	N	%	(Calib. % − Census %)
Mathematics, Grade 3					
All Students	69314	100.0%	68017	100.0%	
Gender					
Male	35268	50.9%	34893	51.3%	-0.4%
Female	34046	49.1%	33124	48.7%	0.4%
Race/Ethnicity					
White	48818	70.4%	48332	71.1%	-0.7%
Black	11594	16.7%	11245	16.5%	0.2%
Hispanic	4342	6.3%	4283	6.3%	0.0%
Asian/Pacific Islander	1500	2.2%	1492	2.2%	0.0%
American Indian/Alaska Native	276	0.4%	239	0.4%	0.0%
Other	2784	4.0%	2426	3.6%	0.4%
Mathematics, Grade 4					
All Students	67869	100.0%	67029	100.0%	
Gender					
Male	34770	51.2%	34257	51.1%	0.1%
Female	33099	48.8%	32772	48.9%	-0.1%
Race/Ethnicity					
White	48348	71.2%	48299	72.1%	-0.9%
Black	11142	16.4%	10766	16.1%	0.3%
Hispanic	4264	6.3%	4097	6.1%	0.2%
Asian/Pacific Islander	1506	2.2%	1497	2.2%	0.0%
American Indian/Alaska Native	240	0.4%	274	0.4%	0.0%
Other	2369	3.5%	2096	3.1%	0.4%
Mathematics, Grade 5					
All Students	66846	100.0%	66431	100.0%	
Gender					
Male	34137	51.1%	33974	51.1%	0.0%
Female	32709	48.9%	32457	48.9%	0.0%
Race/Ethnicity					
White	48207	72.1%	48329	72.8%	-0.7%
Black	10558	15.8%	10743	16.2%	-0.4%
Hispanic	4146	6.2%	3821	5.8%	0.4%
Asian/Pacific Islander	1506	2.3%	1496	2.3%	0.0%
American Indian/Alaska Native	277	0.4%	293	0.4%	0.0%
Other	2152	3.2%	1749	2.6%	0.6%

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6:30: Summary of Calibration and Census Data: Mathematics (cont.)

	Calibration Sample		Census Data (2015)		Difference
	N	%	N	%	(Calib. % − Census %)
Mathematics, Grade 6					
All Students	66355	100.0%	66019	100.0%	
Gender					
Male	33896	51.1%	33782	51.2%	-0.1%
Female	32459	48.9%	32237	48.8%	0.1%
Race/Ethnicity					
White	48209	72.7%	48270	73.1%	-0.4%
Black	10627	16.0%	10636	16.1%	-0.1%
Hispanic	3878	5.8%	3705	5.6%	0.2%
Asian/Pacific Islander	1486	2.2%	1444	2.2%	0.0%
American Indian/Alaska Native	276	0.4%	274	0.4%	0.0%
Other	1879	2.8%	1690	2.6%	0.2%
Mathematics, Grade 7					
All Students	65148	100.0%	65038	100.0%	
Gender					
Male	33252	51.0%	33415	51.4%	-0.4%
Female	31896	49.0%	31623	48.6%	0.4%
Race/Ethnicity					
White	47540	73.0%	47499	73.0%	0.0%
Black	10477	16.1%	10679	16.4%	-0.3%
Hispanic	3747	5.8%	3657	5.6%	0.2%
Asian/Pacific Islander	1369	2.1%	1305	2.0%	0.1%
American Indian/Alaska Native	263	0.4%	299	0.5%	-0.1%
Other	1752	2.7%	1599	2.5%	0.2%
Mathematics, Grade 8					
All Students	52692	100.0%	52842	100.0%	
Gender					
Male	27502	52.2%	27293	51.7%	0.5%
Female	25190	47.8%	25549	48.3%	-0.5%
Race/Ethnicity					
White	37817	71.8%	38117	72.1%	-0.3%
Black	9220	17.5%	9480	17.9%	-0.4%
Hispanic	3102	5.9%	2916	5.5%	0.4%
Asian/Pacific Islander	908	1.7%	858	1.6%	0.1%
American Indian/Alaska Native	274	0.5%	246	0.5%	0.0%
Other	1371	2.6%	1225	2.3%	0.3%

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.31: ELA population ability estimates across multiple groups on all items

Estimates	Grade					
	3	4	5 (base)	6	7	8
N-count	69190	67673	66635	63889	63470	63468
Mean theta	-1.28	-0.58	-0.22	0.00	0.21	0.54
Theta SD	1.25	1.26	1.29	1.13	1.35	1.3

Table 6.32: Math population ability estimates across multiple groups on all items

Estimates	Grade					
	3	4	5 (base)	6	7	8
N-count	69314	67869	66846	66355	65148	52692
Mean theta	-1.16	-0.57	-0.12	0.04	0.18	0.07
Theta SD	1.15	1.43	1.18	1.17	1.36	1.68

Table 6.33: Transformation constants for ELA and Mathematics base grades

Content Area and Grade	Desired scale properties in scale score metric		Estimated population ability in theta metric		Transformation constants	
	Mean	SD	Mean	SD	M1	M2
ELA 5	500	50	-0.21	1.275	39.21569	508.23529
Math 5	500	50	-0.11	1.169	42.77160	504.70488

Table 6.34: ELA items flagged for poor fit

Item Position In Calib.	Item ID	Item Grade	Model	ChiSqr	Chi DF	Total N	Z Score	Z Obsd	Z Pred	Obsd-Pred
41	692626	3	2PPC	928.53	8	69186	230.13	0.4433	0.4436	-0.0003
108	755385	4	2PPC	883.05	8	67647	218.76	0.5099	0.5112	-0.0014
113*	693154	5	2PPC	1201.33	8	73934	298.33	0.4039	0.4068	-0.0029
119	701085	5	2PPC	752.26	8	66630	186.06	0.1707	0.1769	-0.0062
173*	693930	6	3PL	1554.65	7	71331	413.63	0.0606	0.0658	-0.0053
198	699807	6	2PPC	1142.38	8	63885	283.60	0.1210	0.1252	-0.0042
264	755285	7	2PPC	1813.21	8	63468	451.30	0.4567	0.4580	-0.0013
303	698853	8	2PPC	890.99	8	63458	220.75	0.5858	0.5851	0.0007
312	755270	8	2PPC	1245.79	8	63458	309.45	0.2666	0.2694	-0.0027

Note: * in the first column indicates a linking item.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.35: Mathematics items flagged for poor fit

Item Position In Calib.	Item ID	Item Grade	Model	ChiSqr	Chi DF	TotalN	Z Score	Z Obsd	Z Pred	Obsd-Pred
106	686764	5	3PL	2568.49	7	66812	684.59	0.6408	0.5470	0.0938
126	689401	5	2PPC	5349.15	8	66812	1335.29	0.4007	0.3977	0.003
139*	686030	6	3PL	9267.35	7	90262	2474.93	0.1433	0.2528	-0.1095
182	686972	7	3PL	679.28	7	65138	179.68	0.1579	0.1621	-0.0042

Note: * in the first column indicates a linking item.

Table 6.36: ELA scale score means and standard deviations

Grade level	Scale statistics		Mean difference between grades (in scale score points)
	Mean	SD	
3	458.41	50.02	
4	485.86	49.93	27.64
5	499.62	51.08	13.76
6	508.66	44.40	9.04
7	517.14	53.54	8.48
8	529.75	51.18	12.60

Table 6.37: ELA scale scores at different percentiles across grades

Grade	Percentile				
	10th	25th	50th	75th	90th
3	396	427	460	491	520
4	421	455	489	519	545
5	435	468	503	534	559
6	452	480	510	538	564
7	447	482	520	554	583
8	466	498	531	564	593

Table 6.38: Mathematics scale score means and standard deviations

Grade level	Scale statistics		Mean difference between grades (in scale score points)
	Mean	SD	
3	456.01	48.59	
4	481.93	55.94	25.92
5	499.97	49.86	18.05
6	507.16	48.02	7.19
7	514.69	50.24	7.53
8	515.07	49.06	0.38

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 6.39: Mathematics scale scores at different percentiles across grades

Grade	Percentile				
	10th	25th	50th	75th	90th
3	395	427	460	488	514
4	420	456	489	517	541
5	437	470	504	533	559
6	446	477	510	540	566
7	454	486	518	549	574
8	451	491	522	547	569

Table 6.40: Spring 2015 Grade 7 Mathematics scale score summary

Student Groups	N count	Mean scale score	Scale score standard deviation
Total Population (Students who took Grade 7 Mathematics test in Spring 2015)	64965	2529.79	97.26
Matched Students (Students who took Grade 7 Mathematics test in Spring 2015 and Grade 8 Mathematics test in Spring 2016)	49385	2509.50	87.33
Unmatched Students (Students who took Grade 7 Mathematics test in Spring 2015 and are not in the Spring 2016 Mathematics data)	15580	2594.12	99.11

Table 6.41: ELA and Mathematics Lowest and Highest Obtainable Scores

Grade	ELA		Mathematics	
	LOSS	HOSS	LOSS	HOSS
3	230	730	290	650
4	240	740	320	680
5	250	780	340	710
6	260	790	350	730
7	280	810	360	740
8	290	820	390	770

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure 6.1: ELA Test Characteristic Curves

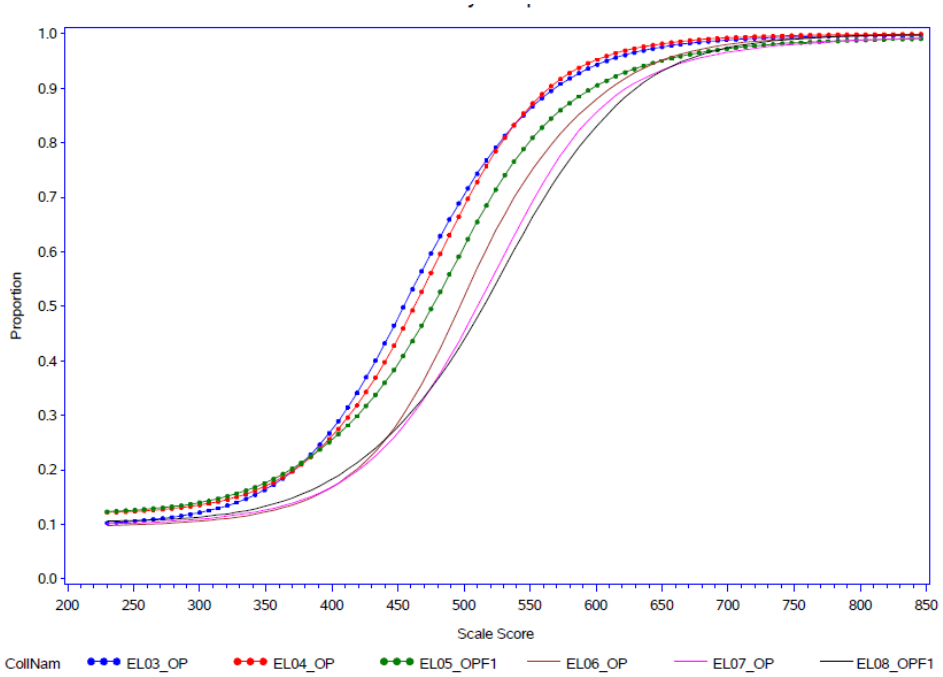
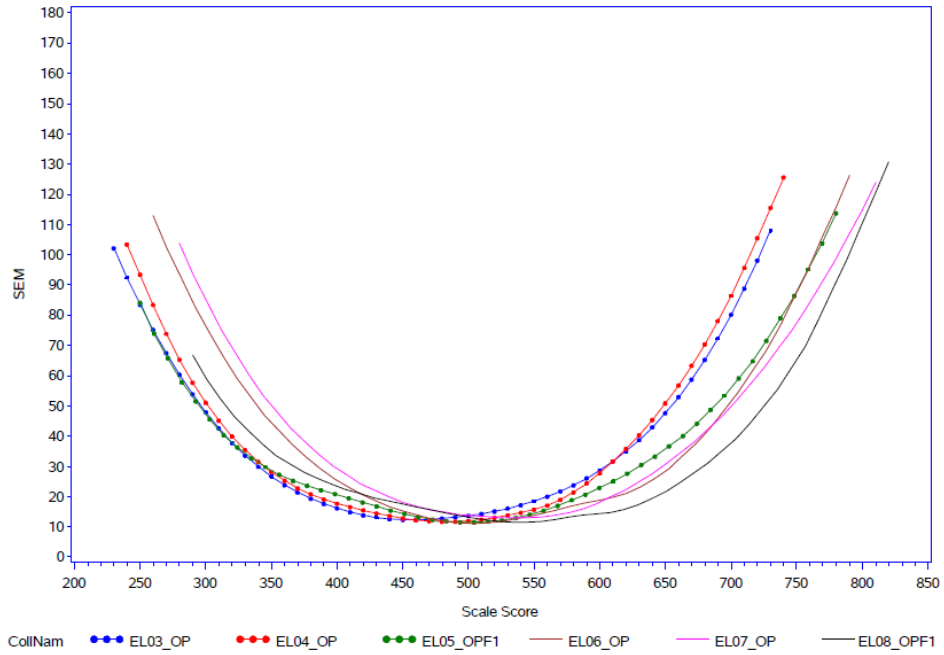


Figure 6.2: ELA Standard Error Curves



Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure 6.3: ELA Test Characteristic Curves for Grade 5 forms

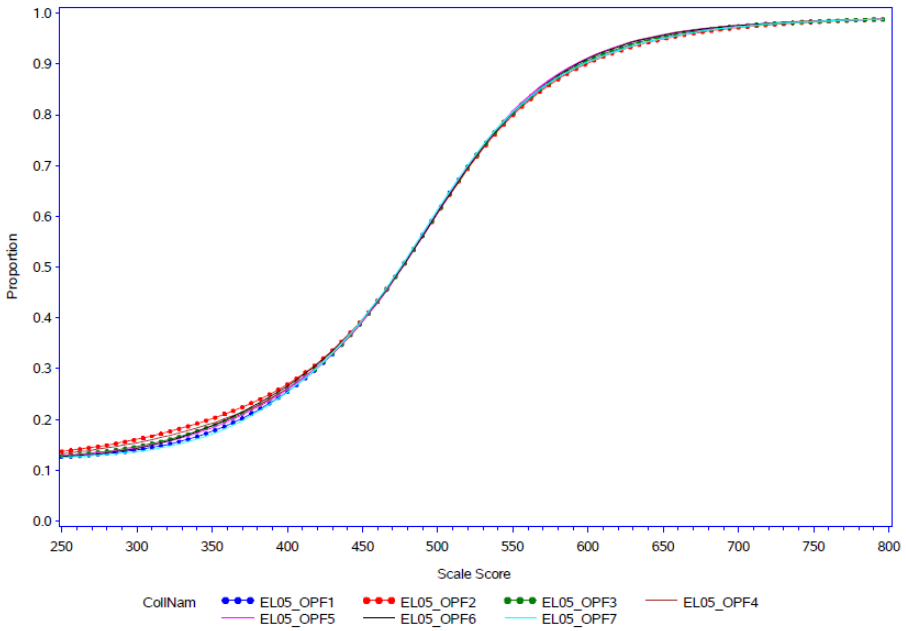
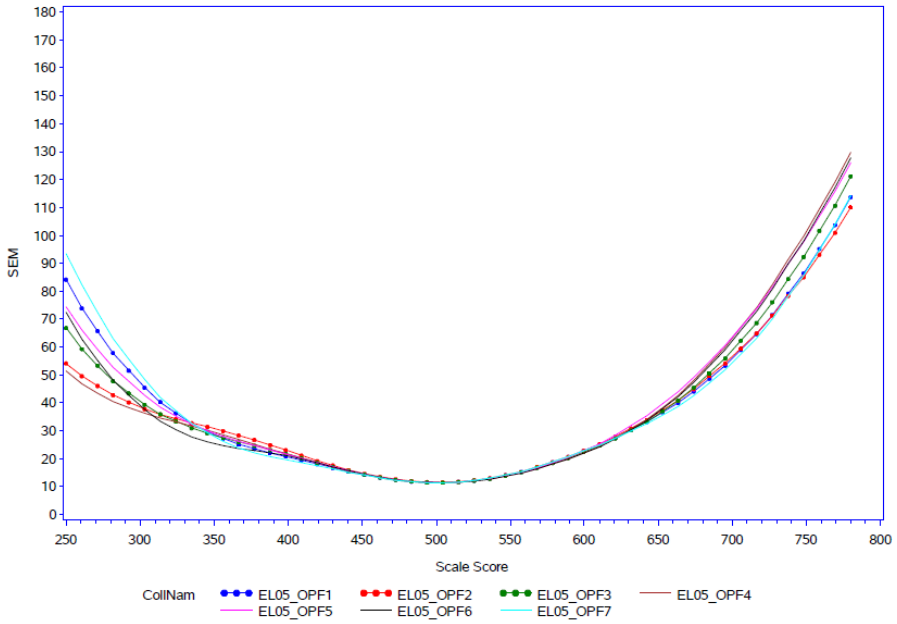


