



# 2024 RICAS Technical Report

---

Prepared by Cognia and the  
Rhode Island Department of Education

**3/13/2025**



**RIDE** Rhode Island  
Department  
of Education

<https://www.ride.ri.gov/>

# Table of Contents

<b>CHAPTER 1. INTRODUCTION TO THE RHODE ISLAND COMPREHENSIVE ASSESSMENT SYSTEM .....</b>	<b>7</b>
1.1 PURPOSES OF THE RICAS AND THIS REPORT.....	7
1.2 ORGANIZATION OF THIS REPORT .....	8
1.2.1 MCAS AND RICAS COMPARISON .....	8
1.3 UPDATES FOR THE 2024 ADMINISTRATION .....	10
1.4 INTENDED INTERPRETATIONS AND USES OF THE RICAS TEST SCORES.....	10
1.4.1 INTENDED INTERPRETATIONS OF THE RICAS TEST SCORES .....	11
1.4.2 INTENDED USES OF THE RICAS TEST SCORES .....	11
1.4.3 VALIDATION ARGUMENTS FOR RICAS .....	11
<b>CHAPTER 2. TEST DESIGN AND DEVELOPMENT .....</b>	<b>14</b>
2.1 APPROPRIATENESS OF USING MASSACHUSETTS STANDARDS.....	14
2.2 CONTENT STANDARDS.....	14
2.3 PERFORMANCE STANDARDS .....	14
2.4 ELA .....	15
2.4.1 ELA PASSAGE SELECTION .....	15
2.4.2 ELA ITEM DESIGN .....	15
2.4.3 ELA FORM DESIGN .....	17
2.4.4 ELA REFERENCE MATERIALS .....	18
2.5 MATHEMATICS .....	18
2.5.1 MATHEMATICS ITEM DESIGN.....	18
2.5.2 MATHEMATICS FORM DESIGN .....	20
2.5.3 MATHEMATICS REFERENCE MATERIALS.....	21
2.6 ITEM AND TEST DEVELOPMENT PROCESS .....	22
2.6.1 ITEM REVIEW AND REFINEMENT .....	24
2.6.2 OPERATIONAL FIELD-TESTING OF ITEMS .....	25
2.6.3 FINAL FORM CONSTRUCTION.....	27
2.6.4 SPECIAL EDITION TEST FORMS.....	28
2.7 QUALITY CHECK .....	29
2.8 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING TEST DESIGN AND DEVELOPMENT .....	29
<b>CHAPTER 3. TEST ADMINISTRATION.....</b>	<b>32</b>
3.1 TEST ADMINISTRATION SCHEDULE .....	32
3.2 SECURITY REQUIREMENTS .....	33
3.3 PARTICIPATION REQUIREMENTS.....	33
3.4 SPANISH EDITION TEST FORMS.....	34
3.5 ADMINISTRATION PROCEDURES .....	34
3.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING TEST ADMINISTRATION .....	35
<b>CHAPTER 4. SCORING .....</b>	<b>36</b>
4.1 PREPARATION OF STUDENT RESPONSE BOOKLETS .....	36
4.2 PREPARATION FOR SCORING CONSTRUCTED-RESPONSE ITEMS.....	37
4.3 BENCHMARKING MEETINGS .....	37

4.4 MACHINE-SCORED ITEMS .....	37
4.5 HAND-SCORED ITEMS .....	38
4.5.1 SCORING PLAN AND STAFF.....	38
4.5.2 SCORER RECRUITMENT AND QUALIFICATIONS .....	39
4.5.3 SCORER TRAINING .....	40
4.5.4 LEADERSHIP TRAINING.....	41
4.5.5 METHODOLOGY FOR SCORING HAND-SCORED POLYTOMOUS ITEMS.....	41
4.5.6 MONITORING OF SCORING QUALITY .....	41
4.5.7 SCORING WITH THE INTELLIGENT ESSAY ASSESSOR (IEA).....	43
4.5.8 MONITORING OF SCORING QUALITY .....	45
4.5.9 INTERRATER CONSISTENCY .....	46
4.6 QUALITY CHECK .....	48
4.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING SCORING .....	49
<b>CHAPTER 5. REPORTING .....</b>	<b>50</b>
5.1 REPORTING OF RESULTS .....	50
5.2 PARENT/GUARDIAN REPORT.....	50
5.3 REPORTING BUSINESS REQUIREMENTS .....	51
5.4 QUALITY ASSURANCE .....	51
5.5 ADDITIONAL RESOURCES.....	52
5.5.1 STUDENTS AND THEIR FAMILIES.....	52
5.5.2 EDUCATORS AND ADMINISTRATORS.....	52
5.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING REPORTING .....	53
<b>CHAPTER 6. CLASSICAL ITEM ANALYSIS.....</b>	<b>56</b>
6.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES .....	56
6.2 DIFFERENTIAL ITEM FUNCTIONING .....	58
6.3 DIMENSIONALITY ANALYSIS .....	59
6.4 QUALITY CHECK .....	61
6.5 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING CLASSICAL ITEM ANALYSES.....	62
<b>CHAPTER 7. ITEM RESPONSE THEORY ANALYSIS .....</b>	<b>62</b>
7.1 OVERVIEW.....	63
7.2 IRT .....	63
7.3 IRT RESULTS.....	66
7.4 EQUATING .....	67
7.5 REPORTED SCALE SCORES AND ACHIEVEMENT STANDARDS .....	69
7.6 QUALITY CHECK .....	72
7.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING ITEM RESPONSE THEORY ANALYSES .....	72
<b>CHAPTER 8. RELIABILITY.....</b>	<b>74</b>
8.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT.....	74
8.2 SUBGROUP RELIABILITY .....	75
8.3 REPORTING SUBCATEGORY RELIABILITY .....	75
8.4 RELIABILITY OF ACHIEVEMENT LEVEL CATEGORIZATION .....	76
8.5 DECISION ACCURACY AND CONSISTENCY RESULTS.....	77

8.6 QUALITY CHECK .....	79
8.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING RELIABILITY .....	79
<b>CHAPTER 9. VALIDITY ARGUMENTS SUPPORTING INTENDED INTERPRETATIONS AND USES OF TEST SCORES .....</b>	<b>80</b>
9.1 RATIONALE FOR VALIDITY ARGUMENT-CENTERED TECHNICAL REPORTING.....	80
9.2 VALIDITY ARGUMENT FOR INTERPRETATION AND USE OF RICAS TEST SCORES .....	81
9.2.1 CLAIMS SUPPORTING INTENDED INTERPRETATIONS OF RICAS TEST SCORES .....	82
9.2.2 CLAIMS SUPPORTING INTENDED USES OF RICAS TEST SCORES .....	86
9.3 VALIDATION SUMMARY.....	89
<b>REFERENCES.....</b>	<b>90</b>
<b>APPENDICES.....</b>	<b>93</b>

APPENDIX A	ACCOMMODATIONS
APPENDIX B	PARTICIPATION RATES
APPENDIX C	INTERRATER CONSISTENCY
APPENDIX D	ACHIEVEMENT LEVEL DISTRIBUTIONS
APPENDIX E	SAMPLE REPORTS
APPENDIX F	REPORTING BUSINESS REQUIREMENTS
APPENDIX G	ITEM-LEVEL CLASSICAL STATISTICS
APPENDIX H	SCORE DISTRIBUTIONS
APPENDIX I	DIFFERENTIAL ITEM FUNCTIONING RESULTS
APPENDIX J	2023-24 MCAS EQUATING REPORT
APPENDIX K	RELIABILITY

## List of Tables

Table 1-1 Relationship between 2024 RICAS and MCAS Tests on Critical Test Components	9
Figure 1-1 Logic of Validity Arguments for Tests	12
Figure 1-2 Chapelle (2020)'s Framework: The Arguments and the Inferential Steps	13
Table 2-1 Target (and Actual) Distribution of ELA Common Item Points by Reporting Category	16
Table 2-2 ELA Item Types and Score Points	16
Table 2-3 Distribution of ELA Common and Matrix Items by Grade and Item Type—Computer-based Test (CBT)	18
Table 2-4 Distribution of ELA Common and Matrix Items by Grade and Item Type—Paper-based Test (PBT) <sup>1</sup>	18
Table 2-5 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grades 3–5	19
Table 2-6 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grades 6 and 7	19
Table 2-7 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grade 8	19
Table 2-8 Mathematics Item Types and Score Points	20
Table 2-9 Distribution of Mathematics Common and Matrix Items by Grade and Item Type—Computer-based Test (CBT)	21
Table 2-10 Distribution of Mathematics Common and Matrix Items by Grade and Item Type—Paper-based Test (PBT)	21
Table 2-11 Overview of Item and Test Development Process	23
Table 3-1 Grades 3–8 ELA and Mathematics Test Administration Schedule	32
Table 3-2 ELA Recommended Testing Times, Grades 3–8	32
Table 3-3 Mathematics Recommended Testing Times, Grades 3–8	33
Table 4-1 Summary of Scorer and Scoring Leadership Backgrounds (Operational Scoring)	40
Table 4-2 N Counts by Prompt	43
Table 4-3 Metrics for Evaluating Automated Scoring	44
Table 4-4 Comparison of Human and IEA Agreement with Validity Papers—ELA	45
Table 4-5 Summary of Interrater Consistency Statistics Organized across Items by Content Area and Grade—ELA	47
Table 4-6 Summary of Interrater Consistency Statistics Organized across Items by Content Area and Grade—Mathematics	47
Table 4-7 Summary of Validity Statistics <sup>1</sup>	48
Table 6-1 Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade	57
Table 6-2 Multidimensionality Effect Sizes by Grade and Content Area	61
Table 7-1 Number of Cycles Required for Convergence	67
Table 7-2 Stocking and Lord Constants	69
Table 7-3 Scale Score Slopes and Intercepts by Content Area and Grade	71
Table 7-4 Cut Scores on the Theta Metric and Reporting Scale by Content Area and Grade	71
Table 8-1 Raw Score Descriptive Statistics, Cronbach's Alpha, and SEMs by Content Area and Grade	75
Table 8-2 Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Achievement Level	78
Table 8-3 Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint	78
Figure 9-1 Validity Argument Logic Model	81
Figure 9-2 Chapelle (2020)'s Framework: The Arguments and the Inferential Steps	82

# Chapter 1. Introduction to the Rhode Island Comprehensive Assessment System

Chapter 1 provides an overview of the purpose and organization of this report, including the comparison between the Massachusetts Comprehensive Assessment System (MCAS) and the Rhode Island Comprehensive Assessment System (RICAS). Also provided are updates for the 2024 administrations, the intended interpretations and uses of the RICAS test scores, and the framework for constructing the validity arguments in this report.

## 1.1 PURPOSES OF THE RICAS AND THIS REPORT

The RICAS is Rhode Island’s state assessment program in English language arts (ELA) and mathematics at grades 3–8 and is designed to meet the federal requirements of the Every Student Succeeds Act (ESSA). In addition to fulfilling ESSA assessment requirements, the specific purposes of the RICAS tests are

- to provide information to parents/guardians and students on Rhode Island student achievement on the state’s ELA and mathematics content standards,
- to provide information to support program evaluation and improvement at the school and district level, and
- to provide academic achievement and growth information used as part of the state’s school accountability program to inform parents/guardians and the public about the performance of Rhode Island schools.

Beginning in the 2017–2018 school year, the Rhode Island Department of Education (RIDE) adopted the MCAS ELA and mathematics tests as its state assessments in ELA and mathematics at grades 3–8. The tests are administered in Rhode Island under a licensing agreement with Massachusetts DESE and labeled RICAS for their use in Rhode Island. The use of the MCAS tests at grades 3–8 is part of Rhode Island’s transition from the use of the Partnership for the Assessment of Readiness for College and Careers (PARCC) tests at grades 3–8 and high school as its state assessments. In high school, the PARCC tests have been replaced by the SAT.

The adoption of the MCAS tests reflects a continuation of Rhode Island’s policy to partner with other states to offer a high-quality state assessment. With the increased assessment requirements of the No Child Left Behind Act in 2001, RIDE determined that it would not be feasible to develop and sustain a high-quality assessment program on its own. From 2003–2014, Rhode Island partnered with New Hampshire, Vermont, and Maine in the New England Common Assessment Program (NECAP). With the adoption of the Common Core State Standards (CCSS) and the creation of national assessment consortia, Rhode Island joined PARCC, administering the PARCC tests from 2015–2017.

As Massachusetts and other states left the PARCC consortium, it was no longer clear that PARCC would be able to offer long-term stability in assessment to support the state’s improvement efforts. MCAS, in

contrast, has been regarded as a model for high-quality and stable state assessment since its inception in 1998. In 2017, Massachusetts developed MCAS tests to fully align with college- and career-ready content standards and established rigorous performance standards consistent with those established by PARCC. With the updated tests and performance standards in place, Rhode Island began administration of the Massachusetts tests in spring 2018.

The main purpose of this 2024 RICAS Technical Report is to document the technical quality and essential design characteristics of the 2024 RICAS ELA and mathematics tests in grades 3–8, to present evidence of the validity, reliability, and fairness of the use of the tests as part of the Rhode Island state assessment program.

Because the RICAS tests administered in Rhode Island are the MCAS ELA and mathematics tests, much of the information related to their technical quality is provided by the MCAS Technical Reports produced by the Massachusetts DESE. That information has been reproduced in this report for the purpose of clarity; consequently, DESE, Massachusetts, and MCAS are all referenced in this report. Additionally, MCAS Technical Reports are available directly on the DESE website: [doe.mass.edu/mcas/tech/?section=techreports](https://doe.mass.edu/mcas/tech/?section=techreports).

This report contains information specific to the administration of the tests in Rhode Island intended to augment the information reproduced from the MCAS Technical Report, to document any differences in the assessment policies and procedures between Rhode Island and Massachusetts, and to provide additional background information about the RICAS program.

The information contained in this report, prepared by Cognia for RIDE, in conjunction with information provided by Massachusetts, demonstrates that MCAS grades 3–8 ELA and mathematics tests are technically sound, function well for students in Rhode Island, and are appropriate instruments to assess the performance of Rhode Island students on the state’s content standards.

This report is primarily intended for users with a working understanding of psychometrics and educational measurement. It assumes knowledge of measurement concepts such as reliability and validity as well as statistical concepts of correlation and central tendency. For some sections, the reader is presumed to have basic familiarity with advanced topics in measurement and applied statistics such as item response theory (IRT).

## 1.2 ORGANIZATION OF THIS REPORT

This report provides information regarding the spring 2024 administration of the RICAS tests in ELA and mathematics, including a description and results of analyses conducted to provide evidence of the technical quality and design characteristics of those tests.

### 1.2.1 MCAS and RICAS Comparison

The RICAS tests were administered, scored, and processed by Cognia, the state’s assessment contractor for the RICAS tests. Cognia is also the Massachusetts assessment contractor for the MCAS tests. Unless noted in this report, all processes and procedures used in administering, processing, scoring, and reporting of the results of the spring 2024 RICAS tests were identical to the corresponding procedures used for the MCAS tests. Table 1-1 provides a summary of the relationship between key aspects of the RICAS and MCAS testing programs.



**Table 1-1 Relationship between 2024 RICAS and MCAS Tests on Critical Test Components**

Test Component	RICAS and MCAS
Test Content	Identical
Test Design	Identical
Test Administration	Identical
Mode of Administration	RI offers Spanish language forms in mathematics.
Administration Platform	Identical
<b>Scoring</b>	
Machine-scored items	Identical
Hand-scored items	Identical
Psychometric Procedures	Identical
<b>Reporting</b>	
Scaled scores	Identical
Achievement levels	Identical

Cognia conducted all the analyses described in this report. The analyses described and presented here are consistent with the types of analyses conducted for the MCAS tests.

***All analyses are based only on Rhode Island students, unless otherwise specified.***

The specific analyses of Rhode Island students included in this report were identified by the Rhode Island Technical Advisory Committee (RI-TAC) as necessary and useful to provide evidence of the validity, reliability, and fairness of the use of the MCAS tests as the Rhode Island state assessments in ELA and mathematics in grades 3–8.

This information includes the following:

- Chapter 2: Test Design and Development – information related to the MCAS design and development of the tests used for RICAS.
- Chapter 3: Test Administration – information related to test administration policies and procedures, including protocols to monitor test security.
- Chapter 4: Scoring – information on machine-scored items and hand-scoring procedures for short-answer, constructed-response, and essay items, including information on the level of interrater agreement among raters.
- Chapter 5: Reporting – detailed information on the type of student-level test scores reported to parents/guardians and a description of the quality assurance procedures used to ensure the accuracy of the reporting of those results.
- Chapter 6: Classical Item Analysis – a description of and summary results from the classical item analyses conducted with Rhode Island students as part of the full analyses provided to demonstrate technical quality of the test. Analyses include classical item statistics, Differential Item Functioning, and dimensionality.
- Chapter 7: Item Response Theory Analysis – a description and results from the IRT analyses conducted with Massachusetts students as part of the full analyses provided to demonstrate technical quality of the test. Results of calibration, scaling, equating, and setting of performance standards are provided.
- Chapter 8: Reliability – a description of and summary results from the reliability analyses conducted with Rhode Island students as part of the full analyses provided to demonstrate the technical quality of the test. Results of reliability, subgroup reliability, and decision consistency and accuracy are provided.

- Chapter 9: Validity – information related to validity evidence supporting the intended uses and interpretations of RICAS test scores.

Additionally, a set of appendices is provided, containing the following information:

- Appendix A – Accommodations
- Appendix B – Participation Rates
- Appendix C – Interrater Consistency
- Appendix D – Achievement Level Distributions
- Appendix E – Sample Reports
- Appendix F – Reporting Business Requirements
- Appendix G – Item-Level Classical Statistics
- Appendix H – Score Distributions
- Appendix I – Differential Item Functioning Results
- Appendix J – 2023–24 MCAS Equating Report
- Appendix K – Reliability

## 1.3 UPDATES FOR THE 2024 ADMINISTRATION

The Pearson Intelligent Essay Assessor (IEA) was used to provide a 10% double-blind score for ELA essays beginning in 2020 and continuing through subsequent years. In 2024, the use of IEA was expanded to score 100% of the ELA essay items, with a 10% human double-blind. This applied to the scoring of all RICAS ELA essays for grades 3-8, as detailed in section 4.5.7.

## 1.4 INTENDED INTERPRETATIONS AND USES OF THE RICAS TEST SCORES

The purposes for administering RICAS include measuring student proficiency relative to standards. Because these standards did not change across administrations, individual student scores can be interpreted in a similar way to previous administrations. Another stated purpose of RICAS is the use of assessment results for state and federal accountability and reporting. Related to the recovery from the COVID-19 pandemic, instruction and assessment trended toward a return to standard in-person practices, although the instructional impact of COVID-19 continues to be monitored.

The RICAS is designed, developed, and implemented to elicit student performances whose qualities are then evaluated and quantified as items and test scores supporting a predefined set of intended interpretations. The resulting test score interpretations are, in turn, applied to inform a predefined set of intended uses. These intended interpretations and uses of test scores and a structure for their validation are described in the sections that follow.

## 1.4.1 Intended Interpretations of the RICAS Test Scores

For grades 3 through 8, in ELA and mathematics, RICAS scores provide reliable and valid information about student knowledge and ability as defined by the content standards for the grade and content area being assessed.

## 1.4.2 Intended Uses of the RICAS Test Scores

Interpretations of RICAS test scores are intended for the following uses:

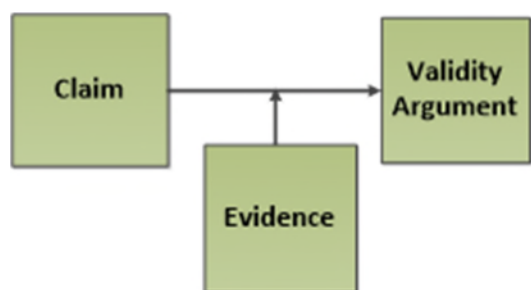
- Parents/guardians and students can use test scores and their underlying interpretations to monitor academic achievement and participate in decisions regarding student learning to support student growth.
- Educators can use test scores and their underlying interpretations to support curricular planning and identify instructional needs at both the classroom and individual student level.
- School- and district-level administrators can use test scores and their underlying interpretations to support program evaluation and improvements at the school and district levels.
- State-level administrators can use test scores and their underlying interpretations to monitor academic achievement and growth as required by state accountability programs and inform parents/guardians and the public of schools' performances on these metrics.
- Federal administrators can use test scores and their underlying interpretations to verify that ESSA federal accountability requirements are met.

## 1.4.3 Validation Arguments for RICAS

This technical report describes all essential components of the design, implementation, scoring, psychometric analyses, and reporting procedures of the RICAS program. These processes contribute to the accumulation of validity evidence supporting the intended interpretations and uses of RICAS test scores. Because the interpretation and uses of test scores, rather than the tests themselves, are evaluated for validity, this report presents documentation to substantiate these intended interpretations and uses of test scores (AERA, APA, & NCME, 2014, p. 11).

Each chapter in this report contributes important information about the RICAS program: test design and development, standards alignment, test administration, scoring, classical item analyses, IRT linking and scaling, and reporting. The information to support validity arguments for intended interpretations and uses of RICAS test scores, summarized in the last section of each chapter then compiled and fully summarized in Chapter 9, is presented as claims: elements that underlie the interpretations and uses articulated within the validity argument. Strength of the validity argument is established by providing evidence supporting each of these claims. The logic of the validity argument structure is shown in Figure 1-1.

**Figure 1-1 Logic of Validity Arguments for Tests**



The phrase “intended score interpretations for uses” appears several times in the Standards for Educational and Psychological Testing (“Standards” for short in the following chapters/sections; AERA et al., 2014) and is the core of the field’s views on validity and validation.

For RICAS (and assessment programs more generally), the phrase refers broadly to information related to test performance (e.g., total scores/scale scores, aggregations of total/scale scores, the percentage of students at or above a given level) supported by supplementary information (e.g., achievement level descriptors for achievement level classifications, item design information for marker items on the scale).

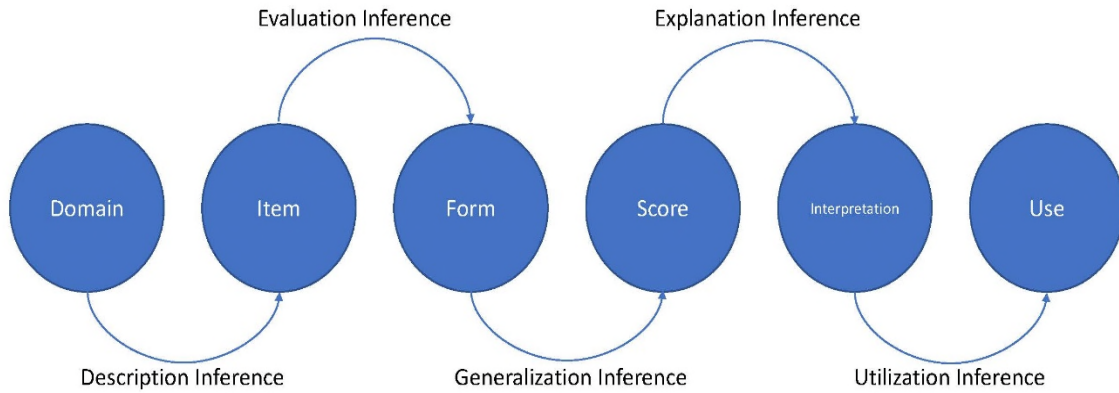
The Standards also provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. These sources address different aspects of supporting evidence for validity arguments but are not considered distinct types of validity. Instead, each contributes to a body of evidence about the individual validity arguments and overall arguments for the validity of intended score interpretations and uses. Moreover, these sources represent only a partial list of potential sources of evidence that informed RICAS design, development, test administration, analysis, and reporting processes that are relevant to the overall validity arguments for intended interpretations and uses of RICAS test scores and related information. Hence, this document will use Chapelle’s (2020) framework based on Kane’s work.

Validity arguments are crafted to not just provide evidence that all steps in the test design, development, and implementation process are taken correctly, but that they are working together to ensure that the resulting scores validly support intended interpretations and uses. The arguments and the logical inferential steps they provide are structured based upon the framework developed by Chapelle (2020) and can be summarized as follows (also seen in Figure 1-2):

- 1) Description Inference: Items sample from the target domain appropriately such that high quality forms can be produced. (Domain to Item)
- 2) Evaluation Inference: Forms sample from items appropriately such that observed scores reflective of the domain can be produced. (Item to Form)
- 3) Generalization Inference: Observed scores on individual forms are reliable such that they are reflective of expected scores across forms. (Form to Score)
- 4) Explanation Inference: Expected scores are associated with classification cuts such that classification decisions are interpretable. (Score to Interpretation)
- 5) Utilization Inferences: Interpretations of scores and classifications are used as intended and only in ways considered appropriate and fair. (Interpretation to Use)

Evidence for these inferences and the claims that comprise them is provided at the end of each chapter. It identifies the specific inference and claims and describes the relevant evidence. This evidence is then gathered and organized according to the structure of inferences presented above.

**Figure 1-2 Chapelle (2020)'s Framework: The Arguments and the Inferential Steps**



# Chapter 2. Test Design and Development

There were no changes in test design or development for the 2024 administration of the RICAS program. The adherence to previous years' blueprints allows for defensible comparisons of where students are relative to grade-level expectations as outlined in the grades 3–8 ELA and mathematics standards despite COVID-related learning disruptions. Chapter 2 is primarily drawn from MCAS technical reporting and relates to the RICAS use of those assessments.

## 2.1 APPROPRIATENESS OF USING MASSACHUSETTS STANDARDS

Before adopting the MCAS tests as its state assessment, it was necessary to determine the appropriateness of the Massachusetts content and performance standards for use in Rhode Island.

To meet the requirements of the Every Student Succeeds Act (ESSA) and provide valid and useful information to Rhode Island parents/guardians, students, and schools, the state assessments must be aligned to the state's content standards.

In addition, to support the state's commitment to ensure that Rhode Island's educational system holds high expectations for all students and that Rhode Island graduates are well prepared for postsecondary education, work, and life, the state must establish rigorous performance standards that signal whether students are on track for success in high school and college and career readiness as they progress through elementary and middle school.

The following sections describe the steps taken by Rhode Island to make the appropriate determinations for content and performance standards followed by descriptions of the test designs for ELA and mathematics.

## 2.2 CONTENT STANDARDS

In 2010, Rhode Island adopted the Common Core State Standards (CCSS) as its state content standards in ELA and mathematics. In July 2010, the Massachusetts Board of Elementary and Secondary Education also adopted the CCSS in ELA and mathematics as the core of its PK–12 content standards.

In March 2011, Massachusetts adopted revised Curriculum Frameworks in ELA and mathematics, which are the state's academic content standards. As described at the time by Mitchell Chester, Massachusetts Commissioner of Elementary and Secondary Education, the 2011 Curriculum Framework "merges the Common Core State Standards for Mathematics with additional Massachusetts standards and other features." Rhode Island transitioned to the Rhode Island Core Standards from the CCSS on March 9, 2021. The Rhode Island Core Standards mirror the Massachusetts Curriculum Frameworks.

## 2.3 PERFORMANCE STANDARDS

In addition to the alignment of the tests to Rhode Island's academic content standards, and for the MCAS tests to be appropriate for Rhode Island, it was essential that the performance standards established for

those tests were consistent with the rigorous performance standards that Rhode Island adopted when it began administering the PARCC tests in 2015. More technical details on the determination of the cut point are provided in section 7.5 in Chapter 7.

## 2.4 ELA

### 2.4.1 ELA Passage Selection

Passages used in the ELA tests are authentic published passages that possess the characteristics required for use in ELA tests; no passages were specifically written for the RICAS tests. They are identified and reviewed by test developers, including DESE test developers. Passages must

- be of interest to and appropriate for students in the grade being addressed;
- have a clear beginning, middle, and end;
- contain appropriate content;
- support the development of a sufficient number of unique assessment items; and
- be free of bias and sensitivity issues.

Passages ranged in length from approximately 600 to 2500 words per passage set. Word counts were slightly reduced at lower grades. Passage sets consisted of either a single passage or paired/tripled passages. Passages are categorized into one of two types:

- **Literary passages**—Literary passages represent a variety of genres: poetry, drama, fiction, biographies, memoirs, folktales, fairy tales, myths, legends, narratives, diaries, journal entries, speeches, and essays. Literary passages are not necessarily fictional passages.
- **Informational passages**—Informational passages are reference materials, editorials, encyclopedia articles, and general nonfiction. Informational passages are drawn from a variety of sources, including magazines, newspapers, and books.

In grades 3–8, there is one common form per grade. Each common form includes three passage sets, with forms in some grades containing two literary passage sets and one informational passage set. Forms in other grades contain one literary passage set and two informational passage sets. Across the forms, sets may be single, paired, or tripled selections.

The RICAS ELA test is designed to include a selection of passage sets with a balanced representation of male and female characters; races and ethnicities; and urban, suburban, and rural settings. Another important consideration is that passages are of interest to the age group being tested.

The main difference among the passages used for grades 3–8 is their degree of complexity, which results from increasing levels of sophistication in language and concepts, as well as passage length. Test developers use a variety of readability formulas to aid in the selection of passages appropriate at each grade level. In addition, subject-matter experts use their grade-level expertise when participating in passage selection as members of the Assessment Development Committees (ADCs).

### 2.4.2 ELA Item Design

The 2024 RICAS grades 3–8 ELA tests, including all matrix items, measured the following learning standards as articulated within the Rhode Island Core Standards.

- Anchor Standards for Reading

- Key Ideas and Details (Standards 1–3)
- Craft and Structure (Standards 4–6)
- Integration of Knowledge and Ideas (Standards 7–9)
- Anchor Standards for Language
  - Conventions of Standard English (Standards 1 and 2)
  - Knowledge of Language (Standard 3)
  - Vocabulary Acquisition and Use (Standards 4–6)
- Anchor Standards for Writing
  - Text Types and Purposes (Standards 1–3)
  - Production and Distribution of Writing (Standards 4–6)

Table 2-1 shows the target and actual percentages of common item points by reporting category. Reporting categories are based on the Rhode Island Core Standards.

**Table 2-1 Target (and Actual) Distribution of ELA Common Item Points by Reporting Category**

Reporting Category	% of Points at Each Grade (+/-5%)					
	3	4	5	6	7	8
Language	25 (27)	25 (27)	25(29)	25 (20)	25 (22)	25 (22)
Reading	65 (64)	65 (64)	55 (54)	55 (60)	55 (58)	55 (58)
Writing	10 (09)	10 (09)	20(17)	20 (20)	20 (20)	20 (20)
Total	100	100	100	100	100	100

The grades 3–8 ELA tests used several item types, as shown in Table 2-2.

**Table 2-2 ELA Item Types and Score Points**

Item Type	Possible Raw Score Points	Grade Levels
Multiple-choice (SR)	0 or 1	3–8
Two-part, multiple-choice (SR)	0, 1, or 2	3–8
Technology-enhanced (SR)	0, 1, or 2	3–8
Constructed-response (CR)	0, 1, 2, or 3	3–4
Essay (ES)	0 to 7	3–5
	0 to 8	6–8

*SR = selected-response, CR = constructed-response, ES = essay*

Each item on the ELA tests is assigned a cognitive level based on Norman Webb’s Depth of Knowledge (DoK) Framework. Cognitive levels are not synonymous with item difficulty. The cognitive level provides information about each item based on the complexity of the mental processing a student must use to answer the item correctly. Levels are assigned by developers and reviewed by an assessment development committee. The three cognitive levels used in ELA tests are described below.

- Level I (Identify/Recall)—Level I items require that the student recognize basic information presented in the text. Examples of skills at this level include identifying main ideas/facts/details; recalling and locating details; identifying genre or setting; and identifying definitions, parts of speech, or functions of punctuation. Key words include identify, list, match, recognize, describe, and distinguish.
- Level II (Infer/Analyze)—Level II items require that the student understand a given text by making inferences and drawing conclusions related to the text. Examples of skills at this level include understanding the whole text (Big Picture)/generalizing; interpreting, making connections, visualizing, and forming questions; explaining a character’s role/motives; determining whether an



idea is fact or opinion; filtering important information and key concepts; and determining the meaning of a word in context. Key words include infer, analyze, describe, interpret, determine, conclude, explain, summarize, and classify.

- Level III (Evaluate/Apply)—Level III items require that the student understand multiple points of view and be able to project his or her own judgments or perspectives on the text. Examples of skills at this level include understanding another point of view; analyzing/evaluating an author's purpose, style, and message; arguing/defending a point of view with evidence from the text; using reasoning to determine an outcome; applying information from the text; and synthesizing elements of text(s) to create a whole. Key words include critique, evaluate, analyze, predict, agree/disagree, argue/defend, apply, synthesize, judge, compare, and contrast.

Each cognitive level is represented in the ELA tests.

### 2.4.3 ELA Form Design

All items are coded to ELA framework standards. There are no stand-alone items on the tests; all vocabulary, grammar, and mechanics questions are associated with a passage set; more details on the number of items at different grades is provided below.

Students read a passage set and answer questions that follow. Question types include selected-response items, constructed-response items (grades 3 and 4 only), and essay items. Approximately 20% of the items were technology-enhanced items such as inline choice, hot spots, and drag and drop that require the student to choose from a range of options presented.

#### Common Portion

##### Grades 3–4

The common portion of each test at grades 3 and 4 included three passage sets, and the matrix portion included two passage sets. One of the common passage-sets included ten or twelve 1- or 2-point selected-response items plus one 7-point text-based essay item; one of them included eleven or twelve 1- or 2-point selected-response items and one 3-point constructed-response item, and one of them included seven or eight 1- or 2-point selected-response items.

Each test contained a total of 44 common points distributed across two testing sessions.

##### Grade 5

The common portion of each test at grade 5 included three passage sets, and the matrix portion included two passage sets. Passage sets included eleven 1- or 2-point selected-response items and one 7-point text-based essay item or seven 1- or 2-point selected-response items.

The test contained a total of 48 common points distributed across two testing sessions.

##### Grades 6–8

The common portion of each test at grades 6–8 included three passage sets, and the matrix portion included two passage sets. Passage sets included eleven or twelve 1- or 2-point selected-response items and one 8-point text-based essay item or seven or eight 1-point items.

Each test contained a total of 50 common points distributed across two testing sessions.

#### Matrix Portion

For grades 3–8, the matrix portion included two passage sets. In grades 3–4, the matrix passage set included eight to eleven 1- or 2-point selected-response items, and either two constructed-response items

or one essay. The other matrix passage set included seven 1- or 2-point machine-scored items. In grades 5–8, the matrix passage set included eight or nine 1- or 2-point selected-response items, and one essay item.

Tables 2-3 (for the computer-based forms) and 2-4 (for the paper-based forms) list the distribution of common and matrix items in each 2024 ELA test, by grade.

**Table 2-3 Distribution of ELA Common and Matrix Items by Grade and Item Type—Computer-based Test (CBT)**

Grade	# of Forms	Items per Form							
		Common				Matrix			
		SR (1 pt.)	SR (2 pt.)	CR	ES	SR (1 pt.)	SR (2 pt.)	CR	ES
3	1	24	5	1	1	14	3	0-2	0-1
4	1	24	5	1	1	14	3	0-2	0-1
5	1	24	5	0	2	14	3	0	1
6	1	24	5	0	2	14	3	0	1
7	1	24	5	0	2	14	3	0	1
8	1	24	5	0	2	14	3	0	1

**Table 2-4 Distribution of ELA Common and Matrix Items by Grade and Item Type—Paper-based Test (PBT)<sup>1</sup>**

Grade	# of Forms	Items per Form							
		Common				Matrix			
		SR (1 pt.)	SR (2 pt.)	CR	ES	SR (1 pt.)	SR (2 pt.)	CR	ES
3	1	24	5	1	1	14	3	0-2	0-1
4	1	24	5	1	1	14	3	0-2	0-1
5	1	24	5	0	2	14	3	0	1
6	1	24	5	0	2	14	3	0	1
7	1	24	5	0	2	14	3	0	1
8	1	24	5	0	2	14	3	0	1

<sup>1</sup> The paper form is derived from Form 1 of the CBT.

## 2.4.4 ELA Reference Materials

The use of bilingual word-to-word dictionaries was allowed during ELA tests only for current and former English language learners (ELLs). No other reference materials were allowed during the ELA tests.

## 2.5 MATHEMATICS

### 2.5.1 Mathematics Item Design

The 2024 RICAS grades 3–8 mathematics tests, including all field-test items, measured the learning standards as articulated within the Rhode Island Core Standards.

- Domains for grades 3–5
  - Operations and Algebraic Thinking
  - Number and Operations in Base Ten
  - Number and Operations—Fractions
  - Geometry

- Measurement and Data
- Domains for grades 6 and 7
  - Ratios and Proportional Relationships
  - The Number System
  - Expressions and Equations
  - Geometry
  - Statistics and Probability
- Domains for grade 8
  - The Number System
  - Expressions and Equations
  - Functions
  - Geometry
  - Statistics and Probability

Tables 2-5 through 2-7 show the target and actual percentages of common item points by reporting category. Reporting categories are based on the Rhode Island Core Standards.

**Table 2-5 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grades 3–5**

Domain	% of Points at Each Grade (+/-5%)		
	3	4	5
Operations and Algebraic Thinking	30 (29)	20 (19)	15 (15)
Number and Operations in Base Ten	15 (15)	20 (20)	30 (30)
Number and Operations – Fractions	20 (21)	30 (30)	25 (24)
Geometry	10 (10)	10 (11)	10 (13)
Measurement and Data	25 (25)	20 (20)	20 (19)
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>

**Table 2-6 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grades 6 and 7**

Domain	% of Points at Each Grade (+/-5%)	
	6	7
Ratios and Proportional Relationships	20 (20)	20 (20)
The Number System	20 (20)	20 (20)
Expressions and Equations	30 (30)	25 (24)
Geometry	15 (15)	15 (15)
Statistics and Probability	15 (15)	20 (20)
<b>Total</b>	<b>100</b>	<b>100</b>

**Table 2-7 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grade 8**

Domain	% of Points (+/-5%)
The Number System and Expressions and Equations	40 (39)
Functions	20 (20)
Geometry	30 (26)
Statistics and Probability	10 (15)
<b>Total</b>	<b>100</b>

The 2024 mathematics tests included several item types, as shown in Table 2-8.

**Table 2-8 Mathematics Item Types and Score Points**

Item Type	Possible Raw Score Points	Grade Levels
Multiple-choice (SR)	0 or 1	3–8
Multiple-select (SR)	0 or 1	3–8
Technology-enhanced (SA)/(SR)/(CR)	0 or 1 0, 1, or 2	3 4–8
Short-answer (SA)	0 or 1	3–8
Constructed-response (CR)	0, 1, 2, or 3 0, 1, 2, 3, or 4	3 4–8

Each item on the mathematics test is assigned a cognitive level based on Norman Webb’s Depth of Knowledge (DoK) Framework. Cognitive levels are not synonymous with difficulty. The cognitive level provides information about each item based on the complexity of the mental processing a student must use to answer the item correctly. The three cognitive levels used in the mathematics tests are described below.

- **Level I (Recall and Recognition)**—Level I items require that the student recall mathematical definitions, notations, simple concepts, and procedures, and apply common, routine procedures or algorithms (that may involve multiple steps) to solve a well-defined problem.
- **Level II (Analysis and Interpretation)**—Level II items require that the student engage in mathematical reasoning beyond simple recall in a more flexible thought process, and in enhanced organization of thinking skills. These items require a student to make a decision about the approach needed, to represent or model a situation, or to use one or more non-routine procedures to solve a well-defined problem.
- **Level III (Judgment and Synthesis)**—Level III items require that the student perform more abstract reasoning, planning, and evidence-gathering. To answer questions of this cognitive level, a student must engage in reasoning about an open-ended situation with multiple decision points, represent or model unfamiliar mathematical situations, and solve more complex, non-routine, or less well-defined problems.

Cognitive Levels I and II are represented by items in all grades and across item types. Cognitive Level III is best represented by constructed-response items; Cognitive Level III items were included at each grade, whenever possible.

## 2.5.2 Mathematics Form Design

### Form Design by Grade

#### Grade 3

The common portion of the grade 3 test included thirty-six 1-point selected-response or short-answer items and four 3-point constructed-response items.

The matrix portion included three 1-point selected-response or short-answer items and one 3-point constructed-response item.

The test contained a total of 48 common points distributed across two testing sessions.

## Grades 4–6

The common portion of the grades 4–6 tests included thirty-four 1-point selected-response or short-answer items, two 2-point selected-response items, and four 4-point constructed-response items.

The matrix portion included four to five 1-point selected-response or short-answer items, up to one 2-point selected-response or short-answer item, and one 4-point constructed-response item.

Each test contained a total of 54 common points distributed across two testing sessions.

## Grades 7–8

The common portion of the grades 7–8 tests included thirty-four 1-point selected-response or short-answer items, two 2-point selected-response items, and four 4-point constructed-response items.

The matrix portion included five 1-point selected-response or short-answer items, one 2-point selected-response or short-answer item, and two 4-point constructed-response items.

Each test contained a total of 54 common points distributed across two testing sessions. Items in session 2 were developed to assess content where the students may need a calculator. These items were either calculator-neutral (calculators are permitted but not required to answer the question) or calculator-active (students are expected to use a calculator to answer the question).

Tables 2-9 (for the computer-based forms) and 2-10 (for the paper form) show the distribution of common and matrix item types.