Figure 6.4: ELA Test Characteristic Curves for Grade 5 forms



Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure 6.5: ELA Test Characteristic Curves for Grade 8 forms

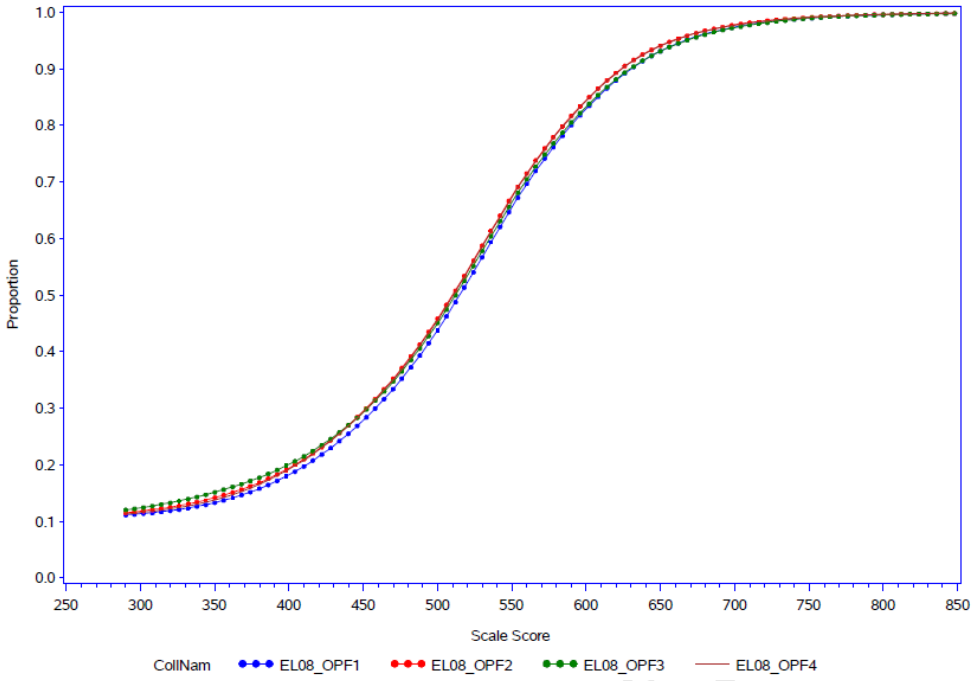
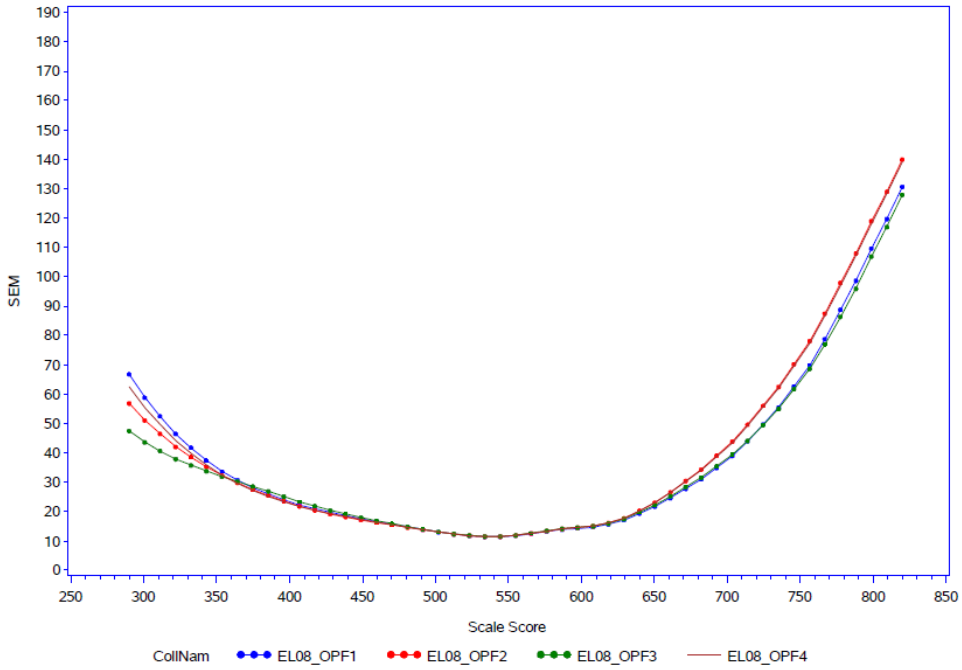


Figure 6.6: ELA Test Characteristic Curves for Grade 8 forms



Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure 6.7: Mathematics Test Characteristic Curves

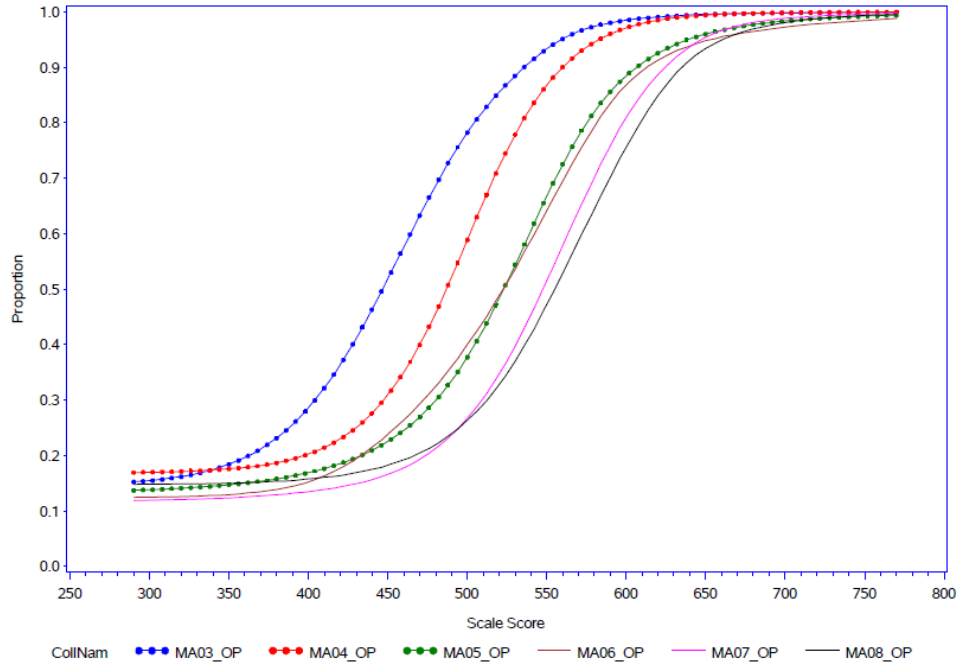
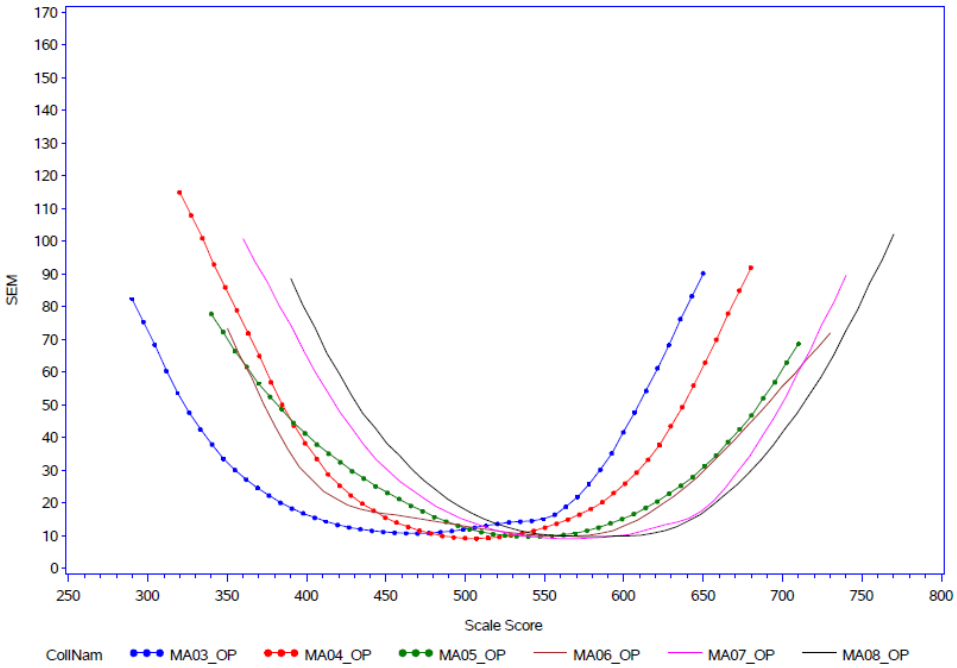


Figure 6.8: Mathematics Standard Error Curves



Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

CHAPTER 7: TEST RESULTS

This section of the document contains information on the results of the spring 2016 administration of the MAP ELA and Mathematics. The scale score results are presented here. Achievement-level information is also provided. Presenting the results by achievement level translates the quantitative scale provided through scale scores into a qualitative description of student achievement: *Below Basic*, *Basic*, *Proficient*, and *Advanced.*

While the scale score provides an essential quantitative reference to student achievement, the achievement-level information plainly outlines the meaning of the scores to parents, students, and educators. When combined, scale scores and achievement levels provide a comprehensive set of tools to assess Missouri student achievement by content and grade level.

It is often desirable to examine the scores of students across time and monitor group performance. This is possible if the test content and the construct measured by the test are comparable from year to year and if the scores are reported on the same scale in multiple years. This was not the case for 2015–16 ELA and Mathematics assessments.
While the 2015–16 ELA and Mathematics assessments are comparable content- and construct-wise to the assessments administered in 2014-15 year, new scales were developed for both ELA and Mathematics after the 2015-16 test administration. Therefore, the test scores for ELA and Mathematics are not directly comparable with the previous year scores and the cross-year scale score summary is not presented for ELA and Mathematics in this report. The 2015-16 ELA and Mathematics scale score summaries are presented in Tables 7.1 and 7.2, respectively. Note that the Spring 2016 results are based on the calibration sample data.

DRAFT

Table 7.3 shows the percentage of students in each achievement level from 2005–06 through 2013–14 on the Communication Arts test and the percentage of students in each achievement level after the 2014–15 and 2015-16 ELA test administrations. It should be noted that the ELA test scores were reported on the Smarter Balanced Assessment Consortium (SBAC) scale after the 2014–15 test administration and students were classified into the achievement levels based on the cut scores established after the 2014 field test SBAC on their item bank. The ELA scores after 2015-16 test administration were reported on a new Missouri scale and the 2014-15 cut scores were revised by Missouri educators in a process of standard setting in the Summer of 2016 to better reflect Missouri student performance on the new assessments. Therefore, the percentages of students in each achievement level after the 2014–15 test administration are not directly comparable to the percentages of students in each achievement level after the 2013–14 test administration; and the percentages of students in each achievement level after the 2015–16 test administration may differ significantly from the percentages of students in each achievement level after the 2014–15 test administration. The past data are provided in the table for reference purpose only and separated from this year's data by grey horizontal bars.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 7.4 shows the percentage of students in each achievement level from 2005–06 through 2015–16 on the Mathematics test. Similarly to ELA, the Mathematics test scores were reported on the SBAC scale after the 2014–15 test administration and students were classified into the achievement levels based on the cut scores established after the 2014 field test by SBAC on their item bank. The Mathematics scores after 2015-16 test administration were reported on a new Missouri vertical scale and the 2014-15 cut scores were revised by Missouri educators in a process of standard setting in Summer of 2016. The past data are provided in the table for reference purpose only and separated from this year data by grey horizontal bars. It is worth noting that not all Grade 8 students participate in Mathematics assessment. As stated in the previous chapter of this document, approximately 20% of Grade 8 students take Algebra 1 assessment instead of Mathematics. Exclusion of these students from Mathematics assessment may affect the state level student performance in Mathematics.

Note that the Spring 2016 results in Tables 7.3 and 7.4 are based on the calibration sample data while the previous test administration results are based on the census data.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 7.1: State-Level Scale Score Statistics: English Language Arts

Grade	N	Mean SS	S.D. SS	Percentile				
				10th	25th	50th	75th	90th
3	69,190	458.41	50.02	396	427	460	491	520
4	67,673	485.86	49.93	421	455	489	519	545
5	66,635	499.62	51.08	435	468	503	534	559
6	63,889	508.66	44.40	452	480	510	538	564
7	63,470	517.14	53.54	447	482	520	554	583
8	63,468	529.75	51.18	466	498	531	564	593

Table 7.1: State-Level Scale Score Statistics: Mathematics

Grade	N	Mean SS	S.D. SS	Percentile				
				10th	25th	50th	75th	90th
3	69,314	456.01	48.59	395	427	460	488	514
4	67,869	481.93	55.94	420	456	489	517	541
5	66,846	499.97	49.86	437	470	504	533	559
6	66,355	507.16	48.02	446	477	510	540	566
7	65,148	514.69	50.24	454	486	518	549	574
8	52,692	515.07	49.06	451	491	522	547	569

*Algebra I students had the option of taking Algebra EOC instead of MAP Mathematics in Grade 8

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 7.3: Comparison of Percentage of Students in Each Achievement Level, English Language Arts 2006 through 2016 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
	2009	67,357	0.3	9.6	49.8	25.1	15.2	40.3
	2010	66,947	0.3	8.2	48.4	26.9	16.2	43.1
	2011	66,487	0.4	7.6	48.4	27.0	16.6	43.6
	2012	66,323	0.3	8.0	46.5	27.2	18.1	45.3
	2013	66,754	0.3	7.8	44.2	27.7	20.1	47.8
	2014	67,211	0.3	9.8	48.3	25.5	16.0	41.6
	2015	67,998	0.2	19.4	23.3	24.0	33.1	57.1
	2016	69,190	n/a	18.8	20.4	42.2	18.6	60.8
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
	2009	66,709	0.3	7.6	45.8	33.6	12.7	46.3
	2010	67,510	0.3	8.6	40.2	31.2	19.7	50.9
	2011	67,049	0.4	8.2	39.5	31.6	20.2	51.9
	2012	65,996	0.3	8.3	39.3	31.2	20.9	52.2
	2013	66,085	0.3	8.2	38.8	31.6	21.2	52.8
	2014	66,647	0.3	7.8	46.4	31.5	14.0	45.5
	2015	67,013	0.2	21.8	19.7	25.3	33.1	58.3
	2016	67,673	n/a	15.1	21.5	42.8	20.7	63.4
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
	2009	67,307	0.3	6.3	44.6	33.9	14.9	48.8
	2010	66,730	0.3	7.1	41.5	32.1	18.9	51.0
	2011	67,461	0.6	6.9	41.4	32.4	18.7	51.1
	2012	66,675	0.3	7.0	40.9	32.3	19.6	51.8
	2013	65,980	0.3	7.1	40.3	32.2	20.1	52.3
	2014	66,153	0.3	6.2	43.5	33.2	16.8	50.0
	2015	66,416	0.2	18.9	21.9	35.6	23.3	58.9
	2016	66,635	n/a	15.0	22.7	41.9	20.4	62.3

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 7.3: Comparison of Percentage of Students in Each Achievement Level, English Language Arts 2006 through 2016 Data (cont.)

Grade	Year	N	No Level	Below Basic	Basic	Proficient	Advanced	Prof. & Adv.
6	2006	67,709	1.1	11.9	44.8	31.6	10.6	42.2
	2007	67,045	1.2	11.2	44	31.8	11.7	43.6
	2008	65,830	0.2	9.0	43.5	34	13.4	47.4
	2009	65,908	0.3	8.6	43.4	33.8	13.9	47.7
	2010	67,476	0.3	7.8	42.3	33.9	15.7	49.6
	2011	66,633	0.3	7.3	41.9	34.3	16.2	50.5
	2012	67,342	0.3	7.5	42.0	34.7	15.5	50.2
	2013	66,731	0.4	7.2	41.4	34.9	16.1	51.0
	2014	66,019	0.3	8.5	43.8	32.9	14.5	47.5
	2015	66,059	0.2	19.6	25.3	35.0	19.8	54.9
	2016	63,889	n/a	17.0	22.8	42.9	17.3	60.2
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
	2009	66,531	0.3	8.7	40.3	37.2	13.6	50.8
	2010	66,279	0.4	9.8	38.1	35.2	16.5	51.7
	2011	67,517	0.4	9.0	36.9	36.0	17.8	53.8
	2012	66,845	0.3	8.7	35.8	36.6	18.7	55.2
	2013	67,319	0.3	9.0	35.7	36.5	18.4	55.0
	2014	66,893	0.4	8.2	36.0	36.9	18.6	55.4
	2015	66,000	0.3	18.4	24.1	38.6	18.5	57.2
	2016	63,470	n/a	21.7	18.3	40.2	19.8	60.0
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
	2009	67,077	0.5	5.3	44.5	33.4	16.3	49.7
	2010	66,463	0.5	4.9	42.8	34.3	17.4	51.8
	2011	66,205	0.5	4.6	42.5	33.9	18.5	52.5
	2012	67,037	0.4	4.3	42.0	34.3	19.0	53.3
	2013	66,710	0.5	4.1	41.5	34.9	19.0	53.9
	2014	67,168	0.5	4.5	44.6	34.1	16.3	50.4
	2015	66,528	0.2	14.7	27.6	40.4	17.1	57.5
	2016	63,468	n/a	17.7	21.2	39.6	21.5	61.1

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 7.4: Comparison of Percentage of Students in Each Achievement Level, Mathematics 2006 through 2016 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
	2009	67,357	0.2	6.8	48.5	35.6	8.8	44.4
	2010	66,947	0.2	6.2	46.6	37.0	10.1	47.1
	2011	66,487	0.3	5.6	44.7	38.1	11.3	49.4
	2012	66,323	0.2	5.4	42.6	39.9	11.9	51.9
	2013	66,754	0.2	5.3	43.8	39.2	11.4	50.7
	2014	67,211	0.2	6.0	43.7	36.6	13.5	50.2
	2015	68,012	0.0	21.4	26.5	30.8	21.2	52.0
	2016	69,314	n/a	17.9	29.4	32.1	20.6	52.7
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
	2009	66,709	0.2	7.3	48.2	36.6	7.8	44.4
	2010	67,510	0.2	6.1	45.4	39.3	9.1	48.4
	2011	67,049	0.3	5.6	43.7	39.9	10.5	50.5
	2012	65,996	0.1	5.7	43.7	40.5	10.0	50.5
	2013	66,085	0.1	5.5	44.2	40.7	9.4	50.1
	2014	66,647	0.2	6.6	51.1	34.5	7.6	42.1
	2015	67,023	0.0	16.8	33.6	29.9	19.6	49.6
	2016	67,869	n/a	15.5	31.6	30.6	22.3	52.9
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
	2009	67,307	0.2	7.5	45.1	35.6	11.6	47.2
	2010	66,730	0.2	6.2	41.9	36.7	15.1	51.7
	2011	67,461	0.5	6.1	40.9	36.3	16.2	52.5
	2012	66,675	0.2	5.8	39.7	35.9	18.4	54.3
	2013	65,980	0.2	5.9	40.1	35.9	18.0	53.9
	2014	66,153	0.2	7.2	40.5	35.5	16.7	52.2
	2015	66,429	0.0	28.5	31.6	20.1	19.7	39.8
	2016	66,846	n/a	20.7	32.6	28.5	18.2	46.7

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 7.4: Comparison of Percentage of Students in Each Achievement Level, Mathematics 2006 through 2016 Data (cont.)

Grade	Year	N	No Level	Below Basic	Basic	Proficient	Advanced	Prof. & Adv.
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
	2009	65,908	0.2	8.9	40.7	37.5	12.6	50.1
	2010	67,476	0.2	7.8	36.6	40.3	15.0	55.4
	2011	66,633	0.2	7.5	35.4	40.5	16.4	56.9
	2012	67,342	0.2	7.4	36.7	39.7	16.0	55.7
	2013	66,731	0.3	7.1	36.4	39.9	16.3	56.2
	2014	66,019	0.3	7.2	36.9	40.3	15.3	55.6
	2015	66,014	0.1	28.7	33.1	21.6	16.5	38.1
	2016	66,355	n/a	20.4	36.2	27.9	15.4	43.4
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
	2009	66,531	0.3	12.5	35.2	37.6	14.3	51.9
	2010	66,279	0.3	10.8	34.3	38.8	15.7	54.5
	2011	67,517	0.3	10.5	33.5	39.2	16.6	55.8
	2012	66,845	0.3	9.8	30.3	40.0	19.6	59.6
	2013	67,319	1.5	10.1	31.1	39.1	18.2	57.3
	2014	66,893	1.6	9.6	32.0	38.6	18.2	56.7
	2015	65,036	0.1	31.4	33.2	21.1	14.1	35.3
	2016	65,148	n/a	22.4	35.5	26.8	15.4	42.2
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	18.0	37.7	29.9	13.9	43.8
	2009	67,077	0.5	16.4	36.8	31.5	14.9	46.4
	2010	66,463	0.4	14.9	33.3	32.1	19.2	51.3
	2011	66,205	0.4	15.0	33.9	31.0	19.8	50.8
	2012	67,037	0.3	14.1	33.6	31.8	20.2	52.0
	2013*	52,335	1.4	17.1	41.2	30.2	10.1	40.3
	2014*	52,818	1.6	17.5	38.7	30.9	11.3	42.2
	2015*	52,840	0.2	39.3	32.3	18.1	10.1	28.2
	2016*	52,692	n/a	27.9	43.6	19.4	9.1	28.4

*Algebra I students had the option of taking Algebra EOC instead of MAP Mathematics in Grade 8

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

CHAPTER 10: FAIRNESS

As noted in the *Standards* (AERA, APA, & NCME, 2014), there are various elements of fairness. In this chapter, we examine fairness as it relates to minimizing bias on a test. We then look at test performance among different subgroups assessed by the MAP. It should be noted that differences in test performance among subgroups do not inherently 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 on the test as a whole, or on individual items, as a result of factors other than bias.

10.1. Evaluating Bias through Differential Item Functioning (DIF) 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, after controlling for overall performance on the test. The DIF procedures used and the results of these analyses are detailed in this section. Importantly, all items included on the MAP have been thoroughly reviewed for content and bias by Missouri educators and DRC Test Development experts prior to test administration, to help ensure that they do not present challenges to the relevance of the construct that the test intends to measure. Therefore DIF flags do not necessarily indicate that an item is biased; rather, DIF flags indicate that the item functions differently for equally able members of different groups (Camilli & Shepard, 1994). For this reason, items are not necessarily suppressed from operational scoring if they are flagged for DIF.

The position of DRC 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 into 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 such biases, DRC strives to minimize the role of extraneous elements, thereby increasing the number of students for whom the test is appropriate. Careful

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

attention is given during the test development and test construction processes to lessen the influence of these elements for large numbers of students. 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.

Two DIF statistics that are commonly used for this purpose are the Mantel-Haenszel (MH) statistic (1959) and the Standardized Mean Difference (SMD) between the reference and focal groups, proposed by Dorans and Schmitt (1991).

The MH statistic is computed as follows (Zwick, Donoghue, & Grima, 1993):

$$\text{Mantel } \chi^2 = \frac{\left(\sum_k F_k - \sum_k E(F_k)\right)^2}{\sum_k \text{Var}(F_k)},$$

where F_k is the sum of scores for the focal group at the kth level of the matching variable. Note that the MH statistic is sensitive to N such that larger sample sizes increase the value of chi square.

In addition to the MH chi-square statistic, the delta statistic (MH-D DIF) was computed for all items. Educational Testing Service (ETS) first developed the MH-D DIF statistic. To compute delta, alpha (the odds ratio) is first computed:

$$\alpha_{MH} = \frac{\sum_{k=1}^{K} N_{r1k} N_{f0k} / N_k}{\sum_{k=1}^{K} N_{f1k} N_{r0k} / N_k},$$

where N_{r1k} is the number of correct responses in the reference group at ability level k, N_{f0k} is the number of incorrect responses in the focal group at ability level k, N_k is the total number of responses, N_{f1k} is the number of correct responses in the focal group at ability level k, and N_{r0k} is the number of incorrect responses in the reference group at ability level k. MH-D DIF is then computed:

$$\text{MH-D DIF} = -2.35 \ln(\alpha_{MH}) .$$

For selected-response items, the MH (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

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

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: Significant MH chi-square statistic ($p < 0.05$) and $1.0 \leq |\text{MH D-DIF}| < 1.5$
- Large DIF: Significant MH chi-square statistic ($p < 0.05$) and $|\text{MH D-DIF}| \geq 1.5$

For constructed-response items, an effect size (ES) statistic based on the MH chi-square will be used. The ES is obtained by dividing the SMD statistics by the standard deviation of the item. The SMD is an effect size index of DIF, which is relatively easy to interpret (Zwick et al., 1993). The SMD compares the mean of the reference and focal group, adjusting for the distribution of reference and focal group members on the conditioning variable (Zwick et al., 1993), which for these analyses is the MAP raw score. SMD is computed as follows (Zwick et al., 1993):

$$SMD = p_{Fk}\left(\sum_k m_{Fk} - \sum_k m_{Rk}\right),$$

where p_{Fk} = proportion of the focal group members at the kth level of the matching variable, $m_{Fk} = 1/N_{F1k}$, and $m_{Rk} = 1/N_{R1k}$. Items are flagged using the same rules that are used in NAEP:

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

- ***Race/Ethnicity***: Focal groups are students whose race/ethnicity is reported as Black, Hispanic, Asian/Pacific Islander, American Indian/Alaska Native, or Other; reference group is students whose race/ethnicity is reported as White.

- ***Accommodations/Universal Tools***: Focal group is students who had one or more testing accommodations or Universal Tools indicated by a teacher; 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.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

The minimum case count for the focal group was set at 200 and the minimum case count for the reference group was set at 400. The DIF analyses are not performed for subgroups of fewer than 200 students. In these cases, the statistical procedures do not have sufficient power to detect differences should they exist.

Tables 10.1 and 10.2 summarize the number of DIF flags by grade for each focal group. For example, consider Grade 3 ELA. In this form no items were flagged for gender DIF. One item was flagged for DIF against the Asian/Pacific Islander subgroup and two items were flagged against the Hispanic subgroup - these items exhibited moderate negative DIF. No items were flagged for the Black subgroup. One item was flagged for the American Indian/Alaska Native subgroup and display moderate positive DIF, and one item was flagged for the accommodated subgroup, also displaying a moderate 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 DRC Test Development experts.

10.2. 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.

10.2.1. Effect Size

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:

$$d = \frac{\overline{x_a} - \overline{x_b}}{\sqrt{\frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{(n_a + n_b) - 2}}},$$

where $\overline{x_a}$ is the mean score of group A, $\overline{x_b}$ 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 that 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.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Using Cohen's (1988) guidelines, certain trends become apparent in Tables 10.3 through 10.8.

On the English Language Arts, there are small differences in mean test scores between girls and boys where girls outperform boys in grades 4 through 8 (no difference in Grade 3).

There is a medium difference in mean English Language Arts test scores of Black students compared to White students where Black students underperform White students in all grades. There is a small difference between the mean test scores of Hispanic and White students where Hispanics underperform White students on English Language Arts in all grades. Similarly, there is a small difference between the mean test scores of Native Americans and White students where Native Americans underperform White students on English Language Arts in all grades. There is also a small difference between the mean English Language Arts test scores of Asian/Pacific Islander students and White students where Asian/Pacific Islander students outperform White students in all grades except for Grade 4 (no difference).

There is a medium difference in mean English Language Arts test scores of students using testing accommodations or universal tools compared to students not using testing accommodation or universal tools for all grades. In all grade levels, students not using testing accommodations or universal tools outperform their peers who use accommodations/universal tools.

DRAFT

In Mathematics, there is a medium difference between the mean tests scores of Black and White students where Black students underperform White students in Grades 3, 5, 6, and 8. There is a large difference between the mean tests scores of Black and White students where Black students underperform White students in grades 4 and 7. There is a small difference in mean Mathematics test scores of Hispanic students compared to White students in all grades where Hispanic students underperform White students. There is a small difference between the mean test scores of Native American students and White students where Native American students underperform White students in all grades except for Grade 8 (no difference). Finally, there is a small difference between the mean Mathematics test scores of Asian/Pacific Islander students and White students in all grades except Grade 5 where there is a medium difference in mean test scores. Asian/Pacific Islander students outperform White students in all grades.

There is a medium difference between the mean Mathematics test scores of students not needing testing accommodations/universal tools and students using testing accommodation or universal tools in all grades except for Grade 7 where there is a large difference. Students not using testing accommodations or universal tools outperform students using testing accommodations/universal tools in all grades.

The mean scale score differences trend observed in the Spring 2016 data are very similar to the trends observed in the Spring 2015.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

10.2.2. Reliability

Tables 10.9 through 10.20 show the regular test form reliability coefficients and SEM by student ethnicity, gender, and whether or not students are using any testing accommodations or universal tools.