**Table 2-9 Distribution of Mathematics Common and Matrix Items by Grade and Item Type—Computer-based Test (CBT)**

Grade	#of Forms	Common				Matrix			
		SR/MS/SA/TE (1 pt.)	(2 pt.)	CR (3 pt.)	(4 pt.)	Totals # (pt.)	SR/MS/SA/TE (1 or 2 pt.)	CR (3 or 4 pt.)	Totals # (pt.)
3	1	36	0	4	0	40 (48)	3	1	4 (6)
4	1	34	2	0	4	40 (54)	5	1	4 (8–10)
5	1	34	2	0	4	40 (54)	5	1	4 (8–10)
6	1	34	2	0	4	40 (54)	5	1	4 (8–10)
7	1	34	2	0	4	40 (54)	6	2	6 (13–15)
8	1	34	2	0	4	40 (54)	6	2	6 (13–15)

**Table 2-10 Distribution of Mathematics Common and Matrix Items by Grade and Item Type—Paper-based Test (PBT)**

Grade	#of Forms	Common				Matrix			
		SR/MS/SA (1 pt.)	(2 pt.)	CR (3 pt.)	(4 pt.)	Totals # (pt.)	SR/MS/SA (1 or 2 pt.)	CR (3 or 4 pt.)	Totals # (pt.)
3	1	36	0	4	0	40 (48)	3	1	4 (6)
4	1	34	2	0	4	40 (54)	5	1	4 (8–10)
5	1	34	2	0	4	40 (54)	5	1	4 (8–10)
6	1	34	2	0	4	40 (54)	5	1	4 (8–10)
7	1	34	2	0	4	40 (54)	6	2	6 (13–15)
8	1	34	2	0	4	40 (54)	6	2	6 (13–15)

## 2.5.3 Mathematics Reference Materials

Rulers were provided to students in grades 3–8. Handheld rulers were provided to students taking the paper version of the mathematics test. Students taking the computer-based mathematics test had access to two separate computer-based rulers: a centimeter ruler and a 1/8-inch ruler; students were not permitted to use handheld rulers on the computer-based test.

Reference sheets were provided to students in grades 5–8. These sheets contain information, such as formulas, that students may need to answer certain items.

The second session of the grades 7–8 mathematics tests was a session where calculator use was permitted. All items included in this session were either calculator-neutral (calculators are permitted but not required to answer the question) or calculator-active (students are expected to use a calculator to answer the question). Each student taking the computer-based grade 7 mathematics test had access to a five-function calculator during session 2 of the mathematics test. Each student taking the computer-based grade 8 mathematics test had access to a scientific calculator during session 2. Students taking the paper-based mathematics tests in grades 7–8 had access to comparable handheld calculators.

## 2.6 ITEM AND TEST DEVELOPMENT PROCESS

Table 2-11 provides a detailed view of the item and test development process, in chronological order, along with parties involved in reviewing content in each phase.

**Table 2-11 Overview of Item and Test Development Process**

Phase	Development Step	Detail of the Process	Cognia Test Developers	DESE Test Developers	ADC	BSC	External Experts
Initial Item Design	Selection of reading passages (for ELA only)	Cognia's test developers find potential passages and present them to DESE for initial approval; DESE-approved passages go to Assessment Development Committees (ADCs) composed of experienced educators, and then to a Bias and Sensitivity Committee (BSC) for review and recommendations. ELA items are not developed until passages have been reviewed by an ADC and a BSC. With the ADC and BSC recommendations, DESE makes the final determination as to which passages will be developed and used on a future MCAS/RICAS test.	X	X	X	X	
	Development of initial item versions	Cognia's test developers generate items and edit items from subcontractors that are aligned to Massachusetts standards and specifications.	X				
Item Review and Refinement	Review of initial item versions by DESE and educators	<ol style="list-style-type: none"> <li>Cognia sends draft items to DESE test developers for review.</li> <li>DESE test developers review and edit items prior to presenting the items to ADCs.</li> <li>ADCs review items and make recommendations.</li> <li>BSC reviews items and makes recommendations.</li> </ol>	X	X	X	X	
	Revision of initial items based on review	DESE test developers edit & revise items based on recommendations from ADC & BSC.	X	X			
Item Review and Refinement	Independent expert review of revised items	Experts from higher education and practitioners review all field-tested items for content accuracy. Each item is reviewed by at least two independent expert reviewers. Comments and suggested edits are provided to DESE staff for review.	X	X			X
Operational Field Testing	Benchmark paper selection for CR and essay scoring	DESE and Cognia test developers meet to determine appropriate benchmark papers for training of scorers of field-tested constructed-response items and essays. Scoring rubrics and notes are reviewed and edited during benchmarking meetings. During the scoring of field-tested items, Cognia contacts DESE test developers with any unforeseen issues.	X	X			
	Item performance review based on data from field tests	ADCs review field-test statistics and recommend items for the common-eligible status, for re-field-testing (with edits, for mathematics, since ELA is passage-based), or for rejection. BSC also reviews items and recommends items to become common-eligible or to be rejected.	X	X	X	X	
Final Form Construction	Test form construction I: Proposition of items for operational use	DESE provides target performance-level cut scores to Cognia's test developers. Cognia proposes sets of common items (items that count toward student scores) and matrix items. Matrix items consist of field-test and equating items, which do not count toward student scores. Each common set of items is delivered with proposed cut scores, including test characteristic curves (TCCs) and test information functions (TIFs).	X	X			
	Test form construction II: Selection of items for operational use	DESE test developers and editorial staff review and edit proposed sets of items. Cognia and DESE test developers and editorial staff meet to review edits and changes to tests. Psychometricians are available to provide statistical information for changes to the common form. Approved common-eligible items become part of the common item set and are used to determine individual student scores.	X	X			
Public Release	Public release of select common items	Approximately 50% of common items in grades 3–8 are released to the public, and the remaining items are returned to the common-eligible pools to be used on future MCAS/RICAS tests. An item description (a statement specifying the content of the item) is released for each common item (both released and non-released).	X	X			

## 2.6.1 Item Review and Refinement

### DESE Initial Item Review

All passages, items, and scoring guides are reviewed by DESE test developers before presentation to the ADCs for review. The DESE test developers evaluate new items for the following as well as other characteristics:

- **Alignment:** Are the items aligned to the standards?
- **Content:** Is the content accurate? Does the item elicit a response that shows a depth of understanding of the content area?
- **Contexts:** Are contexts grade-level appropriate? Are they realistic? Are they interesting to students?
- **Grade-level appropriateness:** Are the content, language, and contexts appropriate for the grade level?
- **Creativity:** Does the item demonstrate creativity regarding approaches to items and contexts?
- **Distractors:** Have the distractors for selected-response items been chosen based on plausible construct-related errors? What are the distractor rationales?
- **Mechanics:** How well are the items written? Are they grammatically correct? Do they follow the conventions of item writing? Is the wording grade-level appropriate and accessible for all students?
- **Technology:** Are the items scored appropriately? Is the item making the best use of the technology? Is there another type of item that is more appropriate?

After initial review, DESE and the contractor’s test developers discuss and revise the proposed item sets in preparation for ADC review.

### Assessment Development Committee (ADC) and Bias & Sensitivity Committee (BSC) Reviews

ADCs and the BSCs are each composed of approximately 10–12 Massachusetts educators from across the state. Each ADC and BSC meeting is co-facilitated by DESE and Cognia’s test developers. There is an ADC for each content area and grade (e.g., ELA grade 3), and there is one BSC for all content areas and grades. All ADC and BSC recommendations remain with each item. ADC and BSC members meet several times a year to review new passages and items, and to review data from field-test items. Members review items using Pearson’s online platform ABBI. Each participant enters his or her “vote” and recommendations, and the facilitators record the consensus of the committee. DESE takes the recommendations of the ADCs and the BSCs into consideration and makes the final decision to approve items to become field-test eligible.

#### *ADC Passage Review (ELA Only)*

ELA ADCs review passages before any corresponding items are written. Committee members consider all the elements noted in section 2.4.1. Committee members are also asked to consider whether a passage is well known or comes from a book that is widely taught, since such a passage is likely to provide an unfair advantage to those students who are familiar with it. Committee members vote to accept or reject each passage, and the facilitators record the consensus of the group.

For each passage recommended for acceptance, committee members can provide suggestions for item development. They also can provide recommendations for the presentation of the passage, including



suggestions for the purpose-setting statement, words to be footnoted or redacted, and graphics, illustrations, or photographs to be included with the text.

#### *ADC Item Review*

Once DESE test developers have reviewed and edited new items and scoring guides, the items are reviewed by the ADCs. Committees review items for the characteristics noted above. Members vote to accept, accept with edits (members may include suggested edits), or reject each item. The meeting facilitators record the consensus of the group.

#### *BSC Passage and Item Review*

After passages and items have been approved by the ADCs, they are also reviewed by a separate BSC. The role of the committee is to identify whether a passage or item contains material that is likely to significantly favor or disadvantage one group of students for reasons that are not educationally relevant. The purpose of the committee's review is to ensure that the ability to answer an item correctly reflects a student's learning, not cultural opportunities or life experiences. Specifically, a passage or item should be flagged by the committee if it is insensitive or disrespectful to a student's ethnic, religious, or cultural background (including disability, socio-economic status, and regional differences). The BSC votes to accept, accept with edits (including suggested edits), or reject (including their reasoning) each passage or item. The meeting facilitators record the consensus of the group.

#### **External Content Expert Item Review**

When items are selected to be included on the field-test portion of the RICAS, they are submitted to expert reviewers for their feedback. The task of the expert reviewer is to consider the accuracy of the content of items. Each item is reviewed by two independent expert reviewers. All experts hold a doctoral degree (either in the content they are reviewing or in the field of education) and are affiliated with institutions of higher education in either teaching or research positions. Each expert reviewer has been approved by the DESE. The External Content Experts recommend either accepting or rejecting the item, including their reasoning. Expert reviewers' comments remain with each item.

#### **Editing of Recommended Items**

DESE test developers review the recommendations of the ADC, BSC, and expert reviewers and determine whether to revise or reject an item based on the suggested edits. The items are also reviewed and edited by DESE and Cognia editors to ensure adherence to style guidelines in *The Chicago Manual of Style*, *American Heritage Dictionary*, RICAS Style Guidelines, and to sound testing principles. According to these principles, all items should

- demonstrate correct grammar, punctuation, usage, and spelling;
- be written in a clear, concise style;
- contain unambiguous descriptions of what is required for a student to attain a maximum score;
- be written at a reading level that allows students to demonstrate their knowledge of the content area being tested.

Items that pass the reviews listed in this section are approved to be field-tested.

## **2.6.2 Operational Field-Testing of Items**

Only Massachusetts student data are used for field-test analyses. Rhode Island field-test data are not used for item evaluation. Field-tested items appear in the matrix portions of the tests. Each matrix item is

typically answered by a minimum of 1,500 students, resulting in enough responses to yield reliable performance data.

### **Scoring of Field-Tested Items**

All field-tested items, except for constructed-response items and essays, are machine-scored. These items include multiple-choice, multiple-select, short-answer, and technology-enhanced items.

All field-tested constructed-response items and essays are hand-scored. To train scorers, DESE works closely with the scoring staff to refine rubrics and scoring notes, and to select benchmark papers that exemplify the score points and variations within each score point. Approximately 2,000 student responses are scored per field-tested constructed-response item or essay. See Chapter 4 for additional information on scorers and scoring.

### **Data Review of Field-Tested Items**

#### *Data Review by DESE*

DESE test developers review all item statistics prior to making them available for review by the ADCs and BSCs. An item displaying statistics that indicate it did not perform as expected is closely reviewed and, if it is found to be flawed, it is rejected from the pool of items. After ADC and BSC reviews of item statistics, DESE test developers make final decisions regarding any recommendations.

#### *Data Review by ADCs*

The ADCs meet to review the field-test items with their associated statistics. ADCs review the following item statistics:

- item difficulty/mean item score,
- item discrimination,
- Differential Item Functioning (DIF) for the following subgroups:
  - Female compared with male [gender/sex]
  - African American/Black compared with White [ethnicity I]
  - Hispanic or Latino/a compared with White [ethnicity II]
  - Current or former ELLs compared with non-ELLs [language status]
- distribution of scores across answer options or score points,
- distribution of answer options or score points across quartiles, and
- distribution of unique student responses (for some items).

The ADCs make one of the following recommendations for each field-tested item:

- accept
- edit and field-test again (this recommendation is made for mathematics items only, since ELA items are passage-based)
- reject (not eligible for operational use).

### *Data Review by BSCs*

The BSC also reviews the statistics for the field-tested items. The committee reviews only the items that the ADCs have accepted. The BSC pays special attention to items that show DIF when comparing the following subgroups of test takers:

- Female compared with male [gender/sex]
- African American/Black compared with White [ethnicity I]
- Hispanic or Latino/a compared with White [ethnicity II]
- Current or former ELLs compared with non-ELLs [language status]

## **2.6.3 Final Form Construction**

Cognia’s test developers propose a set of previously field-tested or common, non-released items to be used in the common portion of the test. Test developers work closely with psychometricians to ensure that the proposed tests meet the statistical requirements set forth by DESE. In preparation for meeting with the DESE test developers, Cognia’s test developers consider the following criteria in selecting items to propose for the common portion of the test:

- **Content coverage/match to test design and blueprints.** The test designs and blueprints stipulate a specific number of items per item type and per reporting category for each content area. A broad coverage of standards and cognitive skills is expected. The previous year’s common test should also be considered, and items should not be duplicated.
- **Item difficulty and complexity.** Item statistics drawn from the data analysis of items are used to ensure similar levels of difficulty and complexity from year to year as well as high-quality psychometric characteristics. Items can be “reused” if they have not been released and not used the previous year. When an item is reused in the common portion of the test, the latest usage statistics accompany that item.
- **“Clueing” items.** Items are reviewed for any information that might “clue” or help the student answer another item.
- **Item types.** A variety of item types, including approximately 20–30% technology-enhanced items, are selected to populate the common slots.

Field-test items are also selected during form construction. Field-test items are drawn from the field-test eligible pools and should mirror the operational test to the extent needed. If a standard or reporting category is lacking in the common eligible item pool, items should be chosen to fill this need.

During assembly of the test forms, the following criteria are considered:

- **Key patterns.** The sequence of keys (correct answers) is reviewed to ensure that the key order appears random.
- **Option balance.** Items are balanced across forms so that each form contains a roughly equivalent number of key options (As, Bs, Cs, and Ds).
- **“Clueing” items.** Items are reviewed for any information that might “clue” or help the student answer another item.
- **Item types.** A variety of item types should populate the matrix slots.

The proposed operational test is posted for DESE to review. DESE test developers consider the proposed items, make recommendations for changes, and then meet with Cognia's test developers to construct the final forms of the tests.

After form construction meetings, the test forms enter several rounds of review by test developers and editors. Items are checked to ensure that the requested changes were made after the test construction meetings, and to ensure that all items are scoring correctly. In addition, items are checked again for any grammatical or “fatal flaw” errors, and these are corrected before the test forms are published, if any are found.

## 2.6.4 Special Edition Test Forms

### Students with Disabilities

RICAS is accessible to students with disabilities through the universal design of test items, provision of special edition test forms, and the availability of a range of accommodations and accessibility features for students taking the standard tests. To be eligible to receive a special edition test form, a student must have a disability that is documented either in an individualized education program (IEP) or in a 504 plan. All RICAS operational tests were available in the following special editions for students with disabilities:

- **Paper**—Form 1 of the operational CBT was produced to appear on paper. Items which used interactions not assessable on paper (typically technology enhanced items) were replaced with items that asked similar questions in a paper assessable manner. The grades 3–8 tests were administered to most students on the computer and to some students with accommodations on a paper form.
- **Large-print**—Form 1 of the operational test was translated into a large-print edition. The large-print edition contains all common and matrix items found in the Paper form.
- **Braille**—This form included only the common items found in the operational test with the following characteristics:
  - If an item indicates bias toward students with visual disabilities (e.g., if it includes a complex graphic that a student taking the Braille test could not reasonably be expected to comprehend as rendered), then simplification of the graphic is considered, with appropriate rewording of the item text, as necessary.
  - If a graphic such as a photograph cannot be rendered in Braille, or if the graphic is not needed for the student to respond to the item, the graphic is replaced with descriptive text or a caption or eliminated altogether.
  - Three-dimensional shapes that are rendered in two dimensions in print are rendered on the Braille test as “front view,” “top view,” and/or “side view,” and are accompanied where necessary by a three-dimensional wooden or plastic manipulative wrapped in a Braille-labeled plastic bag.

Modifications to original test items for the Braille version of the test are made only when necessary, as determined by the Braille test subcontractor and DESE staff, and only when they do not provide clues or assistance to the student or change what the item is measuring. When successful modification of an item or graphic is not possible, all or part of the item is omitted, and may be replaced with a similar item.

- **Screen reader**—This accommodation was available only for those students who are blind or have a visual disability. Students who used a screen reader were also given a separate hard-

copy Braille edition test to have the appropriate Braille graphics. All answers are entered onscreen, either by the student using a Braille writing device, or by the test administrator.

- **Text-to-speech**—This functionality was embedded in the grades 3–8 CBTs. Students typically use headphones with this format but may also be tested individually in a separate setting to minimize distractions to other students (from hearing what is being read aloud).

Appendix A details other accommodations that did not require a special edition test form and lists accessibility features that were available to all students, such as screen magnification and highlighting. After testing was completed, RIDE received a list with the number of students who participated in the 2024 RICAS with each accommodation, based on information compiled in the Personal Needs Profile in PearsonAccess Next.

### **Spanish Edition Test Forms**

Spanish editions of the spring grades 3–8 mathematics tests were available to any ELL student with a low level of English proficiency who was receiving or had received mathematics instruction in Spanish. More details can be found in section 3.4.

## **2.7 QUALITY CHECK**

At each point in the review process, including but not limited to Cognia’s initial work, reviews with the state of Massachusetts, reviews with educators, reviews for bias and sensitivity, expert reviews, and psychometric reviews checklists are used. These allow a variety of parties with multiple perspectives to confirm that details such as Universal Design principles, content alignment to state frameworks, cognitive complexities, and psychometric rigor meet the highest expectations.

## **2.8 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING TEST DESIGN AND DEVELOPMENT**

1.1 **Description Inference:** Observations of performance on the RICAS reflect the knowledge and abilities articulated in the RICAS content standards with appropriate assessment tasks representing the full breadth and depth of the domain as articulated within these standards.

1.1.1 **Claim:** Expected knowledge and abilities are thoroughly articulated and considered appropriate to the grade and content area being assessed.

**Evidence:** The appropriateness and official adoption of the content standards is articulated specifically in subsections 2.1 and 2.2. Subsequent sections in this chapter then describe how these standards are used to guide test design, development, and implementation processes for all grades and content areas.

1.1.2 **Claim:** Assessment tasks are developed to provide evidence of the expected knowledge and abilities for each grade and content area being assessed.

**Evidence:** Subsections 2.4.2 for ELA and 2.5.1 for mathematics explicitly state that items across all grades within those content areas “measured the ... learning standards as articulated in the

Rhode Island Core Standards,” detailing the specific standards addressed by items available for RICAS assessments. Subsections 2.4.1 and 2.4.2 describe passage types, item types, and cognitive levels for items on the ELA assessments. Subsection 2.5.1 describes the item types and cognitive levels for items on the mathematics assessments. Subsection 2.6.1 describes item review and refinement procedures, and subsection 2.6.2 describes item field testing and subsequent scoring and review processes. Together, these subsections describe an overall process of item development that ensures items effectively target the expected knowledge and abilities of the grades being assessed.

1.2 **Evaluation Inference:** Each test form (an organized sampling of assessment tasks) results in an observed score that reflects a student’s knowledge and abilities in the content area being assessed through appropriate test assembly, administration, and scoring procedures.

1.2.1 **Claim:** Each form is constructed to draw from available items such that the underlying domain of knowledge and abilities is adequately sampled.

**Evidence:** Subsection 2.4.3 describes the blueprints and test design specifications for ELA, while subsection 2.5.2 covers the same aspects for mathematics. Subsection 2.6.3 describes the processes for item selection and test form review, and Subsection 2.6.4 details the special edition test forms and modifications to the original test items. The procedures outlined in both subsections aim to ensure design and blueprint specifications are met, and they work to prevent elements of test construction that could potentially confound interpretability. Together, these processes ensure that each form draws a sampling of high-quality items representing the underlying knowledge and abilities defined within the content standards.

1.2.4 **Claim:** Items on the assessment demonstrate appropriate statistical quality.

**Evidence:** Subsection 2.6.2 describes the review process for evaluating items flagged by field-test analyses.

1.3 **Generalization Inference:** The observed score from any specific form testing a given grade and content area is reflective of the expected score on any potential form of the test for that grade and content area.

1.3.1 **Claim:** Task specifications adequately inform production or selection of items with similar content and statistical characteristics.

**Evidence:** Claim 1.1.2, with evidence from throughout Chapter 2, establishes that the task specifications and resulting item development efforts result in assessment tasks representative of expected knowledge and ability being assessed. Subsection 2.6.3 describes,

among other criteria, the need to meet the broad requirements of expected standards and cognitive skills while avoiding unnecessary duplication of items from previous years' forms. Subsection 2.6.3 also describes the rigorous process of form review to ensure that these requirements are met on forms that are accepted for operational administration. The form construction processes, applied to items meeting Claim 1.1.2, provide evidence that task specifications are adequately informing production and selection of items with similar content and statistical characteristics.

1.3.2 **Claim:** Test specifications result in forms of similar length and task distribution.

**Evidence:** Claim 1.2.1, again gathering evidence from Chapter 2, establishes that test construction processes are designed to implement specifications that result in forms of similar length and task distribution. Subsection 2.6.3 describes the application of those processes to realize those specifications while avoiding unnecessary duplication of items. Subsection 2.6.3 also describes the rigorous review process that verifies that these specifications are met prior to acceptance of the form for operational administration.

1.4 **Explanation Inference:** Expected scores are attributable to proficiency in the target knowledge and abilities.

1.4.1 **Claim:** Cut scores are established through defensible standard-setting methods.

**Evidence:** Section 2.3 summarizes the process by which performance standards were established for RICAS (more technical details in section 7.5). Standard-setting activities conducted for the MCAS in 2017 were observed by RIDE staff and technical advisors and rigorously evaluated for consistency with RICAS performance expectations.

# Chapter 3. Test Administration

Rhode Island test administrations were standardized, and the design was not significantly changed from previous years. As COVID-19 recovery progressed in 2023–2024, assessment and instructional practices likewise returned to standard, in-person protocols.

There were no major irregularities in test administration reported by schools or districts. Though the overall participation rate was high, important differences in participation rates and population demographics were observed between the SY 2020–2021, 2021–2022, 2022–2023, and 2023–2024 administrations. For more information, consult Appendix B of the 2021, 2022, and 2023 RICAS technical reports. The participation rates in SY 2020–2021 were an anomaly due to COVID-19. As a result, comparing aggregated student results across years and comparison of historical trends should only be used when the context of those differences is studied and known.

## 3.1 TEST ADMINISTRATION SCHEDULE

The standard grades 3–8 RICAS tests were administered in two modes, computer-based and paper-based, during two overlapping periods in spring 2024, as shown in Table 3-1.

**Table 3-1 Grades 3–8 ELA and Mathematics Test Administration Schedule**

Content Area	Complete the Student Registration/ Personal Needs Profile (SR/PNP) Process	Receive Test Administration Materials	Test Administration Windows	Deadline for Return of Materials to Contractor (for PBT Only)
ELA	4/26	3/13	3/25-4/19	4/29
Mathematics	5/24	4/12	4/22-5/17	5/28

### Testing Times

Table 3-2 shows the recommended testing times for ELA grades 3–8. As RICAS tests are untimed, the recommended times for scheduling test sessions are based on analysis of student testing times from prior computer-based testing administrations. The times shown in the table are approximate.

**Table 3-2 ELA Recommended Testing Times, Grades 3–8**

Grade	Session 1 Recommended Testing Time (min)	Session 2 Recommended Testing Time (min)	Total Recommended Testing Time (min)
3	120–150	120–150	240–300
4	120–150	120–150	240–300
5	120–150	120–150	240–300
6	120–150	120–150	240–300
7	120–150	120–150	240–300
8	120–150	120–150	240–300

Table 3-3 shows the recommended testing times for the 2024 mathematics tests. Since RICAS tests are untimed, the times shown are approximate.



**Table 3-3 Mathematics Recommended Testing Times, Grades 3–8**

Grade	Session 1	Session 2	Total
	Recommended Testing Time (min)	Recommended Testing Time (min)	Recommended Testing Time (min)
3	90	90	180
4	90	90	180
5	90	90	180
6	90	90	180
7	90	90	180
8	90	90	180

## 3.2 SECURITY REQUIREMENTS

Principals were responsible for ensuring that all test administrators complied with the requirements and instructions contained in the Test Coordinator’s Manual and Test Administrator’s Manuals. In addition, other administrators, educators, and staff within the school were responsible for complying with the same requirements. Schools and school staff who violated the test security requirements were subject to numerous sanctions and penalties, including employment consequences, delays in reporting of test results, the invalidation of test results, the removal of school personnel from future RICAS administrations, and possible licensure consequences for licensed educators.

If test content is breached, quick identification and resolution of the breach are critical to the integrity of a testing program. Full security requirements, including details about responsibilities of principals and test administrators, examples of testing irregularities, guidance for establishing and following a document tracking system, and lists of approved and unapproved resource materials, can be found in the *Spring 2024 Test Coordinator’s Manual (TCM), Grades 3–8* and the *2024 Test Administrator’s Manuals (TAMs)*. In spring 2024, there was one TAM for grades 3–8 CBTs and one TAM for grades 3–8 PBTs. The primary delivery mode was computer-based, with paper-based delivery as accommodation only for students with disabilities.

## 3.3 PARTICIPATION REQUIREMENTS

Students in grades 3–8 are expected to participate in RICAS tests for the grade in which they are enrolled and reported to RIDE through the enrollment census.

Participation requirements and guidelines for ELL students and students with significant disabilities are provided in the sections that follow; the participation rates are presented in Appendix B.

See Part III of the *Test Coordinator’s Manual* for information about scheduling test administration, including make-up sessions for students who are absent on the day of testing.

A very small number of students educated with Rhode Island public funds were not required to take the standard RICAS tests. These students were strictly limited to the following categories:

- First-year ELL students who enrolled in U.S. schools after April 1, 2022, for whom ELA testing is not required. (First-year ELL students must participate in RICAS or Dynamic Learning Maps [DLM] mathematics tests.) See the *RICAS Accessibility and Accommodations Manual, 2024* for details on how ELL students participate in spring 2024 RICAS.
- Students with significant cognitive disabilities who are eligible for the alternate assessment, the DLM assessment. For more information, refer to the DLM page of the RIDE website: <https://www.ride.ri.gov/InstructionAssessment/Assessment/DLMAssessments.aspx>.

- Rare and unique situations in which a student is unable to participate in statewide assessments due to a documented, significant, and incapacitating medical emergency that extends across the entire (or remaining) test window.

More details about test administration policies and participation requirements for non-disabled students, for students with disabilities, for ELL students, and for students educated in alternate settings can be found in the *Test Coordinator's Manual*. Data concerning the number of students tested with accommodations are available in Appendix A of this document.

## 3.4 SPANISH EDITION TEST FORMS

### Spanish-Speaking Students

Spanish editions of the spring grades 3–8 mathematics tests were available to any ELL student with a low level of English proficiency who was receiving or had received mathematics instruction in Spanish. The Spanish edition of the grades 3–8 mathematics tests contained all common and matrix items found in Form 1 of the operational test.

Cognia employed two independent translators to complete the translation of the grades 3–8 mathematics test to Spanish. The translation process was as follows:

- A set of translation rules or parameters was generated, taking the following into consideration: vocabulary, usage, and consistency over the years. These rules were provided to both translators.
- The first translator translated from English to Spanish. The second translator proofread the work of the first translator. Discrepancies between the two translations were resolved by the first translator.
- The Publishing Department reviewed the graphics in Spanish.
- The script that the teacher read when administering the test was also translated into Spanish and was included as Appendix A of the *Test Administrator's Manual*.

The Spanish editions of the grades 3–8 mathematics tests were available in both paper and online formats. Human read-aloud in Spanish was also available to students.

## 3.5 ADMINISTRATION PROCEDURES

It is the test coordinator's responsibility to coordinate the school's RICAS test administration. This coordination responsibility includes the following:

- understanding and enforcing the test security requirements and test administration protocols
- ensuring that students participate in testing according to the requirements in section 3.3 of this report
- coordinating the school's test administration schedule and ensuring that tests are scheduled during the prescribed testing window, and in the prescribed order
- ensuring that accommodations are properly administered and that transcriptions, if required for any accommodation, are properly completed
- completing the Principal's Certification of Proper Test Administration (PCPA) and ensuring the accuracy of information provided on the form

- providing RIDE with the school's correct contact information
- ensuring that all students have access to the appropriate grade level reference materials for the mathematics assessment as referenced in Chapter 2, subsection 2.5.3
- ensuring that all students who require supplemental supports (mathematics supplemental reference sheets or ELA graphic organizers) have access to them during testing

More details about test administration procedures, including ordering test materials, scheduling test administration, designating and training qualified test administrators, identifying testing spaces, meeting with students, providing accurate student information, and accounting for and returning test materials, can be found in the *Test Coordinator's Manual*.

The RICAS program is supported by the RICAS Service Center, which includes a toll-free telephone line and email answered by staff members who provide support to schools and districts. The RICAS Service Center operates weekdays from 7:00 a.m. to 5:00 p.m. (Eastern Time), Monday through Friday.

## 3.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING TEST ADMINISTRATION

1.2 **Evaluation Inference:** Each test form as an organized sampling of assessment tasks, results in an observed score that reflects a student's knowledge and abilities in the content area being assessed through appropriate test assembly, administration, and scoring procedures.

1.2.2 **Claim:** The assessment is administered under appropriate conditions.

**Evidence:** Chapter 3 describes the standardized test administration processes for the RICAS. This includes schedules, security requirements, administration procedures, and practices for non-standard administrations. Chapter 3 further references *Test Administrators Manuals* and *Test Coordinators Manuals* for more details of administration procedures, administrator responsibilities, and irregularity tracking.

# Chapter 4. Scoring

All grades and contents requiring human scoring were completed applying a virtual/synchronous scoring model. This scoring model means that the scorers completed the work from their homes as a team of scorers working the same time schedule and communicating via tools like Teams.

This virtual/synchronous model maintained the same stringent quality control measures that were applied in previous years. This continuity of proven methods also applied to rater training and scoring operational in general and showed that the results of that scoring are comparable to previous scoring events and RICAS administrations.

## 4.1 PREPARATION OF STUDENT RESPONSE BOOKLETS

Scoring of the 2024 RICAS tests was conducted by Pearson.

For paper-based tests, Cognia scanned each RICAS student-answer booklet. Images for operational items were transferred via a secure FTP site to Pearson for uploading into the ePEN scoring platform. For computer-based tests, images were uploaded into the appropriate scoring platform so that all scoring was conducted in a similar manner, regardless of the method of test administration.

A set of quality-control procedures was enacted for scanning paper test forms. These procedures included:

- checks of the answer booklet codes against the grade level, to ensure that the correct answer booklets were scanned in each batch;
- counting checks, to ensure that all booklets were accounted for; and
- spot checks, in which the scanned results were checked against randomly selected answer booklets to ensure that the scanners were working as intended.

For computer-based test takers, DESE had previously reviewed all items in the online item bank (ABBI) and approved all selected-response answer keys during test construction. The item scoring specifications (in Question and Test Interoperability [QTI]) were configured using the test maps and keys provided for the tests. Once the scoring system was configured, a quality-assurance group verified that the selected responses entered by the student for an item as shown in the uploaded image corresponded to the response recorded in the database, for both the pre-score and the scored student data files.

Scoring for selected-response items was verified against the specific DESE requirements for the item; the requirement of the test map, which includes the QTI response; and the keys and validations made for an individual student's derived scores per level of the test. This process included a review of all score-value-related fields—such as raw scores, object scores (part one and part two of multi-part items), strand scores, achievement levels, pass/fail indicators, attempt rules, and scaled scores—against the tables provided by Pearson psychometrics.

## 4.2 PREPARATION FOR SCORING CONSTRUCTED-RESPONSE ITEMS

Scoring responses to short-answer, constructed-response, and essay items began by first preparing the documents for scoring. Student identification information, demographic information, and school contact information was converted to alphanumeric format. Digitized student responses to constructed-response items were sorted into specific content areas, grade levels, and items before being scored.

Scoring consistency across scoring departments on all item types was established by conducting the following activities:

- For field-test items, Cognia facilitated benchmarking meetings in meeting rooms at Cognia headquarters in Portsmouth, New Hampshire. This activity is required to determine items' suitability for inclusion in operational assessments. Cognia provided annotated anchor, practice, and qualification sets for all existing items to Pearson for review in advance of scoring. Content specialists at Pearson and Cognia consulted with each other to address any questions and ensure clarity of training materials.
- For operational ELA items that needed to be re-benchmarked due to modifications, content specialists from Cognia, Pearson, and DESE collaborated on the establishment of final scoring decisions.
- Weekly meetings between the Cognia and Pearson scoring departments were held to address any issues and questions before and during scoring.

## 4.3 BENCHMARKING MEETINGS

Samples of student responses to field-test items were read, scored, and discussed by members of Cognia's Scoring Services and Content Development and Publishing (CDP) Departments and by DESE test developers. Each benchmarking meeting is content and grade specific (e.g., grade 6 ELA). All decisions were recorded and considered final upon DESE signoff.

The primary goals of the field-test benchmarking meetings were to

- revise, as necessary, an item's scoring guide and/or scoring rubric;
- revise, as necessary, an item's scoring notes based on student responses—these, along with scoring guides, provide detailed information about how to score an item;
- assign final score points to a given set of student responses; and
- approve anchor and training sets of responses that are used to train scorers.

## 4.4 MACHINE-SCORED ITEMS

Student responses to selected-response and short-answer items were machine-scored by PearsonAccess Next (PAN) Scoring. PAN is a next-generation, web-based technology platform for end-to-end administration of large-scale assessments. Student responses with multiple marks (possible only on paper-based tests) and blank responses were assigned zero points.

## 4.5 HAND-SCORED ITEMS

The hand-scored items were sorted into item-specific groups for scoring. All student responses to a particular item (asset ID) were scored together, before moving on to the next item. However, if there was a need to see a particular student's responses for all hand-scored items together, scoring leadership had the ability to do so. Details on the procedures used to hand-score student responses are provided later in this chapter.

### 4.5.1 Scoring Plan and Staff

The following scoring plan summarizes the approach to the scoring of RICAS administrations for all grades and contents:

- All scoring was conducted applying a virtual/synchronous scoring model maintaining stringent quality control measures applied in a center-based, regional scoring environment.
- Prior to the start of scoring, scorers attended connectivity sessions to support their readiness for virtual/synchronous scoring and to answer any technology-related questions.
- Scorers evaluated student work on a fixed daily schedule under constant supervision of leadership.
- Training and all interaction between leadership and scorers occurred live via Teams (Pearson) and/or via pre-recorded training module or a recording of live training.
- Breakout rooms were used to facilitate scorer training and individualized coaching.
- DESE had remote access to the scoring systems and Teams links were provided to observe training sessions and scoring.
- Scorers worked in a non-public setting and were required to be on camera during training, scoring, and any one-on-one or group coaching sessions.
- A post-scoring survey was sent out to all MCAS and RICAS scoring associates to elicit feedback on their scoring experience. The results were shared with DESE.

The following staff members were involved with scoring the 2024 RICAS responses:

- Cognia Staff
  - The *Scoring Project Manager* was responsible for the communication and coordination of RICAS scoring between Cognia and Pearson, and between Cognia and RIDE.
  - *Scoring Content Specialists* facilitated all benchmarking meetings to ensure consistency of content area benchmarking and field-test scoring across all grade levels. Scoring content specialists prepared training materials for all operational scoring of ELA and mathematics grades 3–8 prior to scoring by Pearson. They also fielded any questions between Pearson and Cognia to ensure a consistent scoring approach across the scoring groups and years.
  - *Scoring Supervisors* were responsible for the training and qualification of both scorers and Scoring Team Leaders, and for ensuring quality targets for their assigned items during field testing.
  - *Scoring Team Leaders* provided support and direction to scorers on quality, accuracy, and timely completion of scoring during field testing.

- Pearson Staff
  - The *Scoring Portfolio Manager* was responsible for the coordination, management, and oversight of RICAS scoring for Pearson.
  - The *Scoring Project Manager* oversaw communication and coordination of RICAS scoring between Pearson and Cogna.
  - *Scoring Content Specialists* ensured consistency of content area scoring across all grade levels. Scoring content specialists monitored the quality of scoring and worked closely with a group of scoring directors to ensure the accurate and timely completion of scoring. Scoring content specialists also coordinated communication with their counterparts at Cogna regarding the training materials.
  - *Scoring Directors* were responsible for the training and qualification of both scorers and scoring supervisors and ensuring quality targets for their assigned items.
  - *Scoring Supervisors* provided support and direction to scorers on quality, accuracy, and timely scoring completion.
  - *Automated Scoring Team Members* were responsible for training and monitoring the scoring performance of the Intelligent Essay Assessor (IEA) on the subset of the ELA prompts selected for automated scoring.

## 4.5.2 Scorer Recruitment and Qualifications

RICAS scorers, a diverse group of individuals with a wide range of backgrounds, ages, and experiences, were recruited to meet contract requirements. These requirements included successful completion of at least two years of college, although hiring preference was given to individuals with a four-year college degree. Those scoring high school students' responses must have at least a 4-year degree and must either have a degree related to the content they were working on OR have at least two classes related to the content and have prior experience in the content area.

Teachers, tutors, and administrators (e.g., principals, guidance counselors) currently under contract or employed by or in Massachusetts schools, and people under 18 years of age were not eligible to score RICAS responses. Potential scorers were required to apply and submit documentation of qualifications, such as résumés and transcripts, which were carefully reviewed. Regardless of their qualifications, potential scorers who did not clearly demonstrate content area knowledge or have at least two college courses with average or above-average grades in the content area they wished to score were eliminated from the applicant pool. A summary of scorers' backgrounds is provided in Table 4-1.

**Table 4-1 Summary of Scorer and Scoring Leadership Backgrounds (Operational Scoring)**

Pearson Education	Scorer		Leadership	
	Number	Percent	Number	Percent
Bachelor's degree or higher	734	100.00%	81	100.00%
Master's degree or higher	326	44.41%	39	48.15%
<b>Teaching Experience</b>				
I have no teaching experience	180	24.52%	23	28.40%
Less than a year	73	9.95%	8	9.88%
1-2 years	98	13.35%	14	17.28%
3-5 years	127	17.30%	9	11.11%
6-10 years	91	12.40%	10	12.35%
11 years or more	165	22.48%	17	20.99%
<b>Scoring Experience</b>				
Less than a year	115	15.67%	0	0.00%
1 year scoring	227	30.93%	4	4.94%
2 years scoring	141	19.21%	17	20.99%
3 years scoring	42	5.72%	6	7.41%
4 years scoring	32	4.36%	8	9.88%
5 years scoring	28	3.81%	10	12.35%
6 years scoring	22	3.00%	5	6.17%
7 years scoring	15	2.04%	2	2.47%
8 years scoring	19	2.59%	5	6.17%
9 years scoring	19	2.59%	4	4.94%
10 years scoring	10	1.36%	2	2.47%
11 years scoring	15	2.04%	2	2.47%
12 years scoring	8	1.09%	1	1.23%
13 years scoring	13	1.77%	4	4.94%
14 years scoring	13	1.77%	6	7.41%
15 years scoring	8	1.09%	1	1.23%
16 years scoring	5	0.68%	2	2.47%
17 years scoring	0	0.00%	0	0.00%
18 years scoring	1	0.14%	2	2.47%
19 years scoring	1	0.14%	0	0.00%

### 4.5.3 Scorer Training

Scoring content specialists had overall responsibility for ensuring that responses were scored consistently, fairly, and according to the approved scoring guidelines. Scoring materials were carefully compiled and checked for consistency and accuracy. Student identification information, demographic information, and school contact information were not visible to scorers. The sequence and manner in which the materials were presented to scorers was standardized to ensure that all scorers had the same training environment and scoring experience, regardless of content, grade level, or item scored.

Three training methods were used to train scorers of RICAS hand-scored items:

- 1) live group training via Teams
- 2) recording of live group training
- 3) pre-recorded interactive modules

Scorers started the training process by receiving an overview of RICAS; this general orientation included the purpose and goal of the testing program and any unique features of the test and the testing population. Scorer training for a specific item to be scored always started with a thorough review and discussion of the scoring guide, which consisted of the task, the scoring rubric, and any specific scoring notes for that task. All scoring guides were previously approved by the DESE during field-test benchmarking meetings and used without any additions or deletions.



As part of training, prospective scorers carefully reviewed three different sets of student responses, some of which had been used to train scorers when the item was a field-test item on MCAS:

- **Anchor sets** are DESE-approved sets consisting of two or three sample responses at each score point. Each response represents a typical response, rather than an unusual or uncommon one; is solid and has a true score, meaning that this response has a precise score. Anchor sets are used to exemplify each score point.
- **Practice sets may** include unusual, discussion-provoking responses, illustrating the range of responses encountered in operational scoring (including exceptionally creative approaches; extremely short or disorganized responses; responses that demonstrate attributes of both higher-score anchor papers and lower-score anchor papers; and responses that show traits of multiple score points). Practice sets are used to refine the scorers' understanding of how to apply the scoring rules across a wide range of responses.
- **Qualifying sets** consist of 10 responses that are clear, typical examples of each of the possible score points. Qualifying sets are used to determine whether scorers can score consistently according to the RIDE-approved scoring standards.

Meeting or surpassing the minimum acceptable standard on an item's qualifying set was an absolute requirement for scoring student responses to that item. An individual scorer must have attained a scoring accuracy rate of 70% exact and 90% exact-plus-adjacent agreement<sup>1</sup> (at least 7 out of the 10 were exact score matches and at least 9 out of the 10 were either zero or one discrepant) on either of two potential qualifying sets. For multi-trait ELA items, each scorer had to meet the 70% and 90% passing the threshold for each individual trait.

#### 4.5.4 Leadership Training

Scoring content specialists also had overall responsibility for ensuring that scoring leadership (Cognia scoring supervisors and Pearson scoring directors) continued their history of scoring consistently, fairly, and according to the approved scoring guidelines. Once they had completed their item-specific training, scoring leadership was required to meet or surpass a qualification standard of at least 80% exact and 90% exact-plus-adjacent scoring accuracy. For multi-trait ELA items, scoring leadership had to meet the 80% and 90% passing threshold for each individual trait.