The reliability coefficients for English Language Arts forms ranged from 0.87 to 0.94. For Mathematics the reliability coefficients ranged from 0.79 to 0.93. Lower reliability coefficients were generally observed for Grade 8 Mathematics subgroups compared to other Mathematics grades. Overall, this analysis shows that the test reliability is of acceptable magnitude for all of the subgroups. Note that the reliability coefficients are not reported for subgroups smaller than 50 students.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10. 1: 2016 MAP DIF Statistics: Number of Flagged Items, English Language Arts

English Language Arts		Grade	3	4	5	6	7	8
		# of Items	52	52	69	52	52	60
Group	**DIF Magnitude**	**DIF Direction**	**Number of Flagged Items**					
Female	Moderate	Negative	0	0	0	3	1	2
		Positive	0	0	5	0	0	7
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	7	0	0	4
Asian/ Pacific Islander	Moderate	Negative	1	2	1	2	0	1
		Positive	0	0	2	0	0	1
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	1	1	0	0
Black	Moderate	Negative	0	0	2	1	1	0
		Positive	0	0	0	0	0	0
	Large	Negative	0	0	0	1	0	0
		Positive	0	0	0	0	0	0
Hispanic	Moderate	Negative	2	0	2	1	0	0
		Positive	0	0	0	0	0	0
	Large	Negative	0	1	0	0	0	0
		Positive	0	0	0	0	0	0
American Indian/ Alaska Native	Moderate	Negative	0	1	0	0	0	0
		Positive	1	0	0	0	1	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
Ethnicity: Other	Moderate	Negative	0	0	0	0	0	0
		Positive	0	0	2	0	0	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
Accommo-dations/ Universal Tools	Moderate	Negative	0	1	3	0	0	3
		Positive	1	1	0	0	0	0
	Large	Negative	0	0	14	0	0	8
		Positive	0	0	0	0	0	0

Note: For Grades 5 and 8, the three components of the writing prompts were analyzed as separate items.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10. 2: 2016 MAP DIF Statistics: Number of Flagged Items, Mathematics

Mathematics		Grade	3	4	5	6	7	8
		# of Items	42	42	42	46	46	46
Group	**DIF Magnitude**	**DIF Direction**	**Number of Flagged Items**					
Female	Moderate	Negative	1	1	0	1	1	1
		Positive	2	0	0	1	1	0
	Large	Negative	1	0	0	0	0	0
		Positive	0	0	0	0	0	0
Asian/ Pacific Islander	Moderate	Negative	0	1	1	2	2	0
		Positive	2	1	0	1	2	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
Black	Moderate	Negative	0	0	1	0	0	0
		Positive	0	0	0	1	0	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
Hispanic	Moderate	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
American Indian/ Alaska Native	Moderate	Negative	0	0	0	1	0	0
		Positive	0	0	0	0	0	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
Ethnicity: Other	Moderate	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0
Accommo-dations/ Universal Tools	Moderate	Negative	0	0	0	1	0	0
		Positive	0	0	1	0	0	0
	Large	Negative	0	0	0	0	0	0
		Positive	0	0	0	0	0	0

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.3: Impact Analysis, Grade 3

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size
English Language Arts	Race/ Ethnicity	White (not Hispanic)	48830	465.45	48.07	
		Asian/Pacific Islander	1429	477.83	52.78	-0.26
		Black (not Hispanic)	11576	430.52	48.51	0.73
		Hispanic	4297	446.82	47.38	0.39
		Am. Indian/Alaska N.	274	452.97	46.11	0.26
		Other	2784	459.45	47.40	0.12
	Gender	Male	35206	454.22	50.82	
		Female	33984	462.75	48.79	-0.17
	Accommo-dations / Universal Tools	No	45091	468.93	47.18	
		Yes	24099	438.73	49.23	0.63
Mathematics	Race/ Ethnicity	White (not Hispanic)	48818	462.72	46.31	
		Asian/Pacific Islander	1500	478.73	51.56	-0.34
		Black (not Hispanic)	11594	429.4	48.03	0.71
		Hispanic	4342	445.3	47.43	0.38
		Am. Indian/Alaska N.	276	446.37	47.97	0.35
		Other	2784	454.59	46.90	0.18
	Gender	Male	35268	456.81	50.63	
		Female	34046	455.19	46.37	0.03
	Accommo-dations / Universal Tools	No	44120	467.73	43.09	
		Yes	25194	435.5	50.82	0.7

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.4: Impact Analysis, Grade 4

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size
English Language Arts	Race/ Ethnicity	White (not Hispanic)	48327	493.17	47.52	
		Asian/Pacific Islander	1455	502.13	54.78	-0.19
		Black (not Hispanic)	11077	456.27	49.08	0.77
		Hispanic	4209	473.64	46.97	0.41
		Am. Indian/Alaska N.	239	482.07	49.31	0.23
		Other	2366	487.08	47.55	0.13
	Gender	Male	34659	481.89	51.2	
		Female	33014	490.02	48.21	-0.16
	Accommo-dations / Universal Tools	No	45151	496.85	45.31	
		Yes	22522	463.82	51.46	0.7
Mathematics	Race/ Ethnicity	White (not Hispanic)	48348	490.41	51.12	
		Asian/Pacific Islander	1506	509.5	58.75	-0.37
		Black (not Hispanic)	11142	446.48	60.04	0.83
		Hispanic	4264	468.61	55.77	0.42
		Am. Indian/Alaska N.	240	473.3	59.24	0.33
		Other	2369	482.83	55.08	0.15
	Gender	Male	34770	482.85	58.08	
		Female	33099	480.96	53.58	0.03
	Accommo-dations / Universal Tools	No	43963	495.89	47.28	
		Yes	23906	456.25	61.28	0.75

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.5: Impact Analysis, Grade 5

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size
English Language Arts	Race/ Ethnicity	White (not Hispanic)	48170	505.74	49.19	
		Asian/Pacific Islander	1456	521.77	54.3	-0.32
		Black (not Hispanic)	10497	472.61	50.61	0.67
		Hispanic	4089	488.59	48.49	0.35
		Am. Indian/Alaska N.	275	491.05	49.34	0.3
		Other	2148	501.23	49.13	0.09
	Gender	Male	34021	493.44	52.45	
		Female	32614	506.05	48.79	-0.25
	Accommo-dations / Universal Tools	No	45827	511.16	46.02	
		Yes	20808	474.19	52.46	0.77
Mathematics	Race/ Ethnicity	White (not Hispanic)	48207	505.78	47.8	
		Asian/Pacific Islander	1506	531.66	54.53	-0.54
		Black (not Hispanic)	10558	473.47	48.9	0.67
		Hispanic	4146	489.65	47.81	0.34
		Am. Indian/Alaska N.	277	489.95	50.15	0.33
		Other	2152	498.99	48.92	0.14
	Gender	Male	34137	499.05	52.02	
		Female	32709	500.94	47.49	-0.04
	Accommo-dations / Universal Tools	No	44536	511.13	45.63	
		Yes	22310	477.69	50.48	0.71

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.6: Impact Analysis, Grade 6

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size
English Language Arts	Race/ Ethnicity	White (not Hispanic)	46410	514.73	42.49	
		Asian/Pacific Islander	1467	525.95	47.08	-0.26
		Black (not Hispanic)	10267	482.83	42.88	0.75
		Hispanic	3676	497.44	42.6	0.41
		Am. Indian/Alaska N.	270	502.58	42.1	0.29
		Other	1798	509.03	43.7	0.13
	Gender	Male	32313	504.03	45.03	
		Female	31575	513.4	43.25	-0.21
	Accommo-dations / Universal Tools	No	49827	514.99	42.34	
		Yes	14061	486.23	44.29	0.67
Mathematics	Race/ Ethnicity	White (not Hispanic)	48208	513.71	45.56	
		Asian/Pacific Islander	1486	533.17	52.94	-0.42
		Black (not Hispanic)	10627	478.23	47	0.77
		Hispanic	3878	496.33	45.97	0.38
		Am. Indian/Alaska N.	276	502.45	47.88	0.25
		Other	1879	505.28	46.46	0.18
	Gender	Male	33895	505.75	49.87	
		Female	32459	508.63	45.96	-0.06
	Accommo-dations / Universal Tools	No	49134	516.87	43.43	
		Yes	17220	479.45	49.67	0.83

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.7: Impact Analysis, Grade 7

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size
English Language Arts/ Literacy	Race/ Ethnicity	White (not Hispanic)	46435	524.03	51.13	
		Asian/Pacific Islander	1439	544.1	56.21	-0.39
		Black (not Hispanic)	10080	485.5	52.42	0.75
		Hispanic	3570	506.1	49.97	0.35
		Am. Indian/Alaska N.	244	513.31	52.75	0.21
		Other	1702	517.38	54.01	0.13
	Gender	Male	32056	511.98	55.24	
		Female	31414	522.41	51.22	-0.2
	Accommo-dations / Universal Tools	No	50639	523.54	51.28	
		Yes	12831	491.87	54.77	0.61
Mathematics	Race/ Ethnicity	White (not Hispanic)	47540	522.07	47.19	
		Asian/Pacific Islander	1369	543.27	54.3	-0.45
		Black (not Hispanic)	10477	482.06	49.41	0.84
		Hispanic	3747	504.22	47.38	0.38
		Am. Indian/Alaska N.	263	507.2	50.7	0.31
		Other	1753	510.73	50.98	0.24
	Gender	Male	33253	514.38	52.9	
		Female	31896	515.01	47.3	-0.01
	Accommo-dations / Universal Tools	No	49114	522.88	47.06	
		Yes	16035	489.61	51.35	0.69

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.8: Impact Analysis, Grade 8

Content Area	Category	Group	N	Mean	Std. Dev.	Effect Size
English Language Arts	Race/ Ethnicity	White (not Hispanic)	46339	536.48	48.69	
		Asian/Pacific Islander	1454	554.12	58.15	-0.36
		Black (not Hispanic)	10134	500.25	50.44	0.74
		Hispanic	3587	516.66	49.36	0.41
		Am. Indian/Alaska N.	303	522.82	48.13	0.28
		Other	1652	530.04	48.39	0.13
	Gender	Male	32251	522.23	52.65	
		Female	31218	537.52	48.42	-0.3
	Accommo-dations / Universal Tools	No	50905	536.25	48.55	
		Yes	12564	503.41	53.11	0.66
Mathematics	Race/ Ethnicity	White (not Hispanic)	37817	521.51	46.06	
		Asian/Pacific Islander	908	535.1	53.93	-0.29
		Black (not Hispanic)	9221	490.04	51.82	0.67
		Hispanic	3102	505.46	48.91	0.35
		Am. Indian/Alaska N.	274	514.03	43.4	0.16
		Other	1371	514.18	49.18	0.16
	Gender	Male	27503	512.11	51.14	
		Female	25190	518.29	46.49	-0.13
	Accommo-dations / Universal Tools	No	38979	523.57	45.35	
		Yes	13714	490.9	51.14	0.7

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.9: Grade 3 English Language Arts Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
3	Race/Ethnicity	White (not Hispanic)	48824	0.90	3.08
		Asian/Pacific Islander	1429	0.91	3.01
		Black (not Hispanic)	11576	0.89	3.10
		Hispanic	4297	0.89	3.12
		Am. Indian/Alaska N.	274	0.89	3.10
		Other	2784	0.90	3.09
	Gender	Male	35200	0.91	3.09
		Female	33984	0.90	3.08
	Accommodations/ Universal Tools	No	45088	0.90	3.07
		Yes	24096	0.89	3.11

Table 10.10: Grade 4 English Language Arts Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
4	Race/Ethnicity	White (not Hispanic)	48320	0.91	2.98
		Asian/Pacific Islander	1455	0.92	2.87
		Black (not Hispanic)	11076	0.90	3.10
		Hispanic	4209	0.90	3.09
		Am. Indian/Alaska N.	239	0.92	3.01
		Other	2366	0.91	3.01
	Gender	Male	34656	0.92	3.03
		Female	33008	0.91	2.98
	Accommodations/ Universal Tools	No	45144	0.90	2.96
		Yes	22520	0.91	3.10

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.11: Grade 5 English Language Arts Reliability and SEM by Subgroup

Form	Category	Group	N Count	Cronbach's Alpha	SEM
F1	Race/Ethnicity	White (not Hispanic)	5651	0.9	3.07
		Asian/Pacific Islander	182	0.88	2.88
		Black (not Hispanic)	990	0.91	3.24
		Hispanic	366	0.90	3.14
		Am. Indian/Alaska N.	23	-	-
		Other	289	0.9	3.1
	Gender	Male	3686	0.91	3.14
		Female	3815	0.90	3.05
	Accommodations/ Universal Tools	No	7001	0.90	3.08
		Yes	500	0.90	3.25
F2	Race/Ethnicity	White (not Hispanic)	5674	0.90	3.06
		Asian/Pacific Islander	148	0.90	2.9
		Black (not Hispanic)	1042	0.90	3.17
		Hispanic	387	0.89	3.12
		Am. Indian/Alaska N.	24	-	-
		Other	231	0.89	3.1
	Gender	Male	3737	0.91	3.1
		Female	3769	0.90	3.05
	Accommodations/ Universal Tools	No	7013	0.89	3.07
		Yes	493	0.91	3.2
F3	Race/Ethnicity	White (not Hispanic)	17136	0.91	3.17
		Asian/Pacific Islander	551	0.93	3.1
		Black (not Hispanic)	5002	0.89	3.24
		Hispanic	1987	0.90	3.22
		Am. Indian/Alaska N.	109	0.91	3.21
		Other	772	0.91	3.19
	Gender	Male	13592	0.91	3.21
		Female	11964	0.91	3.16
	Accommodations/ Universal Tools	No	7474	0.90	3.08
		Yes	18080	0.90	3.23
F4	Race/Ethnicity	White (not Hispanic)	5655	0.90	3.07
		Asian/Pacific Islander	164	0.90	2.87
		Black (not Hispanic)	954	0.90	3.2
		Hispanic	374	0.90	3.14
		Am. Indian/Alaska N.	29	-	-
		Other	254	0.90	3.1
	Gender	Male	3740	0.91	3.13
		Female	3690	0.90	3.05
	Accommodations/ Universal Tools	No	6933	0.90	3.08
		Yes	497	0.92	3.22

Note: Reliability and SEM not computed for groups smaller than 50 students.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.11: Grade 5 English Language Arts Reliability and SEM by Subgroup (cont.)

Form	Category	Group	N Count	Cronbach's Alpha	SEM
F5	Race/Ethnicity	White (not Hispanic)	5621	0.89	3.13
		Asian/Pacific Islander	191	0.89	2.99
		Black (not Hispanic)	1001	0.89	3.25
		Hispanic	378	0.89	3.18
		Am. Indian/Alaska N.	31	-	-
		Other	224	0.91	3.17
	Gender	Male	3721	0.90	3.16
		Female	3725	0.90	3.12
	Accommodations/ Universal Tools	No	6970	0.89	3.13
		Yes	476	0.89	3.30
F6	Race/Ethnicity	White (not Hispanic)	5596	0.90	3.07
		Asian/Pacific Islander	149	0.91	2.91
		Black (not Hispanic)	1013	0.91	3.19
		Hispanic	417	0.90	3.12
		Am. Indian/Alaska N.	35	-	-
		Other	259	0.89	3.12
	Gender	Male	3660	0.91	3.11
		Female	3809	0.90	3.07
	Accommodations/ Universal Tools	No	6942	0.90	3.08
		Yes	527	0.89	3.23
F7	Race/Ethnicity	White (not Hispanic)	2837	0.90	3.06
		Asian/Pacific Islander	71	0.87	2.76
		Black (not Hispanic)	495	0.91	3.22
		Hispanic	180	0.91	3.12
		Am. Indian/Alaska N.	24	-	-
		Other	119	0.90	3.07
	Gender	Male	1885	0.91	3.12
		Female	1841	0.90	3.04
	Accommodations/ Universal Tools	No	3494	0.90	3.07
		Yes	232	0.91	3.27

Note: Reliability and SEM not computed for groups smaller than 50 students.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.12: Grade 6 English Language Arts/Literacy Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
6	Race/Ethnicity	White (not Hispanic)	46408	0.89	3.06
		Asian/Pacific Islander	1467	0.91	2.96
		Black (not Hispanic)	10266	0.88	3.16
		Hispanic	3676	0.89	3.13
		Am. Indian/Alaska N.	270	0.89	3.13
		Other	1798	0.89	3.09
	Gender	Male	32312	0.90	3.10
		Female	31572	0.89	3.06
	Accommodations/ Universal Tools	No	49824	0.89	3.06
		Yes	14060	0.89	3.14

Table 10.13: Grade 7 English Language Arts/Literacy Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
7	Race/Ethnicity	White (not Hispanic)	46432	0.90	3.16
		Asian/Pacific Islander	1439	0.92	3.02
		Black (not Hispanic)	10080	0.89	3.19
		Hispanic	3570	0.90	3.21
		Am. Indian/Alaska N.	244	0.91	3.17
		Other	1702	0.91	3.15
	Gender	Male	32056	0.91	3.16
		Female	31412	0.90	3.16
	Accommodations/ Universal Tools	No	50632	0.90	3.16
		Yes	12830	0.91	3.17

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.14: Grade 8 English Language Arts/Literacy Reliability and SEM by Subgroup

Form	Category	Group	N Count	Cronbach's Alpha	SEM
F1	Race/Ethnicity	White (not Hispanic)	18540	0.91	3.08
		Asian/Pacific Islander	587	0.94	3.02
		Black (not Hispanic)	4663	0.89	3.15
		Hispanic	1608	0.90	3.15
		Am. Indian/Alaska N.	131	0.91	3.09
		Other	611	0.91	3.11
	Gender	Male	13662	0.92	3.13
		Female	12480	0.91	3.06
	Accommodations/ Universal Tools	No	15042	0.91	3.05
		Yes	11100	0.91	3.16
F2	Race/Ethnicity	White (not Hispanic)	11116	0.90	3.06
		Asian/Pacific Islander	344	0.93	2.87
		Black (not Hispanic)	2205	0.90	3.17
		Hispanic	787	0.91	3.14
		Am. Indian/Alaska N.	66	0.90	3.14
		Other	418	0.91	3.08
	Gender	Male	7447	0.91	3.11
		Female	7490	0.91	3.03
	Accommodations/ Universal Tools	No	14360	0.91	3.07
		Yes	576	0.90	3.21
F3	Race/Ethnicity	White (not Hispanic)	11102	0.90	3.02
		Asian/Pacific Islander	370	0.92	2.89
		Black (not Hispanic)	2162	0.90	3.14
		Hispanic	815	0.91	3.08
		Am. Indian/Alaska N.	76	0.89	3.14
		Other	400	0.90	3.07
	Gender	Male	7382	0.91	3.08
		Female	7543	0.91	3.00
	Accommodations/ Universal Tools	No	14320	0.91	3.04
		Yes	605	0.89	3.17
F4	Race/Ethnicity	White (not Hispanic)	5577	0.91	3.06
		Asian/Pacific Islander	153	0.91	2.94
		Black (not Hispanic)	1104	0.90	3.19
		Hispanic	377	0.89	3.15
		Am. Indian/Alaska N.	30	-	-
		Other	223	0.90	3.08
	Gender	Male	3760	0.91	3.12
		Female	3704	0.90	3.04
	Accommodations/ Universal Tools	No	7181	0.91	3.08
		Yes	283	0.90	3.22

Note: Reliability and SEM not computed for groups smaller than 50 students.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.2: Grade 3 Mathematics Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
3	Race/Ethnicity	White (not Hispanic)	48816	0.90	2.63
		Asian/Pacific Islander	1500	0.92	2.47
		Black (not Hispanic)	11594	0.88	2.74
		Hispanic	4342	0.90	2.71
		Am. Indian/Alaska N.	276	0.89	2.74
		Other	2784	0.90	2.67
	Gender	Male	35264	0.91	2.65
		Female	34040	0.90	2.66
	Accommodations/ Universal Tools	No	44120	0.90	2.60
		Yes	25192	0.90	2.74

Table 10.3: Grade 4 Mathematics Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
4	Race/Ethnicity	White (not Hispanic)	48344	0.91	2.72
		Asian/Pacific Islander	1506	0.93	2.55
		Black (not Hispanic)	11142	0.87	2.75
		Hispanic	4264	0.90	2.76
		Am. Indian/Alaska N.	240	0.91	2.76
		Other	2369	0.91	2.73
	Gender	Male	34768	0.92	2.72
		Female	33096	0.91	2.74
	Accommodations/ Universal Tools	No	43960	0.91	2.71
		Yes	23904	0.90	2.76

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.4: Grade 5 Mathematics Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
5	Race/Ethnicity	White (not Hispanic)	48200	0.89	2.72
		Asian/Pacific Islander	1506	0.92	2.61
		Black (not Hispanic)	10558	0.84	2.65
		Hispanic	4146	0.86	2.71
		Am. Indian/Alaska N.	277	0.88	2.71
		Other	2152	0.89	2.71
	Gender	Male	34136	0.90	2.70
		Female	32708	0.88	2.72
	Accommodations/ Universal Tools	No	44536	0.89	2.72
		Yes	22308	0.86	2.67

Table 10.18: Grade 6 Mathematics Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
6	Race/Ethnicity	White (not Hispanic)	48208	0.90	2.71
		Asian/Pacific Islander	1486	0.93	2.63
		Black (not Hispanic)	10626	0.86	2.66
		Hispanic	3878	0.88	2.69
		Am. Indian/Alaska N.	276	0.89	2.73
		Other	1879	0.90	2.70
	Gender	Male	33888	0.91	2.71
		Female	32456	0.90	2.69
	Accommodations/ Universal Tools	No	49128	0.89	2.71
		Yes	17220	0.88	2.68

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table 10.5: Grade 7 Mathematics Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
7	Race/Ethnicity	White (not Hispanic)	47536	0.90	2.85
		Asian/Pacific Islander	1369	0.93	2.79
		Black (not Hispanic)	10476	0.82	2.70
		Hispanic	3747	0.87	2.80
		Am. Indian/Alaska N.	263	0.87	2.82
		Other	1753	0.89	2.82
	Gender	Male	33248	0.90	2.82
		Female	31896	0.89	2.82
	Accommodations/ Universal Tools	No	49112	0.90	2.84
		Yes	16034	0.86	2.74

Table 10.20: Grade 8 Mathematics Reliability and SEM by Subgroup

Grade	Category	Group	N Count	Cronbach's Alpha	SEM
8	Race/Ethnicity	White (not Hispanic)	37816	0.84	2.90
		Asian/Pacific Islander	908	0.91	2.87
		Black (not Hispanic)	9220	0.80	2.79
		Hispanic	3102	0.81	2.86
		Am. Indian/Alaska N.	274	0.80	2.88
		Other	1371	0.84	2.88
	Gender	Male	27500	0.85	2.86
		Female	25188	0.83	2.90
	Accommodations/ Universal Tools	No	38976	0.84	2.90
		Yes	13714	0.79	2.81

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

References

Altman, D.G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

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. (2002). PARDUX [Computer program]. Monterey, CA: CTB/McGraw-Hill.

Camilli, G., & Shepard, A. L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA :Sage Publication.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20(1), 37–46.

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.

Dorans, N.J., & Schmitt, M.P. (1991). *Constructed response and differential item functioning: A pragmatic approach*. Princeton: Educational Testing Service.

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.

Landis, J.R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.

Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement,* 7, 201–210.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

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.

Thissen, D. (1982). Marginal maximum-likelihood estimation for the one-parameter logistic model. *Psychometrika*, 47, 175–186.

Thompson, S., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement,* 30*,* 187–213.

Yen, W. M. (1981). Using simulation results to choose a latent trait model. *Applied Psychological Measurement,* 5*,* 245–262.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233–251.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Appendix A

The methodology and results of vertical linking using alternative approaches are presented in this section of the document. In addition to linking assessments using items from below- and above-grade level, the Missouri TAC recommended conducting a study in which the assessments are linked using only below-grade level items and only above-grade level items.

A.1. Separate calibrations and chain linking

This study was conducted using separate calibration and chain linking method. Similarly to the concurrent calibration, described in the main body of this document, two IRT models were used to calibrate the operational and vertical scaling test items. The three-parameter logistic (3PL) model was used to estimate parameters for the multiple-choice items. The two-parameter partial credit (2PPC) model was used to estimate parameters for the constructed response items. The Pardux (Burket, 2002) program was used for all calibrations and equating (linking) in the study. The Stocking and Lord (1983) Test Characteristic Curve (TCC) method was used for equating.

The separate calibration and chain linking method was accomplished in two steps. The first step was separate calibration of each grade level's data. The second step was grade-by-grade chain linking using common items. The separate calibrations resulted in the establishment of a unique theta metric scale for each grade. Next, the common items used for linking adjacent grades allowed for the development of a common scale. The item parameter estimates for common items were used to estimate scale transformation constants that would allow placement of item parameters from each adjacent grade onto the base grade scale using the Stocking and Lord equating technique. This step was then repeated for each adjacent grade until all grades were placed on the common scale.

The separate calibration and chain linking for each content area was performed first with linking items from the below-grade level and then again with linking items from the above-grade level.

In each case, for both of the content areas, Grade 5 was identified as the base grade. The test data for Grade 5 (operational and vertical linking items administered in this grade) were calibrated first and item parameters in the theta metric were estimated.

Next, item parameters for items administered in Grade 5 were used as anchors to equate Grade 4 test data to the base grade scale. The Grade 4 assessment was equated to the Grade 5 scale in a separate equating run. A set of K1 and K2 equating constants was obtained for Grade 4 item parameters during the Stocking and Lord transformation.

In the next step, item parameters for items administered in Grade 4 (already on the same scale as the base Grade 5) were used to link the Grade 4 assessment to the base grade

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

scale. This equation was also conducted in a separate equating run and another set of K1 and K2 equating constants was obtained for Grade 3 item parameters. Grade 3 items were placed on the common scale.

Equating of grades 6, 7 and 8 to the base grade was conducted in the same manner. After linking, all grades' obtained item parameter estimates were on the same scale and were expressed in a theta scale metric.

In addition, during the calibration of Grade 5 data, the sample mean ability estimates were obtained for the base Grade 5. These estimates were used to identify transformation constants (M1 and M2) that would permit transformation of item parameter estimates from the theta metric into the scale score metric and thus, produce a scale with a desired mean and standard deviation for the base grade. As in the case of concurrent calibration, the mean in scale score metric for the base grade was set to be 500 and the standard deviation was set to be 50, for both ELA and Math.

The formulae used to compute transformation constants for the transformation of the item parameter estimates from the theta metric to the scale score metric were provided in Section 4 of the main body of the report.

Vertical scales established using only below- or only above-grade level items were evaluated using the same criteria as for the scale established using concurrent calibrations with all linking items as described in Section 4 of this report. The scale evaluation included examination of the pattern of grade-to-grade growth (means), grade-to grade variability (standard deviations), and separation of scale score distributions across grades); as well as the test characteristic curves (TCC's) and standard error (SE) curves. Only on-grade level operational test items were used in computation of statistics used in scale evaluation.

DRAFT

A.2. ELA Scale Developed Using Below-Grade Level Linking Items

Item calibration: The following item sets were calibrated separately on-grade level: base Grade 5 data with linking items from Grade 4; Grade 4 data with linking items from Grade 3; Grade 3 data (no linking items); Grade 6 data with linking items from Grade 5; Grade 7 data with linking items from Grade 6; and Grade 8 data with linking items from Grade 7. All items converged in the calibration.
Item fit was evaluated using the criteria described in Section 6.4.3 of the report. There were 9 items flagged for poor fit across all grades. Table A.1 shows the statistics for each flagged item in separate calibrations with below-grade level linking items.

Adjacent Grade Linking: The Stocking and Lord (1983) Test Characteristic Curve (TCC) method was used for linking consecutive grades starting from the base Grade 5. First, Grade 4 was linked to Grade 5 using the common items (from Grade 4) between Grades 4 and 5. This linking placed Grade 4 assessment onto the base grade scale. Next, Grade 3

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

was linked to Grade 4 using the common items (from Grade 3) between Grades 3 and 4. This linking placed Grade 3 assessment onto the common scale. Grade 6 was then linked to the common scale using items from Grade 5 administered to Grade 6 students. Grade 7 was linked to Grade 6 using common items between these two grades (Grade 6 items administered to Grade 7 students). And last, but not least, Grade 8 assessment was linked to the common scale using a common set of Grade 7 items administered to Grade 8 students. There were 10 or 11 linking items between each pair of adjacent grades.

The linking results were evaluated and are presented in Table A.2. The information in the table includes K1 and K2 linking constants determined by minimizing the quadratic loss function (F), number of iterations, a value of the F function, correlations between anchor item input and estimate a-parameter, b-parameter, and c-parameter as well as the number of outliers anchor items identified by plotting the input and estimated item parameters along with the line of best fit. Items with an absolute difference of parameters greater than two times the root mean squared difference were considered to be outliers. Overall, the linking results were acceptable and no anchor items were excluded from linking.

Scale Evaluation: The ELA scale developed using only below-grade level linking items as anchor was evaluated using the same criteria as the scale developed using a concurrent calibration with all linking items included and described in the main body of the report. The M1 and M2 transformation constants were derived to place the parameters in the theta metric onto the scale score metric and yield a scale score mean of 500 and standard deviation of 50 for Grade 5. The M1 and M2 constants were: 45.78755 and 500.45788, respectively.

DRAFT

The grade-level scale score means, standard deviations, mean difference between adjacent grades, and grade-level specific scale bounds are presented in Table A.3. As seen in Table A.3, the scale score means increase as the grade level increases but the mean differences between grades were not uniform across grade levels. Most growth across grades is observed between Grades 3 and 4, followed by growth between Grades 4 and 5, and between Grades 7 and 8. Less growth is observed between Grades 5 and 6, and between Grades 6 and 7. This growth pattern was found to be similar to the growth pattern observed for the scale established via concurrent calibration with all linking items included. However, the grade separation as indicated by the scale score means was smaller compared to the grade separation observed in concurrent calibration with all linking items included. The standard deviations were relatively uniform across grades ranging from 45.6 for Grade 3 to 52.5 for Grade 7.

Looking at the scale score distribution at selected percentiles, it was observed that the scale scores increase as the percentile and grade level increase showing continuous progress upward from Grades 3 through 8 at all selected percentiles except for the 10th and 25th percentiles for Grades 6 and 7. Higher scale scores for Grade 6 at the lower ability end indicate that lower ability Grade 6 students may perform better on the ELA assessment compared to lower ability Grade 7 students.
This pattern was similar to the one observed for the scale established via concurrent calibration with all linking items included.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure A.1 shows TCCs and SE curves for the scale established using only below-grade level linking items. For clarity of the graph, only one Grade 5 form and only one Grade 8 form were presented in the graph. As demonstrated in Figures 6.3, 6.4, 6.5, and 6.6 of this report the TCCs and SE curves of the multiple forms in Grades 5 and 8 were well aligned.

As demonstrated in Figure A1, the TCCs for Grades 3 and 4 were overlapping at the lower and middle ability range and crossing at the upper end of ability scale indicating that the Grade 3 and 4 tests were of similar difficulty for lower and medium-ability students, and the Grade 3 test was more difficult than Grade 4 at the upper end of the ability scale. Grades 6, 7, and 8 TCCs, while in the middle range of the ability scale, were overlapping at the lower and higher ends of the scale, indicating that the test may be of similar difficulty for lower and higher ability students. This TCC pattern was similar to that obtained from the concurrent calibration with all linking items included.

The SE curves in Figure A2 are U-shaped as expected indicating a lower standard error at the middle of the ability scale and a higher standard error at the lower and upper end of the scale.

DRAFT

A.3. ELA Scale Developed Using Above-Grade Level Linking Items

Item calibration: The following item sets were calibrated separately: base Grade 5 data with linking items from Grade 6; Grade 4 data with linking items from Grade 5; Grade 3 data with linking items from Grade 4; Grade 6 data with linking items from Grade 7; Grade 7 data with linking items from Grade 8; and Grade 8 data without any linking items. All items, except for one, converged in the calibration. One item in Grade 8 (ID=755209C; Writing Prompt, Component C) did not converge and was hand-fitted in Pardux. This item is an operational test item and item suppression from scoring was not recommended. Item fit was evaluated using the criteria described in Section 6.4.3 of the report. There were 12 items flagged for poor fit across all grades. Table A.5 shows the statistics for each flagged item in separate calibrations with above-grade level linking items.