#### 4.5.5 Methodology for Scoring Hand-Scored Polytomous Items

All operational items in grades 3–8 ELA and mathematics were selected from items which had been field tested in previous years. For operational scoring, 100% of the ELA essay responses were scored via automated scoring using Pearson's Intelligent Essay Assessor (IEA) with human double-blind scoring at a rate of 10%. All other 3–8 ELA and mathematics responses were done by human scorers at a rate of 10% human-to-human double-blind scoring. Information on how the IEA works and how it was used on the RICAS essay scoring is provided in section 4.5.7 below.

#### 4.5.6 Monitoring of Scoring Quality

The 2024 RICAS tests included constructed-response items and essays (in addition to selected-response and short-answer items) that were scored by hand. Hand-scored items included the following:

---

<sup>1</sup> "Adjacent agreement" means that a pair of scores (for the same response) are only off by one point. "Exact-plus-adjacent agreement" means that a pair of scores are either the same or off by only one point.

- constructed-response items with assigned scores of 0–3 (ELA grades 3 and 4 only)
- constructed-response items with assigned scores of 0–3 (mathematics grade 3) and 0–4 (mathematics grades 4–8)
- essays with assigned scores for two traits, Idea Development and Language Conventions. In ELA grades 3–5 the Idea Development score ranged from 0–4 and for ELA grades 6–8 the scores ranged from 0–5. For all grades of ELA essays, the trait Language Conventions was scored on a range from 0–3 points.

For each of these hand-scored items, a scoring guide was created. For examples of item-specific scoring guides, see the RICAS Student Work/Scoring Guides webpage at <http://ricas.pearsonsupport.com/released-items/>.

Non-numeric scores assigned by Cognia and Pearson could be designated as:

- Blank: The written response form is completely blank.
- Unreadable: The response cannot be read because of poor penmanship, or spelling cannot be deciphered, or writing is too small, too faint to see, or only partially visible.
- Non-English: Response was written entirely in a language other than English or without enough English or numbers to provide a score.
- Off Topic: Response does not address the topic or task for the item. The response is irrelevant to the item prompt, or the response states that the student is refusing to participate in testing.
- Direct Copy: Direct copy of text from the passage or item prompt.

Scorers could also flag a response as a “Crisis” response, which would be sent to scoring leadership for immediate attention.

A response would be flagged as a “Crisis” response if it indicated

- perceived, credible desire to harm self or others;
- perceived, credible, and unresolved instances of mental, physical, or sexual abuse;
- presence of language or thoughts that may require professional intervention;
- sexual knowledge well beyond the student’s developmental age;
- ongoing, unresolved misuse of legal/illegal substances (including alcohol);
- knowledge of or participation in real, unresolved criminal activity; or
- direct or indirect request for adult intervention/assistance (e.g., crisis pregnancy, doubt about how to handle a serious problem at home).

## Scoring Approach

### *Single-Scoring*

All ELA essays were scored using the Intelligent Essay Assessor. All mathematics student responses received at least one human score. These were the only scores unless the response was slotted for double-blind scoring or routed for human scoring by IEA.

### *Double-Blind Scoring*

In double-blind scoring, a response is independently read and scored by two human scorers. These scorers were not aware that double-blind scoring was taking place. For a double-blind response that received adjacent scores (i.e., two scores within one point of each other), the higher score was used. Any double-blind response with discrepant scores greater than one point was sent to an arbitration queue and was read by scoring leadership, in which case the expert score that resolved the scoring discrepancy was used. 10% of the student responses were given a double-blind score. The IEA scoring platform was used to score 100% of the ELA essay items and 10% of those scores were double-blind scored by a human for validity.

A description of how the IEA functions and how it was used is provided in section 4.5.7.

#### *Read-Behind Scoring*

In addition to the double-blind scoring, scoring leadership, at random points throughout the scoring shift, engaged in read-behind (back-read) scoring for each scorer assigned to their team. In this process, scoring leadership viewed responses recently scored by a particular scorer and assigned a score to that same response. Scoring leadership then compared scores and advised/counseled the scorer as necessary.

### **4.5.7 Scoring with the Intelligent Essay Assessor (IEA)**

The Intelligent Essay Assessor (IEA) is used to score student responses to essay prompts.<sup>2</sup> Like human scorers, IEA evaluates the content and meaning of text, as well as grammar, style, and mechanics. IEA learns to score via a range of machine learning and natural language processing technologies. The engine is trained individually on each prompt and trait using hundreds or thousands of human-scored student responses.

IEA measures the content and quality of responses by determining the features human scorers evaluate when scoring a response. Given a set of human-scored responses to a prompt, IEA computes hundreds of different metrics that characterize each response in various ways. Some examples of these metrics include:

- number of grammar errors
- types of grammar errors
- variety of words
- maturity of words
- variety of sentence types
- coherence of the response
- similarity of the response to other responses and/or source materials

All these different metrics are fed to machine learning algorithms that determine which of them best predict the scores assigned by human scorers.

IEA was trained before the operational assessment was administered using responses collected during the field test and scored by trained human scorers. For each prompt, IEA was trained using approximately 1,300 responses per prompt and then evaluated using approximately 640 responses.

---

<sup>2</sup> Additional information about IEA can be found in Foltz, P. W., Streeter, L. A., Lochbaum, K. E., & Landauer, T. K. (2013). *Implementation and applications of the Intelligent Essay Assessor. Handbook of Automated Essay Evaluation, M. Shermis & J. Burstein (Eds.). Pp. 68-88. Routledge, NY, NY.*

Table 4-2 includes the specific N counts for each prompt. The responses were randomly assigned to each set (training or evaluation). Performance on the evaluation set was measured using a variety of criteria comparing IEA with human scoring using the standard metrics shown in Table 4-3.

**Table 4-2 N Counts by Prompt**

Grade	Prompt	Training Set Size	Evaluation Set Size
3	EL206642581	1,242	619
4	EL909132428	1,268	635
5	EL206057533	1,219	607
	EL207135604	1,277	640
6	EL206144523	1,243	623
	EL028556260	1,254	628
7	EL206144041	1,285	644
	EL217467654	1,269	635
8	EL206761648	1,268	634
	EL810463548	1,300	650

**Table 4-3 Metrics for Evaluating Automated Scoring<sup>3</sup>**

Measure	Threshold
Pearson R	≥ 0.70
Quadratic Weighted Kappa (QWK)	≥ 0.70
Kappa	≥ 0.40
Exact Agreement	≥ 65% (or better than human-human agreement)
Per score point agreement	≥ 50% (or better than human-human agreement)
Standardized Mean Difference (SMD)	Within  0.15

Ten prompts met the required performance criteria and were approved by DESE to be scored by IEA as the double-blind score to monitor quality during the operational assessment. Scoring performance on the operational assessment is described in the next section.

Table 4-4 shows a comparison of IEA to human scoring on the validity papers, by exact score point (validity papers are student responses with known scores interspersed among the other student responses; these papers are used to check scoring accuracy). As shown below, IEA scoring accuracy on these validity papers is similar to or slightly higher than the human scoring accuracy at most score points. IEA accuracy tends to be higher than human accuracy at the highest score point, as seen in the Idea Development agreement statistics for grades 3–8. An exception to this trend appears at times in the higher score points (4 or 5) when there is a low percentage of responses at these score points, making it difficult to identify responses that solidly meet the criteria for validity responses and limiting scorers' opportunity to score such. As a result, the validity pool for a score point of 4 or 5 is a smaller sample size than other score points.

<sup>3</sup> Williamson, D. M., Xi, X., & Breyer, F. J. (2012). *A framework for evaluation and use of automated scoring. Educational Measurement: Issues and Practices, 31, 2.*

**Table 4-4 Comparison of Human and IEA Agreement with Validity Papers—ELA**

Grade	UIN	Trait	Validity	N	Exact Agreement	Exact Agreement by Score Point					
						0	1	2	3	4	5
3	EL206642581	Idea Development	IEA	44	86%	100%	89%	93%	100%	33%	--
			Human		84%	93%	97%	79%	72%	63%	--
		Conventions	IEA		96%	100%	90%	94%	100%	--	--
			Human		82%	78%	92%	74%	84%	--	--
4	EL909132428	Idea Development	IEA	33	97%	100%	100%	80%	100%	100%	--
			Human		88%	95%	89%	82%	72%	54%	--
		Conventions	IEA		100%	100%	100%	100%	100%	--	--
			Human		92%	94%	93%	82%	92%	--	--
	EL206057533	Idea Development	IEA	38	90%	100%	80%	90%	86%	100%	--
			Human		83%	95%	91%	75%	60%	100%	--
		Conventions	IEA		95%	90%	90%	100%	100%	--	--
			Human		82%	100%	88%	75%	59%	--	--
5	EL207135604	Idea Development	IEA	18	94%	100%	100%	100%	75%	100%	--
			Human		80%	76%	91%	78%	67%	69%	--
		Conventions	IEA		100%	100%	100%	100%	100%	--	--
			Human		82%	80%	83%	75%	91%	--	--
	EL206144523	Idea Development	IEA	33	85%	100%	80%	100%	100%	60%	50%
			Human		80%	93%	86%	76%	69%	48%	62%
		Conventions	IEA		97%	100%	91%	100%	100%	--	--
			Human		88%	97%	86%	82%	95%	--	--
6	EL028556260	Idea Development	IEA	17	100%	100%	100%	100%	100%	100%	100%
			Human		90%	98%	80%	82%	73%	64%	92%
		Conventions	IEA		100%	100%	100%	100%	100%	--	--
			Human		93%	98%	84%	82%	89%	--	--
	EL206144041	Idea Development	IEA	39	97%	100%	100%	100%	86%	100%	100%
			Human		82%	88%	95%	88%	71%	58%	72%
		Conventions	IEA		92%	60%	100%	91%	100%	--	--
			Human		87%	81%	90%	80%	91%	--	--
7	EL217467654	Idea Development	IEA	21	100%	100%	100%	100%	100%	100%	100%
			Human		89%	98%	94%	88%	76%	50%	53%
		Conventions	IEA		100%	100%	100%	100%	100%	--	--
			Human		92%	95%	94%	87%	90%	--	--
	EL206761648	Idea Development	IEA	20	90%	100%	75%	100%	100%	50%	100%
			Human		81%	98%	80%	87%	59%	36%	60%
		Conventions	IEA		100%	100%	100%	100%	100%	--	--
			Human		92%	98%	85%	91%	89%	--	--
8	EL810463548	Idea Development	IEA	55	93%	100%	92%	100%	78%	86%	100%
			Human		84%	94%	93%	85%	67%	45%	76%
		Conventions	IEA		100%	100%	100%	100%	100%	--	--
			Human		91%	97%	94%	82%	89%	--	--

### 4.5.8 Monitoring of Scoring Quality

Once RICAS scorers met or exceeded the minimum standard on a qualifying set and were allowed to begin scoring, they were constantly monitored throughout the entire scoring window to ensure they scored student responses as accurately and consistently as possible. If a scorer fell below the minimum standard on any of the quality-control indicators, some form of intervention occurred, ranging from counseling to retraining to dismissal. Scorers were required to meet or exceed the minimum standard of 70% exact and 90% exact-plus-adjacent agreement on the following quality control methods listed and further defined below:

- validity responses,
- read-behind scoring (RBs)/back-reading,
- double-blind scoring (DBs), and
- compilation reports (summary of scoring agreement statistics).

**Validity responses** were used to monitor the scorer's accuracy of scoring. These responses were approved by scoring leadership and distributed to scorers based on a percentage of their total number of responses scored. For the first two days, validity responses routed to scorers comprised 6% of their responses for ELA and 3% for mathematics. Starting with the third day of live scoring, these rates were reduced to 4% for ELA and 2% for mathematics. At the third-day rate, a full shift of scoring was expected to result in 6 to 19 validity responses per day in ELA and around 8 validity responses per day in mathematics, based on expected read rates.

Alert messages were issued to scorers who did not meet minimum validity metrics after 10 validity responses. If, after an additional five validity responses, the scorer had not improved, ePEN, the scoring platform, automatically blocked that scorer, and launched a 10-response targeted calibration set. The scorer was required to attain at least 70% exact agreement and 90% exact-plus-adjacent agreement on this calibration set to continue scoring the item for which the calibration set was administered. If the scorer passed the targeted calibration, ePEN was unblocked and the scorer regained admission to operational responses. The scorer was required to continue maintaining scoring standards for validity, as validity statistics continued to be checked every 10 validity responses. If validity fell below scoring standards at any of these subsequent intervals, the scorer was released from the project and all scores assigned immediately reset.

**Read-behinds** involved responses that were first read and scored by a scorer, then read and scored by a member of scoring leadership. Scoring leadership would, at various points during the scoring shift, conduct a review of submitted scorer work. After the scorer scored the response, scoring leadership would give his or her own score to the response and then compare his or her score to the scorer's score. Read-behinds were performed at least 10 times for each full-time day shift scorer and at least five times for each evening shift and partial-day shift scorer. Scorers who fell below the 70% exact and 90% exact-plus-adjacent score agreement standard were counseled, given extra monitoring assignments such as additional read-behinds, and allowed to resume scoring if they demonstrated the ability to meet the scoring standards after the intervention.

**Double-blinds** involved responses scored independently by two different scorers. For ELA essays, IEA was considered a scorer. Double-blind scoring served as an indicator for agreement of scoring between two scorers. Responses given discrepant scores by two independent scorers were read and scored by scoring leadership.

**Compilation reports** were generated daily. Compilation reports displayed all the statistics for each scorer, including the percentage of exact, adjacent, and discrepant scores on the backreads as well as the percentage of exact, adjacent, and discrepant scores on validity sets. As scoring leadership conducted backreads, the scorers' overall percentages on the compilation report were automatically calculated and updated. If the compilation report at the end of the scoring shift listed any individuals who were still below the 70% exact and 90% exact-plus-adjacent standard, their scores for that day were voided. Responses with voided scores were returned to the scoring queue for other scorers to score.

## 4.5.9 Interrater Consistency

Interrater consistency statistics are the result of the processes implemented to ensure valid and reliable hand-scoring of items and, as such, provide evidence of scoring stability. Double-blind scoring was one of the processes used to monitor the quality of the hand-scoring of student responses for constructed-response items. For student constructed-response questions in grades 3–8, 10% were randomly selected and scored independently by two different scorers. Results of the double-blind scoring were used during the scoring process to identify scorers who required retraining or other intervention, and they are presented here as evidence of scoring consistency on the RICAS tests.

A third score was required for any score category in which there was not an exact agreement between scorer one and scorer two. A third score was also required as a confirmation score when either scorer one and/or scorer two provided a score of M for Demonstration of Skills and Concepts and Independence or a score of 1 for Level of Complexity.

Summaries of the interrater consistency results are presented in Tables 4-5 for ELA and 4-6 for mathematics by grade. The tables show the number of score categories, the number of included scores, the percent exact agreement, the percent adjacent agreement, the correlation between the first two sets of scores, the percent of responses that required a third score, and linearly weighted (LW) Kappa as a measure of agreement scorer consistency by accounting for chance agreement. Interrater consistency data are available at the item level in Appendix C.

**Table 4-5 Summary of Interrater Consistency Statistics Organized across Items by Content Area and Grade—ELA**

Grade	Number of		Percent			% of Third Scores	LW Kappa
	Score Categories	Included Scores	Exact	Adjacent	Correlation		
3	4	1,782	79.57	18.46	0.78	2.13	0.72
	5	867	82.01	17.65	0.84	0.58	0.77
4	4	1,916	78.29	19.47	0.80	2.71	0.73
	5	945	68.68	30.37	0.82	1.59	0.69
5	4	1,888	74.52	23.78	0.83	2.65	0.73
	5	1,888	77.12	21.61	0.87	2.65	0.78
6	4	1,859	76.55	22.49	0.88	3.87	0.78
	6	1,859	74.93	21.79	0.90	3.87	0.80
7	4	1,858	78.15	20.67	0.88	2.64	0.80
	6	1,858	69.70	28.31	0.89	2.64	0.77
8	4	1,886	81.50	17.97	0.91	1.27	0.84
	6	1,886	79.53	19.51	0.93	1.27	0.86

Note. LW = linearly weighted

**Table 4-6 Summary of Interrater Consistency Statistics Organized across Items by Content Area and Grade—Mathematics**

Grade	Number of		Percent			% of Third Scores	LW Kappa
	Score Categories	Included Scores	Exact	Adjacent	Correlation		
3	4	3,798	89.18	10.32	0.95	0.50	0.90
4	5	3,936	83.79	15.57	0.96	0.64	0.88
5	5	3,949	78.50	19.95	0.91	1.54	0.80
6	5	3,858	83.72	14.46	0.94	1.81	0.86
7	5	3,900	91.18	8.49	0.97	0.33	0.93
8	5	3,853	83.88	14.82	0.94	1.30	0.87

Note. LW = linearly weighted

Table 4-7 provides a summary of the “validity” statistics. These statistics denote accuracy in scoring; they provide an average of the human and IEA agreement with the validity responses (e.g., agreement with the true scores for each essay).

**Table 4-7 Summary of Validity Statistics<sup>1</sup>**

Content Area	Grade	Number of Score Categories <sup>2</sup>	Number of Validity Responses <sup>3</sup>	Exact Agreement	Agreement by Score Point					
					0	1	2	3	4	5
ELA	3	4 (SR)	3,510	89.9%	94.3%	95.8%	67.8%	73.3%		
		4 (Conv)	431	82.1%	78.2%	91.9%	74.3%	83.5%		
		5 (ID)	431	84.4%	92.6%	97.5%	79.1%	71.8%	62.7%	
	4	4 (SR)	3,591	88.1%	94.6%	93.5%	79.0%	78.8%		
		4 (Conv)	397	91.6%	94.1%	93.2%	81.6%	91.6%		
		5 (ID)	397	87.9%	95.1%	89.4%	81.7%	71.6%	54.0%	
	5	4 (Conv)	790	82.2%	90.2%	85.2%	75.0%	78.8%		
		5 (ID)	790	81.3%	87.3%	91.6%	76.7%	64.3%	70.3%	
	6	4 (Conv)	827	90.4%	98.5%	85.8%	82.2%	92.5%		
	7	6 (ID)	827	84.3%	96.8%	84.7%	77.6%	71.7%	53.5%	75.0%
		4 (Conv)	906	89.6%	91.9%	92.4%	83.4%	90.4%		
	8	6 (ID)	906	85.4%	93.8%	94.5%	87.9%	79.4%	55.9%	60.8%
4 (Conv)		750	91.5%	97.8%	89.9%	86.4%	88.8%			
Mathematics	3	6 (ID)	750	82.0%	96.2%	87.3%	85.9%	62.9%	41.9%	64.8%
		4	7,118	93.7%	96.4%	94.6%	91.7%	92.8%		
	4	5	7,793	94.8%	95.9%	96.1%	94.3%	91.6%	95.8%	
	5	5	7,460	90.9%	97.4%	89.9%	85.7%	90.7%	90.3%	
	6	5	7,617	91.9%	96.6%	91.2%	91.4%	85.6%	94.0%	
	7	5	8,253	97.1%	99.0%	97.8%	95.4%	94.9%	98.1%	
8	5	8,089	93.8%	97.8%	92.4%	91.7%	89.1%	95.6%		

<sup>1</sup>Includes all operational and equating items for ELA and mathematics.

<sup>2</sup>SR= Short response; Conv= Conventions; ID=Idea Development

<sup>3</sup>This column displays the number of validity reads (how many times all the responses were scored against validity papers) that occurred, not the number of validity papers used.

## 4.6 QUALITY CHECK

Numerous quality control measures are employed to ensure the quality and consistency of scoring. Leadership training is conducted prior to scorer training to provide a full understanding of the content. Members of the leadership team must qualify at a higher standard than scorers. All training materials are carefully chosen to expose scorers to the range of student responses they will encounter during live scoring. Scoring is monitored through proven quality control methods, including validity sets, double-blind scoring, and read-behind scoring. If a scorer fails to meet quality control standards, their work is voided and returned to the queue to be rescored. All scoring is conducted in a virtual/synchronous environment with designated hours. The leadership team consists of expert scorers who are always available to provide feedback and coaching and to address questions. These proven methods of quality control are applied to ensure that the consistency of scoring student work is maintained.



## 4.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING SCORING

1.2 **Evaluation Inference:** Each test form, as an organized sampling of assessment tasks, results in an observed score that reflects a student's knowledge and abilities in the content area being assessed through appropriate test assembly, administration, and scoring procedures.

1.2.3 **Claim:** The scoring procedures and models produce scores accurately reflective of targeted knowledge and abilities.

**Evidence:** Chapter 4 has detailed sections describing the scoring process for machine-scored and hand-scored items on RICAS assessments. This includes detailed descriptions of preparation, benchmark meetings, recruitment and training of scorers, monitoring of scoring quality, and interrater consistency.

# Chapter 5. Reporting

## 5.1 REPORTING OF RESULTS

Results on the RICAS were reported in terms of achievement levels that describe student achievement in relation to established state standards. There are four achievement levels for ELA and mathematics for students in grades 3–8:

*Level 1: Not Meeting Expectations*

*Level 2: Partially Meeting Expectations*

*Level 3: Meeting Expectations*

*Level 4: Exceeding Expectations*

Students were given a separate achievement-level classification in each content area. Reports are generated at the student level. The achievement level distributions are provided in Appendix D.

Parent/Guardian Reports were printed and mailed to districts for distribution to parents/guardians and schools. The Parent/Guardian Report is also available to schools in PearsonAccess Next (PAN). Individual Parent/Guardian Report PDFs were run and collated by grade and school and posted online for school, district, and state access.

## 5.2 PARENT/GUARDIAN REPORT

The Parent/Guardian Report (based on the MCAS report design) or “Individual Student Report” was generated for each student who participated in the RICAS tests. The report is designed as a stand-alone 11" x 17" color report that is folded to create four pages; see Appendix E for a sample report. Two full-color copies of each student’s report were printed: one for the parent/guardian and one for the school’s records. The report is designed to present parents/guardians with a detailed summary of their child’s RICAS performance on ELA and mathematics and to enable comparisons with other students at the school, district, and state levels.

The front page of the report provides student identification information, including student name, grade, date of birth, Student ID (SASID), school name, and district name.

The front page also presents general information about the test, website information for parent/guardian resources, and a summary of the student’s results for each content area. This summary provides important information for each content area at a glance, including the student’s achievement level, scaled score, possible range of scores based on the standard error of measurement, and student growth percentile.

The back page of the report is provided by RIDE and contains contextual information about the educational mission and strategic roadmap of the department.

The inside pages of each report contain the achievement level, scaled score, and standard error of the scaled score for each content area. If the student does not receive a scaled score, the reason is displayed after “Your Child’s Achievement Level.” Each achievement level has its own distinct color, and that color is used throughout the report to highlight important report elements based on the student’s

achievement level and score. These report elements include the student's earned achievement level, scaled score, the visual scale's achievement-level title and achievement-level cut scores, and the comparison of the student's scaled score to the average scaled score at the student's school, district, and the state levels.

If the student received a score previously, their earned scaled score from that year's test is also displayed along with the current year scaled score for each content area tested. The previous scaled score is displayed in the color corresponding to the achievement level earned that year. If available, up to 3 years of scores including the current year score is displayed in a table. A student growth percentile (SGP) for each content area tested is displayed, if available, with a comparison to the average SGP for the student's school and district. An SGP describes the student's learning over time compared to his or her academic peers (peers are other students with similar scores on previous state tests).

The student's performance in each content area's reporting categories is also displayed using pictographs and text that indicates the points earned by the student versus the total points possible in that reporting category. For each reporting category, the average number of points earned by students scoring close to 500 (low meeting expectations) is also displayed for comparison purposes. The student's performance on individual test questions is reported at the bottom of the results page in a simplified item response grid. The grid indicates the points earned and points possible for each test question included on the grid. Essay questions are indicated on the grid.

## 5.3 REPORTING BUSINESS REQUIREMENTS

To ensure that RICAS results are processed and reported accurately, a document specifying business requirements is prepared before reporting results. The business requirements are adhered to in the processing and analyses of the RICAS test data and in preparation of the reporting results. These rules specify which, if any, student data needs to be excluded from school-, district-, and state-level summary computations. At an individual student level, the business requirements document describes how any special cases should be treated for reporting purposes. Each year the document is reviewed and approved by RIDE.

## 5.4 QUALITY ASSURANCE

Quality assurance measures are implemented throughout the process of analysis and reporting at Cognia. The data processors and data analysts perform routine quality-control checks of their computer programs. When data are handed off to different units within the data team, the sending unit verifies that the data are accurate before handoff. Additionally, when a unit receives a data set, the first step is to verify the accuracy of the data. Once new report designs were approved by RIDE, reports were run using the previous year's data to test the application of the approved current year business rules (see Appendix F). The populated reports were then reviewed and approved by RIDE.

Another type of quality assurance measure used at Cognia is parallel processing. One data analyst is responsible for writing all programs required to populate the student-level and aggregate reporting tables for the administration. Each reporting table is assigned to a second data analyst who uses the decision rules to independently program the reporting table. The production and quality-assurance tables are compared; when there is 100% agreement, the tables are released for report generation.

The third aspect of quality control involves procedures to check the accuracy of reported data. Using a sample of schools and districts, the quality assurance group verifies that the reported information is

correct. There are two sets of samples selected that may not be mutually exclusive. The first set includes samples that satisfy all the following criteria:

- one-school district
- two-school district
- multi-school district
- private school
- special school (e.g., a charter school)
- small school that does not have enough students to report aggregations
- school with excluded (not tested) students

The second set of samples includes districts or schools that have unique reporting situations that require the implementation of a decision rule. This set is necessary to ensure that each rule is applied correctly.

The quality-assurance group uses a checklist to implement its procedures. Once the checklist is completed, sample reports are circulated for review by psychometric and program management staff. The appropriate sample reports are then sent to RIDE for review and signoff.

## 5.5 ADDITIONAL RESOURCES

In addition to the resources provided within the score reports, RIDE provides online resources to assist students, families, teachers, administrators, and the public to interpret the meanings of test scores and apply their interpretations toward appropriate and valid uses of the test results. Most of these resources are available through web pages linked to the RICAS Assessments page of the RIDE website (<https://www.ride.ri.gov/InstructionAssessment/Assessment/RICASAssessments.aspx>). Stakeholder-specific resources are described in further detail below.

### 5.5.1 Students and Their Families

For students and their families, a “Resources for Families” page (<https://www.ride.ri.gov/InstructionAssessment/Assessment/ResourcesforFamilies.aspx>) provides general information about Rhode Island’s content standards and the RICAS program. Links to more detailed resources are found throughout the page and include:

- An Assessment Results page (<https://www.ride.ri.gov/Assessment-Results>), which provides access to the Rhode Island Assessment Data Portal and guides its use
- A Family Guide to Understanding RICAS (<https://ride.ri.gov/media/34886/download>), which provides guidance not only for properly interpreting RICAS results but also for appropriate use of the results
- A RICAS Individual Student Report (ISR) Guide (<https://ride.ri.gov/media/34876/download>), which explains the different components of the ISR each student receives, how to interpret them, and how to use them to work with teachers to help their child succeed

### 5.5.2 Educators and Administrators

The Assessment Results page of the RIDE Website (<https://www.ride.ri.gov/Assessment-Results>) provides Educator and Administrator Access to the Rhode Island Assessment Data Portal, providing data

at the state, district, school, grade, and student levels. There are also short videos that walk educators through the Student Data Portal and explain each of the data elements.

A Student Data Portal User Guide (<https://www3.ride.ri.gov/StudentDataPortal/docs/UserGuide.pdf>) is linked on this page, describing types of data and reports that are available, guidance for interpreting and using these data and reports and descriptions of access and permissions for different user types.

## 5.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING REPORTING

1.5 **Utilization Inference 1:** RICAS score reports provide students and their families with classification and score information that is useful, fair, and appropriate for monitoring academic achievement and participating in decisions regarding student learning.

1.5.1 **Claim:** Students and their families understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to monitoring academic achievement and participating in decisions regarding student learning.

**Evidence:** Chapter 5 describes how results are reported to students and their families, section 5.2 describes details of the information included in the score reports. This includes important score and classification information, but also explanations of what this information means. Section 5.5 and subsection 5.5.1 describe additional resources that students and their families may use to improve their understanding of this score information.

1.5.2 **Claim:** Interpretations of scores and classifications are genuinely useful to students and their families for the purposes of monitoring academic achievement and participating in decisions regarding their learning.

**Evidence:** Section 5.2 describes details of the information included in the score reports. This includes information about how families can help improve their child's learning. Section 5.5 and subsection 5.5.1 describe resources available to students and families that can be used to apply test results to take appropriate actions toward furthering the student's education.

1.6 **Utilization Inference 2:** RICAS score reports provide educators with classification and score information that is useful, fair, and appropriate for supporting curricular planning and identifying instructional needs at both the classroom and individual student level.

1.6.1 **Claim:** Educators understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to curricular planning and identification of instructional needs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe the reporting tools that educators may use to access the score results of individual students and the group of students that they teach, as well as resources available to educators providing guidance for accurately interpreting scores.

1.6.2 **Claim:** Interpretations of scores and classifications are genuinely useful to educators for the purposes of curricular planning and identification of instructional needs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe resources available to educators that provide guidance for applying test scores and interpretations of test scores to their instruction.

1.7 **Utilization Inference 3:** RICAS score reports provide school- and district-level administrators with classification and score information that is useful, fair, and appropriate for supporting program evaluations and improvements at school and district levels.

1.7.1 **Claim:** School and district-level administrators understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to program evaluations and improvements at school and district levels.

**Evidence:** Section 5.5 and subsection 5.5.2 describe the reporting tools that administrators may use to access the score results of individual students and group-level data of students in their schools and districts, as well as resources available to administrators providing guidance for accurately interpreting scores.

1.7.2 **Claim:** Interpretations of scores and classifications are genuinely useful to school- and district-level administrators for the purposes of program evaluations and improvements.

**Evidence:** Section 5.5 and subsection 5.5.2 describe resources available to administrators that provide guidance for applying test scores and interpretations of test scores to program evaluation and improvement.

1.8 **Utilization Inference 4:** RICAS score reports provide state administrators with classification and score information that is useful, fair, and appropriate for monitoring academic achievement and growth as required by state accountability programs and informing the public of schools' performances on these metrics.

1.8.1 **Claim:** State and federal administrators understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to monitoring academic achievement and growth as required by state and federal accountability programs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe the reporting tools that administrators may use to access the score results of individual students and group-level data of students in schools, districts, and the state, as well as resources available to administrators providing guidance for accurately interpreting scores.

1.8.2 **Claim:** Interpretations of scores and classifications are genuinely useful to state and federal administrators for the purposes of monitoring academic achievement and growth as required by state and federal accountability programs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe resources available to administrators that provide guidance for applying test scores and interpretations of test scores to federal accountability programs.

# Chapter 6. Classical Item Analysis

There were no substantial changes to classical item analysis procedures in current versus previous years. Interpretations of differences using classical item analyses are always difficult because statistics are population dependent. The disruptions due to COVID-19 and the non-uniform instructional delivery make comparison of aggregated classical test statistics to SY 2020-2021 especially inappropriate for the purposes of quantifying the differences between testing populations. However, it is still appropriate to use classical item statistics to flag items for potential issues in item quality, especially as these issues are further investigated by content experts for additional analysis.

A complete evaluation of a test's quality must include an evaluation of each item. Items should predominantly assess the knowledge and skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. Items should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students—in particular, racial, ethnic, or gender groups (AERA et al., 2014).

Both qualitative and quantitative analyses have been conducted to ensure that 2024 RICAS items meet these standards. This section presents statistical evaluations in four parts: (1) difficulty indices, (2) item-test correlations, (3) DIF statistics, and (4) dimensionality analyses. The item analyses presented here are based on the statewide administration of the RICAS assessments in spring 2024. Note that the information presented in this section is based only on the operational items, since student scores are calculated on those items.

## 6.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

All selected-response and constructed-response items are evaluated in terms of item difficulty according to standard classical test theory (CTT) practices. Difficulty is commonly defined as the average proportion of points achieved on an item and is measured by obtaining the average score on an item and dividing it by the maximum possible score for the item.

Selected-response items are scored dichotomously (correct vs. incorrect), so, for these items, the difficulty index is simply the proportion of students who correctly answered the item. Constructed-response items and essay items are scored polytomously, meaning that a student can achieve scores other than just 0 or 1 (e.g., 0, 1, 2, 3, or 4 for a 4-point constructed-response item). By computing the difficulty index as the average proportion of points achieved, the indices for the different item types are placed on a similar scale, ranging from 0.0 to 1.0 regardless of the item type.

Although this index is traditionally described as a measure of difficulty, it is properly interpreted as an easiness index, because larger values indicate easier items. An index of 0.0 indicates that all students earned 0% of the item points, and an index of 1.0 indicates that all students received full credit for the item (i.e., all the item points).

A summary of the distributions of item difficulty and item discrimination statistics for each grade and content area combination is presented in Table 6-1. Note that these are presented in the aggregate for all items combined as well as separately by item type: selected-response (SR), open response (OR), and essay (ES). The mean difficulty and discrimination values as well as their standard deviations shown in the table are within generally acceptable and expected ranges. Note that an “item” is defined as a



scorable opportunity for psychometric purposes. For example, each trait is treated as a separate item for an essay scored on multiple traits.

**Table 6-1 Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade**

Content Area	Grade	Item Type	Number of Items	Difficulty		Discrimination		
				Mean	Standard Deviation	Mean	Standard Deviation	
ELA	3	ALL	32	0.56	0.15	0.44	0.10	
		SR	24	0.59	0.13	0.41	0.08	
		OR	6	0.53	0.17	0.46	0.09	
		ES	2	0.31	0.04	0.69	0.04	
	4	ALL	32	0.56	0.14	0.41	0.12	
		SR	24	0.59	0.13	0.37	0.09	
		OR	6	0.52	0.11	0.51	0.10	
		ES	2	0.31	0.10	0.60	0.06	
	5	ALL	33	0.60	0.14	0.48	0.13	
		SR	24	0.64	0.12	0.42	0.08	
		OR	5	0.56	0.13	0.56	0.04	
		ES	4	0.40	0.08	0.72	0.01	
	6	ALL	33	0.56	0.16	0.45	0.13	
		SR	24	0.60	0.14	0.41	0.06	
		OR	5	0.50	0.11	0.42	0.10	
		ES	4	0.34	0.13	0.73	0.06	
	7	ALL	33	0.55	0.14	0.45	0.14	
		SR	24	0.58	0.13	0.40	0.07	
		OR	5	0.47	0.12	0.43	0.08	
		ES	4	0.47	0.12	0.77	0.02	
	8	ALL	33	0.56	0.11	0.47	0.15	
		SR	24	0.59	0.09	0.41	0.09	
		OR	5	0.50	0.09	0.53	0.08	
		ES	4	0.44	0.11	0.80	0.02	
	Mathematics	3	ALL	40	0.57	0.18	0.47	0.11
			SR	15	0.63	0.14	0.43	0.07
			OR	25	0.54	0.19	0.50	0.12
		4	ALL	40	0.55	0.16	0.47	0.11
SR			17	0.56	0.17	0.40	0.08	
OR			23	0.55	0.15	0.52	0.10	
5		ALL	40	0.53	0.17	0.45	0.12	
		SR	17	0.57	0.16	0.40	0.09	
		OR	23	0.50	0.17	0.49	0.13	
6		ALL	40	0.47	0.17	0.45	0.13	
		SR	17	0.50	0.18	0.36	0.09	
		OR	23	0.46	0.17	0.52	0.12	
7		ALL	40	0.41	0.13	0.47	0.15	
		SR	15	0.46	0.09	0.38	0.11	
		OR	25	0.37	0.14	0.52	0.15	
8		ALL	40	0.46	0.14	0.48	0.13	
		SR	22	0.51	0.12	0.41	0.10	
		OR	18	0.40	0.14	0.56	0.11	

Caution should be exercised when comparing indices across grade levels for the purpose of comparing students in different grade levels and content areas. Differences may be due not only to differences in the item difficulties on the test but also may be affected by differences in student abilities and/or differences in the standards and/or curricula taught in each grade. It is reasonable to compare the indices to common benchmarks in the field for the purpose of confirming the items meet industry recognized standards of quality.

Difficulty indices for selected-response items tend to be higher (indicating that students performed better on these items) than the difficulty indices for open-response items because selected-response items can be answered correctly by simply identifying rather than providing the correct answer, or by guessing. Similarly, discrimination indices for those open-response items with more than two points tend to be larger

than those for dichotomous items because of the greater variability in score points of the former (i.e., the partial credit these items allow). The restriction of range (i.e., only two score categories) in dichotomous items tends to make the discrimination indices lower. Note that these patterns are more consistent within item type, so when interpreting classical item statistics, comparisons should be emphasized among items of the same type.

In addition to the item difficulty and discrimination summaries presented above, item-level CTT statistics for all items, on which the distributions in Table 6-1 are based, are provided in Appendix G. Furthermore, item-level score point distributions are provided in Appendix H; for each item, the percentage of students who received each score point is presented.

As with Table 6-1, the individual item difficulty and discrimination indices are within generally acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that students who performed well on individual items tended to perform well overall.

There were only a few items with low discrimination values below 0.20, or very high or very low item difficulty values included on the 2024 RICAS tests. These items were included because their statistical values did not negatively impact the quality of the tests, and their inclusion ensured that content specifications were appropriately covered.

## 6.2 DIFFERENTIAL ITEM FUNCTIONING

Subgroup differences in performance should be examined when sample sizes permit and actions should be taken to ensure that differences in performance are attributable to construct-relevant, rather than irrelevant, factors (AERA et al., 2014; Joint Committee on Testing Practices, 2004). As part of the effort to identify such problems, psychometricians evaluated the 2024 RICAS items in terms of DIF statistics. One application of the DIF statistics is to use them to evaluate item quality in the item review process.

For the 2024 RICAS, the standardization DIF procedure (Dorans & Kulick, 1986) was employed to evaluate subgroup differences, which denote significant group-level differences in performance for examinees with equivalent achievement levels on the test. The standardization DIF procedure is designed to identify items for which subgroups of interest perform differently and beyond the impact of differences in overall achievement. The DIF procedure calculates the difference in item performance for two groups of students (at a time) matched for achievement on the total test. Specifically, average item performance is calculated for students at every total score. Then an overall average is calculated, weighting the total score distribution so that it is the same for the two groups. DIF statistics were calculated for all subgroups with at least 75 students.

DIF for MCAS items is evaluated initially at the time of field-testing. When differential performance between two groups occurs on an item (i.e., a DIF index in the “low” or “high” categories, explained below), it may or may not indicate actual item bias. Consequently, all items with either high or low DIF are examined by content experts and educators to try to identify the cause. If subgroup differences in performance can be traced to differential experience such as geographical living conditions or access to technology, the inclusion of such items is reconsidered during the item review process. If content experts do not identify a source of bias on the item, the item may be eligible for operational form construction.

The main DIF index produced under the standardization procedure has a theoretical range from -1.0 to 1.0 for selected-response items and open-response items; the latter uses an adjusted index. Dorans and Holland (1993) suggested that index values between -0.05 and 0.05 denote either a negligible amount of DIF or the absence of DIF. The majority of 2024 RICAS items fell within this range. Dorans and Holland further stated that items with values between -0.10 and -0.05 and between 0.05 and 0.10 (i.e., “low” DIF)

should be inspected to ensure that no possible effect is overlooked, and that items with values outside the -0.10 to 0.10 range (i.e., “high” DIF) are more unusual and should be examined very carefully before being used operationally.

DIF analyses were conducted for all subgroups defined in the No Child Left Behind Act, for which the sample size was at least 75 students for both the focal and reference groups separately. Six subgroup comparisons were evaluated for DIF:

- male compared with female
- not ELL compared with ELL<sup>4</sup>
- not economically disadvantaged compared with economically disadvantaged
- White compared with African American or Black
- White compared with Hispanic or Latino
- students with disabilities compared with students without disabilities

After the 2024 spring administration, DIF analyses were conducted again as a post-hoc quality check based on the operational data. The tables in Appendix I present the number of items classified as either “low” or “high” DIF, in total and by group favored. Following Dorans and Holland’s recommendation, items with DIF indices between -0.10 and -0.05 and between 0.05 and 0.10 were categorized as “low” DIF, and values outside the -0.10 to 0.10 range were categorized as “high” DIF. Very few items exhibited high DIF in the operational data, which suggested that the item review that occurred after the MCAS field-testing effectively ruled out items displaying large DIF for the RICAS operational spring tests.

## 6.3 DIMENSIONALITY ANALYSIS

Because tests are constructed with multiple content area subcategories and their associated knowledge and skills, the potential exists for the invocation of multiple dimensions beyond the common primary dimension. Generally, the subcategories are highly correlated with each other; therefore, a primary dimension typically explains the majority of variance in test scores. The presence of one dominant primary dimension is the primary psychometric assumption to support the use of the unidimensional IRT models that are used for calibrating and scaling the items administered on the 2024 RICAS assessments.

The purpose of dimensionality analysis is to investigate (a) whether violation of the assumption of test unidimensionality is statistically detectable and, if so, (b) quantify the degree to which unidimensionality is violated, and (c) specify the structure of the multidimensionality. Dimensionality analyses were performed on the operational items for all RICAS test forms used during the spring 2024 administration. A total of 12 computer-based test forms were analyzed; the results for these analyses are reported below.

The dimensionality analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999). Nonparametric techniques are often preferred because they avoid strong parametric modeling assumptions while still adhering to the fundamental principles of IRT.

Both DIMTEST and DETECT methods use the estimated average conditional covariances for item pairs as their basic statistical building block. A conditional covariance is the covariance between two items conditioned on true score (expected value of observed score) for the rest of the test, and the average conditional covariance is obtained by averaging across all possible conditioning scores. When a test is

---

<sup>4</sup> ELL = English Language Learner (includes current and former English Language Learners).

strictly unidimensional, all conditional covariances are expected to take on values within random noise of zero, indicating statistically independent item responses for examinees with equal expected scores. Nonzero conditional covariances are essentially violations of the principle of local independence, and such local dependence implies multidimensionality. Thus, nonrandom patterns of positive and negative conditional covariances are indicative of multidimensionality.

DIMTEST is a hypothesis-testing procedure for detecting violations of local independence. The data are first randomly divided into a training sample and a cross-validation sample. Then an exploratory analysis of the conditional covariances is conducted on the training sample data to find the cluster of items that displays the greatest evidence of local dependence. The cross-validation sample is then used to test whether the conditional covariances of the selected cluster of items display local dependence, conditioning on total score from the nonclustered items. The DIMTEST statistic follows a standard normal distribution under the null hypothesis of unidimensionality.

DETECT is an effect-size measure of multidimensionality. As with DIMTEST, the data are first randomly divided into a training sample and a cross-validation sample (these samples are drawn independently of those used with DIMTEST). The training sample is used to find a set of mutually exclusive and collectively exhaustive clusters of items that best fit a systematic pattern of positive conditional covariances for pairs of items from the same cluster and negative conditional covariances for pairs composed of items from different clusters. Next, the clusters from the training sample are used with the cross-validation sample data to average the conditional covariances: within-cluster conditional covariances are summed; from this sum, the between-cluster conditional covariances are subtracted. This difference is divided by the total number of item pairs, and this average is multiplied by 100 to yield an index of the average violation of local independence for an item pair. DETECT values less than 0.2 indicate very weak multidimensionality (or near unidimensionality); values of 0.2 to 0.4, weak to moderate multidimensionality; values of 0.4 to 1.0, moderate to strong multidimensionality; and values greater than 1.0, very strong multidimensionality (Roussos & Ozbek, 2006).

DIMTEST and DETECT were applied to the operational items of the RICAS tests administered during spring 2024. The data for each grade were split into a training sample and a cross-validation sample. For all grades, there were over 9,400 student examinees per test form in both ELA and mathematics, so every training sample and cross-validation sample had at least 4,700 students. After randomly splitting the data into training and cross-validation samples, DIMTEST was applied to each data set to see if the null hypothesis of unidimensionality would be rejected. DETECT was then applied to each data set for which the DIMTEST null hypothesis was rejected to estimate the effect size of the multidimensionality.

The results of the DIMTEST analyses indicated that the null hypothesis was rejected at a significance level of 0.01 for every data set. Because strict unidimensionality is an idealization that almost never holds exactly for a given data set, the statistical rejections in the DIMTEST results were not surprising. Indeed, because of the large sample sizes involved in the data sets, DIMTEST would be expected to be sensitive to even quite small violations of unidimensionality.

DETECT was then used to estimate the effect size for the violations of local independence for the 2024 tests. Table 6-2 displays the multidimensionality effect-size estimates from DETECT across three operational years.

**Table 6-2 Multidimensionality Effect Sizes by Grade and Content Area**

Content Area	Grade	Multidimensionality Effect Size		
		2022	2023	2024
ELA	3	0.15	0.14	0.18
	4	0.21	0.17	0.21
	5	0.21	0.17	0.37
	6	0.25	0.22	0.34
	7	0.30	0.34	0.43
	8	0.30	0.32	0.44
	<b>Average</b>	<b>0.24</b>	<b>0.23</b>	<b>0.33</b>
Mathematics	3	0.20	0.17	0.19
	4	0.18	0.14	0.13
	5	0.18	0.10	0.09
	6	0.13	0.15	0.15
	7	0.13	0.13	0.11
	8	0.18	0.14	0.17
	<b>Average</b>	<b>0.17</b>	<b>0.14</b>	<b>0.14</b>

The DETECT values indicate weak ( $0.2 < \text{DETECT} < 0.4$ ) or very weak ( $\text{DETECT} < 0.2$ ) multidimensionality for all the 2024 RICAS test forms, except for ELA grades 7 and 8 ( $\text{DETECT} = 0.43\text{-}0.44$ ) indicating low-end moderate multidimensionality with more dimensionality distinctiveness of writing prompt items relative to the rest of the test. The patterns are consistent with the previous year’s results.

The way in which DETECT divided the tests into clusters was investigated to determine whether there were any discernible patterns with respect to the selected-response and open-response item types. Inspection of the DETECT clusters indicated that selected-response/open-response separation generally occurred much more strongly with ELA than with mathematics, a pattern that has been consistent across all previous years. Specifically, for the ELA test forms, every grade had one set of clusters dominated by selected-response items and another set of clusters dominated by writing prompt items. On the mathematics test forms, there was less clear evidence of consistent separation of selected-response and open-response items.

In summary, for the 2024 dimensionality analyses, the violations of local independence, as evidenced by the DETECT effect sizes, were either weak or very weak in all test forms. The patterns with respect to the selected-response and open-response items were consistent with those in the previous year, with ELA tending to display more separation than mathematics. However, this separation did not result in an effect size that would suggest use of a unidimensional IRT model is inappropriate.

## 6.4 QUALITY CHECK

Conducting a key verification and adjudication is an essential step to ensuring the accuracy of the scoring applied to each item. Both activities are conducted to ensure the key options and scoring rules are accurately applied to the student responses as intended. Cognia’s psychometricians and test developers collaborate to perform key verification and adjudication before psychometricians conduct any classical item analysis.

Quality check measures are implemented throughout classical item analysis at Cognia. It starts with automated psychometric software tools, such as Cognia’s Measurement Portal (“the Portal”) and dimensionality software (i.e., DIMTEST and DETECT). These tools play a vital role in standardizing the processes for conducting psychometric analyses, including classical item analysis and dimensionality analysis. All software has been rigorously checked throughout its development cycle to ensure accuracy.

A large part of ensuring the accuracy of classical analysis is applying the correct inclusion rules to the dataset used in the calculations. Psychometricians perform routine quality-control checks of the specifications before running the CTT analyses to ensure business rules are followed.

After all analyses are run, psychometricians review the results in detail for reasonableness. The results are primarily compared to those of previous years. Additionally, psychometricians review the results to see if the interpretations are the same as last year. If any changes are identified, the analyses are run numerous times to check for accuracy, and the lead psychometrician consults with other psychometricians to gain insights into the observed changes.

## 6.5 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING CLASSICAL ITEM ANALYSES

1.2 **Evaluation Inference:** Each test form, an organized sampling of assessment tasks, results in an observed score that reflects a student's knowledge and abilities in the content area being assessed through appropriate test assembly, administration, and scoring procedures.

1.2.4 **Claim:** Items on the assessment demonstrate appropriate statistical quality.

**Evidence:** Chapter 6 describes the classical item analysis procedures conducted to ensure that all items meet the standards of quality outlined by the Standards (AERA et al., 2014) and Code of Fair Testing Practices in Education (Joint Committee on Testing Practices, 2004). Differential Item Functioning (DIF) analysis, presented in section 6.2, provides evidence that the items are free of systematic biases.

1.3 **Generalization Inference:** The observed score from any specific form testing a given grade and content area is reflective of the expected score on any potential form of the test for that grade and content area.

1.3.2 **Claim:** Test specifications result in forms of similar length and task distribution.

**Evidence:** Dimensionality analyses, presented in section 6.3, provide evidence that any differences in length or task distribution are small enough that interpretation of the resulting scores is preserved.

1.3.3 **Claim:** Statistical analyses of observed scores on specific forms show that they are good predictors of expected scores on other potential forms.

**Evidence:** DIF analysis and subsequent review of items classified as exhibiting DIF, described in section 6.2, support observed score generalization to expected score by ruling out the items specific to this year's forms as sources of bias in the scores.

# Chapter 7. Item Response Theory Analysis

## 7.1 OVERVIEW

As reported in Chapter 1, RICAS uses the Massachusetts MCAS ELA and mathematics items and test forms. The IRT processes used to link and scale MCAS administrations are managed by DESE and Cognia and are leveraged by RIDE in the sense that the results of those processes are used to establish the RICAS IRT model and subsequent reporting scale.

Chapter 7 describes the procedures used to calibrate, equate, and scale the MCAS tests. During these psychometric analyses, several quality-control procedures and checks on the processes were conducted. These procedures included

- evaluations of the calibration processes (e.g., checking the number of cycles required for convergence for reasonableness);
- checking item parameters and their standard errors for reasonableness;
- examination of test characteristic curves (TCCs) and test information function (TIF) curves for reasonableness;
- evaluation of model fit (e.g., test level, item-level, and person-level);
- evaluation of equating items (e.g., delta analyses, *b-b* analyses, beta analyses);
- examination of a-plots and b-plots for reasonableness; and
- evaluation of the scaling results (e.g., comparing look-up tables to the previous year's).

Chapter 7 is reprinted with minor modifications from the MCAS technical report to provide added clarity within this document.

## 7.2 IRT

All RICAS items were calibrated applying IRT on the MCAS data. IRT uses mathematical models to define a relationship between an unobserved measure of student performance, usually referred to as theta ( $\theta$ ), and the probability [ $P(\theta)$ ] of getting a dichotomous item correct or of getting a particular score on a polytomous item (Hambleton, Swaminathan, & Rogers, 1991; Hambleton & Swaminathan, 1985). In IRT, it is assumed that all items are independent measures of the same construct (i.e., of the same  $\theta$ ). Another way to think of  $\theta$  is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between  $\theta$  and  $P(\theta)$  (van der Linden, 2016; Hambleton & van der Linden, 1997; Hambleton & Swaminathan, 1985). The process of determining the mathematical relationship between  $\theta$  and  $P(\theta)$  is called *item calibration*. After items are calibrated, they are defined by a set of parameters that specify a nonlinear, monotonically increasing relationship between  $\theta$  and  $P(\theta)$ . Once the item parameters are known, an estimate of  $\theta$  for each student can be calculated. This estimate  $\hat{\theta}$  is considered to be an estimate of the student's true score or a general representation of

student performance. IRT has characteristics that may be preferable to those of raw scores for equating purposes because it specifically models examinee responses at the item level and facilitates equating to an IRT-based item pool (Kolen & Brennan, 2014).

For the 2024 RICAS tests, the three-parameter logistic (3PL) model was used for traditional four-option multiple-choice items, and the two-parameter logistic (2PL) model was used for binary-scored open-response and technology-enhanced items (Hambleton & van der Linden, 1997; Hambleton, Swaminathan, & Rogers, 1991). The graded-response model (GRM) was used for polytomous items (Nering & Ostini, 2010), including polytomously scored multi-part items, open-response items, and writing prompts.

The 3PL model for multiple-choice items can be defined as:

$$P_i(\theta_j) = P(U_i = 1|\theta_j) = c_i + (1 - c_i) \frac{\exp[D\alpha_i(\theta_j - b_i)]}{1 + \exp[D\alpha_i(\theta_j - b_i)]}, \quad (\text{Equation 1})$$

where

$U$  represents the scored response on an item,  
 $i$  indexes the items,  
 $j$  indexes students,  
 $\alpha$  represents item discrimination,  
 $b$  represents item difficulty,  
 $c$  is the pseudo guessing parameter,  
 $\theta$  is the student's latent person parameter, and  
 $D$  is a normalizing constant equal to 1.701.

For the 2PL model, this equation reduces to the following:

$$P_i(\theta_j) = P(U_i = 1|\theta_j) = \frac{\exp[D\alpha_i(\theta_j - b_i)]}{1 + \exp[D\alpha_i(\theta_j - b_i)]}. \quad (\text{Equation 2})$$

In the GRM for polytomous items, an item is scored in  $(k + 1)$  graded categories that can be viewed as a set of  $k$  dichotomies. At each point of dichotomization (i.e., at each threshold), a two-parameter model can be used to model the probability that a student's response falls at or above a particular ordered category, given  $\theta$ . This implies that a polytomous item with  $(k + 1)$  categories can be characterized by  $k$  item category threshold curves (ICTCs) of the 2PL form:

$$P_{ik}^*(\theta_j) = P(\theta_j) = \frac{\exp[D\alpha_i(\theta_j - b_i + d_{ik})]}{1 + \exp[D\alpha_i(\theta_j - b_i + d_{ik})]}, \quad (\text{Equation 3})$$

where

$U$  indexes the scored response on an item,  
 $i$  indexes the items,  
 $j$  indexes students,



$k$  indexes threshold,  
 $\theta$  is the student's latent person parameter,  
 $\alpha$  represents item discrimination,  
 $b$  represents item difficulty,  
 $d$  represents threshold, and  
 $D$  is a normalizing constant equal to 1.701.

After computing  $k$  ICTCs in the GRM,  $(k + 1)$  item category characteristic curves (ICCCs), which indicate the probability of obtaining a score assigned to a particular category given  $\theta$ , are derived by subtracting adjacent ICTCs:

$$P_{ik}(\theta_j) = P(\theta_j) = P_{ik}^*(\theta_j) - P_{i(k+1)}^*(\theta_j), \quad (\text{Equation 4})$$

where

$i$  indexes the items,  
 $j$  indexes students,  
 $k$  indexes threshold,  
 $\theta$  is the student ability,  
 $P_{ik}$  represents the probability that the score on item  $i$  falls in category  $k$ , and  
 $P_{ik}^*$  represents the probability that the score on item  $i$  falls at or above the threshold  $k$   
 $(P_{i0}^* = 1$  and  $P_{i(m+1)}^* = 0)$ .

The GRM is also commonly expressed as:

$$P_{ik}(\theta_j) = \frac{\exp[D\alpha_i(\theta_j - b_i + d_k)]}{1 + \exp[D\alpha_i(\theta_j - b_i + d_k)]} - \frac{\exp[D\alpha_i(\theta_j - b_i + d_{k+1})]}{1 + \exp[D\alpha_i(\theta_j - b_i + d_{k+1})]}. \quad (\text{Equation 5})$$

Finally, the item characteristic curve (ICC) for a polytomous item is computed as a weighted sum of ICCCs, where each ICCC is weighted by a score assigned to a corresponding category. The expected score for a student with a given theta,  $\theta_j$ , is expressed as:

$$E(U_i|\theta_j) = \sum_k^{m+1} w_{ik} P_{ik}(\theta_j), \quad (\text{Equation 6})$$

where  $w_{ik}$  is the weighting constant and is equal to the number of score points for score category  $k$  on item  $i$ .

Note that for a dichotomously scored item,  $E(U_i|\theta_j) = P_i(\theta_j)$ . For more information about item calibration and determination, see Lord and Novick (1968), Hambleton and Swaminathan (1985), or Baker and Kim (2004).

TCCs display the expected (average) raw score associated with each  $\theta_j$  value typically between -4.0 and 4.0. Mathematically, the TCC is computed by summing the ICCs of all items that contribute to the raw

score. Using the notation introduced earlier in this section, the expected raw score at a given value of  $\theta_j$  is as follows:

$$E(X|\theta_j) = \sum_{i=1}^n E(U_i|\theta_j), \quad (\text{Equation 7})$$

where

$i$  indexes the items (and  $n$  is the number of items contributing to the raw score),

$j$  indexes students (here,  $\theta_j$  runs from -4 to 4), and

$E(X|\theta_j)$  is the expected raw score for a student of ability  $\theta_j$ .

The expected raw score monotonically increases with  $\theta_j$ , consistent with the notion that students of high ability tend to earn higher raw scores than students of low ability. Most TCCs are “S-shaped”: they are flatter at the ends of the distribution and steeper in the middle.

The TIF displays the amount of statistical information that the test provides at each value of  $\theta_j$ .

Information functions depict test precision across the entire latent trait continuum. There is an inverse relationship between the information of a test and its standard error of measurement (SEM). For long tests, the SEM at a given  $\theta_j$  is approximately equal to the inverse of the square root of the statistical information at  $\theta_j$  (Hambleton, Swaminathan, & Rogers, 1991), as follows:

$$SEM(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}. \quad (\text{Equation 8})$$

Compared to the tails, TIF values are often higher near the middle of the  $\theta$  distribution where most students are located. This is by design. Test items are often selected with middle difficulty levels and high discriminating powers so that test information is maximized for the majority of candidates who are expected to take a test.

## 7.3 IRT RESULTS

IRT calibration was conducted using flexMIRT 3.03 (Cai, 2012) on the CBT items in all the grades. Because paper test forms are treated as accommodated forms, item parameters for computer-based items were applied to their paper counterparts. The tables in Appendix J provide the IRT item parameters and associated standard errors of all operational scoring items on the 2024 RICAS tests. The MCAS equating report in Appendix J contains graphs of the TCCs and TIFs, which are defined in the previous section. While the information provided in Appendix J pertains solely to the MCAS equating process, it should be noted that the RICAS assessment program utilized the MCAS equating results to report the student scores. The RICAS achievement level distributions are available in Appendix D.

The number of cycles required for convergence for each grade and content area during the IRT analysis can be found in Table 7-1. The number of cycles required for convergence fell within acceptable ranges (less than 150) for all tests.

**Table 7-1 Number of Cycles Required for Convergence**

Content Area	Grade	Initial Cycles	FCIP Cycles
ELA	3	26	7
	4	62	7
	5	46	10
	6	34	12
	7	38	9
	8	30	9
Mathematics	3	63	--
	4	68	--
	5	58	--
	6	39	--
	7	76	--
	8	45	--

## 7.4 EQUATING

Section 7.4 summarizes the equating procedure and results to place the 2024 MCAS tests on the same scale as the previous year. An equating report provides complete documentation of the quality-control procedures and results of the 2024 MCAS equating (Appendix J).

The purpose of equating is to ensure that scores obtained from different forms of a test are comparable to one another. Equating may be used if multiple test forms are administered in the same year; or one year’s forms may be equated to those used in the previous year. Equating ensures that students are not given an unfair advantage or disadvantage because the test form they took is easier or harder than that taken by other students. See Chapter 2 for more information about how the test development process supports successful equating. To call out an example, Cognia test developers and psychometricians closely collaborated to ensure that the constructed forms are representative from both content and statistical perspectives. Specifically, the Content team strived to meet the content coverage stipulated in the test design and blueprints while considering the item difficulty and complexity. Then, psychometricians evaluated the forms to ensure that the proposed forms were statistically comparable to the reference form (typically, the previous year’s operational form).

### General Equating Approach

For RICAS, the raw-to-scaled score lookup tables are produced using the on-scale IRT parameters from the MCAS bank. Hence, it is crucial to understand the equating procedure used in maintaining the MCAS item bank.

The 2024 administration of the MCAS used a raw-score-to-theta equating procedure in which test forms were equated to the theta scale established on the reference form (i.e., the form used in the most recent standard setting). The groups of students who take equating items on the different test forms are never strictly equivalent to the groups who took the tests in the reference years. IRT is particularly useful for equating scenarios that involve nonequivalent groups (Allen & Yen, 1979). Equating for the MCAS uses the anchor test–nonequivalent groups design described by Petersen, Kolen, and Hoover (1989). In this equating design, no assumption is made about the equivalence of the examinee groups taking different test forms (i.e., naturally occurring groups are assumed). Comparability is instead evaluated by using a set of anchor items (also called equating items), assuming they perform in the same way in both groups and thus can accurately measure the differences in the two groups.

For mathematics, the item parameter estimates for 2024 test forms were placed on the reference scale using the Stocking-Lord method (SL; Stocking & Lord, 1983). The estimates of the item parameters for the anchor items were used to estimate the SL transformation. The SL method estimates the combination

of slope and intercept values that minimize the squared difference between the test characteristic curves cumulated over the anchor items. Then, the estimated SL constants were applied to linearly transform the freely calibrated parameter estimates to put them on the reference scale.

However, a two-step equating approach was taken for ELA because of the presence of the writing prompts. The first step for ELA involved applying the SL method for all items except the writing prompt items; thus, isolating any dimensionality variability in the writing prompt items from the estimation of the equating relationship across years. Then, the writing prompt items were brought onto the scale established in the first step by applying the fixed common item parameters (FCIP2; Kim, 2006) calibration method. The FCIP2 method is based on the IRT principle of item parameter invariance. According to this principle, the equating items for both tests should have the same item parameters. After the item parameters for the non-writing prompt items were put on the reference scale (the first step), the FCIP2 method was employed to place the writing prompt items onto the operational scale (the second step). This method is performed by fixing the parameters of the “equating” items (in this case, all non-writing prompt items) to their previously obtained on-scale values and then calibrating using flexMIRT to place the remaining items (in this case, the writing prompt items) on scale.

### Parameter Drift Evaluations

Prior to implementing the SL method, two evaluations of the equating items were conducted to check for parameter drift, as follows.

- **Delta method:** compares two years’ delta values (the percent correct transformed into a scale “with an effective range of 6 [very easy item] to 20 [very difficult item]”<sup>5</sup>) for equating items and flags an item if its standardized distance to the principal axis line is at or above 3 in absolute value.
- **b-b method:** compares current year’s freely estimated IRT difficulty parameters with the previous year’s values for equating items, and flags an item if its standardized distance to the principal axis line is at or above 3 in absolute value.

During the implementation of the SL method, a third evaluation of the equating items was conducted to check for parameter drift, as follows.

- **IRT curve-based beta method:** a measure of the weighted average difference between the item response function (IRF) curves between two years for each equating item (Jiang, Roussos, & Yu, 2017; Wang & Roussos, 2018). The current year’s IRF is calculated based on transformed item parameters using the SL constants estimated with all equating items. The difference index is denoted as  $\beta$ , its estimate is denoted as  $\hat{\beta}$ . Mathematically, it can be expressed as  $\beta = \int (P(\theta, R) - P(\theta, F)) f_F(\theta) d\theta$ , where  $P(\theta, R)$  and  $P(\theta, F)$  indicate the IRFs for the reference (e.g., previous administration) and focal (e.g., current year) groups, respectively, and  $f_F(\theta)$  is the density function for  $\theta$  in the focal group. The following threshold is used to categorize an item into negligible, moderate, or large drift:
  - $|\hat{\beta}| < 0.05$ , negligible drift
  - $0.05 \leq |\hat{\beta}| < 0.1$ , moderate drift
  - $|\hat{\beta}| \geq 0.1$ , large drift

Detailed results from each drift analysis, along with Delta and *b*-plots are presented in Appendix J.

---

<sup>5</sup> Walker, M. E. (2014, May 13). *Enhancing the Equating of Item Difficulty Metrics: Estimation of Reference Distribution*. ETS Research Report Series. P. 1. Retrieved 1.10.20 from: <https://onlinelibrary.wiley.com/doi/full/10.1002/ets2.12006>

## Content Review

Following the statistical evaluation, each of the flagged items went through a content review process to further investigate whether there are construct-irrelevant or relevant factors that may have resulted in the item parameter drift. Anything pertaining to the content being measured is considered a construct relevant factor, such as any instructional shift in certain content areas. A list of content irrelevant factors follows:

- changes to item administration mode
- word/graphic changes to any part of the item
- change to option order
- change in position (e.g., beginning of test vs. end of test)
- whether an item experiences “clueing” in one administration but not in the other
- whether there are test security risks associated with the flagged items
- any other difference that may affect the testing experience.

An item is removed from the equating set if a construct-irrelevant reason is identified in the content review. If a content relevant reason is identified, an item is kept as an equating item. If the content review does not find any reason, an item is kept in the equating set.

The equating items that remained following these evaluation procedures were then employed in the SL method, and the linking relationship obtained from the SL method was used to transform the item parameters for all items in the 2024 MCAS administration onto the target scale. The transformed item parameters were then used to build the raw score to theta look-up tables for the 2024 RICAS tests. The SL constants are presented in Table 7-2.

**Table 7-2 Stocking and Lord Constants**

Content Area	Grade	Slope	Intercept
ELA	3	1.12	-0.22
	4	1.06	-0.31
	5	1.16	-0.28
	6	1.48	-0.31
	7	1.29	-0.39
	8	1.40	-0.22
Mathematics	3	1.09	0.05
	4	1.05	0.16
	5	0.99	-0.01
	6	1.06	-0.04
	7	1.11	-0.11
	8	1.10	-0.13

## 7.5 REPORTED SCALE SCORES AND ACHIEVEMENT STANDARDS

Because the  $\theta$  scale used in IRT calibrations is not understood by most stakeholders, reporting scales were developed for the MCAS ELA and mathematics tests in grades 3–8, which then were applied to RICAS. The reporting scales are linear transformations of the underlying  $\theta$  scale. As the three  $\theta$  cutpoints from the standard setting have equal intervals, one single linear transformation was sufficient to transform the  $\theta$  scale from each achievement level category on one reporting scale.

Student scores on the RICAS tests are reported in integer values from 440 to 560. Because the same transformation is applied to all achievement-level categories, and the reported scaled scores preserve the interval scale properties (except for the truncated scaled scores at the lower and upper end of the score scale), it is appropriate to calculate means and standard deviations with scaled scores.

By providing information that is more specific about the position of a student's results, scaled scores supplement achievement-level scores. Students' raw scores (i.e., total number of obtained points) on the 2024 RICAS tests were translated to scaled scores using a data analysis process called *scaling*, which simply converts from one scale to another. In the same way that a given temperature can be expressed on either the Fahrenheit or the Celsius scale, or the same distance can be expressed in either miles or kilometers, student scores on the 2024 RICAS tests can be expressed in raw or scaled scores.

It is important to note that converting from raw scores to scaled scores does not change students' achievement level classifications. Given the relative simplicity of raw scores, it is fair to question why scaled scores for the RICAS are reported instead of raw scores. The answer is that scaled scores make the reporting of results consistent. To illustrate, standard setting typically results in different raw cut scores across content areas. The raw cut score between *Partially Meeting Expectations* and *Meeting Expectations* could be, for example, 35 in grade 3 mathematics but 33 in grade 4 mathematics, yet both raw scores would be transformed to scaled scores of 500. It is this uniformity across scaled scores that facilitates the understanding of student performance. The psychometric advantage of scaled scores over raw scores comes from their being linear transformations of  $\theta$ . Since the  $\theta$  scale is used for equating, scaled scores are comparable from one year to the next. Raw scores are not.

The scaled scores are obtained by a simple translation of ability estimates ( $\hat{\theta}$ ) using the linear relationship between threshold values on the  $\theta$  metric and their equivalent values on the scaled score metric.

Students' ability estimates are obtained by mapping their raw scores through the TCC. Scale scores are calculated using the following linear equation:

$$SS = m\hat{\theta} + b, \quad (\text{Equation 9})$$

where  
 $m$  is the slope and  
 $b$  is the intercept.

A separate linear transformation is used for each grade and content area combination. Table 7-3 shows the slope and intercept terms used to calculate the scaled scores for each grade and content area. Note that the values in Table 7-3 will not change unless the standards are reset.

**Table 7-3 Scale Score Slopes and Intercepts by Content Area and Grade**

Content Area	Grade	Slope	Intercept
ELA	3	18.839	499.785
	4	18.846	499.421
	5	17.686	499.335
	6	18.984	500.202
	7	19.098	499.791
	8	19.900	498.981
Mathematics	3	21.357	499.413
	4	20.938	498.869
	5	19.039	499.525
	6	19.870	500.165
	7	20.758	499.353
	8	20.172	500.170

Massachusetts conducted standard setting activities in August 2017 to establish achievement level cut scores on the MCAS tests using standardized methods consistent with what is used in the professional field. RIDE staff and technical advisors observed those standard setting procedures and analyzed the results of the standard setting process. Although results of the MCAS tests are reported in terms of four achievement levels, *Not Meeting Expectations*, *Partially Meeting Expectations*, *Meeting Expectations*, and *Exceeding Expectations*, rather than the five levels used to report PARCC results, analyses indicate that the MCAS performance standards are consistent with and as rigorous as the PARCC performance standards previously used in Rhode Island.

Across all grade levels 3–8, results from Rhode Island and Massachusetts suggested that performance at the Meeting Expectations level on the MCAS tests (level 3) is roughly equivalent to performance at the Met Expectations level on the PARCC tests (level 4), in terms of the resulting proportions of students classified above and below those levels.

Cutpoints for grades 3–8 ELA and mathematics RICAS tests were set via standard setting in 2017 by DESE and MCAS for grades 3–8 ELA and mathematics tests (see the *2017 Next-Generation MCAS and MCAS-Alt Technical Report* for the 2017 standard setting report). The standard setting establishes the theta cutpoints used for reporting each year. These theta cuts are presented in Table 7-4. The operational cut scores will remain fixed throughout the assessment program unless standards are reset. Also shown in the table are the cutpoints on the reporting score scale.

**Table 7-4 Cut Scores on the Theta Metric and Reporting Scale by Content Area and Grade**

Content Area	Grade	Theta			Scale Score				
		Cut 1	Cut 2	Cut 3	Min	Cut 1	Cut 2	Cut 3	Max
ELA	3	-1.581	0.011	1.604	440	470	500	530	560
	4	-1.561	0.031	1.623	440	470	500	530	560
	5	-1.659	0.038	1.734	440	470	500	530	560
	6	-1.591	-0.011	1.570	440	470	500	530	560
	7	-1.560	0.011	1.582	440	470	500	530	560
	8	-1.456	0.051	1.559	440	470	500	530	560
Mathematics	3	-1.377	0.027	1.432	440	470	500	530	560
	4	-1.379	0.054	1.487	440	470	500	530	560
	5	-1.551	0.025	1.601	440	470	500	530	560
	6	-1.518	-0.008	1.502	440	470	500	530	560
	7	-1.414	0.031	1.476	440	470	500	530	560
	8	-1.496	-0.008	1.479	440	470	500	530	560

## 7.6 QUALITY CHECK

Similar to classical item analysis, quality check measures are implemented throughout item response theory (IRT) analysis at Cognia. Again, conducting a key verification and adjudication is an essential step to ensuring the accuracy of the scoring applied to each item. Both activities serve as a safeguard to ensure the key options and scoring rules are accurately applied to the student responses as intended.

An automated psychometric software tool, Cognia's Measurement Portal ("the Portal"), plays a vital role in standardizing the processes for conducting psychometric analyses, including all steps of IRT analysis. The Portal has been rigorously checked throughout its development cycle to ensure accuracy. Within the Portal, specifications are set up to apply the inclusion rules to the dataset used in the calculations and to specify the IRT procedures. Psychometricians perform routine quality-control checks of the specifications before running the IRT analyses to ensure the business rules are followed accurately.

A "copy match" of equating items is conducted to ensure the stability of the equating item performances. Copy match is performed to ensure that the equating items appear the same on the previous and current administrations. Because psychometricians rely on the equating items to behave similarly across years, copy match is a crucial step before any equating activities occur.

Several analyses are performed to check the stability of equating items during equating. These include delta, b/b, and beta analyses (details are provided in Section 7.4). In addition, a rescore analysis is performed on human-scored items to evaluate whether *scorer drift* occurred across different years (e.g., this year's scorers are generally more lenient than last year's).

After the equating, psychometricians check the reasonableness of results by comparing them to previous years and predicted values from form construction. Similar to classical analysis, all calculated values are reviewed for reasonableness and if the interpretations are the same as last year. If any changes are identified, the analyses are run numerous times to check for accuracy, and the lead psychometrician consults with other psychometricians to gain insights into the observed changes. Below is a list of checks psychometricians perform when reviewing the equating results:

- Reasonableness of item parameters and associated standard errors
- Fit plots
- Normal end (i.e., IRT software convergence) evaluation
- Delta plot
- *a-a* and *b-b* plot
- TCCs and test information functions
- Achievement levels and scaled score distributions
- Watch list (e.g., item flagged during equating item evaluation).

## 7.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING ITEM RESPONSE THEORY ANALYSES

1.2 **Evaluation Inference:** Each test form, an organized sampling of assessment tasks, results in an observed score that reflects a student's knowledge and abilities in the content area being assessed through appropriate test assembly, administration, and scoring procedures.



1.2.3 **Claim:** The scoring procedures and models produce scores accurately reflective of targeted knowledge and abilities.

**Evidence:** Section 7.2 describes the scoring models used for items on the RICAS, describing the models used in detail and citing the references that establish the appropriateness of these models for placing student performances on a common scale for scoring purposes.

1.2.4 **Claim:** Items on the assessment demonstrate appropriate statistical quality.

**Evidence:** Section 7.3 describes IRT results referring to tables within the equating report (Appendix J) that describe quality control checks on items and procedures for making interventions based on items being flagged during these checks.

1.3 **Generalization Inference:** The observed score from any specific form testing a given grade and content area is reflective of the expected score on any potential form of the test for that grade and content area.

1.3.4 **Claim:** Equating and scaling methods accurately place scores from different forms onto a common scale.

**Evidence:** Section 7.4 describes equating procedures in detail and summarizes results from the full equating report, provided in Appendix J. Section 7.5 describes the processes of applying equating and scaling results to place raw scores onto RICAS score scales. These sections demonstrate a high level of rigor in selection, application, and interpretation of equating results, placing scores from this year's forms on the same scales as forms from prior years.

# Chapter 8. Reliability

## 8.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT

Although an individual item's performance is an important factor in evaluating an assessment, a complete evaluation must also address the way items grouped in a set function as a set and complement one another. Tests that function well provide a dependable assessment of a student's level of ability. Just like the measurement of physical properties such as temperature, any measurement tool contains some amount of measurement error, which leads to different results if the measurements were taken multiple times. As the tools to measure latent ability, the quality of items determines the degree to which a given student's score can be higher or lower than his or her true ability on a test.

There are several ways to estimate an assessment's reliability. The approach that was implemented to assess the reliability of the 2024 RICAS tests is the  $\alpha$  coefficient of Cronbach (1951). This approach is most easily understood as an extension of a related procedure, split-half reliability. In the split-half approach, a test is split in half, and students' scores on the two half-tests are correlated. To estimate the correlation between two full-length tests, the Spearman-Brown correction (Spearman, 1910; Brown, 1910) is applied. If the correlation is high, this is evidence that the items complement one another and function well as a group, suggesting that measurement error is minimal. The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation since each different possible split of the test into halves will result in a different correlation.

Cronbach's  $\alpha$  eliminates the item selection impact by comparing individual item variances to total test variance, and it has been shown to be the average of all possible split-half correlations. Along with the split-half reliability, Cronbach's  $\alpha$  is referred to as a coefficient of internal consistency. The term "internal" indicates that the index is measured internal to each test of interest, using data that come only from the test itself (Anastasi & Urbina, 1997).

The formula for Cronbach's  $\alpha$  is given as follows:

$$\alpha = \frac{n}{n-1} \left[ 1 - \frac{\sum_{i=1}^n \sigma_{(Y_i)}^2}{\sigma_x^2} \right], \quad (\text{Equation 10})$$

where

$i$  indexes the item,

$n$  is the total number of items,

$\sigma_{(Y_i)}^2$  represents individual item variance, and

$\sigma_x^2$  represents the total test variance.

Table 8-1 presents descriptive statistics, Cronbach's  $\alpha$  coefficient, and the raw score standard error of measurement (SEM) for each content area and grade. Statistics are based on operational items from online test forms, which were taken by most of the student examinee population. The reliability estimates range from 0.88 to 0.93, which is a generally acceptable range.

**Table 8-1 Raw Score Descriptive Statistics, Cronbach’s Alpha, and SEMs by Content Area and Grade**

Content Area	Grade	Number of Students	Raw Score			Alpha ( $\alpha$ )	SEM
			Maximum	Mean	Standard Deviation		
ELA	3	9,400	44	23	8.89	0.89	2.90
	4	9,690	44	23	8.93	0.88	3.05
	5	9,716	48	26	10.70	0.92	3.09
	6	9,611	50	24	10.70	0.90	3.36
	7	9,737	50	26	10.77	0.90	3.33
	8	9,845	50	26	11.54	0.92	3.33
Mathematics	3	9,579	48	24	11.54	0.93	3.02
	4	9,860	54	28	12.68	0.93	3.40
	5	9,904	54	25	11.96	0.92	3.31
	6	9,763	54	22	12.19	0.92	3.43
	7	9,895	54	20	12.13	0.92	3.37
	8	9,963	54	23	12.48	0.92	3.42

Because of the dependency of the  $\alpha$  coefficients on the test-taking population and the test characteristics, precautions need to be taken when making inferences about the quality of one test by comparing its reliability to that of another test from a different grade or content area. To elaborate, reliability coefficients are highly influenced by test-taking population characteristics such as the range of individual differences in the group (i.e., variability within the population), average ability level of the population that took the exams, test designs, test difficulty, test length, ceiling or floor effect, and influence of guessing. Hence, “the reported reliability coefficient is only applicable to samples similar to that on which it was computed” (Anastasi & Urbina, 1997, p.107). It is reasonable to compare the indices to common benchmarks in the field for the purpose of confirming the tests meet similar industry recognized standards of quality.

## 8.2 SUBGROUP RELIABILITY

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2024 RICAS online forms. Appendix K presents reliabilities for various subgroups of interest for ELA and mathematics, respectively. Cronbach’s  $\alpha$  coefficients were calculated based only on the members of the subgroup in question in the computations; values are calculated only for subgroups with 10 or more students. The reliability coefficients for subgroups range from 0.77 to 0.94 across the tests, with a median of 0.90 and a standard deviation of 0.03, indicating that reliabilities are generally within a reasonable range.

For several reasons, the subgroup reliability results should be interpreted with caution. Reliability coefficients are dependent not only on the measurement properties of a test but also on the statistical distribution of the studied subgroup. For example, subgroup sizes may vary considerably, which results in natural variation in reliability coefficients. Alternatively,  $\alpha$ , which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper & Smith, 1998).

## 8.3 REPORTING SUBCATEGORY RELIABILITY

Reliabilities were calculated for the reporting subcategories within the 2024 RICAS content areas. Results and reporting category descriptions are presented in Appendix K. The reliability coefficients for the reporting subcategories range from 0.51 to 0.91, with a median of 0.73 and a standard deviation of 0.10. Because they are based on a subset of items rather than the full test, subcategory reliabilities were typically lower than were overall test score reliabilities, approximately to the degree expected based on the classical test theory (Haertel, 2006), and interpretations should take this into account. Qualitative

differences among grades and content areas once again preclude valid inferences about the reliability of the full test score based on statistical comparisons among subtests.

## 8.4 RELIABILITY OF ACHIEVEMENT LEVEL CATEGORIZATION

The accuracy and consistency of classifying students into achievement levels are critical components of a standards-based reporting framework (Livingston & Lewis, 1995). For the 2024 RICAS tests, students were classified into one of four achievement levels: *Not Meeting Expectations*, *Partially Meeting Expectations*, *Meeting Expectations*, or *Exceeding Expectations*.

Cognia conducted decision accuracy and consistency (DAC) analyses to determine the statistical accuracy and consistency of the classifications. This section explains the methodologies used to assess the reliability of classification decisions and gives the results of these analyses.

Accuracy refers to the extent to which achievement classifications based on test scores match the classifications that would have been assigned if the scores did not contain any measurement error. Accuracy must be assessed because errorless test scores do not exist. Consistency measures the extent to which classifications based on test scores match the classifications based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are administered to the same group of students. In operational testing programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and the consistency of classifications based on a single administration of a test. The Livingston and Lewis (1995) technique was used for the 2024 RICAS tests because it is easily adaptable to all types of testing formats, including mixed formats.

The DAC estimates reported in Tables 8-2 and 8-3 make use of “true scores” in the classical test theory sense. A true score is the score that would be obtained if a test had no measurement error. True scores cannot be observed and so must be estimated. In the Livingston and Lewis (1995) method, estimated true scores are used to categorize students into their “true” classifications.

For the 2024 RICAS tests, after various technical adjustments (described in Livingston & Lewis, 1995), a four-by-four contingency table of accuracy was created for each content area and grade, where cell  $[i, j]$  represented the estimated proportion of students whose true score fell into classification  $i$  (where  $i = 1$  to 4) and observed score fell into classification  $j$  (where  $j = 1$  to 4). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

To calculate consistency, true scores were used to estimate the joint distribution of classifications on two independent, parallel test forms. Following statistical adjustments (per Livingston & Lewis, 1995), a new four-by-four contingency table was created for each content area and grade and populated by the proportion of students who would be categorized into each combination of classifications according to the two (hypothetical) parallel test forms. Cell  $[i, j]$  of this table represented the estimated proportion of students whose observed score on the first form would fall into classification  $i$  (where  $i = 1$  to 4) and whose observed score on the second form would fall into classification  $j$  (where  $j = 1$  to 4). The sum of the diagonal entries (i.e., the proportion of students categorized by the two forms into the same classification) signified overall consistency.

Cognia also measured consistency on the 2024 RICAS tests using Cohen’s (1960) coefficient  $\kappa$  (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. It is calculated using the following formula:

$$\kappa = \frac{(\text{Observed agreement}) - (\text{Chance agreement})}{1 - (\text{Chance agreement})} = \frac{\sum_i C_{ii} - \sum_i C_i C_i}{1 - \sum_i C_i C_i}, \quad (\text{Equation 11})$$

where

$C_i$  is the proportion of students whose observed achievement level would be level  $i$  (where  $i = 1-4$ ) on the first hypothetical parallel form of the test;

$C_i$  is the proportion of students whose observed achievement level would be level  $i$  (where  $i = 1-4$ ) on the second hypothetical parallel form of the test; and

$C_{ii}$  is the proportion of students whose observed achievement level would be level  $i$  (where  $i = 1-4$ ) on both hypothetical parallel forms of the test.

Because  $\kappa$  is corrected for chance, its values are lower than other consistency estimates.

## 8.5 DECISION ACCURACY AND CONSISTENCY RESULTS

DAC analyses were conducted both for the overall population and for subpopulations at each performance achievement level. Results of the DAC analyses are provided in Tables 8-2 and 8-3 for the 2024 RICAS tests.

Table 8-2 includes overall accuracy indices with consistency indices displayed in parentheses next to the accuracy values, as well as overall kappa values. Overall ranges for accuracy (0.80–0.86), consistency (0.72–0.80), and kappa (0.59–0.68) indicate that most students were classified accurately and consistently with respect to measurement error and chance. Accuracy and consistency values conditional on achievement level are also given. For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, the conditional accuracy value is 0.84 for *Not Meeting Expectations* for the grade 3 ELA test. This figure indicates that among the students whose true scores placed them in this classification, 84% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.76 indicates that 76% of students with observed scores in the *Not Meeting Expectations* level would be expected to score in this classification again if a second, parallel test form was taken.