Adjacent Grade Linking: The Stocking and Lord (1983) Test Characteristic Curve (TCC) method was used for linking consecutive grades starting from the base Grade 5. First, Grade 4 was linked to Grade 5 using the common items (from Grade 5) between Grades 4 and 5. This linking placed Grade 4 assessment onto the base grade scale. Next, Grade 3 was linked to Grade 4 using the common items (from Grade 4) between Grades 3 and 4. This linking placed Grade 3 assessment onto the common scale. Grade 6 was then linked to the common scale using items from Grade 6 administered to Grade 5 students. Grade 7 was linked to Grade 6 using common items between these two grades (Grade 7 items administered to Grade 8 students). Grade 8 assessment was linked to the common scale

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

using a common set of Grade 8 items administered to Grade 7 students. There were 10 or 11 linking items between each pair of adjacent grades.

The linking results were evaluated and are presented in Table A.6. Overall, the linking results were acceptable and no anchor items were excluded from linking.

Scale Evaluation: The ELA scale developed using only above-grade level linking items as anchor was evaluated using the same criteria as the scale developed using a concurrent calibration with all linking items included and described in the main body of the report. The M1 and M2 transformation constants were derived to place the parameters in the theta metric onto the scale score metric and yield a scale score mean of 500 and standard deviation of 50 for Grade 5. The M1 and M2 constants were: 45.74565 and 500.91491, respectively.

The grade-level scale score means, standard deviations, mean difference between adjacent grades, and grade-level specific scale bounds are presented in Table A.7. As seen in Table A.7, the scale score means increase as the grade level increases but the mean differences between grades were not uniform across grade levels. Most growth across grades is observed between Grades 3 and 4, followed by growth between Grades 4 and 5, between Grades 6 and 7, and between Grades 7 and 8. Least growth is observed between Grades 5 and 6. This growth pattern was found to be slightly different from the growth pattern observed for the scale established via concurrent calibration with all linking items included and the scale established using only below-grade level linking items. The grade separation as indicated by the scale score means was larger compared to the grade separation observed in concurrent calibration with all linking items included for Grades 3-to-4, 4-to-5, 6-to-7, and 7-to-8. However, mean difference between Grades 5 and 6 was smaller for the scale established with linking items from the grade above compared to the scale established with all anchor items. In Table A.7, the standard deviations ranged from approximately 51 for Grade 5 to approximately 43 for Grade 8 and were found to be decreasing between grades 5 through 8.

Looking at the scale score distribution at selected percentiles, it was observed that the scale scores increase as the percentile and grade level increase showing continuous progress upward from grades 3 through 8 at all selected percentiles.

As demonstrated in Figure A3, the TCCs for Grades 3 and 4 were very close to each other at the lower and middle ability range and crossing at the upper end of the ability scale indicating that the Grade 4 test was only slightly more difficult than the Grade 3 test for lower and medium-ability students, and the Grade 3 test was more difficult than Grade 4 at the upper end of the ability scale. Grade 4 through 8 TCCs were ordinal for the most part of the ability range indicating increasing difficulty as the grade level increases.

The SE curves in Figure A4 are U-shaped as expected indicating a lower standard error at the middle of the ability scale and a higher standard error at the lower and upper end of the scale.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

A.4. Mathematics Scale Developed Using Below-Grade Level Linking Items

Item calibration: The separate calibrations, starting from the base Grade 5, were conducted as described in Section A.2. All items, except for one, converged in the calibration. One Mathematics Grade 7 item (ID=686910) did not converge in calibration and was hand-fitted in Pardux. This item is an operational test item and item suppression from scoring was not recommended. Item fit was evaluated using the criteria described in Section 6.4.3 of the report. There were 7 items flagged for poor fit across all grades. Table A.9 shows the statistics for each flagged item in separate calibrations with below-grade level linking items.

Adjacent Grade Linking: The Stocking and Lord Test Characteristic Curve (TCC) method was used for linking consecutive grades starting from the base Grade 5 as described in Section A.2 of the Appendix. There were 10 linking items between each pair of adjacent grades, except for the anchor set of 8 items for linking Grades 6 and 7. The linking results are presented in Table A.10. Overall, the linking results were acceptable and no anchor items were excluded from linking.

Scale Evaluation: The Mathematics scale developed using only below-grade level linking items as anchors was evaluated using the same criteria as the scale developed using the concurrent calibration with all linking items included and described in the main body of the report. The M1 and M2 transformation constants were derived to place the parameters in the theta metric onto the scale score metric and yield a scale score mean of 500 and standard deviation of 50 for Grade 5. The M1 and M2 constants were: 41.32231 and 503.30579, respectively.

DRAFT

The grade-level scale score means, standard deviations, mean difference between adjacent grades, and grade-level specific scale bounds are presented in Table A.11. As seen in Table A.11, the scale score means increase as the grade level increases but the amount of growth decreases as the grade level increases. The largest growth is observed between Grades 3 and 4 (approximately 20 scale score points) and the least growth (less than 5 scale score points) is seen between Grades 7 and 8. Less growth is observed between Grades 5 and 6, and between Grades 6 and 7. This growth pattern was found to be slightly different from the growth pattern observed for the scale established via concurrent calibration with all linking items included (which showed about the same amount of growth between Grades 5 and 6, and Grades 6 and 7, and relatively no growth between Grades 7 and 8). In Table A.11 the standard deviations were generally larger in lower grades (52 scale score points in Grades 3 and 4) and smaller in higher grades (approximately 40 scale score points in Grade 8).

Looking at the scale score distribution at selected percentiles, it was observed that the scale scores increase as the percentile and grade level increase showing continuous progress upward from grades 3 through 8 at all selected percentiles except for the 75th

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

percentile and above for Grade 8. Higher scale scores for Grade 7 at the higher ability end indicate that higher ability Grade 7 students may perform better on the Mathematics assessment compared to higher ability Grade 8 students. This pattern was similar to the one observed for the scale established via concurrent calibration with all linking items included. Possible factors contributing to these results were discussed in the main body of the report.

As shown in Figure 6, the TCCs for all grades, except Grades 5 and 6, were ordinal indicating increasing test difficulty as the grade level increases. Grade 5 and 6 TCCs were overlapping indicating that these two tests may be of the same difficulty. Grade 7 and 8 TCCs were very close to each other suggesting that the Grade 8 test was only slightly more difficult than the Grade 7 test. This TCC pattern was similar to that obtained from the concurrent calibration with all linking items included.

A.5. Mathematics Scale Developed Using Above-Grade Level Linking Items

Item calibration: The separate calibrations, starting from the base Grade 5, were conducted as described in Section A.3. All items, except for one Grade 7 item (ID= 686910) converged in the calibration. The non-convergent item was hand-fitted in Pardux (note: the same item did not converge in Grade 7 calibration with linking items from Grade 6). As stated in the previous section, this item is an operational test item and item suppression from scoring was not recommended. Item fit was evaluated using the criteria described in Section 6.4.3 of the report. There were 3 items flagged for poor fit across all grades. Table A.13 shows the statistics for each flagged item in separate calibrations with above-grade level linking items.

DRAFT

Adjacent Grade Linking: The Stocking and Lord (1983) Test Characteristic Curve (TCC) method was used for linking consecutive grades starting from the base Grade 5 as explained in Section A.3. There were 10 linking items between each pair of adjacent grades. The linking results were evaluated and are presented in Table A.14. No anchor items were excluded from linking.

Scale Evaluation: The Mathematics scale developed using only above-grade level linking items as anchor was evaluated using the same criteria as the scale developed using a concurrent calibration with all linking items included and described in the main body of the report. The M1 and M2 transformation constants were derived to place the parameters in the theta metric onto the scale score metric and yield a scale score mean of 500 and standard deviation of 50 for Grade 5. The M1 and M2 constants were: 40.81630 and 504.08163, respectively.

The grade-level scale score means, standard deviations, mean difference between adjacent grades, and grade-level specific scale bounds are presented in Table A.15. As seen in Table A.15, the scale score means increase as the grade level increases from grades 3 to 7 but the amount of growth decreases as the grade level increases. The largest growth is observed between Grades 3 and 4 (approximately 31 scale score points) and the

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

least growth (approximately 13 scale score points) is seen between Grades 6 and 7. No growth is observed between Grades 7 and 8; additionally, the Grade 8 scale score mean was lower by approximately 1 scale score point than the Grade 7 scale score mean. This growth pattern was found to be slightly different from the growth pattern observed for the scale established via concurrent calibration with all linking items included (which showed about the same amount of growth between Grades 5 and 6, and Grades 6 and 7). In Table A.15, the standard deviations were generally smaller in lower grades (39 scale score points in Grades 3) and larger in higher grades (approximately 62 scale score points in Grades 7 and 8). This standard deviation pattern was opposite of the pattern observed for the Mathematics scale developed using only below-grade level linking items.

Looking at the scale score distribution at selected percentiles, it was observed that the scale scores increase as the percentile and grade level increase showing continuous progress upward from grades 3 through 7 at all selected percentiles. The scale scores for Grade 8 were lower than the scale scores for Grade 7 at all percentiles except for the 50th percentile at which the scale scores for Grade 8 were equal to the scale scores for Grades 7. The Grade 7 and 8 scale score distribution pattern seem to suggest that Grade 7 students may perform equally well or slightly better on the Mathematics test compared to Grade 8 students.

As demonstrated in Figure A7, the TCCs for all grades were ordinal for the most part of the ability range indicating increasing difficulty as the grade level increases. Some TCC cross-over was observed for Grades 5 and 6 TCCs and for Grades 7 and 8 TCCs at the lower end of the ability scale.

The SE curves in Figure A8 are U-shaped as expected indicating a lower standard error at the middle of the ability scale and a higher standard error at the lower and upper end of the scale.

DRAFT

A.6. Summary

In summary, the scale properties obtained from a concurrent calibration using linking items from below- and above-grade levels, and the scale properties obtained from separate calibrations/chain linking with either below- or above-grade level linking items show several common patterns and also a few differences.

The mean scale scores increased as grade level increased regardless of the calibration method and the linking set used. An exception to this pattern was little or no growth occurring between Grades 7 and 8 Mathematics and this phenomenon was discussed in Section 6.4.4 of this report.
The scale developed using only above-grade level items tended to result in more growth between grades than the scale from the concurrent calibration with all linking item. The concurrent calibration scale developed with all linking items included, in turn, showed more growth between grades than the scale established using only below-grade level linking items. This pattern was observed for both ELA and Mathematics.

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

When all available linking items were used in concurrent calibration, the resulting standard deviations tended to be more uniform across grades for both ELA and Mathematics. Similar pattern was also observed for separate calibrations with below-grade level linking items for ELA.
On the other hand, when only above-grade level items were used for linking of ELA tests, the standard deviation of the scale scores tended to be smaller for higher grade levels than for lower grade levels. The choice of linking items also appeared to affect the scale standard deviations for Mathematics. A pattern of decreasing standard deviations as the grade level increased was observed for the scale established using only below-grade level linking items. The opposite pattern of increasing standard deviations as the grade level increased was noticed for the Mathematics scale developed using only above-grade level items.

When scale score distributions were examined, a general upward progression of scale scores across grades 3 through 8 for ELA and across grades 3 through 7 for Mathematics was observed for all linking options. Again, the exception was Mathematics Grade 8 scale score distribution which showed little or no separation from the Grade 7 scale score distribution, particularly at the higher end of the ability scale.

When TCCs were examined, the scaling with above-grade level linking items resulted in the most desirable TCCs in regard to their ordinality and separation for both ELA and Mathematics. The results from the concurrent calibration with all linking items included and from the separate calibrations and linking using only below-grade level items were similar to each other. While they resulted in mostly ordinal TCCs, there were exceptions of crossing TCCs (ELA Grades 3 and 4) and overlapping TCCs (Mathematics Grades 5 and 6).

Taken together, multiple factors were taken into consideration when evaluating the scales. While an argument could be made that the scales developed using only above-grade level items may be technically superior, based on the largest amount of growth between the grades as indicated by the scale score means and the best ordinality of TCCs; a counter-argument can be offered that using only items measuring the concepts the students did not have a chance to learn yet for linking may negatively affect the scale validity.

Including items from the below- and above-grade levels for linking purposes provided students with an opportunity to demonstrate their ability on a wider range of the Missouri Learning Standards compared to the linking approach with only below- or only above-grade level items. Given this content-based factor, as well as the satisfactory results of scaling, the concurrent calibration with all linking items included was determined to be the most appropriate method of vertical scales development for Missouri ELA and Mathematics tests.

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table A.1: Items flagged for poor fit: ELA separate calibrations with below-grade level linking items

Item Position In Calib.	Item ID	Item Grade	Model	ChiSqr	Chi DF	Total N	Z Score	Z Obsd	Z Pred	Obsd-Pred
41	692626	3	2PPC	1080.6	8	69186	268.15	0.4433	0.4435	-0.0002
14	698163	4	3PL	799.71	7	67643	211.86	0.7084	0.7102	-0.0017
56	755385	4	2PPC	902.52	8	67643	223.63	0.5098	0.5111	-0.0013
41	693154	5	2PPC	954.29	8	66630	236.57	0.4072	0.4098	-0.0026
15	699807	6	2PPC	1048.72	8	63885	260.18	0.1210	0.1248	-0.0038
36	693930	6	3PL	1104.54	7	63885	293.33	0.0622	0.0670	-0.0049
35	755285	7	2PPC	1823.22	8	63467	453.81	0.4567	0.4581	-0.0015
37	694024	7	3PL	660.92	7	63467	174.77	0.6252	0.6277	-0.0025
38	755270	8	2PPC	1119.31	8	63458	277.83	0.2666	0.2693	-0.0027

Table A.2: Equating summary: ELA linking using below-grade level linking items

Grade	K1	K2	# of Iterations	F Value	Corr A	# of Outliers	Corr B	# of Outliers	Corr C	# of Outliers
3	0.8705	-0.6480	6	0.656179	0.96	0	0.93	0	-0.03	0
4	0.8915	-0.1683	8	0.233596	0.88	0	0.97	0	0.87	0
6	0.8656	0.2557	18	0.443474	0.93	0	0.91	0	0.97	1
7	1.0162	0.3206	21	0.153081	0.84	0	0.50	0	0.53	0
8	0.9954	0.4664	11	0.291158	0.94	1	0.91	0	0.87	0

Table A.3: Scale properties: ELA scaling with below-grade level linking items

Grade	Scale statistics		Mean difference between grades (in scale score points)	LOSS	HOSS
	Mean	SD			
3	467.49	45.55		250	730
4	489.52	46.48	22.03	260	740
5	499.90	50.71	10.38	270	780
6	508.22	45.91	8.32	280	790
7	510.71	52.47	2.49	290	800
8	521.12	50.35	10.41	300	830

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table A.4: Scale score distribution at selected percentiles: ELA scaling with below-grade level linking items

Grade	Percentile				
	10th	25th	50th	75th	90th
3	411	439	468	497	523
4	430	460	491	520	545
5	436	468	502	534	560
6	451	480	510	539	565
7	443	477	513	547	575
8	458	489	522	554	583

Table A.5: Items flagged for poor fit: ELA separate calibrations with above-grade level linking items