Because one use of RICAS tests is the placement of student test scores into achievement levels, an important concern is the accuracy and consistency of decisions around achievement level thresholds. In this case, accuracy at the *Partially Meeting Expectations/Meeting Expectations* threshold is critically important, which summarizes the percentage of students who are correctly classified either above or below the particular cutpoint. Table 8-3 provides the accuracy and consistency estimates and false positive and false negative decision rates at each cutpoint for the 2024 RICAS tests. A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.

In Table 8-3, the accuracy and consistency indices at the *Partially Meeting Expectations/Meeting Expectations* threshold range from 0.91–0.93 and 0.87–0.91, respectively. The false positive and false negative decision rates at the *Partially Meeting Expectations/Meeting Expectations* threshold both range from 3%–5%. These results indicate that nearly all students were correctly classified with respect to being above or below the *Partially Meeting Expectations/Meeting Expectations* cutpoint.

**Table 8-2 Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Achievement Level**

Content Area	Grade	Overall	Kappa	Conditional on Achievement Level			
				Not Meeting Expectations	Partially Meeting Expectations	Meeting Expectations	Exceeding Expectations
ELA	3	0.81 (0.73)	0.60	0.84 (0.76)	0.81 (0.75)	0.79 (0.72)	0.76 (0.55)
	4	0.82 (0.75)	0.61	0.85 (0.77)	0.84 (0.78)	0.76 (0.70)	1.00 (0.29)
	5	0.84 (0.78)	0.65	0.86 (0.78)	0.86 (0.82)	0.79 (0.73)	0.78 (0.57)
	6	0.80 (0.72)	0.59	0.86 (0.81)	0.78 (0.70)	0.73 (0.63)	0.77 (0.60)
	7	0.82 (0.74)	0.61	0.87 (0.80)	0.83 (0.77)	0.73 (0.66)	0.66 (0.43)
	8	0.81 (0.74)	0.63	0.90 (0.84)	0.78 (0.70)	0.77 (0.68)	0.75 (0.61)
Mathematics	3	0.83 (0.76)	0.65	0.88 (0.81)	0.82 (0.76)	0.81 (0.75)	0.78 (0.63)
	4	0.84 (0.77)	0.66	0.87 (0.79)	0.86 (0.81)	0.81 (0.75)	0.70 (0.51)
	5	0.86 (0.80)	0.68	0.86 (0.77)	0.87 (0.84)	0.84 (0.78)	0.82 (0.67)
	6	0.84 (0.78)	0.66	0.85 (0.79)	0.86 (0.81)	0.80 (0.71)	0.81 (0.63)
	7	0.84 (0.77)	0.65	0.86 (0.79)	0.84 (0.79)	0.80 (0.72)	0.79 (0.61)
	8	0.84 (0.77)	0.66	0.88 (0.82)	0.84 (0.78)	0.80 (0.72)	0.81 (0.65)

**Table 8-3 Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint**

Content Area	Grade	Not Meeting Expectations / Partially Meeting Expectations			Partially Meeting Expectations / Meeting Expectations			Meeting Expectations / Exceeding Expectations		
		Accuracy (consistency)	False		Accuracy (consistency)	False		Accuracy (consistency)	False	
			Positive	Negative		Positive	Negative		Positive	Negative
ELA	3	0.93 (0.91)	0.03	0.03	0.91 (0.87)	0.05	0.04	0.96 (0.95)	0.03	0.01
	4	0.93 (0.90)	0.03	0.04	0.91 (0.87)	0.04	0.05	0.98 (0.97)	0.02	0.00
	5	0.94 (0.91)	0.03	0.03	0.92 (0.89)	0.04	0.04	0.98 (0.97)	0.02	0.00
	6	0.91 (0.88)	0.05	0.04	0.92 (0.89)	0.04	0.04	0.97 (0.95)	0.02	0.01
	7	0.92 (0.89)	0.04	0.04	0.92 (0.89)	0.03	0.04	0.97 (0.96)	0.02	0.01
	8	0.93 (0.90)	0.04	0.04	0.92 (0.89)	0.04	0.04	0.96 (0.95)	0.02	0.02
Mathematics	3	0.94 (0.92)	0.03	0.03	0.92 (0.89)	0.04	0.04	0.96 (0.95)	0.02	0.01
	4	0.95 (0.92)	0.03	0.03	0.92 (0.89)	0.03	0.04	0.97 (0.96)	0.02	0.01
	5	0.95 (0.92)	0.02	0.03	0.93 (0.90)	0.04	0.03	0.99 (0.98)	0.01	0.00
	6	0.93 (0.90)	0.04	0.03	0.93 (0.90)	0.04	0.03	0.98 (0.97)	0.01	0.00
	7	0.93 (0.89)	0.04	0.04	0.93 (0.90)	0.04	0.03	0.98 (0.98)	0.01	0.01
	8	0.92 (0.89)	0.04	0.04	0.93 (0.91)	0.03	0.03	0.98 (0.98)	0.01	0.01

The indices in Tables 8-2 and 8-3 are derived from Livingston and Lewis’s (1995) method of estimating DAC. Livingston and Lewis discuss two versions of the accuracy and consistency tables. A standard version performs calculations for forms parallel to the form taken. An “adjusted” version adjusts the results of one form to match the observed score distribution obtained in the data. The tables use the standard version for two reasons: (1) This “unadjusted” version can be considered a smoothing of the data, thereby decreasing the variability of the results; and (2) for results dealing with the consistency of two parallel forms, the unadjusted tables are symmetrical, indicating that the two parallel forms have the same statistical properties. This second reason is consistent with the notion of forms that are parallel (i.e., it is more intuitive and interpretable for two parallel forms to have the same statistical distribution).

As with other methods of evaluating reliability, DAC statistics that are calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Tables 8-2 and 8-3 should be interpreted with caution. In addition, it is important to remember that it might be inappropriate to compare DAC statistics across grades and content areas.

## 8.6 QUALITY CHECK

For reliability and decision accuracy and consistency (DAC) results, automated psychometric software tools, such as Cognia’s Measurement Portal (“the Portal”) and BB-CLASS software, are used. These tools play a vital role in standardizing the processes for conducting psychometric analyses, including reliability and DAC analysis. All software has been rigorously checked throughout its development cycle to ensure accuracy.

Similar to other psychometric analyses, psychometricians review the results in detail for reasonableness after all analyses are run. The results are primarily compared to those of previous years. Additionally, the results are examined to see if the interpretations are the same as last year. If any changes are identified, the analyses are run numerous times to check for accuracy, and the lead psychometrician consults with other psychometricians to gain insights into the observed changes.

## 8.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING RELIABILITY

1.3 **Generalization Inference:** The observed score from any specific form testing a given grade and content area is reflective of the expected score on any potential form of the test for that grade and content area.

1.3.3 **Claim:** Statistical analyses of observed scores on specific forms show that they are good predictors of expected scores on other potential forms.

**Evidence:** Section 8.1 describes the process for analyzing the reliability of RICAS forms and the results of these analyses. These analyses establish the reliability of each form. Subject to the equating and scaling methods placing scores from forms on the same scale, adequate reliability of individual forms establishes them as good predictors of expected score.

1.4 **Explanation Inference:** Expected scores are attributable to proficiency in the target knowledge and abilities.

1.4.2 **Claim:** Tests are assembled with adequate precision near cut points.

**Evidence:** Sections 8.4 and 8.5 describe decision accuracy and consistency analysis procedures and results. Accuracy and consistency rates were reported as being adequately high while false positive and negatives demonstrated strong agreement between true score and observed score classification decisions.

# Chapter 9. Validity Arguments Supporting Intended Interpretations and Uses of Test Scores

## 9.1 RATIONALE FOR VALIDITY ARGUMENT-CENTERED TECHNICAL REPORTING

Chapter 9 presents the primary intended interpretations and uses for RICAS test scores, the assumptions that underlie these score interpretations and uses, and the evidence supporting these assumptions. A validity argument logic model is introduced and applied to the evidence and assumptions to produce a structured argument in support of all intended score interpretations and uses. The structure applied to the validity argument closely follows the Chapelle (2020) framework, which provides a chain of inferences, each building on the previous, to preserve the interpretations as defined by the content standards such that they are realized within the resulting test scores and applicable to the intended uses of the RICAS program.

The Standards (2014) define validity as “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (p. 11). Elaborating on that definition, the Standards assert that “it is the interpretations of test scores for proposed uses that are evaluated, not the test itself” (p. 11) and that “validation logically begins with an explicit statement of the proposed interpretation of test scores, along with a rationale for the relevance of the interpretation to the proposed use” (p. 11). This definition applies specifically to intended interpretations and uses of test scores, rather than to the broader program of curriculum and instruction in which a testing program is embedded or to the surrounding education and school improvement policies and aspirations for student learning.

The Standards further state that “a sound *validity argument* integrates various strands of evidence into a coherent account of the degree to which existing evidence and theory support the intended interpretations of test scores for specific uses” (p. 21; emphasis added). An emerging common practice in state assessment programs is to construct validity arguments based on Toulmin’s model of argumentation (Toulmin, 1958). A model for validity arguments, derived from the Toulmin model, is shown in Figure 9-1.



**Figure 9-1 Validity Argument Logic Model**



## 9.2 VALIDITY ARGUMENT FOR INTERPRETATION AND USE OF RICAS TEST SCORES

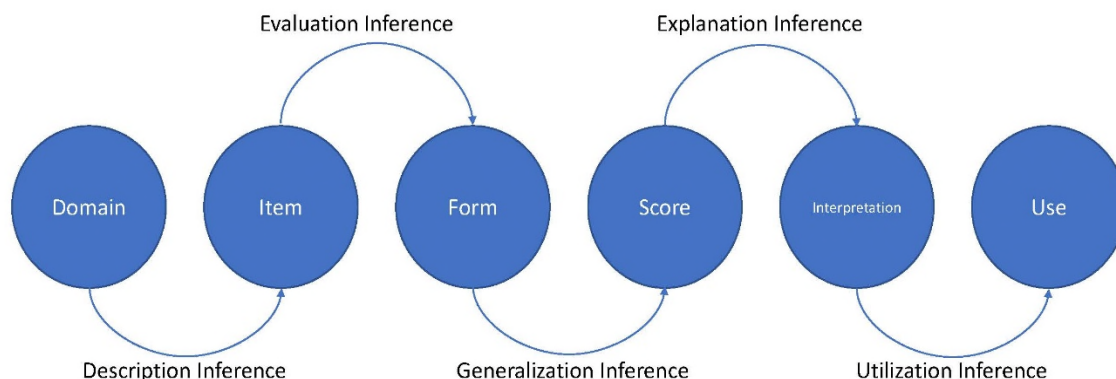
For the RICAS, the overarching validity argument is that the existing design, procedural, and psychometric evidence supports all intended score interpretations and uses of resulting test scores. Each of the interpretation and use inferences is comprised of one or more claims requiring supporting evidence. With all claims backing an inference supported by evidence, the inference is upheld. With all inferences upheld, the argument for the validity of the interpretations and uses is thusly made.

Specifically, the structure of the validation argument in this technical report follows closely Chapelle et al. (2018) and differentiates five layers:

- 1) **Description Inference:** Items sample from the target domain appropriately such that high quality forms can be produced. (Domain to Item)
- 2) **Evaluation Inference:** Forms sample from items appropriately such that observed scores reflective of the domain can be produced. (Item to Form)
- 3) **Generalization Inference:** Observed scores on individual forms are reliable such that they are reflective of expected scores across forms. (Form to Score)
- 4) **Explanation Inference:** Expected scores are associated with classification cuts such that classification decisions are interpretable. (Score to Interpretation)
- 5) **Utilization Inferences:** Interpretations of scores and classifications are used as intended and only in ways considered appropriate and fair. (Interpretation to Use)

See Figure 9-2 for a visual representation of Chapelle's framework.

**Figure 9-2 Chapelle (2020)’s Framework: The Arguments and the Inferential Steps**



It is important for the gathering of information in support of the Generalization Inference (3) to define what is meant by the term “form” in this context. A test form is not just the set of items on which the score is based, but the structure of the exam in terms of all elements that can affect an individual’s performance. This can include, among other things, the raters scoring an exam, the occasion on which the exam is administered, and the setting in which it is administered. Generalization from observed to expected score is optimized when all sources of potential variability of test scores are identified and accounted for such that observed scores maximally reflect a student’s ability and not the influence of unwanted sources of variance.

Evidence in support of these five layers of the validation argument is presented in two main sections:

Section 9.2.1 presents inferences that support the intended interpretations of RICAS test scores, their necessary claims, and evidence supporting those claims (inferences 1.1 to 1.4).

Section 9.2.2 presents separate inferences for an intended use of the RICAS test scores, each presented with its necessary claims and supporting evidence (inferences 1.5 to 1.8).

## **9.2.1 Claims Supporting Intended Interpretations of RICAS Test Scores**

1.1 **Description Inference:** Observations of performance on the RICAS reflect the knowledge and abilities articulated in the RI Core Standards with appropriate assessment tasks representing the full breadth and depth of the domain as articulated within these standards.

1.1.1 **Claim:** Expected knowledge and abilities are thoroughly articulated and considered appropriate to the grade and content area being assessed.

**Evidence:** The need for alignment of the assessments to the content standards is made clear in the introductory paragraph in Chapter 1, referencing the goal of measuring student proficiency relative to these standards. The direct link between the content standards and the assessments throughout the test design, development, and implementation processes for all grades and content areas is thoroughly articulated in Chapter 2.

1.1.2 **Claim:** Assessment tasks are developed to provide evidence of the expected knowledge and abilities for each grade and content area being assessed.

**Evidence:** Subsections 2.4.2 for ELA and 2.5.1 for mathematics detail the specific standards addressed by items available for RICAS assessments. Subsections 2.4.1 and 2.4.2 describe passage types, item types, and cognitive levels for items on the ELA assessments. Subsection 2.5.1 describes the item types and cognitive levels for items on the mathematics assessments. Subsection 2.6.1 describes item review and refinement procedures, and subsection 2.6.2 describes item field testing and subsequent scoring and data review processes. Together, these subsections describe an overall process of item development that ensures items effectively target the expected knowledge and abilities of the grades being assessed.

1.2 **Evaluation Inference:** Each test form, an organized sampling of assessment tasks, results in an observed score that reflects a student's knowledge and abilities in the content area being assessed through appropriate test assembly, administration, and scoring procedures.

1.2.1 **Claim:** Each form is constructed to draw from available items such that the underlying domain of knowledge and abilities is adequately sampled.

**Evidence:** Subsection 2.4.3 describes the blueprints and test design specifications for ELA, while subsection 2.5.2 covers the same aspects for mathematics. Subsection 2.6.3 describes the processes for item selection and test form review, and subsection 2.6.4 details the special edition test forms and modifications to the original test items. The procedures outlined in both subsections aim to ensure design and blueprint specifications are met, and they work to prevent elements of test construction that could potentially confound interpretability. Together, these processes ensure that each form draws a sampling of high-quality items representing the underlying knowledge and abilities defined within the content standards.

1.2.2 **Claim:** The assessment is administered under appropriate conditions.

**Evidence:** Chapter 3 describes test administration processes for the RICAS. This includes schedules, security requirements, administration procedures, and practices for non-standard administrations. Chapter 3 further references *Test Administrators Manuals* and *Test Coordinators Manuals* for more details of administration procedures, administrator responsibilities, and irregularity tracking. Together, the evidence given in Chapter 3 demonstrates that the administration was properly designed and implemented, quality-control procedures worked as intended, and there were no notable threats to validity from the administration.

1.2.3 **Claim:** The scoring procedures and models produce scores accurately reflective of targeted knowledge and abilities.

**Evidence:** Chapter 4 has detailed sections describing the scoring process for machine-scored items and hand-scored, polytomous items on RICAS assessments. These steps, in conjunction with the appropriate item and blueprint design described under sections 2.4 and 2.5, support this claim. The design and implementation of the machine- and hand-scoring procedures is also documented in Chapter 4, which shows that the procedures adhere to industry-accepted practices and standards. Section 7.2 describes the scoring models used for items on the RICAS, describing the models used in detail and citing the references that establish the appropriateness of these models for placing student performances on a common scale for scoring purposes.

1.2.4 **Claim:** Items on the assessment demonstrate appropriate statistical quality.

**Evidence:** Chapter 6 describes the classical item analysis procedures conducted to ensure that all items adhere to industry-accepted practices and standards (AERA et al., 2014). Differential Item Functioning (DIF) analysis, presented in section 6.2, provides evidence that the items are free of systematic biases. Subsection 2.6.2 describes the review process for evaluating items flagged by these and other field-test analyses. Section 7.3 describes IRT results referring to tables within the equating report (Appendix J) that describe quality control checks on items and procedures for making interventions based on items being flagged during these checks.

1.3 **Generalization Inference:** The observed score from any specific form testing a given grade and content area is reflective of the expected score on any potential form of the test for that grade and content area.

1.3.1 **Claim:** Task specifications adequately inform production or selection of items with similar content and statistical characteristics.

**Evidence:** Claim 1.1.2, with evidence from throughout Chapter 2, establishes that the task specifications and resulting item development efforts result in assessment tasks representative of expected knowledge and ability being assessed. Subsection 2.6.3 describes the essential procedural steps taken to meet the broad requirements of expected standards and cognitive skills while avoiding unnecessary duplication of items from previous years' forms. Subsection 2.6.3 also describes the rigorous process of form review to ensure that these requirements are met on forms that are accepted for operational administration. These form construction processes, applied to items meeting Claim 1.1.2, provide evidence that task specifications are adequately informing production and selection of items with similar content and statistical characteristics.

1.3.2 **Claim:** Test specifications result in forms of similar length and task distribution.

**Evidence:** Claim 1.2.1, again gathering evidence from Chapter 2, establishes that test construction processes are designed to implement specifications that result in forms of similar length and task distribution. Subsection 2.6.3 describes the application of those processes to realize those specifications while avoiding unnecessary duplication of items. Subsection 2.6.3 also describes the rigorous review process that verifies that these specifications are met prior to acceptance of the form for operational administration. Dimensionality analyses presented in section 6.3, provide evidence that any differences in length or task distribution are small enough that interpretation of the resulting scores is preserved.

1.3.3 **Claim:** Statistical analyses of observed scores on specific forms show that they are good predictors of expected scores on other potential forms.

**Evidence:** Section 8.1 describes the process for analyzing the reliability of RICAS forms and the results of these analyses. These analyses establish the reliability of each form, which meets professional standards for reliability for tests like RICAS. Subject to the equating and scaling methods placing scores from forms on the same scale, adequate reliability of individual forms establishes them as good predictors of expected scores on other potential forms. Differential Item Functioning (DIF) analyses and subsequent review of items classified as exhibiting DIF, described in section 6.2, support observed score generalization to expected score by ruling out the items specific to this year's forms as sources of bias in the scores.

1.3.4 **Claim:** Equating and scaling methods accurately place scores from different forms onto a common scale.

**Evidence:** Section 7.4 describes equating procedures in detail and summarizes results from the full equating report, provided in Appendix J. Section 7.5 describes the processes of applying

equating results to place raw scores onto RICAS score scales. These sections demonstrate a high level of rigor in selection, application, and interpretation of equating results, placing scores from this year's forms on the same scales as forms from prior years.

1.4 **Explanation Inference:** Expected scores are attributable to proficiency in the target knowledge and abilities.

1.4.1 **Claim:** Cut scores are established through defensible standard setting methods.

**Evidence:** Section 2.3 summarizes the process by which performance standards were established for RICAS (more technical details in Section 7.5). Standard setting activities conducted for the MCAS in 2017 were observed by RIDE staff and technical advisors, rigorously evaluated for consistency with RICAS performance expectations, and deemed sufficient for a sound and technically appropriate implementation in the context of RICAS.

1.4.2 **Claim:** Tests are assembled with adequate precision near cut points.

**Evidence:** Sections 8.4 and 8.5 describe decision accuracy and consistency analysis procedures and results. Accuracy and consistency rates were reported that most students were classified accurately and consistently with respect to measurement error and chance; false positive and negatives demonstrated strong agreement between true score and observed score classification decisions.

## 9.2.2 Claims Supporting Intended Uses of RICAS Test Scores

With evidence provided in support of RICAS scores preserving intended interpretations of the content standards, validation of the primary intended uses of these scores requires evidence that these interpretations can be applied to each use in an appropriate, fair, and just way.

Evidence for each use should show that the intended audience (i.e., those using the scores)

- 1) understands the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to the intended use, and
- 2) finds the scores and classifications genuinely useful for that intended use.

The evidence described in this section pertains to the aspects that relate to activities performed by Cognia/Pearson and RIDE. Evidence regarding the resulting utility of the information is outside the scope of this report.

1.5 **Utilization Inference 1:** RICAS score reports provide students and their families with classification and score information that is useful, presented fairly, and appropriate for monitoring academic achievement and participating in decisions regarding student learning.

1.5.1 **Claim:** Students and their families understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to monitoring academic achievement and participating in decisions regarding student learning.

**Evidence:** Chapter 5 describes how results are reported to students and their families, Section 5.2 describes details of the information included in the score reports. This includes important score and classification information, and explanations of what this information means. Section 5.5 and subsection 5.5.1 describe additional resources that students and their families may use to improve their understanding of this score information. Cognia/Pearson and RIDE provided materials and other implementation supports (e.g., town halls, professional development/educational sessions) that put all stakeholders in a strong position to be able to understand the *intended* meanings and uses of the RICAS scores.

1.5.2 **Claim:** Interpretations of scores and classifications are genuinely useful to students and their families for the purposes of monitoring academic achievement and participating in decisions regarding their learning.

**Evidence:** Section 5.2 describes details of the information included in the score reports. This includes information about how families can help improve their child's learning. Section 5.5 and subsection 5.5.1 describe resources available to students and families that can be used to apply test results to take appropriate actions toward furthering the student's education.

1.6 **Utilization Inference 2:** RICAS score reports provide educators with classification and score information that is useful, presented fairly, and appropriate for supporting curricular planning and identifying instructional needs at both the classroom and individual student level.

1.6.1 **Claim:** Educators understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to curricular planning and identification of instructional needs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe the reporting tools that educators may use to access the score results of individual students and the group of students that they teach, as well as resources available to educators providing guidance for accurately interpreting scores.

1.6.2 **Claim:** Interpretations of scores and classifications are genuinely useful to educators for the purposes of curricular planning and identification of instructional needs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe resources available to educators that provide guidance for applying test scores and interpretations of test scores to their instruction.

1.7 **Utilization Inference 3:** RICAS score reports provide school- and district-level administrators with classification and score information that is useful, presented fairly, and appropriate for supporting program evaluations and improvements at school and district levels.

1.7.1 **Claim:** School- and district-level administrators understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to program evaluations and improvements at school and district levels.

**Evidence:** Section 5.5 and Subsection 5.5.2 describe the reporting tools that administrators may use to access the score results of individual students and group-level data of students in their schools and districts, as well as resources available to administrators providing guidance for accurately interpreting scores.

1.7.2 **Claim:** Interpretations of scores and classifications are genuinely useful to school- and district-level administrators for the purposes of program evaluations and improvements.

**Evidence:** Section 5.5 and subsection 5.5.2 describe resources available to administrators that provide guidance for applying test scores and interpretations of test scores to program evaluation and improvement.

1.8 **Utilization Inference 4:** RICAS score reports provide state administrators with classification and score information that is useful, presented fairly, and appropriate for monitoring academic achievement and growth as required by state accountability programs and informing the public of schools' performances on these metrics.

1.8.1 **Claim:** State and federal administrators understand the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to monitoring academic achievement and growth as required by state and federal accountability programs.

**Evidence:** Section 5.5 and subsection 5.5.2 describe the reporting tools that administrators may use to access the score results of individual students and group-level data of students in schools, districts, and the state, as well as resources available to administrators providing guidance for accurately interpreting scores.

1.8.2 **Claim:** Interpretations of scores and classifications are genuinely useful to state and federal administrators for the purposes of monitoring academic achievement and growth as required by state and federal accountability programs.



**Evidence:** Section 5.5 and subsection 5.5.2 describe resources available to administrators that provide guidance for applying test scores and interpretations of test scores to federal accountability programs.

## 9.3 VALIDATION SUMMARY

Validity arguments for the RICAS are crafted to not just provide evidence that all steps in the test design, development, and implementation process are taken correctly, but that they are working together to ensure that the resulting scores validly support intended interpretations and uses. In other words, each argument should not only be considered individually, but also considered as part of the whole. The reader should consider the chain of evidence and whether it provides a compelling argument to support the way test scores are being used.

The arguments and the logical inferential steps they provide can be summarized as follows. The Description and Evaluation Inferences concern the selection of appropriate items and their arrangement into forms that accurately reflect the domain being measured. The Generalization Inference ensures that scores obtained from individual forms are reliable indicators of the expected scores across all forms. The Explanation Inference links expected scores to classification decisions, ensuring that these decisions are interpretable. Finally, the Utilization Inferences emphasize the importance of using scores and classifications appropriately and fairly, in ways consistent with the intended interpretations of the test. Together, the evidence described in this technical report supports the key claims across the five layers of the framework outlined in Chapelle et al. (2018):

- Following the Chapelle (2020) framework, we have provided a chain of inferences, each building on the previous, to preserve the interpretations as defined by the content standards such that they are realized within the resulting test scores and applicable to the intended uses of the RICAS program. By establishing the description inference, providing evidence that the items used in the assessment target the domain as defined by the standards, we argue that we can create individual forms that produce a test score reflective of achievement on that domain.
- We present evidence we have created such individual forms that these forms do elicit test scores reflecting achievement on the intended domain, which in turn is able to provide a classification for a student achievement level and that such classification decisions are interpretable.
- Finally, we provide evidence to support that the test score and classification interpretations are clearly enough explained as to be used as intended and only in ways considered appropriate and fair. We use this chain of evidence to assert the scores and classifications resulting from RICAS tests are interpretable and used in a way that is intended and fair.

# References

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Belmont, CA: Wadsworth, Inc.
- 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. Available from [https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards\\_2014edition.pdf](https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf)
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Baker, F. B., & Kim, S. H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York, NY: Marcel Dekker, Inc.
- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth, TX: Holt, Rinehart and Winston.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology* 3, 296–322.
- Cai, L. (2012). flexMIRT: Flexible multilevel item factor analysis and test scoring [Computer software]. Chapel Hill, NC: Vector Psychometric Group, LLC.
- Chapelle, C. A. (2020). *Argument-based validation in testing and assessment*. Sage Publications.
- Chapelle, C. A., Schmidgall, J., Lopez, A., Blood, I., Wain, J., Cho, Y., Hutchinson, A., Lee, H-W., & Dursun, A. (2018). *Designing a prototype tablet-based learning-oriented assessment for middle school English learners: An evidence-centered design approach* (ETS RR – 18-46). Princeton, NJ: Educational Testing Service. Available from <https://files.eric.ed.gov/fulltext/EJ1202815.pdf>
- Chicago Manual of Style* (16th ed.). (2003). Chicago: University of Chicago Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70, 213–220.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334.
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description. In P. W. Holland & H. Wainer (Eds.), *Differential Item Functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement* 23, 355–368.
- Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York, NY: John Wiley and Sons, Inc.
- Haertel, E. H. (2006). Reliability. In R.L. Brennan (Ed). *Educational measurement* (pp. 65–110). Westport, CT: Praeger Publishers.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and Applications*. Boston, MA: Kluwer Academic Publishers.

- Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc.
- Hambleton, R. K., & van der Linden, W. J. (1997). *Handbook of modern item response theory*. New York, NY: Springer-Verlag.
- Jiang, J., Roussos, L. A., & Yu, L. (2017, April). An iterative procedure to detect item parameter drift in equating items. Paper presented at the National Council on Measurement in Education Conference, San Antonio, TX.
- Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: Joint Committee on Testing Practices. Available from <https://www.apa.org/science/programs/testing/fair-testing.pdf>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1–73.
- Kim, S. (2006). A comparative study of IRT fixed parameter calibration methods. *Journal of Educational Measurement* 43(4), 355–381.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). New York, NY: Springer-Verlag.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Nering, M., & Ostini, R. (2010). *Handbook of polytomous item response theory models*. New York, NY: Routledge.
- Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 221–262). New York, NY: Macmillan Publishing Company.
- Roussos, L. A., & Ozbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement* 43, 215–243.
- Spearman, C. C. (1910). Correlation calculated from faulty data. *British Journal of Psychology* 3, 271–295.
- Stocking, M. L. & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied psychological measurement* 7, 201–210.
- Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika* 52, 589–617.
- Stout, W.F., Froelich, A. G., & Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duijn, & T. A. B. Snijders (Eds.), *Essays on Item Response Theory* (pp. 357–375). New York, NY: Springer-Verlag.
- Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: University Press.
- van der Linden, W. J. (2016). *Handbook of item response theory, Volume one: Models*. New York, NY: Chapman and Hall/CRC.

- Wang, X., & Roussos, L. A. (2018, April). *A Simple Parametric Procedure for Detecting Drift in Anchor Items*. Paper presented at the National Council on Measurement in Education Conference, New York, NY.
- Zhang, J., & Stout, W. F. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika* 64, 213–249.

# Appendices

**APPENDIX A**  
**ACCOMMODATIONS**

**Table A-1. Numbers of Students Tested with and Without Accommodations by Content Area and Grade**

Content Area	Grade	Number of Students Tested	
		With Accommodations	Without Accommodations
ELA	3	1,092	8,308
	4	1,269	8,421
	5	1,314	8,402
	6	1,208	8,403
	7	1,285	8,452
	8	1,233	8,612
Mathematics	3	2,262	7,317
	4	2,406	7,454
	5	2,179	7,725
	6	1,719	8,044
	7	1,775	8,120
	8	1,725	8,238

**Table A-2. Numbers of Students Tested with Accommodations by Accommodation Type and Grade—ELA**

Description	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Color Contrast	0	0	0	0	0	0
Black on Cream	0	0	0	0	0	0
Black on Light Blue	0	0	0	0	0	0
Black on Light Magenta	0	0	0	0	0	0
White on Black	0	0	0	0	0	0
Yellow on Blue	0	0	0	0	0	0
Dark Gray on Pale Green	0	0	0	0	0	0
Answer Masking	0	0	0	0	0	0
Large Print Test Edition	0	0	1	0	0	1
Screen Reader Edition	0	2	0	0	0	1
Assistive Technology	8	13	12	3	2	5
Braille Test Edition	0	1	0	0	0	0
Human Read Aloud as a Standard Accommodation	0	0	0	0	0	0
Human Read Aloud as a Non-Standard Accommodation	37	32	25	21	19	12
Human Signer as a Standard Accommodation	6	5	2	1	3	3
Human Signer as a Non-Standard Accommodation	0	0	0	0	0	0
Text-to-Speech	169	163	145	134	129	100
Human Scribe as a Standard Accommodation	0	0	0	0	0	0
Human Scribe as a Non-Standard Accommodation	62	42	52	17	12	10
Speech-to-Text as a Standard Accommodation	0	0	0	0	0	0
Speech-to-Text as a Non-Standard Accommodation	86	116	97	66	48	34
Typed Responses	0	0	0	0	1	0
Spell-checker	56	47	48	62	52	34
Word Prediction	57	61	67	56	46	35
Graphic Organizer/Reference Sheet	846	1,028	1,074	915	976	927
Any Other Accommodation	87	118	92	65	61	43
Bilingual Dictionary and Glossary	139	159	160	235	269	282

**Table A-3. Numbers of Students Tested with Accommodations by Accommodation Type and Grade—Mathematics**

Description	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Color Contrast	0	0	0	0	0	0
Black on Cream	0	0	0	0	0	0
Black on Light Blue	0	0	0	0	0	0
Black on Light Magenta	0	0	0	0	0	0
White on Black	0	0	0	0	0	0
Yellow on Blue	0	0	0	0	0	0
Dark Gray on Pale Green	0	0	0	0	0	0
Answer Masking	0	0	0	0	0	0
Large Print Test Edition	0	0	1	1	0	1
Screen Reader Edition	0	2	0	0	0	0
Assistive Technology	2	5	8	3	2	4
Braille Test Edition	0	0	0	0	0	0
Human Read Aloud as a Standard Accommodation	83	87	69	20	20	13
Human Read Aloud as a Non-Standard Accommodation	0	0	0	0	0	0
Human Signer as a Standard Accommodation	6	2	2	1	3	3
Human Signer as a Non-Standard Accommodation	0	0	0	0	0	0
Text-to-Speech	1,846	1,940	1,642	1,124	1,043	916
Human Scribe as a Standard Accommodation	49	28	42	16	9	12
Human Scribe as a Non-Standard Accommodation	0	0	0	0	0	0
Speech-to-Text as a Standard Accommodation	49	68	91	37	31	15
Speech-to-Text as a Non-Standard Accommodation	0	0	0	0	0	0
Typed Responses	0	0	0	1	0	0
Calculation Device	90	99	147	118	245	241
Graphic Organizer/Reference Sheet	758	951	1,019	831	842	782
Any Other Accommodation	0	0	0	0	0	0
Spanish	198	197	201	191	236	261
Bilingual Dictionary and Glossary	257	313	254	297	343	318



**APPENDIX B**  
**PARTICIPATION RATES**

**Table B-1. Summary of Participation by Student Subgroup English Language Arts, Grades 3–8**

Description	Number Tested	Percent Tested
All Students	57,999	100.00
ELL	9,903	17.07
Economically Disadvantaged	26,668	45.98
African American	5,200	8.97
Asian	1,914	3.30
Hispanic	17,360	29.93
Native American/Alaska Native	442	0.76
White	29,859	51.48
Pacific Islander/Hawaiian	100	0.17
Multiracial	3,124	5.39
Male	29,725	51.25
Female	28,258	48.72
Special Education	10,276	17.72

**Table B-2. Summary of Participation by Student Subgroup Mathematics, Grades 3–8**

Description	Number Tested	Percent Tested
All Students	58,964	100.00
ELL	10,968	18.60
Economically Disadvantaged	27,266	46.24
African American	5,339	9.05
Asian	1,960	3.32
Hispanic	18,065	30.64
Native American/Alaska Native	462	0.78
White	29,901	50.71
Pacific Islander/Hawaiian	110	0.19
Multiracial	3,127	5.30
Male	30,246	51.30
Female	28,703	48.68
Special Education	10,258	17.40

## **APPENDIX C**

### **INTERRATER CONSISTENCY**

**Table C-1. Item-Level Interrater Consistency Statistics—ELA Grade 3**

Item Number	Number of		Percent		Percent of		LW
	Score Categories	Responses Scored Twice	Exact	Adjacent	Correlation	Third Scores	Kappa
EL206455922	4	915	82.73	13.66	0.73	3.61	0.71
EL206642581#SCORE_TRAIT_Conv	4	867	76.24	23.53	0.82	0.58	0.73
EL206642581#SCORE_TRAIT_Ideadev	5	867	82.01	17.65	0.84	0.58	0.77

**Table C-2. Item-Level Interrater Consistency Statistics—ELA Grade 4**

Item Number	Number of		Percent		Percent of		LW
	Score Categories	Responses Scored Twice	Exact	Adjacent	Correlation	Third Scores	Kappa
EL909132428#SCORE_TRAIT_Conv	4	945	69.42	29.95	0.77	1.59	0.64
EL909132428#SCORE_TRAIT_Ideadev	5	945	68.68	30.37	0.82	1.59	0.69
EL911875859	4	971	86.92	9.27	0.84	3.81	0.82

**Table C-3. Item-Level Interrater Consistency Statistics—ELA Grade 5**

Item Number	Number of		Percent		Percent of		LW
	Score Categories	Responses Scored Twice	Exact	Adjacent	Correlation	Third Scores	Kappa
EL206057533#SCORE_TRAIT_Conv	4	949	66.49	30.98	0.77	4.00	0.64
EL206057533#SCORE_TRAIT_Ideadev	5	949	72.60	25.50	0.84	4.00	0.73
EL207135604#SCORE_TRAIT_Conv	4	939	82.64	16.51	0.89	1.28	0.82
EL207135604#SCORE_TRAIT_Ideadev	5	939	81.68	17.68	0.89	1.28	0.82

**Table C-4. Item-Level Interrater Consistency Statistics—ELA Grade 6**

Item Number	Number of		Percent		Percent of		LW
	Score Categories	Responses Scored Twice	Exact	Adjacent	Correlation	Third Scores	Kappa
EL028556260#SCORE_TRAIT_Conv	4	935	69.30	29.20	0.85	6.42	0.71
EL028556260#SCORE_TRAIT_Ideadev	6	935	68.98	25.56	0.88	6.42	0.74
EL206144523#SCORE_TRAIT_Conv	4	924	83.87	15.69	0.92	1.30	0.86
EL206144523#SCORE_TRAIT_Ideadev	6	924	80.95	17.97	0.92	1.30	0.85

**Table C-5. Item-Level Interrater Consistency Statistics—ELA Grade 7**

Item Number	Number of		Percent		Percent of		LW
	Score Categories	Responses Scored Twice	Exact	Adjacent	Correlation	Third Scores	Kappa
EL206144041#SCORE_TRAIT_Conv	4	936	74.79	24.47	0.86	2.35	0.76
EL206144041#SCORE_TRAIT_Ideadev	6	936	66.56	31.52	0.88	2.35	0.74
EL217467654#SCORE_TRAIT_Conv	4	922	81.56	16.81	0.90	2.93	0.83
EL217467654#SCORE_TRAIT_Ideadev	6	922	72.89	25.05	0.90	2.93	0.79

**Table C-6. Item-Level Interrater Consistency Statistics—ELA Grade 8**

Item Number	Number of		Percent		Percent of		LW
	Score Categories	Responses Scored Twice	Exact	Adjacent	Correlation	Third Scores	Kappa
EL206761648#SCORE_TRAIT_Conv	4	941	86.29	13.18	0.94	1.38	0.88
EL206761648#SCORE_TRAIT_Ideadev	6	941	82.57	16.37	0.94	1.38	0.87
EL810463548#SCORE_TRAIT_Conv	4	945	76.72	22.75	0.89	1.16	0.80
EL810463548#SCORE_TRAIT_Ideadev	6	945	76.51	22.65	0.93	1.16	0.84

**Table C-7. Item-Level Interrater Consistency Statistics—Mathematics Grade 3**

Item Number	Number of Score Categories	Number of Responses Scored Twice	Percent		Percent of		LW
			Exact	Adjacent	Correlation	Third Scores	Kappa
MA297399A	4	935	89.09	10.70	0.94	0.21	0.90
MA297399A_ES	4	20	100.00	0.00	1.00	0.00	--
MA306360A	4	927	82.63	16.29	0.90	1.08	0.83
MA306360A_ES	4	20	100.00	0.00	1.00	0.00	--
MA900579739	4	929	91.28	8.18	0.96	0.54	0.93
MA900579739_ES	4	20	100.00	0.00	1.00	0.00	--
MA901132847	4	929	92.79	7.00	0.96	0.22	0.93
MA901132847_ES	4	18	100.00	0.00	1.00	0.00	--

Note. “\_ES” denotes Spanish items / “\_PA” denotes Paper items.

**Table C-8. Item-Level Interrater Consistency Statistics—Mathematics Grade 4**

Item Number	Number of Score Categories	Number of Responses Scored Twice	Percent		Percent of		LW
			Exact	Adjacent	Correlation	Third Scores	Kappa
MA302496A	5	969	76.47	22.81	0.91	0.72	0.82
MA302496A_ES	5	20	100.00	0.00	1.00	0.00	--
MA304983A	5	958	83.09	16.28	0.96	0.63	0.89
MA304983A_ES	5	20	100.00	0.00	1.00	0.00	--
MA803833224	5	962	93.66	6.24	0.97	0.10	0.94
MA803833224_ES	5	20	100.00	0.00	1.00	0.00	--
MA900750814	5	967	80.66	18.20	0.93	1.14	0.85
MA900750814_ES	5	20	100.00	0.00	1.00	0.00	--

Note. “\_ES” denotes Spanish items / “\_PA” denotes Paper items.

**Table C-9. Item-Level Interrater Consistency Statistics—Mathematics Grade 5**

Item Number	Number of Score Categories	Number of Responses Scored Twice	Percent		Percent of		LW
			Exact	Adjacent	Correlation	Third Scores	Kappa
MA002829660	5	967	74.87	24.10	0.91	1.03	0.82
MA002829660_ES	5	21	100.00	0.00	1.00	0.00	--
MA715102421	5	968	67.77	28.41	0.77	3.82	0.64
MA715102421_ES	5	20	100.00	0.00	1.00	0.00	--
MA903746888	5	964	80.60	18.46	0.92	0.93	0.84
MA903746888_ES	5	20	100.00	0.00	1.00	0.00	--
MA903746888_PA	5	3	33.33	66.67	0.87	0.00	--
MA903948538	5	966	89.13	10.35	0.96	0.52	0.91
MA903948538_ES	5	20	100.00	0.00	1.00	0.00	--

Note. “\_ES” denotes Spanish items / “\_PA” denotes Paper items.

**Table C-10. Item-Level Interrater Consistency Statistics—Mathematics Grade 6**

Item Number	Number of Score Categories	Number of Responses Scored Twice	Percent		Correlation	Percent of Third Scores	LW Kappa
			Exact	Adjacent			
MA003843559	5	937	93.60	6.19	0.97	0.21	0.94
MA003843559_ES	5	18	100.00	0.00	1.00	0.00	--
MA136463214	5	948	85.97	12.24	0.93	1.79	0.85
MA136463214_ES	5	15	100.00	0.00	1.00	0.00	--
MA307295	5	949	81.24	16.97	0.92	1.79	0.84
MA307295_ES	5	19	100.00	0.00	1.00	0.00	--
MA311694	5	952	73.11	23.32	0.91	3.57	0.81
MA311694_ES	5	20	95.00	5.00	0.98	0.00	--

Note. “\_ES” denotes Spanish items / “\_PA” denotes Paper items.

**Table C-11. Item-Level Interrater Consistency Statistics—Mathematics Grade 7**

Item Number	Number of Score Categories	Number of Responses Scored Twice	Percent		Correlation	Percent of Third Scores	LW Kappa
			Exact	Adjacent			
MA007839825	5	951	94.85	4.94	0.98	0.21	0.95
MA007839825_ES	5	20	100.00	0.00	1.00	0.00	--
MA289853	5	954	83.54	15.93	0.94	0.52	0.87
MA289853_ES	5	23	100.00	0.00	1.00	0.00	--
MA295758	5	959	93.12	6.57	0.97	0.31	0.94
MA295758_ES	5	21	100.00	0.00	1.00	0.00	--
MA715009356	5	953	92.44	7.24	0.98	0.31	0.95
MA715009356_ES	5	19	100.00	0.00	1.00	0.00	--

Note. “\_ES” denotes Spanish items / “\_PA” denotes Paper items.

**Table C-12. Item-Level Interrater Consistency Statistics—Mathematics Grade 8**

Item Number	Number of Score Categories	Number of Responses Scored Twice	Percent		Correlation	Percent of Third Scores	LW Kappa
			Exact	Adjacent			
MA203081760	5	897	90.08	9.59	0.97	0.33	0.92
MA203081760_ES	5	25	100.00	0.00	1.00	0.00	--
MA284774	5	922	81.02	17.68	0.94	1.30	0.87
MA284774_ES	5	24	100.00	0.00	1.00	0.00	--
MA311452	5	968	76.55	20.45	0.91	3.00	0.83
MA311452_ES	5	25	100.00	0.00	1.00	0.00	--
MA311452_PA	5	1	100.00	0.00	--	0.00	--
MA902364058	5	962	86.49	12.89	0.92	0.62	0.86
MA902364058_ES	5	27	100.00	0.00	1.00	0.00	--
MA902364058_PA	5	2	100.00	0.00	--	0.00	--

Note. “\_ES” denotes Spanish items / “\_PA” denotes Paper items.

**APPENDIX D**

**ACHIEVEMENT LEVEL DISTRIBUTIONS**

**Table D-1. Achievement-Level Distributions by Grade—ELA**

Grade	Achievement Level	Percent in Level		
		2024	2023	2022
3	Not Meeting Expectations	19.43	21.13	19.41
	Partially Meeting Expectations	42.76	42.03	44.03
	Meeting Expectations	32.28	31.54	31.57
	Exceeding Expectations	5.54	5.30	4.98
4	Not Meeting Expectations	21.74	19.91	21.67
	Partially Meeting Expectations	48.18	46.79	49.30
	Meeting Expectations	27.80	29.71	26.49
	Exceeding Expectations	2.27	3.59	2.55
5	Not Meeting Expectations	20.28	20.35	17.74
	Partially Meeting Expectations	50.01	44.55	50.51
	Meeting Expectations	26.56	32.17	28.25
	Exceeding Expectations	3.15	2.94	3.50
6	Not Meeting Expectations	33.73	31.12	32.10
	Partially Meeting Expectations	37.87	36.94	36.72
	Meeting Expectations	22.18	27.16	25.72
	Exceeding Expectations	6.21	4.78	5.47
7	Not Meeting Expectations	29.33	27.50	28.80
	Partially Meeting Expectations	43.86	43.45	42.01
	Meeting Expectations	23.17	24.41	25.74
	Exceeding Expectations	3.64	4.63	3.46
8	Not Meeting Expectations	34.19	31.50	27.54
	Partially Meeting Expectations	33.89	36.28	43.45
	Meeting Expectations	24.83	25.60	24.87
	Exceeding Expectations	7.09	6.62	4.14



**Table D-2. Achievement-Level Distributions by Grade—Mathematics**

Grade	Achievement Level	Percent in Level		
		2024	2023	2022
3	Not Meeting Expectations	22.56	25.20	24.84
	Partially Meeting Expectations	38.30	40.33	40.14
	Meeting Expectations	31.85	28.99	30.99
	Exceeding Expectations	7.29	5.47	4.04
4	Not Meeting Expectations	20.23	22.32	23.68
	Partially Meeting Expectations	44.36	41.67	46.13
	Meeting Expectations	30.94	30.66	27.02
	Exceeding Expectations	4.46	5.36	3.17
5	Not Meeting Expectations	17.27	17.90	23.90
	Partially Meeting Expectations	52.66	52.13	50.16
	Meeting Expectations	27.24	27.49	24.31
	Exceeding Expectations	2.84	2.48	1.64
6	Not Meeting Expectations	24.76	24.09	23.77
	Partially Meeting Expectations	49.32	47.18	48.87
	Meeting Expectations	22.67	25.99	25.65
	Exceeding Expectations	3.26	2.74	1.72
7	Not Meeting Expectations	27.06	31.83	29.61
	Partially Meeting Expectations	47.63	42.48	47.67
	Meeting Expectations	22.29	22.34	20.14
	Exceeding Expectations	3.01	3.36	2.57
8	Not Meeting Expectations	29.83	30.86	28.65
	Partially Meeting Expectations	45.30	46.16	50.59
	Meeting Expectations	21.47	19.92	18.40
	Exceeding Expectations	3.40	3.06	2.36

**APPENDIX E**

**SAMPLE REPORTS**

# Building Brighter Futures



As the academic year progresses, we want to extend our gratitude for your continued support and involvement in your child's education. We believe that every child deserves access to a high-quality education that prepares them for a bright future. From real-world, relevant courses to unique and engaging out-of-school learning opportunities, Rhode Island is working to ensure all students are equipped with what they need to thrive during their educational journey and graduate prepared for college and career. **Please remember that being present is essential for students to learn. Rhode Island defines chronic absenteeism as the percentage of students who miss 10% or more of school days during the year (or two days per month).**

**18 days** makes your child chronically absent.  
**10 days** puts your child at risk.

## Why does it matter?

Children who are chronically absent in kindergarten and first grade are less likely to read on grade level by the third grade.



Chronic absenteeism is the single strongest predictor of dropping out before graduation.



Students who drop out of school are less likely to succeed in a career. In fact, a high school graduate makes, on average, over a lifetime, \$1 million more than a student who dropped out.



Chronic absenteeism is linked to increased suspensions, teen substance use, as well as poor health as adults.

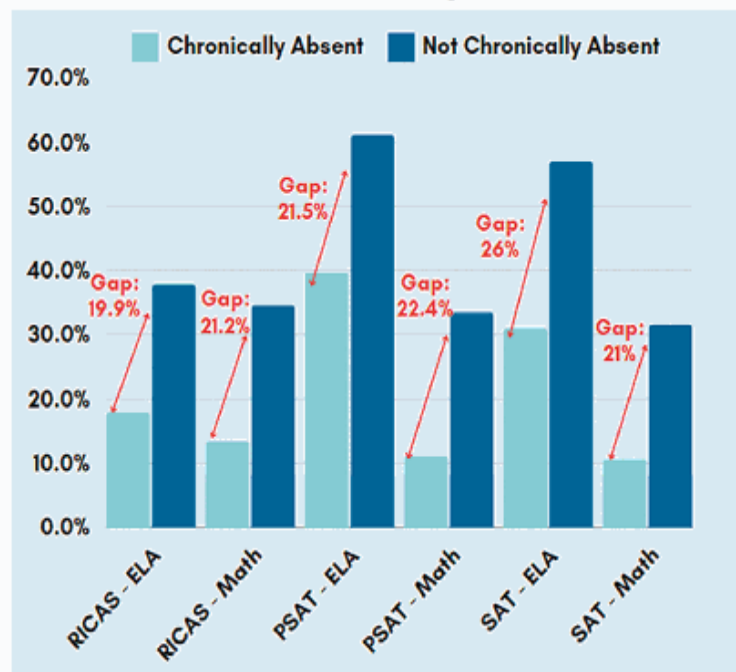


85% of students who drop out of high school were chronically absent. This can be predicted as early as third grade, based on their early elementary school attendance.



Significant performance gaps exist for students who are chronically absent, as seen in the data below:

**2023 Percent Proficient: Chronically Absent Students vs. Not Chronically Absent Students**



# Spring 2024 RICAS Individual Student Report



**Name:** \_\_\_\_\_ **District:** \_\_\_\_\_  
**SASID:** \_\_\_\_\_ **School:** \_\_\_\_\_  
**Date of Birth:** \_\_\_\_\_ **Grade:** 8

This report provides your child's results from the 2024 Rhode Island Comprehensive Assessment System (RICAS) tests in English Language Arts (reading and writing) and mathematics.

Information from the RICAS, in combination with other academic and social measures, will help educators assess grade level placement, design specialized instruction, set learning goals, and monitor progress. These tests will allow schools, districts, and RIDE to identify where we need to take action to improve teaching and learning. Every student deserves an equal chance to demonstrate what they understand, and RIDE extends its gratitude to our school staff and students who administer and show what they know during the RICAS tests each year.

We thank you for your participation in these tests which help guide critical work to improve outcomes for students. We hope understanding your child's comprehension of ELA and mathematics knowledge and skills will empower you as you advocate for your child. You know your child best. For more information on how to understand the results, visit [www.RIDE.ri.gov/Assessment-Results](http://www.RIDE.ri.gov/Assessment-Results).

**For each subject, the report shows:**

- Your child's score between 440 and 560 and their achievement level
- How your child performed in reading and mathematics based on the test reporting categories
- A growth score that shows how your child performed compared to other students who scored similarly

## Your Child's Overall Results

### English Language Arts

Achievement Level

**Exceeding Expectations**

Score

**534**

(Score range: 440-560)

Growth Percentile

**82**

**Details on page 2**

### Mathematics

Achievement Level

**Partially Meeting Expectations**

Score

**499**

(Score range: 440-560)

Growth Percentile

**9**

**Details on page 3**

**Did you know that establishing family routines can help your child succeed?**

Make a habit of setting up designated times for homework, reading, mealtimes, family conversations, bedtime, and leaving for school each day.

**What do I do next?**

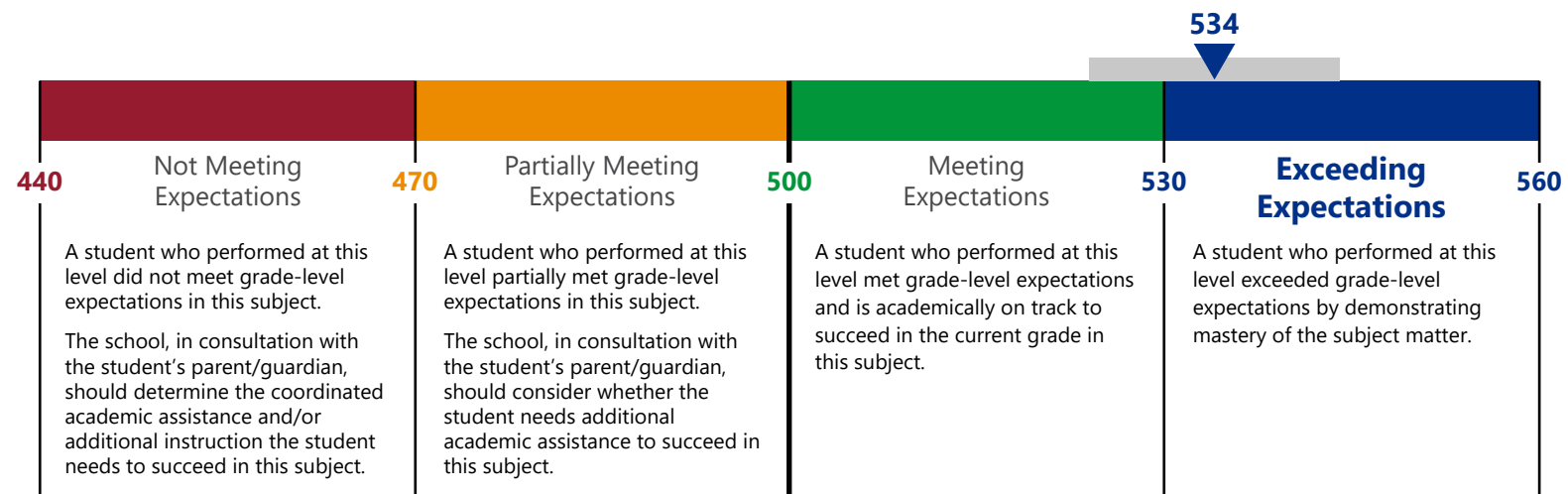
After reviewing this report, it is critical that you attend family-teacher conferences and discuss with your child's teachers your questions and concerns. Don't be afraid to speak up. Children whose families stress the value of education are more likely to find it important, as well.

**How can I support my child's education?**

- School attendance matters, every single day. Missing just two days of school a month means your student is chronically absent, so make it a priority to get your child to school on time daily.
- Establish daily reading routines, let your child see you read, and encourage your child to read for fun all year long.
- Get involved and stay connected to your child's school, however and whenever you can.
- Share your voice! Help improve your child's school by participating in SurveyWorks every year.

Start a conversation. Ask questions. Talk to your child about what they're learning, and show an interest in the subjects that excite them. Remember, you are also your child's teacher, and you play an important role in setting your child up for success.

Your Child's Achievement Level: **Exceeding Expectations**  
 Your Child's Score: **534**



The horizontal gray bar shown in the graphics above and below show the range of likely scores your child would receive if he or she took the test multiple times.

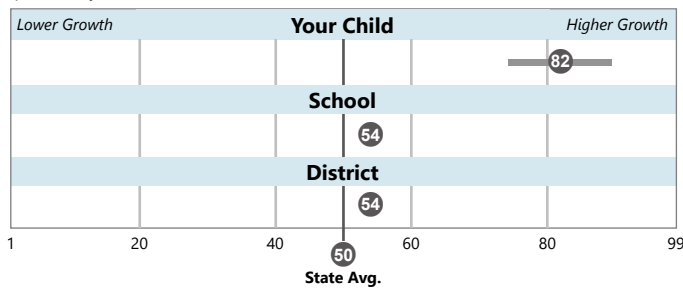
#### Achievement

How your child performed compared to students in their school, district, and state.

Your Child's		Year	Average Score		
Grade	Score		School	District	State
8	534	2024	515	515	486
8	534	2023			
7	516	2022			

#### 2024 Student Growth Percentiles

Your child's score this year is the same as or better than 82 percent of Rhode Island students who had a similar score to your child on the assessment(s) in a previous year(s).



#### How your child performed in each reporting category and on each individual test question

Reporting Category	Points Earned by Your Child	Total Possible Points	Average Points			Average Points Earned by Students Meeting Expectations
			School	District	State	
Reading 	28	29	22.0	21.8	17.7	21.8
Language+ 	10	11	9.1	9.0	6.2	8.0
Writing ‡ 	5	10	6.1	6.0	3.3	4.2

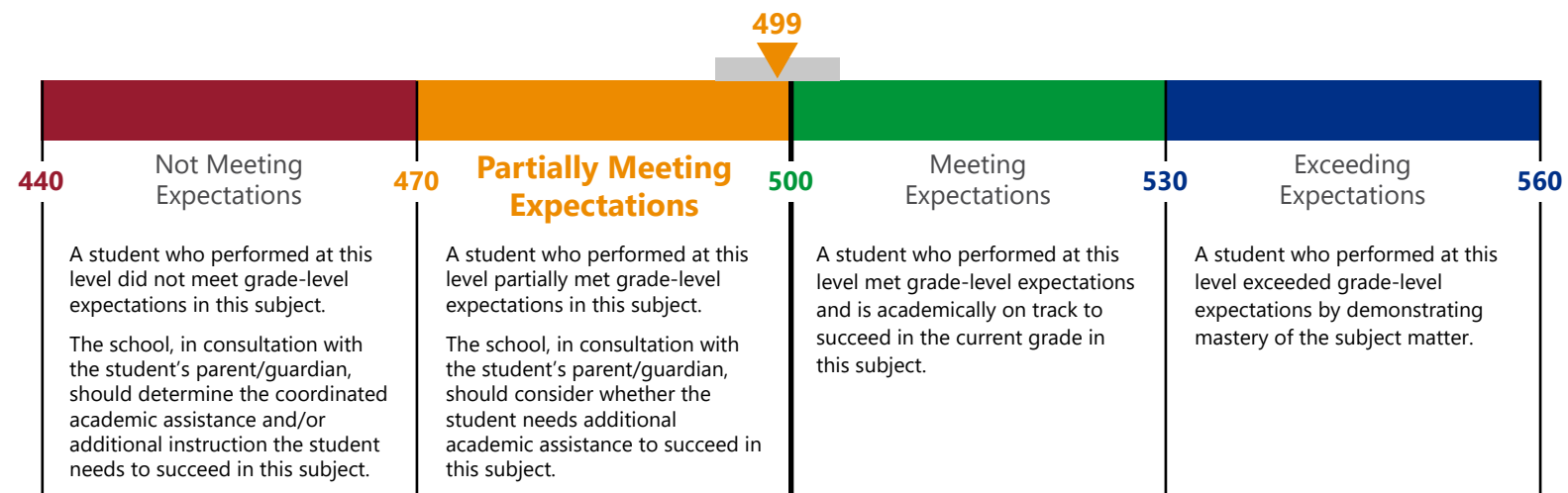
† The Language reporting category includes the standard English convention scores.  
 ‡ The Writing reporting category is based on the idea development scores.

#### Individual Test Questions

Question Number	1	2	3	4	5	6	7	8	9	10	11	12 CV	12 ID	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31 CV	31 ID
Points Earned	1/1	2/2	1/1	1/1	1/1	1/1	1/1	1/1	2/2	1/1	1/1	2/3	2/5	1/1	1/1	1/2	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	2/2	2/2	3/3	3/5

Key x/y = x points earned out of y points possible Blank space = no answer N/A = Item not administered  
 ID = Essay idea development score CV = Essay conventions score

Your Child's Achievement Level: **Partially Meeting Expectations**  
 Your Child's Score: **499**



The horizontal gray bar shown in the graphics above and below show the range of likely scores your child would receive if he or she took the test multiple times.

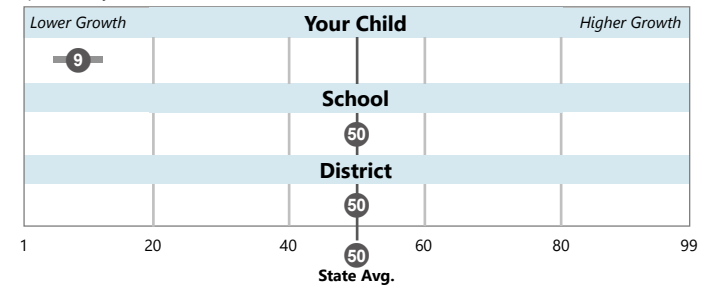
#### Achievement

How your child performed compared to students in their school, district, and state.

Your Child's		Year	Average Score		
Grade	Score		School	District	State
8	499	2024	508	507	484
8	499	2023			
7	514	2022			

#### 2024 Student Growth Percentiles

Your child's score this year is the same as or better than 9 percent of Rhode Island students who had a similar score to your child on the assessment(s) in a previous year(s).



#### How your child performed in each reporting category and on each individual test question

Reporting Category	Points Earned by Your Child	Total Possible Points	Average Points			Average Points Earned by Students Meeting Expectations
			School	District	State	
Number System & Expressions/Equations 	10	20	12.4	12.3	8.2	11.8
Functions 	6	11	7.4	7.4	4.4	6.4
Geometry 	10	16	10.0	9.9	5.7	8.7
Statistics and Probability 	5	7	5.8	5.7	3.9	5.6

#### Individual Test Questions

Question Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Points Earned	0/1	1/1	0/1	0/1	0/4	0/1	0/1	0/1	0/1	1/1	1/1	1/1	2/4	0/1	1/1	1/1	1/1	0/1	2/2	1/1	0/1	1/1	2/4	1/1	1/1	1/1	1/1	1/1	2/2	0/1	1/1	1/1	1/1	0/1	2/4	1/1	1/1	1/1	1/1	0/1

Key x/y = x points earned out of y points possible Blank space = no answer N/A = Item not administered

## **APPENDIX F**

### **2023 REPORTING BUSINESS REQUIREMENTS**

# Reporting Business Requirements (2024)

## Rhode Island Comprehensive Assessment System (RICAS)

This document details rules for reporting the RICAS assessments for grades 3-8 ELA and Mathematics. The final student level data used for reporting is described in the “Data Processing Specifications.” This document is considered a draft until the Rhode Island Department of Education (RIDE) signs off. If there are rules that need to be added or modified after said sign-off, RIDE sign-off will be obtained for each such rule. These rules will be documented in the Addenda section at the end of the document.

### Table of Contents

<b>YEAR TO YEAR CHANGE HIGHLIGHTS:</b> .....	<b>3</b>
<b>I. CONTRACT OVERVIEW</b> .....	<b>3</b>
<b>A. Test Administration(s)</b> .....	<b>3</b>
<b>II. DELIVERABLES</b> .....	<b>3</b>
<b>A. PRELIMINARY REPORTING</b> .....	<b>3</b>
<b>B. FINAL REPORTING</b> .....	<b>3</b>
<b>III. INTERNAL DATA SOURCES</b> .....	<b>3</b>
<b>A. Test Information and Item Banking (Test Map)</b> .....	<b>3</b>
<i>i. Test Design</i> .....	<b>3</b>
<i>ii. Test Item</i> .....	<b>3</b>
<b>B. Item Metadata</b> .....	<b>3</b>
<b>C. School Information (iCore)</b> .....	<b>4</b>
<b>D. Scanned Data</b> .....	<b>4</b>
<b>E. PAN</b> .....	<b>4</b>
<b>F. TestNav</b> .....	<b>4</b>
<b>G. Valid Item Responses</b> .....	<b>4</b>
<b>H. Item Scoring</b> .....	<b>4</b>
<i>i. Machine-Scored Item Scores</i> .....	<b>4</b>
<i>ii. Single Score Open Response</i> .....	<b>4</b>
<i>iii. For the purposes of analyses Technology Enhanced Items (TEIs) are treated as open response items.</i> ....	<b>5</b>
<b>IV. EXTERNAL DATA SOURCES</b> .....	<b>5</b>
<b>A. Exceptions List</b> .....	<b>5</b>
<b>B. Exemptions List</b> .....	<b>5</b>
<b>C. eRIDE (Demographic File)</b> .....	<b>5</b>
<b>D. Student Growth Percentile (SGP)</b> .....	<b>5</b>
<b>V. DATA RECONCILIATION AUDITS</b> .....	<b>5</b>
<b>VI. STUDENT PARTICIPATION AND REPORTING STATUS</b> .....	<b>6</b>
<b>A. Basic Definitions</b> .....	<b>6</b>
<b>B. Participation Status Assignment Hierarchy (by subject)</b> .....	<b>6</b>
<b>C. Participation Status Summary</b> .....	<b>7</b>
<b>VII. CALCULATIONS</b> .....	<b>7</b>
<b>A. Rounding Rules</b> .....	<b>7</b>
<b>B. Student Level Calculations</b> .....	<b>7</b>
<b>C. Aggregate Calculations</b> .....	<b>8</b>
<b>REPORT DELIVERABLES SPECIFICATIONS</b> .....	<b>9</b>
<b>I. STUDENT REPORT</b> .....	<b>9</b>
<b>DATA FILE DELIVERABLES</b> .....	<b>10</b>
<b>I. MEGAFILE</b> .....	<b>10</b>
<b>REPORTING PRODUCTS: INTERNAL TO COGNIA</b> .....	<b>11</b>
<b>APPENDIX</b> .....	<b>11</b>

## Year to Year Change Highlights:

1. West Bay Christian Academy will not be participating in the RICAS in 2024.
2. Spotlight Videos will not be produced this year.

### I. Contract Overview

**Contract Code: 104750**

#### A. Test Administration(s)

Subject	Grade(s)	Mode
English Language Arts (ELA)	03-08	Paper, Online
Mathematics	03-08	Paper, Online

### II. Deliverables

#### A. Preliminary Reporting

1. Preliminary Reporting participation datafile-Used along with discrepancy reports.
2. Preliminary Megafile datafile-Used for growth calculations.

#### B. Final Reporting

1. Final Megafile datafile
2. Lookup Tables datafile
3. Item Statistics datafile
4. Testing Time datafile
5. Response Change Analysis from Pearson
6. Printed Individual Student Results Reports
7. Online Individual Student Results Reports

### III. Internal Data Sources

#### A. Test Information and Item Banking (Test Map)

##### i. Test Design

Grade	Subject	Test Mode	Items included in Raw Score
03-08	ELA	Paper	OP Items
		Paper_Braille	
		Online	
		Online_Accom	
03-08	Mat	Paper	OP Items
		Paper_Braille	
		Paper_Spanish	
		Online	
		Online_Accom	
Online_Spanish			

##### ii. Test Item

- An item consists of one or more student interactions. Items with more than one student interaction are a composite item. A composite item is treated as one item for analysis and reporting.
- Writing prompts are scored on two traits. Each trait score is treated as a separate item for scaling and item statistics. Additionally, the writing sum of the trait scores is included in the results file and statistics.

#### B. Item Metadata

- i. A table for each test contains the item order and item metadata for reporting and analysis. The item order is used to order item scores for student test results data files and student reports.
- ii. Data also includes flags indicating an item will be released.
- iii. Point values for OR and MC items as well as reporting categories are contained in these files.

**C. School Information (iCore)**

*i. School*

- Each school is identified by a unique 5-digit code.
- The district associated with the school is defined by a 2-digit district code.
  - a. Except for districts associated with Outplacement schools, which shall be defined as a 6-digit district code.

*ii. School Type*

- Public Schools are designated as OrgTypeID = 1. All others are considered non-public.

**D. Scanned Data**

Source for accommodations, absence, change of enrollment status, voided answer booklet for paper testers.

**E. PAN**

Both the online platform and paper forms serve as sources for accommodations, not tested reasons, and voided/invalidated student responses.

**F. TestNav**

Test administration platform collecting student’s item and test information for online testers.

**G. Valid Item Responses**

*i. Multiple Choice Scores – Scanning*

Valid multiple-choice scores are A, B, C, D, blank, and \* = multiple responses. All responses except “blank” are considered a response attempt.

*ii. Open Response / Short Answer Scores – Scoring*

Raw Data Value	Reported Value	Description	Point Value	Response Attempted
0-max pts	Final score	Open Response / Short Answer	0-max	ü
N	0	Not Scorable	0	ü
B		Blank response	0	

**H. Item Scoring**

*i. Machine-Scored Item Scores*

Valid machine scored items are scored by QTI. All responses except “blank” are considered a response attempt. The test map describes the scoring method for all items.

*ii. Single Score Open Response*

A student attempted an item if there is evidence of attempting at least one interaction for the item. The evidence depends on the scoring method for the interaction. An item could have multiple scoring methods.

- If a student earned 1 or more points for the item, the student attempted the item.
- If at least one interaction has a human score and the condition code is not scored BL (blank), the student attempted the item.
- If an interaction does not have a human score and the Response is not blank in the student item data, the student attempted the item.



- iii. For the purposes of analyses Technology Enhanced Items (TEIs) are treated as open response items.

Raw Data Value	Description	Reported Value	Point Value	Response Attempted
0-Max Possible Points	Student Points Earned	Numeric Score	0-Max Possible Points	ü
B	Blank response		0	

**Note:** In 2024 there was a typo in the Spanish online and paper versions of a multiple-choice item. RIDE made the decision to credit the affected students for the item.

#### IV. External Data Sources

##### A. Exceptions List

- i. Students with a test irregularity or who are considered a security breach are provided by RIDE in the Exceptions List. Instructions for processing and reporting each security breach student test are provided.
- ii. Data Analysis reviews the Exceptions List and adds necessary Amend flag values (> '1') and instructions to the Amend Code Definition Lookup for each distinct scenario on the Exceptions List.
- iii. Data Processing applies any necessary changes to the raw student record based on RIDE instructions and applies the corresponding Amend Flag value from the lookup to the student for Data Analysis.

##### B. Exemptions List

- i. RIDE to provide a list of students qualifying for test exemptions, i.e., first year EL students, medical exceptions.

##### C. eRIDE (Demographic File)

- i. Student data are provided by RIDE for reporting use following the file layout (also provided by RIDE.)

##### D. Student Growth Percentile (SGP)

- i. Student Growth Percentile are provided by RIDE for reporting using the agreed upon layout.

#### V. Data Reconciliation Audits

*The following cleanup will be performed on student level data prior to analysis once demographic data and reconciled test information are compiled to ensure consistency. Calculations are performed in the order listed below, and audited values are used in each subsequent check and for all analysis, reporting, and deliverables as applicable:*

##### A. ELL

- i. ELL - provided in eRIDE and will not be audited by Cognia.

##### B. Official School and Official District Code

###### i. Terminology:

- Test\_DiscCode = Cognia *Testing* DiscCode from DPRaw (See DP Specifications)
- Test\_SchCode = Cognia *Testing* SchCode from DPRaw (See DP Specifications)
- Resp\_DiscCode = *Responsible* (sending) district, source: eRIDE
- Resp\_Schcode = *Responsible* (sending) school, source: eRIDE

ii. **Official District**

- The official district is the responsible district from eRIDE

iii. **Official School**

The official school is the responsible school code from eRIDE. Missing official school and district information will be cleaned up during discrepancy period.

**VI. Student Participation and Reporting Status**

**A. Basic Definitions**

*The following criteria are defined for use during the participation status assignment hierarchy. Students may meet the criteria for multiple definitions, but during the hierarchy are assigned a single final participation status.*

i. **Standard Test Attemptedness** (by subject)

- A student is considered to have **met attemptedness** if they have a response attempt for at least one common item in each test session.
- A student is considered to have **partially attempted** if they have a response attempt to at least one common item, but they do not have a common item attempted in every session (has at least 1 session with no common items attempted).
- All other students are considered to have **not attempted**.

ii. **Not Tested Indicator – Medically Excused Absence**

The not tested code of “Medically Excused” may be bubbled on the student’s answer booklet or collected from PAN. However, RIDE will be considered the final, official source for this information.

iii. **Transfer**

- The transfer bubbles (Added and Removed) on the SRB are ignored.

iv. **Void** (by subject)

- Students whose only answer booklet (or from data provided in PAN) has been voided (Void[*sub*]= ‘1’) are considered Void. Booklets that remain void post-discrepancy are suppressed along with all corresponding data.

v. **Not Tested – Alternate Assessment**

- The alt flag in eRide is used to flag any students with no or partial attempt as Not Tested Alternate Assessment for a subject.
- Students flagged as Not Tested Alternate Assessment for both subjects will be reported in the Preliminary and final megafiles but do not receive a Student Report or results label.

**B. Participation Status Assignment Hierarchy** (by subject)

i. **Void or Invalidated** (PartStatus = ‘N’)

ii. **Final results:**

- a. If the student *meets attemptedness* then: **Tested** (PartStatus = ‘Z’).
- b. If the student partially attempted or did not attempt:
  - If medically excused then: **Not Tested - Medically Excused** (PartStatus=‘G’).
  - If student is first year ELL, then: **Not Tested – First Year ELL Student** (PartStatus = ‘E’)
  - If alternate assessment, then: **Not Tested – Alt Assessment** (PartStatus = ‘A’)
  - If other, then: **Not Tested – Other** (PartStatus = ‘N’)

### C. Participation Status Summary

Description	Part Status	Test Stat	Assigned a Scaled Score and Achievement Level
Tested	Z	1	Yes
Not Tested - Other	N	3	No
Not Tested – Medical Exception	G	2	No
Not tested – First Year ELL Student	E	4	No
Not Tested – Alternate Assessment	A	5	No

## VII. Calculations

### A. Rounding Rules

Calculation	Rounded (to the nearest)
Student Counts	Whole Number
Percentages	Whole Number
Student Growth Percentile Standard Error	Hundredths
Mean Growth Percentile	Whole Number

### B. Student Level Calculations

#### i. StudentID

- StudentID = rptStudentID from DPRaw (verified SASID).
- For non-demonstration students, if StudentID does not begin with '10' (or '20' for private school students) it was generated by DP for linking purposes and will be set to blank for reporting.

#### ii. Accommodations(tblStudemo)

- If a student did not attempt any items in a subject the corresponding raw accommodation indicator is *ignored* during the determination of accommodations. Otherwise, if a student attempts at least one item in a subject, the corresponding raw accommodation indicators are evaluated.
- Standard Accommodations:
  - a. Accom\_e = '1' if the student received any accommodations (not accessibility features) in ELA Reading Comprehension, otherwise set it to blank.
  - b. Accom\_m = '1' if the student received any accommodations (not accessibility features) in Math, otherwise set it to blank.

#### iii. Attempt Status (Attempt[sub])

- Attempt[sub] indicates if a student fully meets attemptedness, partially attempted, or did not attempt the subject, based on the definition of attemptedness in *section VI.A.i*
- Calculated for all participation statuses, values:
  - a. 'F' = Fully Meets Attemptedness
  - b. 'P' = Partial Attempt
  - c. 'N' = No Attempt

#### iv. ParentLetter

- Class Pack Identifiers (Cognia) for printing the School and Parent version of the Parent/GuardianReport are produced for all students with ParentLetter = '1'.
- If a student has a not tested reason for both subjects, ParentLetter='0'. The student does not receive a student report.

#### v. TestStat

- [e/m] TestStat is populated based on the student's participation status and is not dependent on scores. See the Participation Status Summary table for values.

vi. **Raw Scores**

- Overall Raw Score
  - a. The student's overall raw score is the sum of scores for all scaling items, including the writing composition task total scores (if applicable).
  - b. If a student has a partstatus of Not Tested or if the student did not attempt any items, then the raw score is set to blank after all subsequent calculations are complete.

vii. **Points Earned**

- MC points are based on common, scaling multiple choice or selected response items with point values stored in item metadata tables and OR points are based on non-MC or involving open-response and text entries.
- If a student does not receive reported raw scores these calculations are set to blank.

viii. **Item Responses**

- Re-formatted and re-ordered responses to all common items are reported and stored to support the student report and results file deliverables.
  - a. OR and MC items: the item score is reported.

ix. **Current-Year Reporting Results**

- **Scaled Score**
  - a. Current year scaled score results that are eligible to earn a current year scaled score (PartStatus = 'Z').
  - b. Blank for students not eligible to receive a scaled score based solely on participation status.
- **Performance Level**
  - a. Earned current-year achievement level based on scaledscore (1-4).
  - b. If the student does not receive an achievement level based solely on partstatus PerfLevel is blank.
  - c. Valid Values: numeric achievement levels (1-4) or blank.
- **Student Growth Percentile**
  - a. SGP will be blank for any student that does not receive a performance level in the current year.

**C. Aggregate Calculations**

i. **Aggregation Summary**

*These rules are applied to all aggregate calculations. Any additional rules specific to a particular calculation will be listed under the rules for the calculation.*

- All reporting levels (sch/dis/sta)
- Students are aggregated to their official school and official district.
- Students with public school subtypes are used for school-level aggregations.
- Students with a Home School Accountability status are excluded from all aggregate calculations and shall be reported individually.

ii. **Number and Percent of Students by Achievement Level**

- Calculated by grade and subject at the school, district, and state level.
- Calculations are performed using PerfLevel. All students with a non-blank PerfLevel are eligible to be included in the calculations.
- **Minimum N-Requirement:** if  $N < 10$  for a school or district results are calculated but suppressed from reports.

- School level aggregations are not reported if the school is in a district whose code begins with a “D” (outplacement). The school is instead included in the district level aggregations in the responsible district results.
- iii. Student Growth Percentile (SGP)**
- Students with an SGP value will be included in school, district, and state level.
  - Mean SGP will be calculated at the school and district levels.
  - School level mean SGP is not reported if the school is in a district whose code begins with a “D” (outplacement), should be included in responsible district aggregations.

## Report Deliverables Specifications

### I. Student Report

The following sections discuss the formatting, displays, and delivery for the Student Report. All calculations and aggregation rules can be found in earlier sections of this document.

#### Definitions

The following terms will be used to describe certain formats/behavior:

- i.* Data listed as being taken “from eRIDE” are taken from tblStuInfo, after any necessary audits are complete. If the student does not link to ERIDE the data are blank.
- ii.* Test mode – Displayed is which mode of test the student used.
  - “Paper-based test” or “Computer-based test”
  - If a student is not tested, then test mode will be blank on the student’s report.

#### A. Delivery

- iii.* Printed Student Reports
  - 2 printed copies of the student reports are shipped to the districts.
  - Student Reports are sent to the responsible school and responsible district.
  - Testing school is reported on the report.
  - Private school students are considered outplacement and reports should be sent back to the responsible school district. (190 schools where the Testing District begins with ‘D’)
    - a. There will be no school-level aggregations reported for these students (Blank)
    - b. Aggregations will still occur at the responsible district level.
    - c. Exception: West Bay Christian Academy – Student Reports need to be sent directly to the school (SchoolCode = 23334) since there is no responsible school district (School and District Aggregations shall be reported on Student Reports)
  - Home school students are reported individually and shall not be included in any aggregate reporting (State and District aggregations shall be included on the reports, but School level aggregations should be blanked)
  - Students who completed at least one assessment (either ELA or Math) will receive a Student Report. Otherwise, students who are exempted from or did not complete both tests will not receive a Student Report.
- iv.* Online Student Reports
  - Student Reports are available to the testing school and testing district for all students. For Outplacement students their student reports will be available to their responsible school and district AND their testing school and testing district.
  - Student Reports will be produced by SASID instead of by grade and school.

#### B. Cover Page

- v.* Title – “Spring 20YY RICAS Parent Guardian Report” where 20YY = test year, e.g. 2024.
- vi.* Student Name – Presented as proper case based on *LName*, *FName MI*. (with a period after the middle initial when the middle initial is not blank). Examples: SMITH, JOHN T. or JONES, JENNY

- This section requires special formatting when one or more of the names is missing:
  - a. If Lname is blank and Fname is blank, then section = “Blank Name”
  - b. If Lname is blank and Fname is not blank, then section = “Blank, Fname”
  - c. If Lname is not blank and Fname is blank then section = “Lname, Blank”
- vii. SASID – Student ID from eRIDE, no special formatting applied.
- viii. School Name and District Name – School and District names from iCore based on testing school. No special formatting applied.
- ix. Grade – Student’s tested grade will be used for all 03-08 reports.
- x. DOB – DOB from eRIDE, no special formatting. Must be equal to 10 characters in length (MM/DD/YYYY).

**C. Reporting Category Display**

- xi. Subject – Formatted with the following values:
  - If subject = ‘ela’ then ‘English Language Arts’
  - If subject = ‘mat’ then ‘Mathematics’
- xii. Points earned by your child – RawScore variables from tblStuRepCatPoints, no special formatting with RepCatID indicating the Reporting category RepOrder in daPointsPossible.

**D. Released Item Display**

- xiii. *Subject Title* and *Subject Ordering* follow the same rules as above.
- xiv. Order of rows within each grid
 

The following definitions are used to both describe what appears and also what is printed in item tables.

  - 1 = “Question Number” – this is the released item order number.
  - 2 = “Your Child’s Score” – this is the response provided by the student.
- xv. Formatting of Student Responses
  - a. ELA and Math data are displayed as the number of points earned out the total points possible (Ex. 1/1 or 0/3).
  - b. If a student did not answer an item, that cell will be blank.

**E. Student Achievement Level and Scaled Score Statements**

- xvi. Test Grade reference – “Your Child’s Overall Results in Grade [GG]” (where grade = student’s tested grade)
- xvii. Achievement level – contains either the achievement level text or the not-tested statement. This is set using the PerfLevel.
- xviii. Score – contains the reported scaled score.

- F. Historical scores-** Historical scores are reported in grades 4-8. Up to three years of scores are reported. The years reported in 2024 are 2024, 2023,2022 if available.
- The State level growth mean is fixed at 50.

**Data File Deliverables**

**I. Megafile**

- xix. The megafile contains all the student level results for RICAS. The file follows the ResultsLayout tab in the RICAS\_2324\_ReportingLayout.xlsx
- xx. The file is formatted as comma separated file, csv.

xxi. The layout is used for the participation file produced during the preliminary reporting period as well as the file used to provide scaled scores for the calculation of growth percentiles.