Item Position In Calib.	Item ID	Item Grade	Model	ChiSqr	Chi DF	Total N	Z Score	Z Obsd	Z Pred	Obsd-Pred
41	692626	3	2PPC	1116.93	8	69186	277.23	0.4433	0.4435	-0.0002
14	698163	4	3PL	873.90	7	67642	231.69	0.7084	0.7107	-0.0022
56	755385	4	2PPC	888.71	8	67642	220.18	0.5098	0.5109	-0.0011
61	693154	4	2PPC	100.47	8	7304	23.12	0.3739	0.3777	-0.0038
41	693154	5	2PPC	935.74	8	66629	231.94	0.4072	0.4096	-0.0025
77	693930	5	3PL	128.36	7	7446	32.43	0.0469	0.0506	-0.0038
15	699807	6	2PPC	1065.58	8	63885	264.40	0.1210	0.1248	-0.0038
36	693930	6	3PL	1075.43	7	63885	285.55	0.0622	0.0670	-0.0048
70	694024	6	3PL	120.17	7	7858	30.25	0.5774	0.5817	-0.0043
35	755285	7	2PPC	1834.77	8	63466	456.69	0.4567	0.4579	-0.0012
37	694024	7	3PL	661.58	7	63466	174.94	0.6252	0.6275	-0.0023
38	755270	8	2PPC	1086.24	8	63454	269.56	0.2666	0.2695	-0.0030

DRAFT

Table A.6: Equating summary: ELA linking using above-grade level linking items

Grade	K1	K2	# of Iterations	F Value	Corr A	# of Outliers	Corr B	# of Outliers	Corr C	# of Outliers
3	0.9170	-1.0340	4	0.370172	0.89	0	0.97	0	0.15	1
4	0.9709	-0.4262	9	0.092993	0.98	0	0.99	0	0.45	0
6	0.8909	0.1981	7	0.898629	0.32	0	0.98	0	0.55	1
7	0.8587	0.5234	24	0.372946	0.85	0	0.62	0	-0.65	0
8	0.8377	0.7771	15	0.615348	0.84	0	0.92	0	0.95	1

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table A.7: Scale properties: ELA scaling with above-grade level linking items

Grade	Scale statistics		Mean difference between grades (in scale score points)	LOSS	HOSS
	Mean	SD			
3	448.99	49.03		230	730
4	478.61	50.55	29.62	240	740
5	500.01	50.91	21.40	250	780
6	505.57	47.39	5.56	260	790
7	521.14	44.62	15.57	280	810
8	536.26	42.57	15.12	290	820

Table A.8: Scale score distribution at selected percentiles: ELA scaling with above-grade level linking items

Grade	Percentile				
	10th	25th	50th	75th	90th
3	390	420	451	481	508
4	415	448	481	512	539
5	436	468	502	534	560
6	447	476	507	537	564
7	464	492	523	551	575
8	484	510	538	564	588

Table A.9: Items flagged for poor fit: Mathematics separate calibrations with below-grade level linking items

Item Position In Calib.	Item ID	Item Grade	Model	ChiSqr	Chi DF	Total N	Z Score	Z Obsd	Z Pred	Obsd-Pred
42	689401	5	2PPC	5088.82	8	66811	1270.2	0.4007	0.3981	0.0026
29	689665	6	2PPC	716.03	8	66337	177.01	0.1391	0.1424	-0.0033
58*	690383	6	2PPC	326.68	8	24602	79.67	0.1730	0.1748	-0.0019
3	686972	7	3PL	673.70	7	65135	178.18	0.1579	0.1622	-0.0043
34	688423	7	2PPC	860.01	8	65135	213.00	0.1178	0.1196	-0.0018
57*	686030	7	3PL	2641.12	7	23924	704.00	0.1239	0.2398	-0.1159
1*	687356	8	2PPC	433.25	8	26704	106.31	0.0810	0.0825	-0.0014

Note: * in the first column indicate a linking item

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table A.10: **Equating summary: Mathematics linking using below-grade level linking items**

Grade	K1	K2	# of Iterations	F Value	Corr A	# of Outliers	Corr B	# of Outliers	Corr C	# of Outliers
3	1.0975	-0.8574	19	0.069209	0.89	1	0.98	0	0.82	0
4	1.0621	-0.3782	32	0.279261	0.91	1	0.90	1	0.69	0
6	0.9585	0.0954	28	0.639622	0.80	0	0.99	0	0.98	0
7	0.9138	0.3418	49	1.079371	0.78	1	0.96	0	0.28	1
8	0.6949	0.4695	58	0.095241	0.89	0	0.93	0	0.98	0

Table A.11: Scale properties: Mathematics scaling with below-grade level linking items

Grade	Scale statistics		Mean difference between grades (in scale score points)	LOSS	HOSS
	Mean	SD			
3	465.63	52.19		290	650
4	485.65	52.32	20.02	320	680
5	499.97	50.45	14.32	330	710
6	506.07	44.91	6.10	340	730
7	511.55	48.67	5.48	350	740
8	516.15	39.64	4.60	390	770

Table A.12: Scale score distribution at selected percentiles: Mathematics scaling with below-grade level linking items

Grade	Percentile				
	10th	25th	50th	75th	90th
3	401	434	468	500	528
4	423	456	489	520	546
5	438	470	503	533	559
6	451	479	509	536	560
7	458	488	516	543	566
8	470	498	521	541	559

Table A.13: Items flagged for poor fit: Mathematics separate calibrations with above-grade level linking items

Item Position In Calib.	Item ID	Item Grade	Model	ChiSqr	Chi DF	Total N	Z Score	Z Obsd	Z Pred	Obsd-Pred
13	687133	5	3PL	727.59	7	66811	192.58	0.7247	0.7202	0.0045
42	689401	5	2PPC	5004.90	8	66811	1249.23	0.4007	0.3977	0.0029
34	688423	7	2PPC	847.89	8	65132	209.97	0.1178	0.1196	-0.0019

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Table A.14: Equating summary: Mathematics linking using above-grade level linking items

Grade	K1	K2	# of Iterations	F Value	Corr A	# of Outliers	Corr B	# of Outliers	Corr C	# of Outliers
3	0.8085	-1.3345	25	0.562189	0.61	0	0.49	0	0.88	0
4	1.0017	-0.5505	46	0.344658	0.84	0	0.96	2	0.95	1
6	1.3245	0.4538	7	0.331747	0.94	1	0.95	1	0.33	1
7	1.2610	0.8948	67	0.981519	0.53	1	0.92	0	0.93	0
8	1.2524	0.9110	43	0.232792	0.84	1	0.91	1	0.85	0

Table A.15: Scale properties: Mathematics scaling with above-grade level linking items

Grade	Scale statistics		Mean difference between grades (in scale score points)	LOSS	HOSS
	Mean	SD			
3	448.06	38.63		290	650
4	479.78	49.09	31.72	320	680
5	499.93	50.59	20.15	330	710
6	521.09	59.55	21.16	340	730
7	533.83	62.26	12.74	350	740
8	532.46	62.19	-1.37	390	770

Table A.16: Scale score distribution at selected percentiles: Mathematics scaling with above-grade level linking items

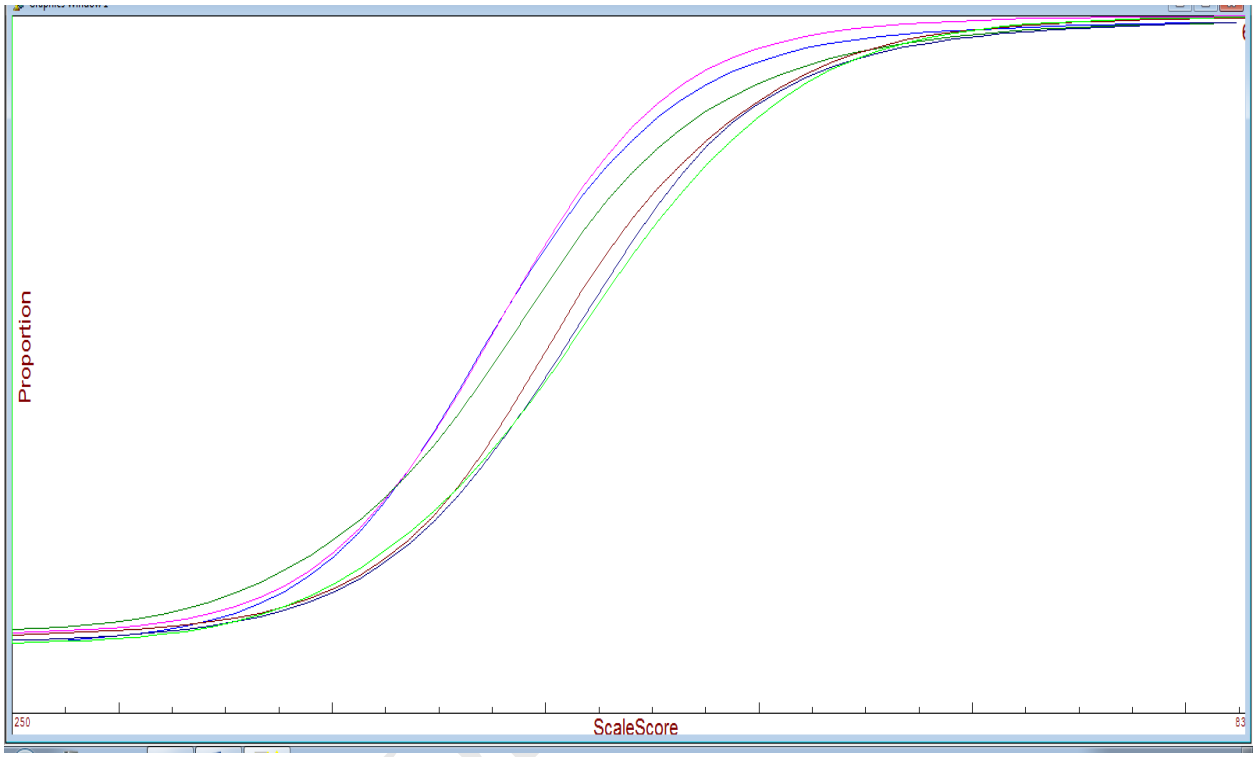
Grade	Percentile				
	10th	25th	50th	75th	90th
3	400	425	450	473	494
4	421	452	483	511	536
5	438	471	503	533	559
6	445	484	525	562	594
7	459	500	539	576	607
8	447	497	539	574	606

DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

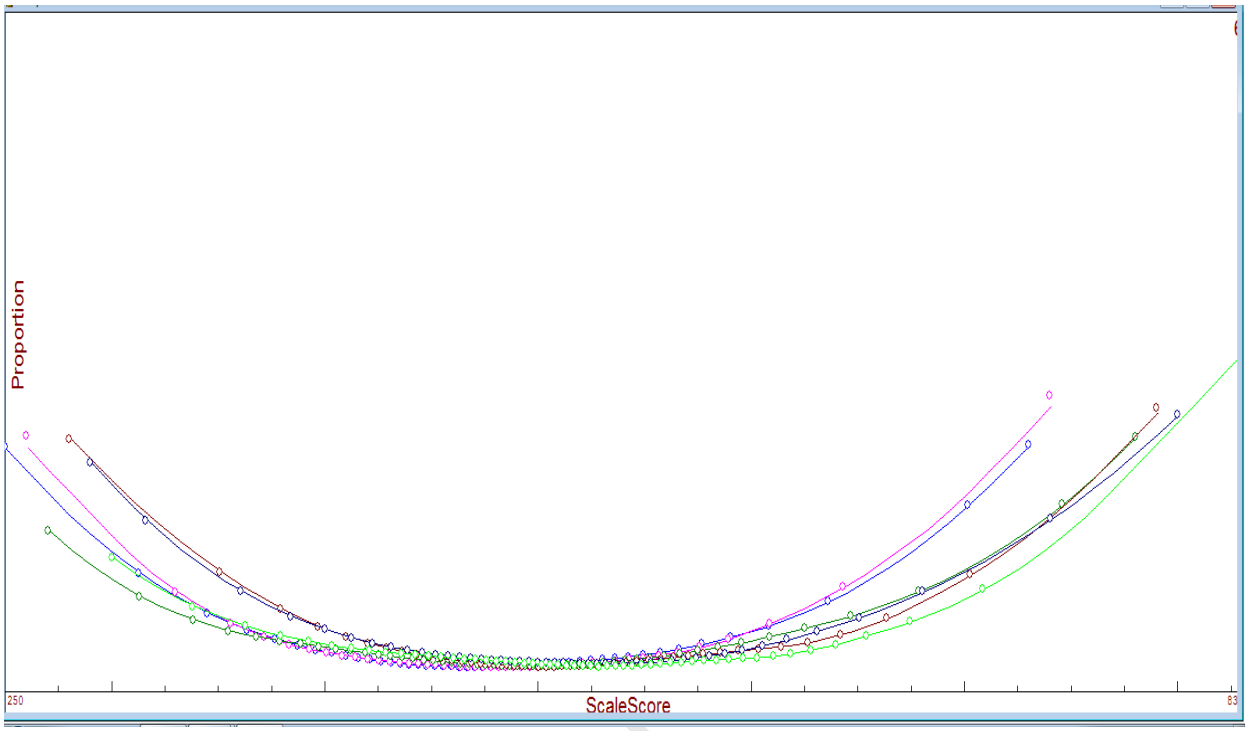
In Figures A1 to A8 the TCCs and the SE curves are color-coded as follows:
Grade 3=Blue; Grade 4=Pink, Grade 5 Form 1=Green; Grade 6=Brown; Grade 7=Dark Blue; Grade 8=Lime.

Figure A1. TCCs: ELA scaling with below-grade level linking items



Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

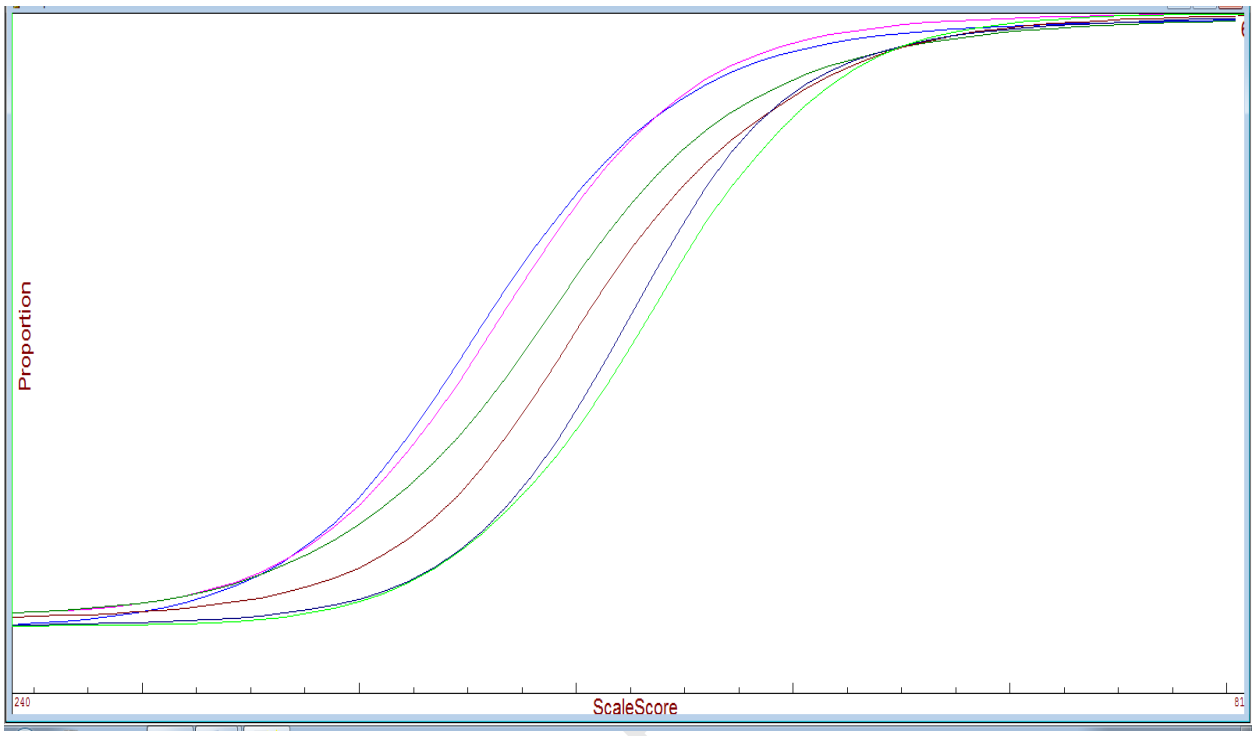
Figure A2. SE Curves: ELA scaling below-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