## Reporting Products: Internal to Cognia

<b>Contract Code: [104750] Description RICAS 2024 Admin ID 1</b>	<b>Report Type</b>	<b>Report For</b>	<b>Grade(s)</b>	<b>Report Subtype</b>	<b>Content Code</b>	<b>Qty</b>
Student Report Parent Copy	07	1	03-08	02	00	1
Student report School Copy	07	1	03-08	01	00	1

## Appendix

### A. RICAS daReportingCategoryLookup

*This table lists the updated item reporting categories by subject. The actual reporting categories that exist do vary by grade and possibly year. These categories are sorted alphabetically and do not necessarily reflect position or sort order within a grade.*

<b>Subject</b>	<b>Grade</b>	<b>RepCat Sort Order</b>	<b>Student Report Text: (RepCatText)</b>
<b>ELA</b>	03-08	1	Reading
		2	Language
		3	Writing
<b>Mathematics</b>	03-05	1	Operations and Algebraic Thinking
		2	Number and Operations in Base Ten
		3	Number and Operations-Fractions
		4	Measurement and Data
		5	Geometry
<b>Mathematics</b>	06-07	1	Ratios and Proportional Relationships
		2	The Number System
		3	Expressions and Equations
		4	Geometry
		5	Statistics and Probability
<b>Mathematics</b>	08	1	The Number System and Expressions and Equations
		2	Functions
		3	Geometry
		4	Statistics and Probability

**APPENDIX G**  
**ITEM-LEVEL CLASSICAL STATISTICS**



**Table G-1. Item-Level Classical Test Theory Statistics—ELA Grade 3**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
EL206632601	SR	0.68	0.53	0
EL206633507	SR	0.66	0.51	0
EL206635137	SR	0.74	0.42	0
EL206635666	SR	0.64	0.46	0
EL206635928	SR	0.72	0.50	0
EL206636183	SR	0.76	0.45	0
EL206636821	SR	0.74	0.48	0
EL206636985	SR	0.70	0.51	0
EL206638372	SR	0.60	0.38	0
EL206639826	OR	0.83	0.50	1
EL206640503	OR	0.49	0.37	1
EL206642581#SCORE_TRAIT_Conv	ES	0.33	0.71	1
EL206642581#SCORE_TRAIT_Ideadev	ES	0.28	0.66	1
EL006852708	SR	0.58	0.34	0
EL006853420	SR	0.43	0.43	0
EL006853709	SR	0.43	0.29	0
EL007171237	SR	0.52	0.41	0
EL007171625	SR	0.43	0.30	0
EL007253254	SR	0.70	0.50	0
EL008648215	OR	0.46	0.56	1
EL206247438	SR	0.33	0.25	0
EL206248298	OR	0.50	0.41	0
EL206252440	SR	0.44	0.34	0
EL206252891	SR	0.49	0.39	0
EL206440619	SR	0.68	0.39	0
EL206455922	OR	0.32	0.56	0
EL206538101	SR	0.75	0.48	0
EL206538400	SR	0.60	0.38	0
EL206539824	SR	0.35	0.36	0
EL206636010	OR	0.59	0.36	0
EL206658239	SR	0.66	0.40	0
EL206658538	SR	0.49	0.30	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-2. Item-Level Classical Test Theory Statistics—ELA Grade 4**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
EL909132428#SCORE_TRAIT_Conv	ES	0.38	0.64	1
EL909132428#SCORE_TRAIT_Ideadev	ES	0.24	0.56	1
EL909139013	SR	0.75	0.34	0
EL909146543	SR	0.44	0.31	0
EL909147325	SR	0.49	0.31	0
EL909150609	SR	0.65	0.39	0
EL909151025	SR	0.40	0.29	0
EL909152457	OR	0.55	0.39	0
EL909153399	OR	0.52	0.42	0
EL909155188	SR	0.67	0.18	0
EL909156595	SR	0.53	0.38	0
EL913042685	SR	0.55	0.34	--
EL913366686	SR	0.78	0.46	0
EL297747	SR	0.58	0.35	--
EL297749	SR	0.67	0.50	0
EL297750	OR	0.56	0.60	0
EL297754	SR	0.53	0.34	--
EL909952647	SR	0.46	0.29	--
EL909966060	SR	0.68	0.38	0
EL911875859	OR	0.36	0.61	0
EL911905412	SR	0.86	0.47	0
EL911907107	OR	0.69	0.58	0
EL911909187	SR	0.43	0.19	0
EL911910431	SR	0.51	0.30	0
EL911910990	SR	0.74	0.40	0
EL911919086	SR	0.75	0.48	0
EL915149134	SR	0.66	0.46	--
EL915167319	SR	0.57	0.46	0
EL915169315	OR	0.44	0.48	--
EL915543104	SR	0.51	0.43	0
EL915543753	SR	0.58	0.53	0
EL915545933	SR	0.46	0.31	--

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-3. Item-Level Classical Test Theory Statistics—ELA Grade 5**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
EL207033428	SR	0.83	0.43	--
EL207035725	SR	0.65	0.32	0
EL207037283	SR	0.62	0.42	--
EL207039028	SR	0.57	0.42	0
EL207041852	SR	0.46	0.31	0
EL207043067	OR	0.51	0.54	0
EL207128330	SR	0.53	0.46	0
EL207128708	OR	0.60	0.54	0
EL207129544	SR	0.60	0.48	0
EL207135604#SCORE_TRAIT_Conv	ES	0.48	0.72	0
EL207135604#SCORE_TRAIT_Ideadev	ES	0.38	0.74	0
EL207984109	SR	0.63	0.52	0
EL208106818	SR	0.64	0.29	--
EL206053008	SR	0.83	0.49	--
EL206057533#SCORE_TRAIT_Conv	ES	0.44	0.72	0
EL206057533#SCORE_TRAIT_Ideadev	ES	0.29	0.71	0
EL206060317	SR	0.48	0.45	0
EL206068188	SR	0.72	0.34	0
EL206085599	SR	0.88	0.48	--
EL206100724	SR	0.72	0.48	--
EL206127119	SR	0.72	0.41	0
EL206130727	SR	0.83	0.47	0
EL206142184	SR	0.69	0.43	0
EL208382001	OR	0.37	0.53	0
EL208451664	SR	0.46	0.37	--
EL208452118	SR	0.65	0.56	0
EL208452922	SR	0.62	0.42	--
EL208571752	SR	0.61	0.51	0
EL208575656	SR	0.57	0.48	0
EL208577435	SR	0.43	0.30	0
EL209368842	OR	0.67	0.56	--
EL209442821	SR	0.63	0.23	0
EL209449008	OR	0.67	0.63	--

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-4. Item-Level Classical Test Theory Statistics—ELA Grade 6**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
EL206054907	SR	0.65	0.44	0
EL206055338	SR	0.77	0.36	0
EL206069631	SR	0.75	0.52	0
EL206071655	SR	0.35	0.32	0
EL206072413	SR	0.80	0.37	--
EL206073117	SR	0.55	0.39	--
EL206131013	SR	0.60	0.41	0
EL206144523#SCORE_TRAIT_Conv	ES	0.51	0.78	1
EL206144523#SCORE_TRAIT_Ideadev	ES	0.31	0.78	1
EL206147428	OR	0.50	0.29	0
EL206156401	SR	0.62	0.43	0
EL209683892	OR	0.33	0.39	0
EL211482852	SR	0.63	0.39	0
EL027853702	SR	0.74	0.46	0
EL027973437	SR	0.41	0.38	0
EL028044247	SR	0.69	0.46	0
EL028171110	SR	0.71	0.44	0
EL028261263	SR	0.67	0.37	0
EL028278272	SR	0.57	0.47	0
EL028283260	SR	0.67	0.55	0
EL028337921	OR	0.64	0.55	0
EL028340179	OR	0.56	0.50	0
EL028464862	SR	0.46	0.33	0
EL028474800	SR	0.28	0.34	--
EL028556260#SCORE_TRAIT_Conv	ES	0.36	0.72	1
EL028556260#SCORE_TRAIT_Ideadev	ES	0.19	0.65	1
EL208346712	SR	0.56	0.36	0
EL208352214	SR	0.74	0.41	0
EL208431592	SR	0.52	0.35	0
EL208956821	SR	0.45	0.29	--
EL215337440	SR	0.66	0.47	0
EL216263506	OR	0.48	0.38	0
EL216269224	SR	0.62	0.47	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-5. Item-Level Classical Test Theory Statistics—ELA Grade 7**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
EL208449230	SR	0.52	0.53	0
EL208449595	SR	0.58	0.39	0
EL208451648	SR	0.40	0.39	0
EL208452611	SR	0.49	0.41	0
EL208475824	SR	0.81	0.40	0
EL208476152	SR	0.54	0.45	0
EL208579240	SR	0.48	0.38	0
EL208580556	OR	0.47	0.45	0
EL208734693	OR	0.53	0.52	0
EL208736539	SR	0.60	0.53	0
EL213654472	SR	0.67	0.41	0
EL217467654#SCORE_TRAIT_Conv	ES	0.54	0.78	1
EL217467654#SCORE_TRAIT_Ideadev	ES	0.34	0.79	1
EL206038392	SR	0.74	0.30	--
EL206040879	SR	0.61	0.42	0
EL206046704	SR	0.55	0.41	0
EL206048853	SR	0.87	0.41	--
EL206049453	SR	0.66	0.37	0
EL206054058	SR	0.46	0.24	0
EL206054673	SR	0.62	0.42	0
EL206073551	SR	0.82	0.45	0
EL206141630	OR	0.28	0.31	0
EL206144041#SCORE_TRAIT_Conv	ES	0.60	0.76	1
EL206144041#SCORE_TRAIT_Ideadev	ES	0.39	0.75	1
EL206145148	OR	0.60	0.41	0
EL206735123	SR	0.64	0.49	0
EL207653069	SR	0.35	0.27	0
EL208035502	SR	0.60	0.39	--
EL208037687	SR	0.55	0.40	--
EL208038191	SR	0.42	0.28	--
EL208038966	SR	0.45	0.33	0
EL213027859	SR	0.53	0.41	0
EL217355629	OR	0.49	0.46	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-6. Item-Level Classical Test Theory Statistics—ELA Grade 8**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
EL809713456	OR	0.56	0.65	0
EL809734614	SR	0.65	0.46	0
EL809863460	SR	0.66	0.53	0
EL810133273	SR	0.63	0.46	0
EL810180368	SR	0.66	0.48	0
EL810222585	OR	0.45	0.48	0
EL810439521	SR	0.60	0.50	0
EL810456981	SR	0.47	0.32	0
EL810463548#SCORE_TRAIT_Conv	ES	0.49	0.82	2
EL810463548#SCORE_TRAIT_Ideadev	ES	0.33	0.80	2
EL812838757	SR	0.54	0.45	0
EL812849329	SR	0.70	0.45	0
EL812849633	SR	0.73	0.48	0
EL007242472	SR	0.53	0.35	0
EL007242520	SR	0.57	0.54	0
EL007242572	SR	0.40	0.38	0
EL007954597	OR	0.37	0.44	0
EL206761648#SCORE_TRAIT_Conv	ES	0.56	0.78	2
EL206761648#SCORE_TRAIT_Ideadev	ES	0.36	0.79	2
EL206773007	SR	0.67	0.37	0
EL206773521	SR	0.71	0.44	0
EL206837355	SR	0.66	0.28	0
EL206837795	SR	0.48	0.22	--
EL206839116	SR	0.65	0.45	0
EL206840059	SR	0.56	0.49	0
EL206840381	SR	0.56	0.35	0
EL206843141	SR	0.64	0.35	0
EL206844860	SR	0.49	0.37	0
EL206847632	OR	0.57	0.53	0
EL208942592	SR	0.43	0.25	0
EL208943016	SR	0.53	0.39	0
EL208944699	SR	0.62	0.38	0
EL208945360	OR	0.55	0.57	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-7. Item-Level Classical Test Theory Statistics—Mathematics Grade 3**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
MA287143	SR	0.55	0.46	0
MA310868	SR	0.78	0.38	--
MA713744192	OR	0.79	0.48	0
MA834632167	SR	0.56	0.39	0
MA900360943	SR	0.62	0.48	0
MA900438114	OR	0.80	0.46	0
MA900439380	OR	0.73	0.46	0
MA900571833	SR	0.70	0.43	0
MA900577259	OR	0.68	0.47	0
MA900579739	OR	0.52	0.73	0
MA902576979	SR	0.40	0.35	0
MA001034993	OR	0.70	0.54	0
MA001330803	OR	0.55	0.59	0
MA001444510	OR	0.58	0.45	0
MA002746011	OR	0.38	0.30	0
MA134765618	OR	0.30	0.39	1
MA134826027	OR	0.37	0.55	1
MA134834397	OR	0.53	0.50	0
MA306330	SR	0.66	0.48	0
MA901132847	OR	0.39	0.71	0
MA297399A	OR	0.54	0.71	0
MA703080328	OR	0.23	0.34	0
MA306360A	OR	0.34	0.65	1
MA734752687	SR	0.59	0.32	0
MA802238054	OR	0.47	0.43	0
MA900368129	SR	0.61	0.33	0
MA900371163	SR	0.68	0.51	0
MA900375580	SR	0.76	0.52	0
MA900734747	OR	0.68	0.58	0
MA001044551	SR	0.35	0.46	--
MA001061117	OR	0.83	0.36	0
MA002740850	OR	0.26	0.39	0
MA002753131	OR	0.37	0.46	0
MA002150122	SR	0.57	0.40	0
MA134772227	OR	0.72	0.52	0
MA134775263	OR	0.50	0.43	1
MA134776400	SR	0.74	0.53	--
MA134935661	OR	0.38	0.57	0
MA735660441	SR	0.86	0.36	--
MA935135473	OR	0.90	0.35	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-8. Item-Level Classical Test Theory Statistics—Mathematics Grade 4**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
MA221898	OR	0.61	0.49	0
MA311538	SR	0.75	0.45	0
MA714226023	OR	0.55	0.48	0
MA736379879	SR	0.75	0.41	--
MA736459765	OR	0.24	0.40	0
MA900750814	OR	0.72	0.63	0
MA900845882	OR	0.64	0.55	0
MA903678056	SR	0.46	0.41	0
MA003543807	OR	0.48	0.44	0
MA003746997	SR	0.82	0.35	0
MA803740205A	OR	0.71	0.35	0
MA002231198	OR	0.52	0.54	0
MA200156568	SR	0.40	0.43	0
MA200257075	SR	0.56	0.45	0
MA301825	SR	0.72	0.43	--
MA304983A	OR	0.47	0.73	0
MA713963514	OR	0.45	0.58	0
MA800635117	OR	0.48	0.65	0
MA800778188	SR	0.73	0.42	0
MA258255	SR	0.51	0.42	0
MA311529	SR	0.44	0.34	0
MA713673616	OR	0.42	0.53	0
MA800765452	SR	0.25	0.26	0
MA803975569	OR	0.43	0.56	0
MA302496A	OR	0.43	0.68	0
MA900669607	SR	0.58	0.46	0
MA900682939	OR	0.65	0.36	0
MA900782440	SR	0.32	0.22	0
MA903054837	OR	0.47	0.51	0
MA903352782	OR	0.39	0.37	0
MA001748629	OR	0.78	0.51	0
MA002143728	SR	0.47	0.49	0
MA002333692	OR	0.82	0.43	0
MA002334822	OR	0.54	0.50	0
MA200150657	OR	0.51	0.53	0
MA311557	SR	0.66	0.39	0
MA714257493	OR	0.43	0.56	0
MA803833224	OR	0.86	0.59	0
MA903767905	SR	0.63	0.52	--
MA286781	SR	0.45	0.33	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.



**Table G-9. Item-Level Classical Test Theory Statistics—Mathematics Grade 5**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
MA624355953	OR	0.49	0.47	0
MA301160	SR	0.27	0.31	0
MA715102062	OR	0.47	0.38	0
MA715102381	OR	0.64	0.45	0
MA301147	SR	0.68	0.58	0
MA900362696	OR	0.52	0.57	--
MA900371480	OR	0.70	0.56	0
MA900885042	OR	0.30	0.20	0
MA000827038	OR	0.30	0.58	0
MA000929280	SR	0.69	0.28	0
MA001143716	OR	0.29	0.46	0
MA306437A	OR	0.30	0.38	0
MA000950721	OR	0.68	0.53	0
MA201546463	SR	0.48	0.43	0
MA306459	SR	0.59	0.42	--
MA311299	SR	0.78	0.39	--
MA715102421	OR	0.41	0.58	0
MA900353683	SR	0.74	0.47	0
MA900643667	SR	0.73	0.46	0
MA903948538	OR	0.57	0.69	0
MA298010	OR	0.89	0.29	0
MA715102279	OR	0.24	0.45	0
MA301831	SR	0.29	0.40	0
MA802358485	OR	0.40	0.65	0
MA804341327	OR	0.38	0.43	0
MA900647188	OR	0.57	0.48	0
MA900660286	SR	0.51	0.30	0
MA903732126	SR	0.54	0.49	0
MA904134306	SR	0.51	0.48	0
MA904247425	SR	0.74	0.42	0
MA904338797	OR	0.61	0.58	0
MA002829660	OR	0.44	0.71	0
MA006260011	OR	0.68	0.45	0
MA001040409	SR	0.54	0.26	0
MA201757235	SR	0.41	0.29	0
MA201858644	OR	0.55	0.39	0
MA201936073	OR	0.75	0.40	0
MA311286	SR	0.66	0.44	0
MA801221646	SR	0.45	0.39	--
MA903746888	OR	0.36	0.64	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-10. Item-Level Classical Test Theory Statistics—Mathematics Grade 6**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
MA624253312	OR	0.44	0.65	0
MA307295	OR	0.26	0.70	1
MA311660	SR	0.37	0.37	0
MA703176747	OR	0.23	0.34	0
MA311694	OR	0.52	0.67	0
MA736359148	SR	0.39	0.31	--
MA900284082	SR	0.32	0.26	0
MA900470149	SR	0.70	0.30	0
MA900540139	SR	0.27	0.46	0
MA001557959	OR	0.63	0.41	0
MA002527442	SR	0.40	0.32	0
MA001551849	SR	0.62	0.37	0
MA001574310	OR	0.63	0.54	0
MA002532873	OR	0.36	0.62	0
MA002536026	SR	0.30	0.32	0
MA003476915	SR	0.75	0.41	--
MA136042342	OR	0.34	0.33	0
MA136079412	OR	0.76	0.35	0
MA136282011	OR	0.64	0.44	0
MA136500757	OR	0.54	0.49	0
MA703178216	OR	0.33	0.41	0
MA272299	SR	0.77	0.46	0
MA736443588	OR	0.47	0.63	0
MA805276878	SR	0.43	0.11	0
MA736361568	SR	0.60	0.44	0
MA800365824	OR	0.45	0.45	0
MA805276792	OR	0.27	0.59	0
MA900437517	OR	0.33	0.40	0
MA900462230	OR	0.54	0.58	0
MA902635513	OR	0.74	0.48	0
MA001528147	SR	0.32	0.45	0
MA001545119	SR	0.56	0.34	0
MA001555259	SR	0.75	0.40	--
MA001555876	OR	0.62	0.52	0
MA003843559	OR	0.19	0.67	1
MA136077359	SR	0.35	0.32	0
MA136463214	OR	0.20	0.71	1
MA219859831	OR	0.48	0.44	0
MA298162	OR	0.57	0.56	0
MA311690	SR	0.53	0.48	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-11. Item-Level Classical Test Theory Statistics—Mathematics Grade 7**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
MA311103	SR	0.49	0.37	0
MA311075	OR	0.48	0.64	0
MA703930934	OR	0.44	0.27	0
MA713848042	OR	0.40	0.65	0
MA804458974	OR	0.35	0.49	0
MA289832	SR	0.52	0.45	0
MA715009356	OR	0.38	0.73	2
MA801082691	OR	0.32	0.26	0
MA900559852	SR	0.39	0.29	0
MA900635073	OR	0.26	0.32	0
MA900736982	OR	0.51	0.52	0
MA907259763	OR	0.17	0.42	0
MA200155755	OR	0.23	0.52	1
MA201077581	OR	0.70	0.40	0
MA289853	OR	0.35	0.71	1
MA801371719	OR	0.47	0.37	0
MA804440567	OR	0.28	0.48	0
MA272872	SR	0.63	0.28	0
MA287437	SR	0.45	0.51	--
MA306621	SR	0.45	0.51	0
MA228036	SR	0.50	0.34	0
MA205602	SR	0.27	0.30	0
MA295758	OR	0.40	0.74	1
MA713848086	OR	0.20	0.54	0
MA801600966	SR	0.33	0.28	--
MA804653792	OR	0.55	0.47	0
MA900647693	OR	0.23	0.58	1
MA904158907	SR	0.56	0.36	0
MA005076023	OR	0.63	0.51	0
MA005169543	SR	0.50	0.46	0
MA007839825	OR	0.26	0.73	1
MA804637729	OR	0.39	0.59	0
MA001385352	OR	0.20	0.60	0
MA200915673	OR	0.40	0.66	0
MA201085794	SR	0.40	0.58	0
MA288411	SR	0.55	0.39	--
MA311086	SR	0.37	0.16	0
MA801379630	OR	0.28	0.49	0
MA904183531	SR	0.49	0.35	0
MA904260053	OR	0.49	0.31	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**Table G-12. Item-Level Classical Test Theory Statistics—Mathematics Grade 8**

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
MA715919523	OR	0.32	0.65	0
MA228246	SR	0.51	0.50	--
MA272928	SR	0.56	0.58	0
MA311419	OR	0.15	0.52	0
MA311424	SR	0.39	0.31	0
MA311452	OR	0.60	0.65	0
MA804152353	OR	0.53	0.61	0
MA800475031	OR	0.62	0.44	0
MA901142533	SR	0.49	0.40	0
MA902303951	SR	0.54	0.53	0
MA001865862	OR	0.71	0.50	0
MA002159901	SR	0.57	0.46	0
MA003777138	SR	0.71	0.39	0
MA001738779	OR	0.50	0.51	0
MA201642904	OR	0.38	0.44	1
MA202931621	SR	0.52	0.57	0
MA800571638	SR	0.51	0.35	0
MA901139952	SR	0.46	0.46	0
MA902264555	OR	0.28	0.57	1
MA902364058	OR	0.31	0.71	0
MA284774	OR	0.41	0.79	2
MA298049	SR	0.80	0.39	0
MA284198	SR	0.61	0.43	0
MA800974248	OR	0.35	0.61	0
MA900545813	SR	0.45	0.29	0
MA902268353	OR	0.28	0.52	0
MA902303258	OR	0.42	0.54	0
MA905906652	OR	0.34	0.47	0
MA004058494	SR	0.52	0.38	0
MA001734156	SR	0.48	0.33	0
MA002239631	OR	0.36	0.35	1
MA002275463	SR	0.30	0.22	0
MA003955575	SR	0.29	0.50	--
MA201656350	SR	0.57	0.36	0
MA202724888	SR	0.67	0.45	0
MA203081760	OR	0.25	0.73	3
MA800977322	SR	0.36	0.39	0
MA804577289	OR	0.45	0.54	0
MA307537	SR	0.38	0.23	0
MA307560	SR	0.54	0.43	0

*Note.* Blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

**APPENDIX H**  
**SCORE DISTRIBUTIONS**

**Table H-1. Item-Level Score Distributions for SR and OR Items and ESs—ELA**

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
3	EL206639826	2	3.73	23.15	71.71			
	EL206640503	2	37.11	25.49	36.73			
	EL206642581#SCORE_TRAIT_Conv	3	33.49	36.62	25.20	3.75		
	EL206642581#SCORE_TRAIT_Ideadev	4	23.40	44.18	26.80	4.46	0.23	
	EL008648215	2	34.25	38.81	26.40			
	EL206248298	2	41.49	17.90	40.56			
	EL206455922	3	25.74	55.07	15.40	3.32		
EL206636010	2	27.95	24.99	46.73				
4	EL909132428#SCORE_TRAIT_Conv	3	18.83	50.85	24.88	4.83		
	EL909132428#SCORE_TRAIT_Ideadev	4	38.64	32.44	20.74	6.31	1.26	
	EL909152457	2	33.81	22.72	43.42			
	EL909153399	2	38.47	18.80	42.72			
	EL297750	2	29.38	29.60	41.00			
	EL911875859	3	29.78	37.17	26.84	5.90		
	EL911907107	2	14.52	32.32	52.76			
EL915169315	2	49.47	13.48	37.05				
5	EL207043067	2	38.21	20.82	40.95			
	EL207128708	2	22.71	34.13	43.12			
	EL207135604#SCORE_TRAIT_Conv	3	18.31	34.21	31.88	15.18		
	EL207135604#SCORE_TRAIT_Ideadev	4	15.71	34.33	33.18	14.39	1.97	
	EL206057533#SCORE_TRAIT_Conv	3	20.59	36.31	32.34	10.29		
	EL206057533#SCORE_TRAIT_Ideadev	4	33.87	29.33	25.15	9.37	1.81	
	EL208382001	2	54.81	16.06	29.10			
EL209368842	2	26.61	13.06	60.32				
EL209449008	2	21.92	23.05	55.03				
6	EL206144523#SCORE_TRAIT_Conv	3	18.58	32.93	23.07	24.52		
	EL206144523#SCORE_TRAIT_Ideadev	5	23.16	29.55	22.10	17.60	6.19	0.51
	EL206147428	2	46.06	7.94	45.98			
	EL209683892	2	47.25	39.60	13.13			
	EL028337921	2	23.60	24.29	52.10			
	EL028340179	2	33.60	20.49	45.89			
	EL028556260#SCORE_TRAIT_Conv	3	31.57	40.88	14.34	12.47		
EL028556260#SCORE_TRAIT_Ideadev	5	62.33	8.41	11.02	9.22	5.55	2.73	
EL216263506	2	31.17	42.54	26.27				
7	EL208580556	2	44.54	17.13	38.29			
	EL208734693	2	31.91	29.22	38.84			
	EL217467654#SCORE_TRAIT_Conv	3	20.81	21.66	29.58	26.64		
	EL217467654#SCORE_TRAIT_Ideadev	5	23.23	18.90	31.41	16.04	7.51	1.59
	EL206141630	2	55.85	32.58	11.56			
	EL206144041#SCORE_TRAIT_Conv	3	12.11	25.11	30.91	30.90		
	EL206144041#SCORE_TRAIT_Ideadev	5	14.35	22.68	29.35	19.31	9.96	3.38
EL206145148	2	14.83	50.64	34.52				
EL217355629	2	32.70	36.24	31.05				
8	EL809713456	2	29.12	30.61	40.22			
	EL810222585	2	27.58	55.21	16.91			
	EL810463548#SCORE_TRAIT_Conv	3	22.23	27.09	27.08	21.83		
	EL810463548#SCORE_TRAIT_Ideadev	5	23.89	25.44	21.93	16.83	7.39	2.73
	EL007954597	2	50.24	25.27	24.48			
	EL206761648#SCORE_TRAIT_Conv	3	18.93	22.71	25.57	30.99		
	EL206761648#SCORE_TRAIT_Ideadev	5	19.90	21.30	26.96	19.04	8.68	2.32
EL206847632	2	26.52	32.34	41.13				
EL208945360	2	27.58	33.81	38.57				

**Table H-2 Item-Level Score Distributions for SR and OR Items—Mathematics**

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
3	MA713744192	1	20.22	79.46				
	MA900438114	1	19.80	80.14				
	MA900439380	1	26.92	72.84				
	MA900577259	1	31.98	67.65				
	MA900579739	3	21.69	26.56	25.33	26.17		
	MA001034993	1	29.69	70.19				
	MA001330803	1	45.19	54.58				
	MA001444510	1	42.35	57.50				
	MA002746011	1	61.12	38.40				
	MA134765618	1	69.58	29.59				
	MA134826027	1	62.91	36.53				
	MA134834397	1	46.88	53.02				
	MA901132847	3	35.50	26.00	23.53	14.58		
	MA297399A	3	14.70	29.80	33.60	21.64		
	MA703080328	1	77.19	22.69				
	MA306360A	3	39.76	28.54	21.40	9.73		
	MA802238054	1	52.52	47.00				
	MA900734747	1	32.04	67.89				
	MA001061117	1	16.85	83.10				
	MA002740850	1	73.99	25.88				
	MA002753131	1	63.34	36.51				
	MA134772227	1	27.73	72.22				
	MA134775263	1	49.31	50.11				
	MA134935661	1	62.24	37.64				
	MA935135473	1	10.33	89.57				
	4	MA221898	1	38.84	61.07			
		MA714226023	2	17.52	55.15	26.96		
		MA736459765	1	75.71	24.18			
MA900750814		4	6.32	7.66	16.04	31.26	38.66	
MA900845882		1	35.87	64.06				
MA003543807		1	51.47	48.21				
MA803740205A		1	29.36	70.56				
MA002231198		1	48.25	51.65				
MA304983A		4	23.44	20.05	18.69	18.31	19.16	
MA713963514		1	54.88	45.00				
MA800635117		1	52.11	47.81				
MA713673616		1	58.37	41.59				
MA803975569		1	56.58	43.16				
MA302496A		4	14.76	33.47	22.95	20.65	7.91	
MA900682939		1	34.46	65.43				
MA903054837		1	52.67	47.25				
MA903352782		1	60.73	39.04				
MA001748629		1	22.03	77.65				
MA002333692		1	18.04	81.91				
MA002334822		1	45.16	54.46				
MA200150657		1	49.23	50.73				
MA714257493		2	27.88	57.61	14.28			
MA803833224		4	2.01	3.55	5.61	24.06	64.74	
5		MA624355953	2	23.88	53.72	22.30		
	MA715102062	1	53.12	46.53				
	MA715102381	1	36.19	63.68				
	MA900362696	1	48.32	51.68				
	MA900371480	1	29.98	69.83				
	MA900885042	1	69.48	30.40				

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
5	MA000827038	1	69.94	30.03				
	MA001143716	1	70.97	28.88				
	MA306437A	1	69.87	30.09				
	MA000950721	1	31.81	67.94				
	MA715102421	4	6.50	51.56	21.56	12.84	7.26	
	MA903948538	4	9.14	12.98	33.19	28.25	16.20	
	MA298010	1	10.55	89.34				
	MA715102279	1	76.08	23.77				
	MA802358485	2	41.31	37.01	21.60			
	MA804341327	1	61.55	38.19				
	MA900647188	1	43.11	56.80				
	MA904338797	1	38.92	61.05				
	MA002829660	4	20.11	23.69	24.96	20.68	10.18	
	MA006260011	1	31.64	68.19				
	MA201858644	1	45.05	54.89				
	MA201936073	1	25.09	74.81				
	MA903746888	4	27.67	29.81	19.73	17.06	5.57	
6	MA624253312	2	39.14	34.06	26.74			
	MA307295	4	44.52	27.72	11.82	6.63	8.42	
	MA703176747	1	76.43	23.42				
	MA311694	4	18.17	18.23	24.22	14.77	24.30	
	MA001557959	1	36.73	63.25				
	MA001574310	1	36.65	63.27				
	MA002532873	1	64.31	35.65				
	MA136042342	1	65.78	34.16				
	MA136079412	1	24.24	75.69				
	MA136282011	1	36.01	63.97				
	MA136500757	1	45.48	54.38				
	MA703178216	1	66.88	32.76				
	MA736443588	2	35.21	35.23	29.48			
	MA800365824	1	54.58	45.32				
	MA805276792	1	72.63	27.18				
	MA900437517	1	67.02	32.93				
	MA900462230	1	45.86	54.07				
	MA902635513	1	26.22	73.53				
	MA001555876	1	37.50	62.42				
	MA003843559	4	61.43	16.16	8.56	7.50	5.53	
MA136463214	4	59.89	18.81	7.39	6.32	6.51		
MA219859831	1	51.92	47.92					
MA298162	1	43.10	56.76					
7	MA311075	1	51.53	48.38				
	MA703930934	2	29.27	53.00	17.69			
	MA713848042	1	60.07	39.77				
	MA804458974	1	64.56	35.16				
	MA715009356	4	35.50	22.99	11.81	9.07	19.08	
	MA801082691	1	67.88	32.03				
	MA900635073	1	73.64	26.21				
	MA900736982	1	48.84	50.99				
	MA907259763	1	82.85	16.77				
	MA200155755	1	76.03	23.24				
	MA201077581	1	30.03	69.83				
	MA289853	4	29.65	21.98	28.62	14.17	4.64	
	MA801371719	1	52.66	47.17				
	MA804440567	1	71.41	28.47				
	MA295758	4	12.95	44.83	19.28	11.61	10.58	
	MA713848086	1	79.62	20.23				
	MA804653792	1	44.15	55.46				

continued



Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
7	MA900647693	1	75.81	23.22				
	MA005076023	1	36.86	63.07				
	MA007839825	4	33.86	40.89	14.21	3.47	6.56	
	MA804637729	1	61.27	38.65				
	MA001385352	1	79.72	20.06				
	MA200915673	2	42.98	33.82	23.02			
	MA801379630	1	71.74	28.20				
MA904260053	1	51.02	48.95					
8	MA715919523	1	67.77	32.09				
	MA311419	1	84.68	15.08				
	MA311452	4	10.49	17.25	19.10	26.73	26.04	
	MA804152353	2	29.55	34.65	35.48			
	MA800475031	1	38.06	61.90				
	MA001865862	1	28.76	71.02				
	MA001738779	1	50.22	49.61				
	MA201642904	1	61.35	38.05				
	MA902264555	1	71.66	27.60				
	MA902364058	4	21.14	50.15	16.75	7.43	4.26	
	MA284774	4	25.57	25.91	17.60	14.93	14.10	
	MA800974248	1	65.13	34.83				
	MA902268353	2	52.28	38.36	9.26			
	MA902303258	1	58.27	41.50				
	MA905906652	1	66.24	33.59				
	MA002239631	1	63.53	35.84				
	MA203081760	4	47.83	20.75	12.44	6.72	9.02	
MA804577289	1	55.15	44.69					

**APPENDIX I**  
**DIFFERENTIAL ITEM FUNCTIONING RESULTS**

**Table I-1. Number of Items Classified as “Low” or “High” DIF, Overall and by Group Favored—ELA**

Grade	Reference	Group	Focal	Item Type	Number of Items	Total	Number “Low” Favoring		Total	Number “High” Favoring	
							Reference	Focal		Reference	Focal
3	Male	Female	MC	24	1	1	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	Not ELL	ELL	MC	24	3	3	0	1	1	0	0
			OR	6	3	3	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	1	1	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	White	African American	MC	24	3	2	1	0	0	0	0
			OR	6	1	1	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
		Hispanic / Latino	MC	24	2	2	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	0	0	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	Online	Paper	MC	24	0	0	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
4	Male	Female	MC	24	3	3	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	2	0	2	0	0	0	0
	Not ELL	ELL	MC	24	3	3	0	1	1	0	0
			OR	6	1	1	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	2	2	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	White	African American	MC	24	5	5	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
		Hispanic / Latino	MC	24	4	4	0	1	1	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	0	0	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	2	2	0	0	0	0	0
	Online	Paper	MC	24	0	0	0	0	0	0	0
			OR	6	0	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0	0
5	Male	Female	MC	24	3	3	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not ELL	ELL	MC	24	3	3	0	0	0	0	0
			OR	5	1	1	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	1	1	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	White	African American	MC	24	1	1	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
		Hispanic / Latino	MC	24	1	1	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	0	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	4	4	0	0	0	0	0
	Online	Paper	MC	24	0	0	0	0	0	0	0

continued

Grade	Reference	Group	Focal	Item Type	Number of Items	Total	Number "Low" Favoring		Total	Number "High" Favoring	
							Reference	Focal		Reference	Focal
5	Online	Paper	OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
6	Male	Female	MC	24	2	2	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	1	0	1	0	0	0	0
	Not ELL	ELL	MC	24	6	6	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	0	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	White	African American	MC	24	1	0	1	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
		Hispanic / Latino	MC	24	1	1	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	0	0	0	0	0	0	0
			OR	5	1	1	0	0	0	0	0
			WP	4	3	3	0	0	0	0	0
Online	Paper	MC	24	0	0	0	0	0	0	0	
		OR	5	0	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	0	
7	Male	Female	MC	24	1	1	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not ELL	ELL	MC	24	5	4	1	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	0	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	White	African American	MC	24	3	2	1	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	1	1	0	0	0	0	0
		Hispanic / Latino	MC	24	0	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	3	3	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	2	2	0	0	0	0	0
Online	Paper	MC	24	0	0	0	0	0	0	0	
		OR	5	0	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	0	
8	Male	Female	MC	24	2	1	1	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not ELL	ELL	MC	24	3	3	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	0	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	White	African American	MC	24	1	1	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
		Hispanic / Latino	MC	24	0	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	2	2	0	0	0	0	0
			OR	5	0	0	0	0	0	0	0
			WP	4	3	3	0	0	0	0	0
Online	Paper	MC	24	0	0	0	0	0	0	0	
		OR	5	0	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	0	

**Table I-2. Number of Items Classified as “Low” or “High” DIF, Overall and by Group Favored—  
Mathematics**

Grade	Reference	Group Focal	Item Type	Number of Items	Total	Number “Low” Favoring		Total	Number “High” Favoring	
						Reference	Focal		Reference	Focal
3	Male	Female	MC	15	1	1	0	1	1	0
			OR	25	2	1	1	0	0	0
	Not ELL	ELL	MC	15	3	1	2	0	0	0
			OR	25	1	1	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	15	0	0	0	0	0	0
			OR	25	1	1	0	0	0	0
	White	African American	MC	15	4	2	2	0	0	0
			OR	25	4	4	0	0	0	0
		Hispanic / Latino	MC	15	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	15	2	1	1	0	0	0
			OR	25	0	0	0	0	0	0
	Online	Paper	MC	15	0	0	0	0	0	0
			OR	25	0	0	0	0	0	0
	4	Male	Female	MC	17	1	1	0	1	1
OR				23	1	1	0	0	0	0
Not ELL		ELL	MC	17	2	2	0	0	0	0
			OR	23	0	0	0	0	0	0
Not Economically Disadvantaged		Economically Disadvantaged	MC	17	0	0	0	0	0	0
			OR	23	0	0	0	0	0	0
White		African American	MC	17	1	1	0	0	0	0
			OR	23	2	2	0	0	0	0
		Hispanic / Latino	MC	17	2	2	0	0	0	0
Students Without Disabilities		Students with Disabilities	MC	17	1	0	1	0	0	0
			OR	23	1	1	0	0	0	0
Online		Paper	MC	17	0	0	0	0	0	0
			OR	23	0	0	0	0	0	0
5		Male	Female	MC	17	3	2	1	0	0
	OR			23	1	1	0	0	0	0
	Not ELL	ELL	MC	17	0	0	0	0	0	0
			OR	23	1	1	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	17	0	0	0	0	0	0
			OR	23	0	0	0	0	0	0
	White	African American	MC	17	1	1	0	0	0	0
			OR	23	4	4	0	0	0	0
		Hispanic / Latino	MC	17	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	17	0	0	0	0	0	0
			OR	23	3	2	1	0	0	0
	Online	Paper	MC	17	0	0	0	0	0	0
			OR	23	0	0	0	0	0	0
	6	Male	Female	MC	17	4	4	0	1	1
OR				23	3	1	2	1	1	0
Not ELL		ELL	MC	17	2	2	0	0	0	0
			OR	23	3	3	0	0	0	0
Not Economically Disadvantaged		Economically Disadvantaged	MC	17	0	0	0	0	0	0
			OR	23	3	3	0	0	0	0
White		African American	MC	17	3	1	2	0	0	0
			OR	23	1	0	1	1	1	0
		Hispanic / Latino	MC	17	1	1	0	0	0	0
OR		23	4	4	0	0	0	0		

continued

Grade	Reference	Group Focal	Item Type	Number of Items	Total	Number "Low" Favoring		Total	Number "High" Favoring	
						Reference	Focal		Reference	Focal
6	Students Without Disabilities	Students with Disabilities	MC	17	4	3	1	0	0	0
			OR	23	2	0	1	1	0	
	Online	Paper	MC	17	0	0	0	0	0	0
7	Male	Female	MC	15	4	3	1	0	0	0
			OR	25	3	1	2	0	0	0
	Not ELL	ELL	MC	15	0	0	0	1	1	0
			OR	25	2	2	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	15	0	0	0	0	0	0
			OR	25	0	0	0	0	0	0
	White	African American	MC	15	3	2	1	0	0	0
			OR	25	2	1	1	0	0	0
			Hispanic / Latino	MC	15	1	1	0	0	0
	Students Without Disabilities	Students with Disabilities	OR	25	3	3	0	0	0	0
			MC	15	0	0	0	0	0	0
	Online	Paper	MC	15	0	0	0	0	0	0
OR			25	0	0	0	0	0	0	
8	Male	Female	MC	22	3	2	1	0	0	0
			OR	18	0	0	0	0	0	0
	Not ELL	ELL	MC	22	1	1	0	0	0	0
			OR	18	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	22	1	1	0	0	0	0
			OR	18	0	0	0	0	0	0
	White	African American	MC	22	1	1	0	0	0	0
			OR	18	0	0	0	0	0	0
			Hispanic / Latino	MC	22	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	OR	18	5	5	0	0	0	0
			MC	22	4	3	1	0	0	0
	Online	Paper	MC	22	0	0	0	0	0	0
			OR	18	0	0	0	0	0	0

**APPENDIX J**  
**2023–2024 MCAS EQUATING REPORT**



# Massachusetts Comprehensive Assessment System

---

2023–2024: EQUATING REPORT



# 2023 - 2024 Massachusetts Comprehensive Assessment System

## Equating Report

The purpose of this document is to summarize the psychometric calibration and equating results obtained from Cognia for Next-Gen MCAS. Presented in this report are various program summary statistics and specific results related to the study.

The results of this report are organized as follows:

1. Aggregate Results
  1. Percentage of Students by Achievement Levels Categories
  2. Raw Scores Associated with Cutpoints
  3. Calibration Report
  4. Equating Item Summary Statistics
2. Grade Subject Results
  1. A/A, B/B, Delta, Test Characteristic Curve, Test Information Function, and Cumulative Scale Score Distribution Plots
  2. Lookup Tables
  3. Cumulative Scale Score Distribution Tables
  4. Rescore Analysis Results
  5. Tabled Delta Analysis Results
  6. Tabled B/B Analysis Results
  7. Tabled Beta Analysis Results
  8. Final Item Parameters
  9. Decision Accuracy and Consistency (DAC)
  10. Fit Plots of Watchlist Items

The final results of this equating will be included as part of the 2023 - 2024 Next-Gen MCAS Technical Manual. If requested, Cognia will distribute and/or present this report at the next MCAS TAC.