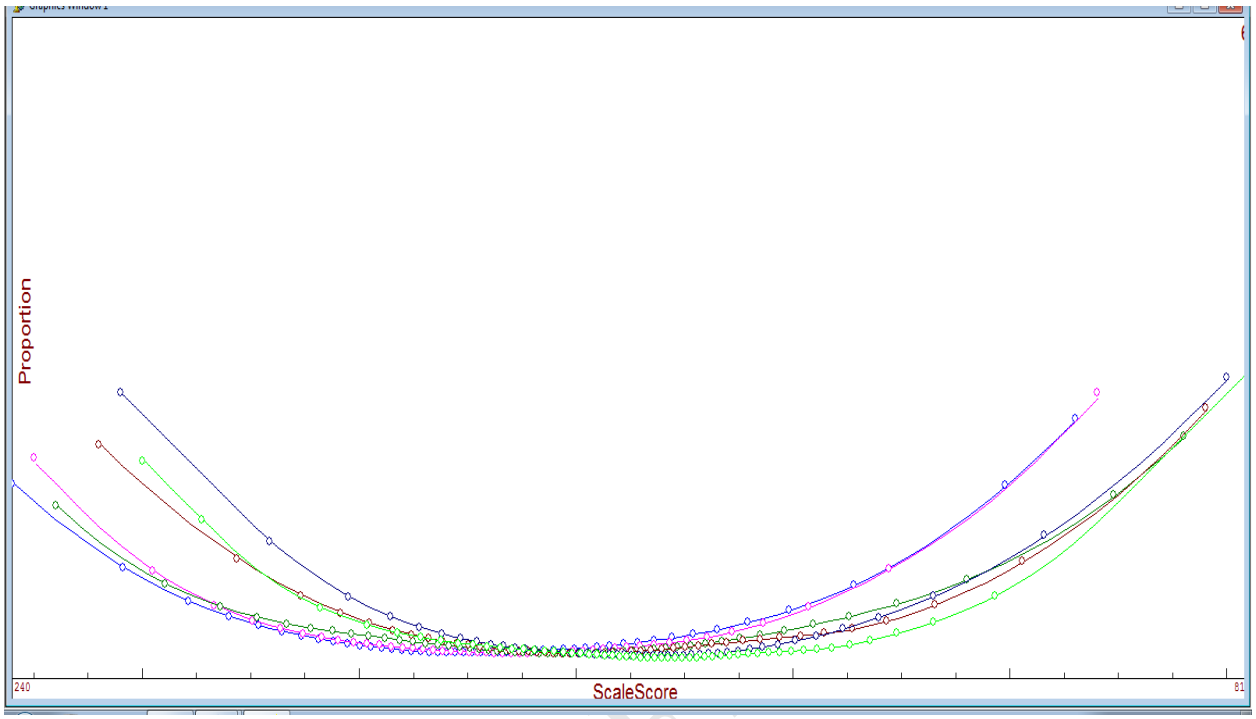
Figure A3. TCCs: ELA scaling with above-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

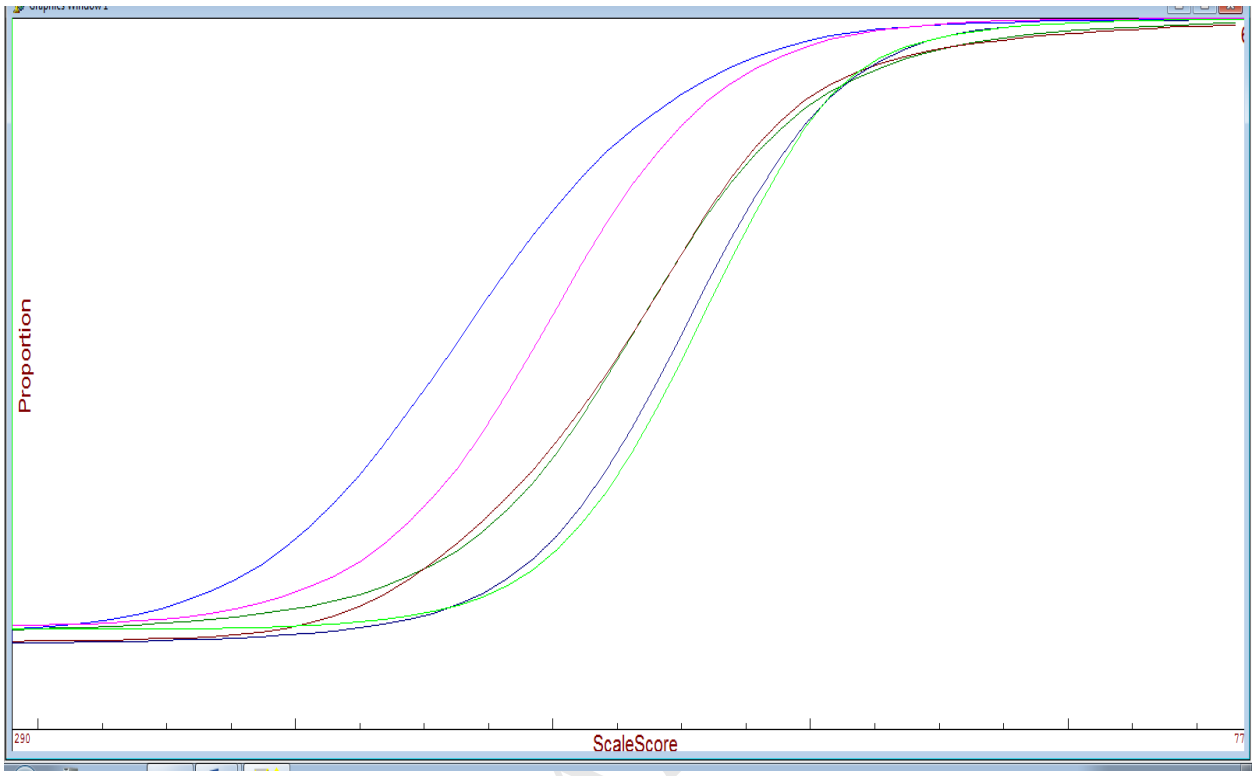
Figure A4. SE Curves: ELA scaling above-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

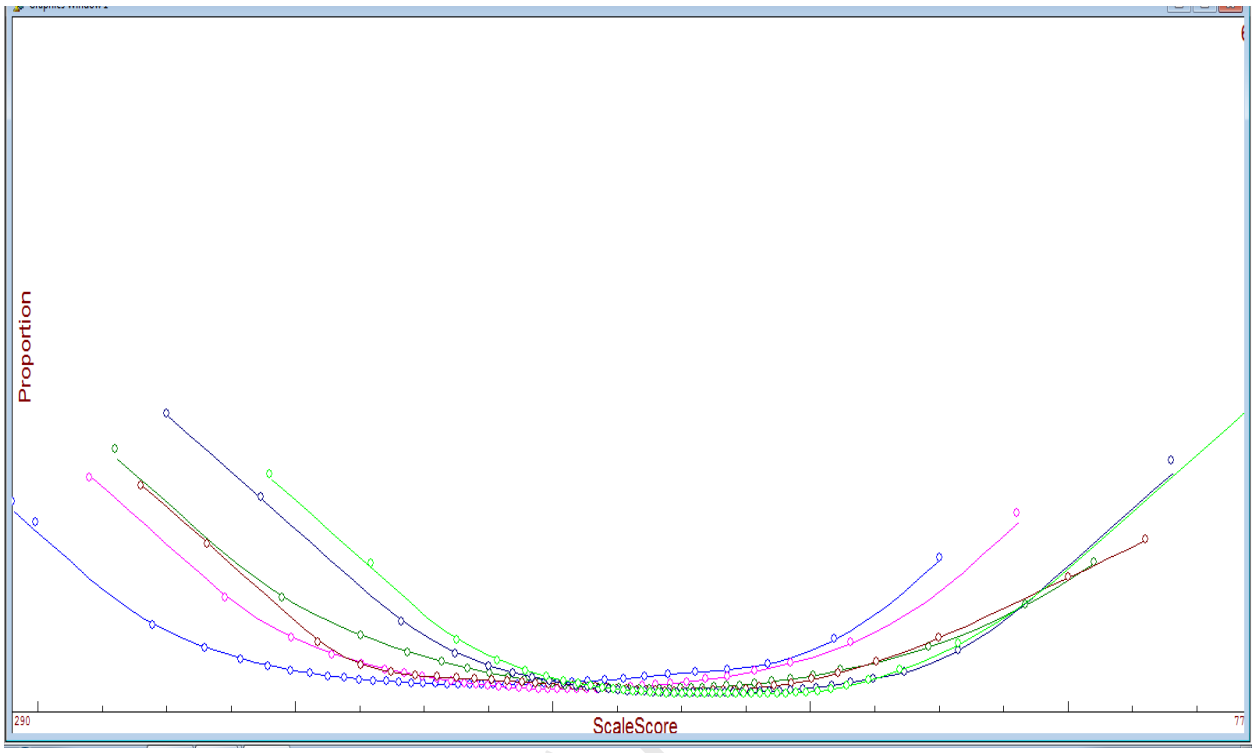
Figure A5. TCCs: Mathematics scaling with below-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

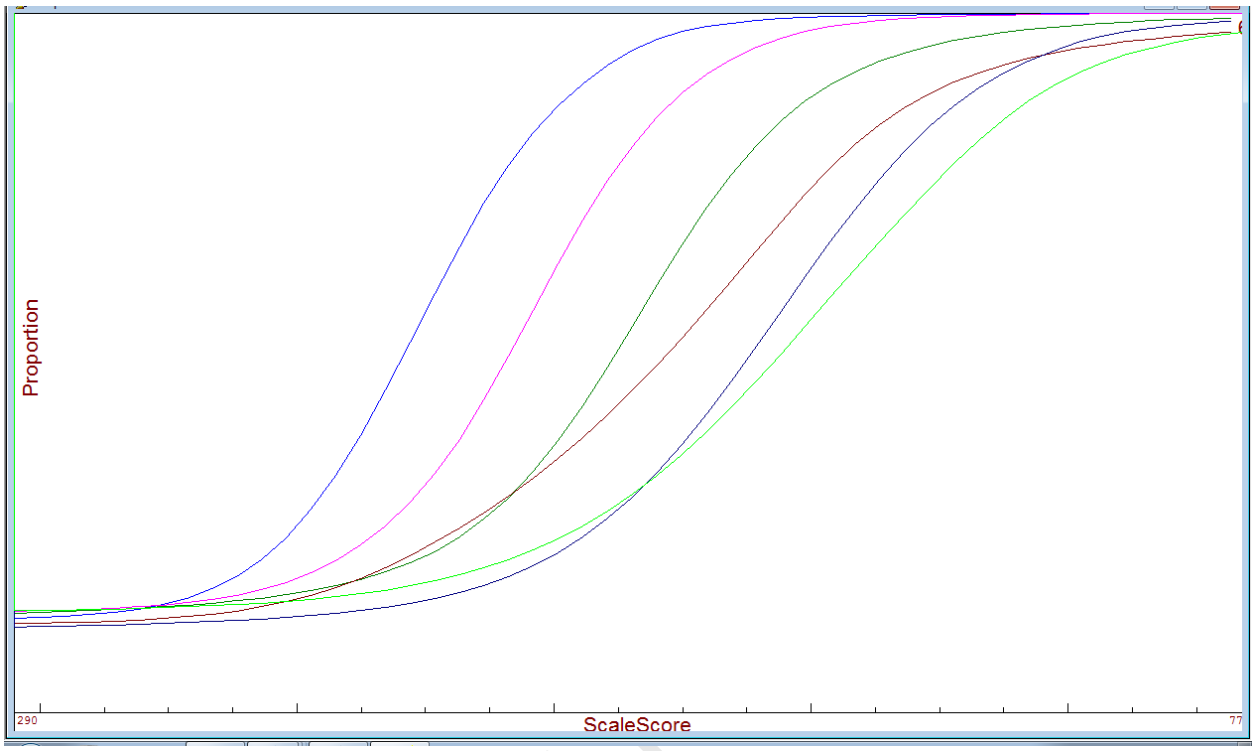
Figure A6. SE Curves: Mathematics scaling below-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

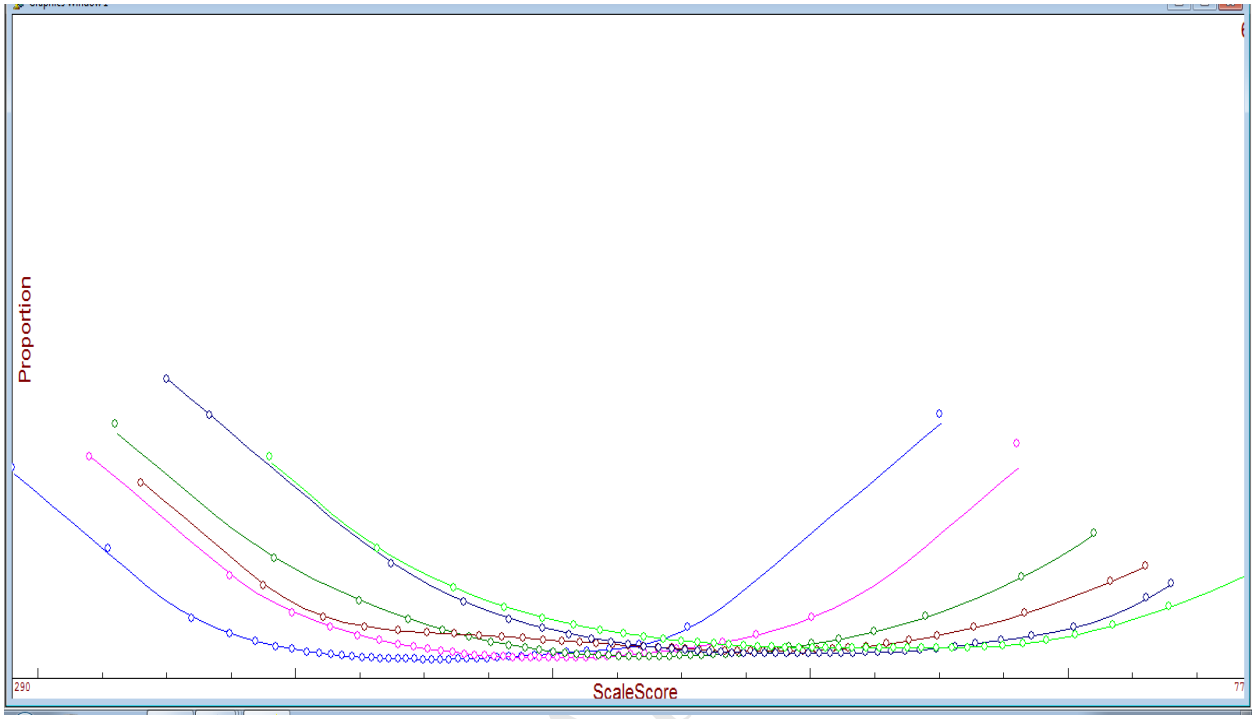
Figure A7. TCCs: Mathematics scaling with above-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.

Figure A8. SE Curves: Mathematics scaling above-grade level linking items



DRAFT

Copyright © 2016 by Missouri Department of Elementary and Secondary Education.