# Section 1.1

## Percentage of Students by Achievement Levels Categories

Table 1.1.1  
 Percentage of Students by Achievement Levels Categories  
 English Language Arts

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS	SD SS
3	2024	59965	15	41	38	7	44	-1.8	495.7	22.9
	2023	60542	14	40	39	7	46	0.3	496.0	23.4
	2022	61648	13	41	39	7	46	-6.9	497.1	22.8
	2021	50011	8	39	43	10	53	-5.7	500.1	22.4
	2019	63602	6	36	48	11	58	5.8	504.8	21.8
	2018	43046	6	41	43	10	53	1.0	501.8	20.6
	2017	26459	7	41	43	8	52		500.1	21.1
4	2024	61103	16	46	34	5	39	-3.2	493.5	22.2
	2023	61836	14	44	36	6	42	1.7	494.7	22.2
	2022	62100	14	46	36	4	40	-11.2	493.8	21.1
	2021	50867	11	38	45	6	51	-2.7	498.9	22.8
	2019	65450	7	39	44	10	54	-0.5	502.6	21.1
	2018	69078	7	38	44	10	55	3.5	502.2	21.6
	2017	63918	8	41	43	8	51		500.1	21.2
5	2024	62601	13	47	34	6	40	-6.0	494.6	21.9
	2023	62316	13	41	41	5	46	3.8	496.0	21.8
	2022	63620	11	47	37	5	42	-7.1	495.6	20.4
	2021	51362	10	41	41	8	49	-5.0	497.9	22.1
	2019	67933	6	39	47	8	54	-1.5	501.9	21.0
	2018	69390	6	38	49	7	56	4.4	502.3	20.0
	2017	28547	7	42	46	5	51		499.9	19.8
6	2024	62688	22	36	30	11	42	-2.5	494.6	28.9
	2023	63574	21	35	36	8	44	1.3	494.0	28.1
	2022	63887	20	37	35	8	43	-7.2	494.0	26.0
	2021	51319	19	31	37	13	50	-5.7	498.4	28.3
	2019	67612	11	33	42	13	56	3.4	502.5	25.6
	2018	53988	10	38	42	11	52	-0.7	501.3	23.5
	2017	29369	8	39	47	6	53		500.3	20.7
7	2024	63626	19	43	31	7	38	-4.4	492.6	25.6
	2023	63711	17	41	34	8	42	-0.3	494.2	25.1
	2022	65584	17	40	37	6	42	-3.2	493.7	23.6
	2021	51120	17	37	39	7	46	-4.8	495.6	24.8
	2019	67462	11	39	42	9	50	3.6	499.8	23.0
	2018	66410	13	40	39	8	47	-6.5	497.4	23.4
	2017	30209	8	38	48	6	53		500.2	20.8
8	2024	64139	21	34	33	12	44	-0.9	494.9	28.5
	2023	65553	20	35	35	10	45	2.2	495.4	28.5
	2022	67919	16	40	36	7	43	-0.3	494.8	24.8
	2021	50822	15	41	37	7	44	-10.0	496.2	25.5
	2019	67350	11	35	42	12	54	1.1	500.6	24.5
	2018	69486	13	34	42	10	52	1.1	499.6	24.9
	2017	65314	9	40	43	8	51		499.5	21.8

Table 1.1.2  
Percentage of Students by Achievement Levels Categories  
Mathematics

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS	SD SS
3	2024	50046	12	35	41	12	53	4.6	500.6	24.5
	2023	51707	12	39	39	10	49	1.2	499.1	24.3
	2022	53433	13	39	40	8	48	9.8	497.5	24.7
	2021	45242	20	42	32	6	38	-18.0	491.2	25.7
	2019	56176	7	37	45	11	56	7.2	503.0	22.4
	2018	43501	11	40	40	9	49	-3.2	499.1	23.2
	2017	26659	11	37	44	8	52		499.2	22.7
4	2024	50882	8	37	45	10	55	1.5	502.5	23.0
	2023	52554	10	37	43	10	54	4.0	501.1	22.9
	2022	53577	10	40	43	7	50	10.9	498.8	23.0
	2021	45553	17	44	34	4	39	-17.8	491.7	23.9
	2019	57629	6	37	47	10	57	7.7	503.0	21.2
	2018	69779	11	40	42	7	49	-1.9	498.0	22.6
	2017	64473	10	39	44	6	51		498.7	22.0
5	2024	53399	7	45	41	7	48	-0.3	499.6	20.2
	2023	54159	7	45	42	6	48	5.7	499.3	20.6
	2022	55635	10	48	38	5	42	3.1	496.5	20.6
	2021	46011	13	47	35	5	39	-15.5	493.5	21.3
	2019	60444	5	40	48	6	55	8.0	501.7	19.8
	2018	70083	9	45	42	5	47	-2.7	497.7	19.7
	2017	29285	8	42	42	8	49		499.4	20.7
6	2024	54954	10	43	39	9	47	-0.1	499.4	22.5
	2023	56389	11	42	40	8	47	-0.3	498.1	21.9
	2022	56939	9	43	42	6	48	8.9	498.2	21.2
	2021	46699	16	45	34	5	39	-18.9	493.4	22.7
	2019	61719	6	37	46	12	58	9.6	504.0	22.0
	2018	54582	9	43	42	6	48	-4.2	498.4	21.9
	2017	29704	9	39	46	6	52		499.7	21.2
7	2024	56448	12	45	34	9	43	-0.1	497.0	24.2
	2023	57234	16	41	35	9	44	1.7	496.1	24.5
	2022	59311	13	45	34	8	42	2.7	495.5	23.5
	2021	46839	13	48	32	7	39	-13.4	494.9	23.5
	2019	62495	9	39	41	12	53	5.3	501.0	23.7
	2018	66925	12	40	40	8	47	-0.9	497.7	23.4
	2017	30144	9	43	40	8	48		498.9	22.0
8	2024	57441	13	43	35	9	44	1.3	497.4	23.6
	2023	59572	14	43	34	9	43	2.3	496.4	23.8
	2022	62311	12	48	32	8	40	4.4	495.8	22.7
	2021	47150	16	48	32	4	36	-15.1	492.0	22.9
	2019	62817	8	41	40	11	51	0.4	501.5	22.5
	2018	70044	11	39	42	8	51	1.1	498.9	22.7
	2017	66077	9	42	40	9	49		500.3	21.6

# Section 1.2

## Raw Scores Associated with Cutpoints

Table 1.2.1  
Raw Scores Associated with Cutpoints

Subject	Grade	Cut Point	2023 Actual	2024 Actual	2024 Pred
English Language Arts	3	NM-PM	13	14	12
		PM-ME	28	27	26
		ME-EE	38	37	36
English Language Arts	4	NM-PM	16	15	14
		PM-ME	29	29	28
		ME-EE	38	39	39
English Language Arts	5	NM-PM	17	16	14
		PM-ME	32	34	31
		ME-EE	42	44	42
English Language Arts	6	NM-PM	19	19	17
		PM-ME	31	32	30
		ME-EE	41	42	40
English Language Arts	7	NM-PM	16	19	16
		PM-ME	31	34	31
		ME-EE	43	44	43
English Language Arts	8	NM-PM	22	20	20
		PM-ME	34	33	33
		ME-EE	43	43	42
Mathematics	3	NM-PM	14	14	14
		PM-ME	32	29	30
		ME-EE	44	42	43
Mathematics	4	NM-PM	16	16	16
		PM-ME	34	34	35
		ME-EE	48	49	49
Mathematics	5	NM-PM	13	13	13
		PM-ME	32	32	32
		ME-EE	49	49	49
Mathematics	6	NM-PM	12	12	12
		PM-ME	30	30	30
		ME-EE	49	48	48
Mathematics	7	NM-PM	11	11	11
		PM-ME	27	28	27
		ME-EE	47	47	47
Mathematics	8	NM-PM	14	14	13
		PM-ME	32	32	31
		ME-EE	49	49	49

# Section 1.3

## Calibration Report

## Calibration Report—Executive Summary

FlexMIRT 3.03 was used for the IRT calibration at Cognia. All command files were set up in a way following general settings. The calibration convergence criterion was set to 0.001.

A 3PLM was used for standard four-option multiple choice (MC) items, a 2PLM was used for dichotomously-scored short response items, multi-select items, and technology-enhanced items, and a Graded Response Model (GRM) was specified for the polytomously-scored multi-part items and open response items. The logistic version of the IRT models was used. The prior distribution for the guessing parameter was set to be  $\text{beta}(5, 17)$ , and  $\text{logNormal}(0, 0.25)$  was used as the prior for the item discrimination parameter. No prior was supplied for the item difficulty parameter.

The calibration went smoothly and was converged in all subjects/grades. In particular, the largest change in parameter values (from one iteration to the next) was decreasing and tended to flatten out towards the end of the calibration process. The IRT model fit was evaluated for each of the items. The resulting parameters demonstrated good model fit for most of the items.

In ELA, a two-stage process was used to bring the item parameters onto the operational scale. First all items except the essays were freely calibrated. Next the items were placed onto scale using the Stocking and Lord procedure. These first two steps are referred to as Stage 1. Next, the essays were brought onto scale with a Fixed Common Item Parameter calibration using FlexMIRT while holding the parameters from Stage 1 fixed. This two-stage process is used to assure that the essay parameter estimation process does not unduly influence the dimensional structure of the initial parameter estimation in Stage 1, providing for greater scale stability.

The first table in this section shows the number of cycles to achieve convergence in Stage 1 of the ELA procedure:

Table 1.3.1.a  
Number of Cycles to Convergence for ELA Calibration with no Essays

Subject	Grade	Initial Cycles
English Language Arts	Grade 3	26
English Language Arts	Grade 4	62
English Language Arts	Grade 5	46
English Language Arts	Grade 6	34
English Language Arts	Grade 7	38
English Language Arts	Grade 8	30

The Stocking and Lord procedure was used to transform non-essay parameter estimates onto the operational scale. These transformation constants were found using the STUIRT program which can be found at the CASMA website: <http://www.education.uiowa.edu/casma/>. The Stocking & Lord transformation constants that were calculated in the second step of Stage 1 are listed in the following table:

Table 1.3.1.b  
Stocking and Lord Constants for ELA Equating with no Essays

Subject	Grade	Slope	Intercept	Num Eq Items	Num Eq Items Rem
English Language Arts	3	1.12	-0.22	18	0
English Language Arts	4	1.06	-0.31	18	0
English Language Arts	5	1.16	-0.28	16	0
English Language Arts	6	1.48	-0.31	16	0
English Language Arts	7	1.29	-0.39	16	0
English Language Arts	8	1.40	-0.22	16	0



The third table shows the number of cycles to achieve convergence in the equating (FCIP) calibration runs for Stage 2:

Table 1.3.1.c  
Number of Cycles to Convergence for ELA FCIP Calibration with Essays Included

Subject	Grade	Initial Cycles	Equating Cycles
English Language Arts	Grade 3	34	7
English Language Arts	Grade 4	23	7
English Language Arts	Grade 5	49	10
English Language Arts	Grade 6	33	12
English Language Arts	Grade 7	37	9
English Language Arts	Grade 8	34	9

The Math and Science tests were equated using a single stage procedure of freely calibrating all items and placing them on the operational scale using the Stocking and Lord procedure. The next table in this section lists the number of cycles to achieve convergence, followed by a table of the Stocking and Lord transformation constants.

Table 1.3.1.d  
Number of Cycles to Convergence for Math

Subject	Grade	Initial Cycles
Mathematics	Grade 3	63
Mathematics	Grade 4	68
Mathematics	Grade 5	58
Mathematics	Grade 6	39
Mathematics	Grade 7	76
Mathematics	Grade 8	45

Table 1.3.1.e  
Stocking and Lord Constants for Math

Subject	Grade	Slope	Intercept	Num Eq Items	Num Eq Items Rem
Mathematics	3	1.09	0.05	20	0
Mathematics	4	1.05	0.16	20	0
Mathematics	5	0.99	-0.01	20	0
Mathematics	6	1.06	-0.04	20	0
Mathematics	7	1.11	-0.11	20	0
Mathematics	8	1.10	-0.13	20	0

Four methods of evaluating the suitability of the equating items were used: delta analysis, b/b analysis, beta analysis, and rescore analysis. As a result of all four analyses, flagged items were reviewed by content personnel and no items were removed from the equating analysis. Results from these analyses are included in Section II of this report.

Items flagged by the delta, b/b, beta, or rescore analyses, or any item that required intervention during the calibration process, were compiled and placed in our item watch list, which includes the final actions taken on these items. The final watch list is presented in the following table:

Table 1.3.2  
Final Items Watch List

Subject	Grade	ItemID	Reason	Action
English Language Arts	3	IA00286 (EL308842)	beta analysis	retained for equating
English Language Arts	3	IA00287 (EL308855)	beta analysis	retained for equating
English Language Arts	8	IA00378 (EL623955555)	beta analysis	retained for equating
Mathematics	3	IA01019 (MA311277)	beta analysis	retained for equating
Mathematics	3	IA07855 (MA900579464)	beta analysis	retained for equating
Mathematics	4	IA00961 (MA307081)	beta analysis	retained for equating
Mathematics	4	IA01055 (MA311572)	beta analysis	retained for equating
Mathematics	4	IA02819 (MA713583365)	beta analysis	retained for equating
Mathematics	5	IA02917 (MA715102107)	beta analysis	retained for equating
Mathematics	6	IA04899 (MA736511626)	beta analysis	retained for equating
Mathematics	7	IA04538 (MA282218)	beta analysis	retained for equating
Mathematics	8	IA00865 (MA297656)	beta analysis	retained for equating
Mathematics	8	IA00905 (MA301702)	beta analysis	retained for equating
Mathematics	8	IA02495 (MA309741)	beta analysis	retained for equating
Mathematics	8	IA04665 (MA307399)	beta analysis	retained for equating
Mathematics	8	IA04719 (MA311462)	beta analysis	retained for equating
Mathematics	8	IA05070 (MA804042487)	beta analysis	retained for equating

# Section 1.4

## Equating Item Summary Statistics

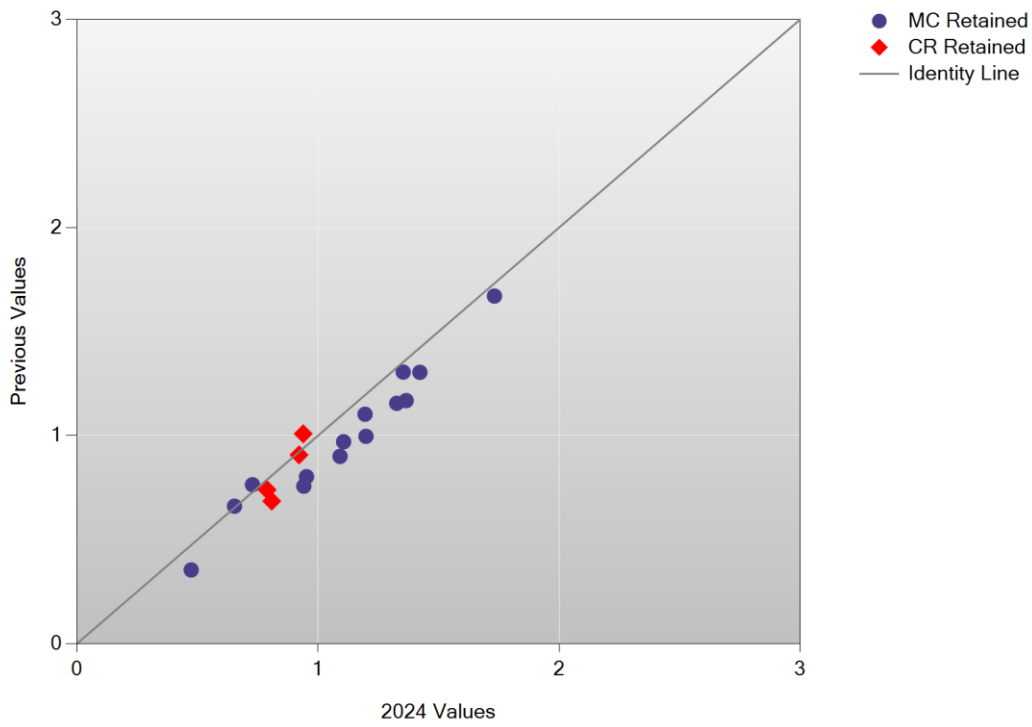
Table 1.4.1  
Equating Item Summary Statistics  
(2024 a and b values are derived from unscaled parameters before equating for all tests)

Subject	Grade	Year	P-Value		Point Biserial		a		b	
			Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
English Language Arts	03	2024	0.57	0.12	0.47	0.09	1.05	0.32	0.02	0.49
		Previous	0.58	0.13	0.49	0.09	0.96	0.30	-0.22	0.59
English Language Arts	04	2024	0.63	0.16	0.42	0.09	0.83	0.20	-0.33	0.80
		Previous	0.63	0.16	0.41	0.09	0.80	0.18	-0.57	0.87
English Language Arts	05	2024	0.67	0.14	0.41	0.08	0.85	0.28	-0.48	0.64
		Previous	0.68	0.14	0.40	0.08	0.75	0.24	-0.78	0.75
English Language Arts	06	2024	0.68	0.11	0.41	0.09	0.89	0.23	-0.44	0.69
		Previous	0.68	0.11	0.41	0.08	0.60	0.15	-0.96	0.97
English Language Arts	07	2024	0.69	0.11	0.43	0.09	0.91	0.24	-0.56	0.65
		Previous	0.70	0.11	0.43	0.08	0.70	0.17	-1.14	0.80
English Language Arts	08	2024	0.63	0.13	0.41	0.07	0.83	0.25	-0.33	0.58
		Previous	0.64	0.12	0.40	0.08	0.63	0.20	-0.62	0.87
Mathematics	03	2024	0.64	0.15	0.48	0.12	0.97	0.25	-0.39	0.73
		Previous	0.63	0.15	0.48	0.12	0.92	0.24	-0.39	0.79
Mathematics	04	2024	0.64	0.12	0.50	0.13	1.06	0.30	-0.33	0.53
		Previous	0.63	0.12	0.50	0.12	0.98	0.29	-0.24	0.55
Mathematics	05	2024	0.59	0.19	0.49	0.11	0.98	0.26	-0.28	0.86
		Previous	0.59	0.19	0.48	0.13	1.00	0.27	-0.31	0.98
Mathematics	06	2024	0.56	0.16	0.50	0.10	1.00	0.30	-0.14	0.72
		Previous	0.55	0.15	0.49	0.11	0.98	0.34	-0.13	0.78
Mathematics	07	2024	0.54	0.18	0.50	0.13	1.21	0.37	0.06	0.77
		Previous	0.53	0.19	0.50	0.13	1.09	0.29	-0.05	0.88
Mathematics	08	2024	0.55	0.14	0.50	0.14	1.19	0.42	0.09	0.59
		Previous	0.54	0.13	0.51	0.13	1.08	0.36	-0.05	0.62

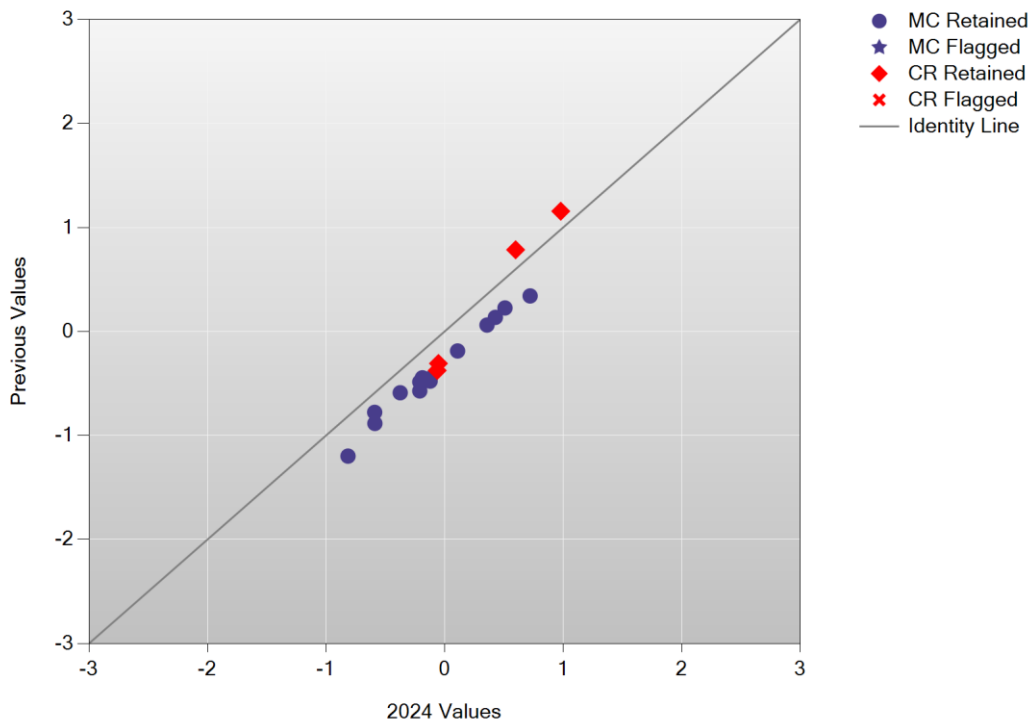
# Section 2.1

A/A, B/B, Delta, Test Characteristic Curve, Test Information Function, and Cumulative Scale Score Distribution Plots

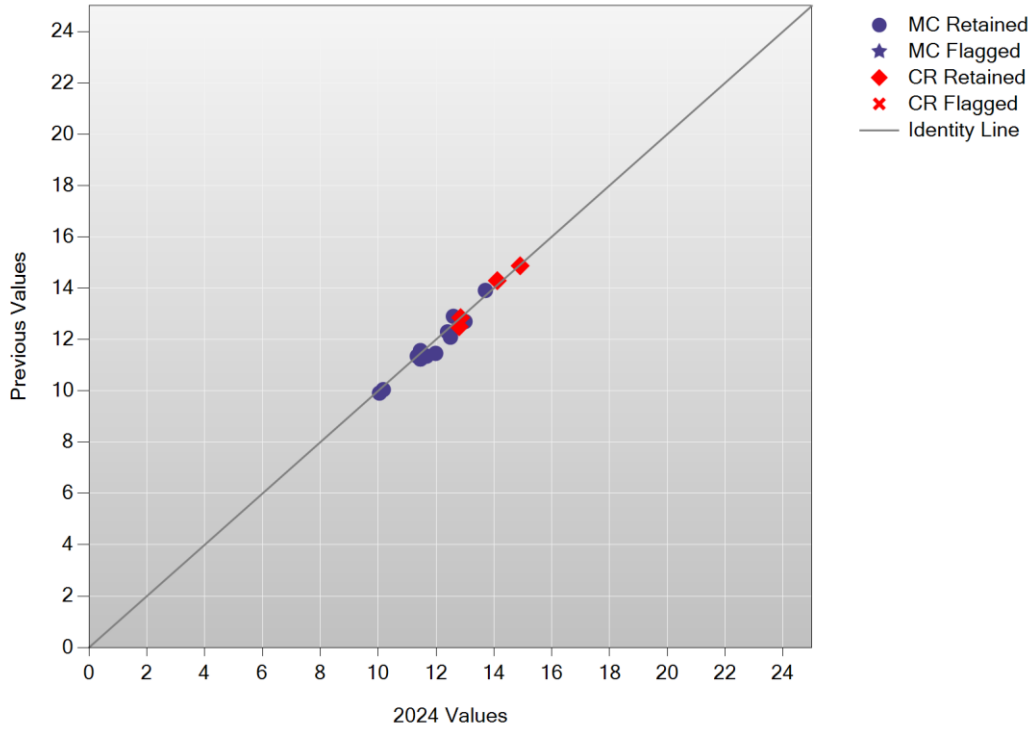
A/A Plot: English Language Arts Grade 3



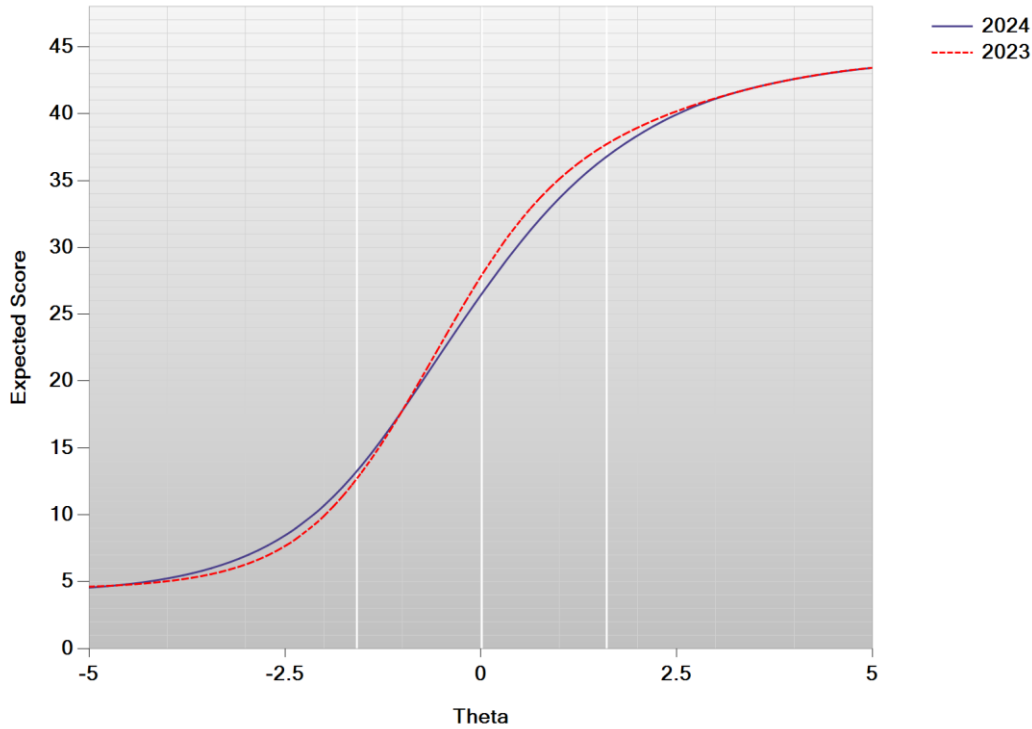
B/B Plot: English Language Arts Grade 3



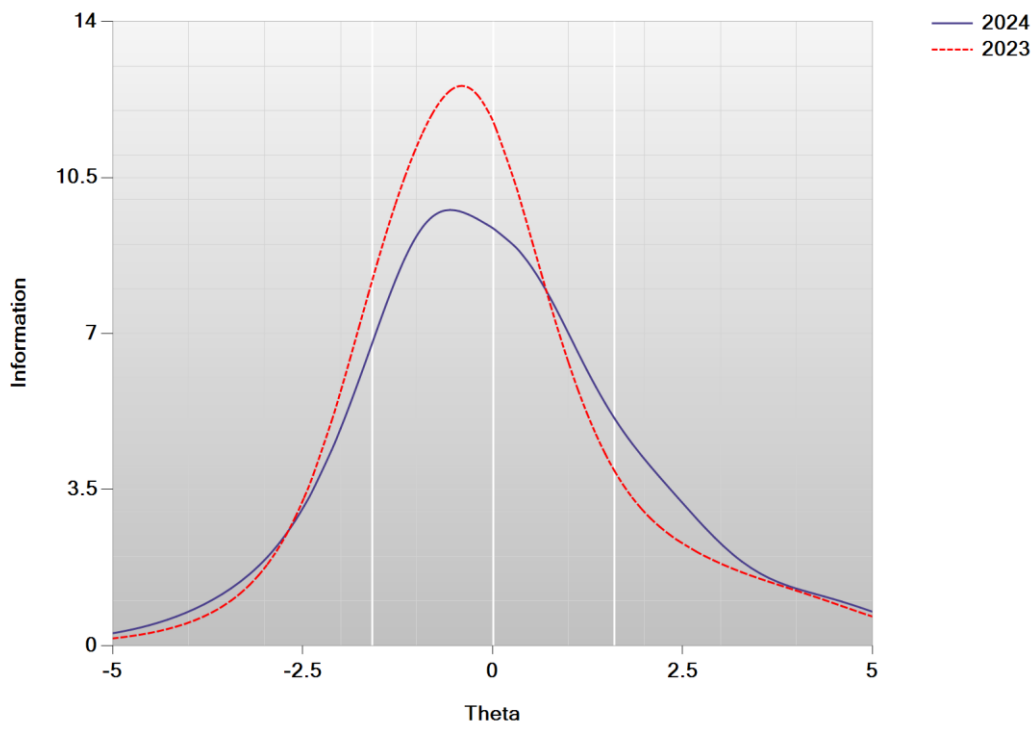
Delta Plot: English Language Arts Grade 3



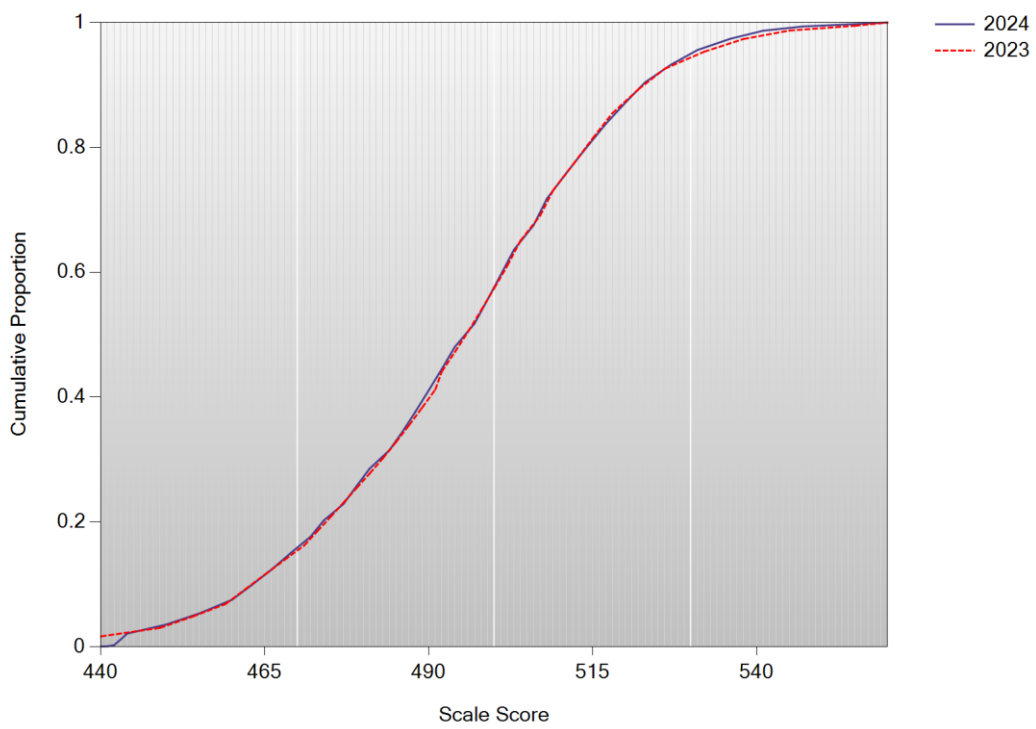
Test Characteristic Curve: English Language Arts Grade 3



### Test Information Function: English Language Arts Grade 3

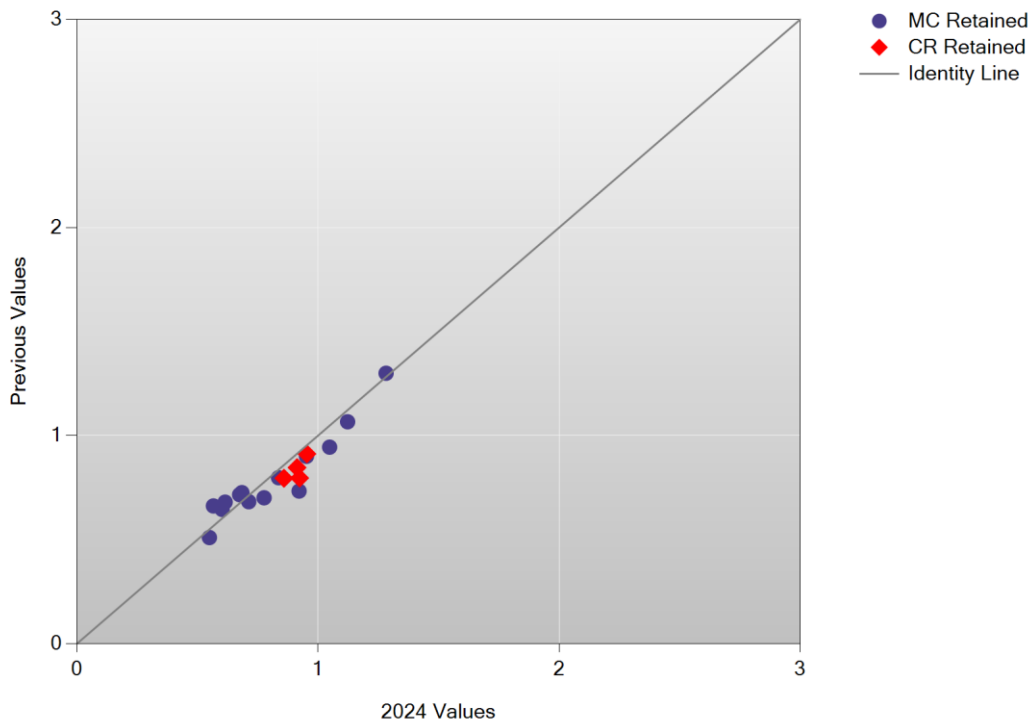


### Cumulative Scale Score Distributions: English Language Arts Grade 3

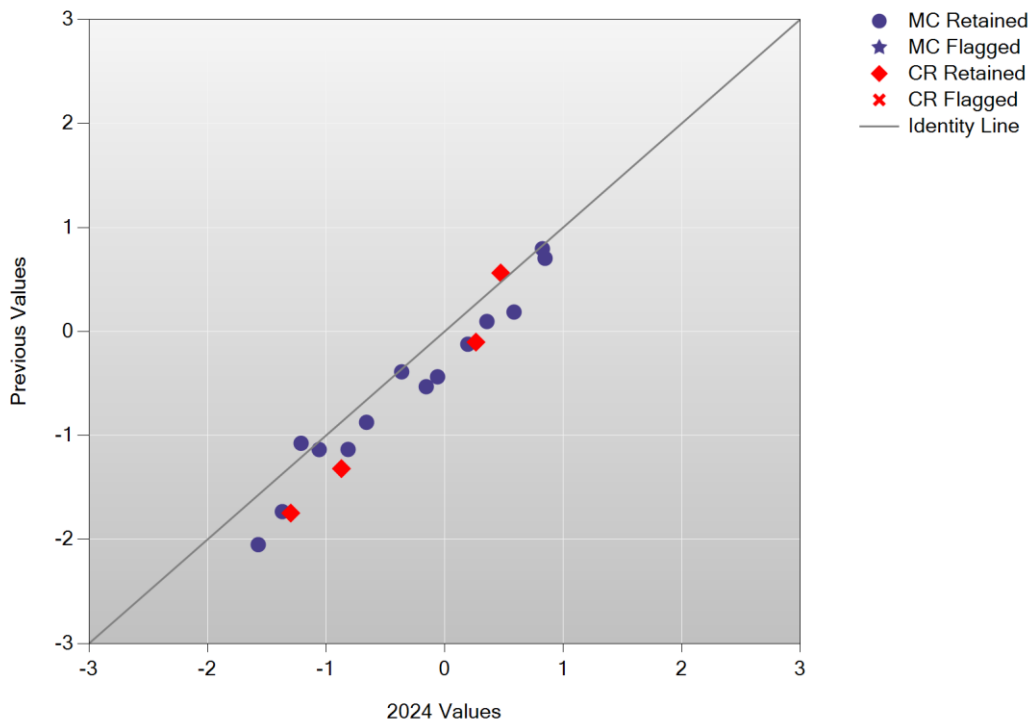




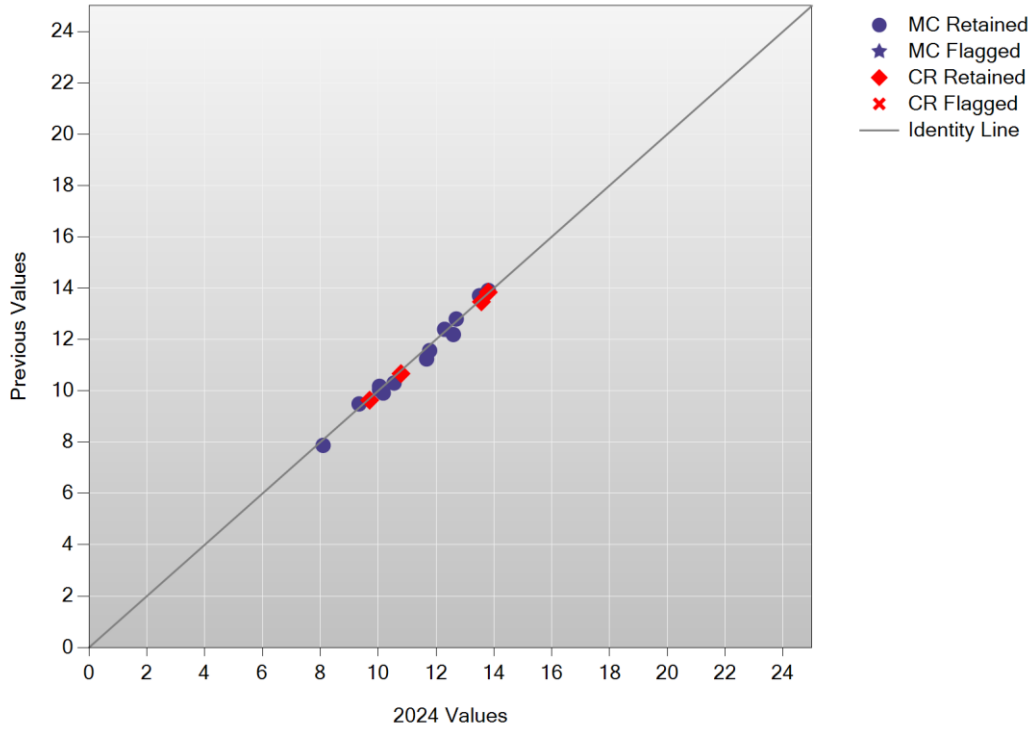
A/A Plot: English Language Arts Grade 4



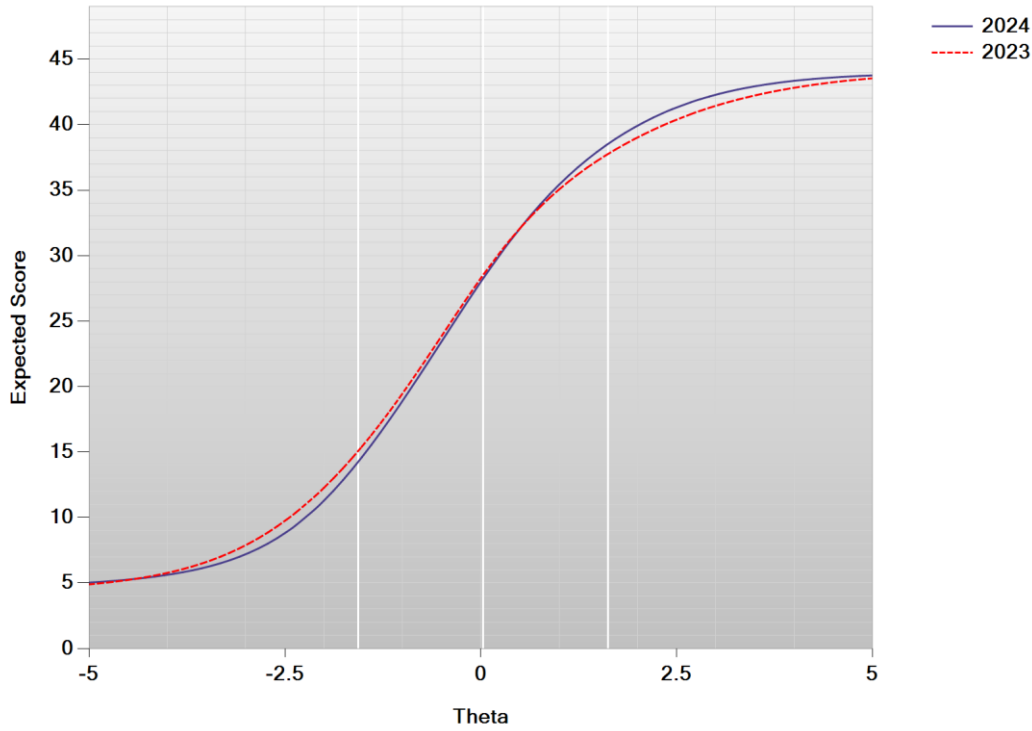
B/B Plot: English Language Arts Grade 4



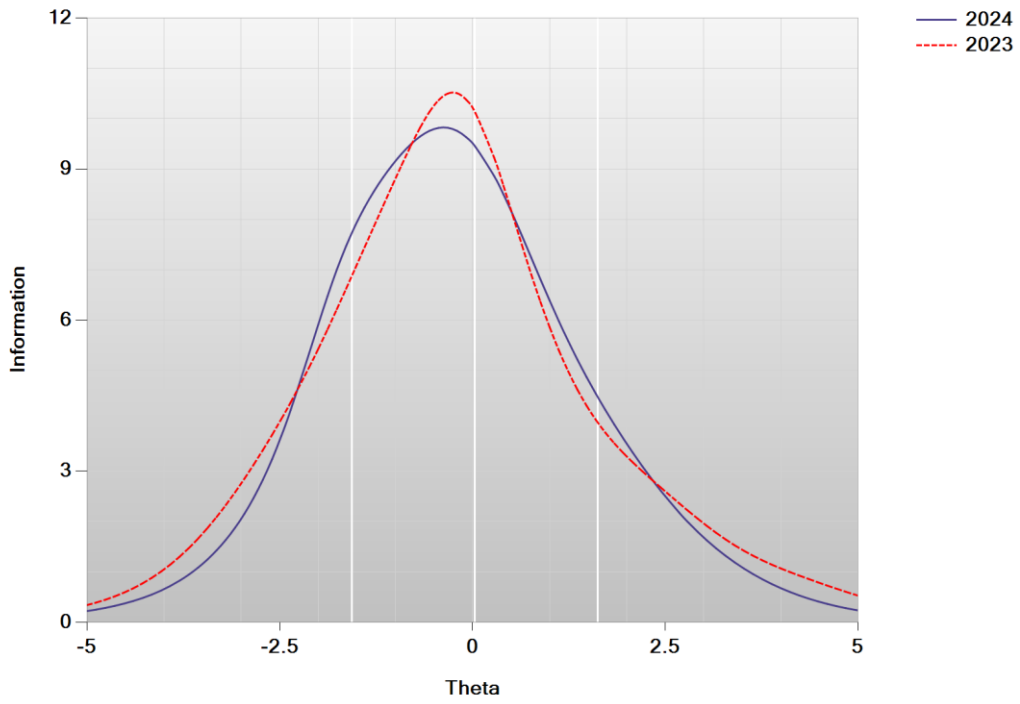
Delta Plot: English Language Arts Grade 4



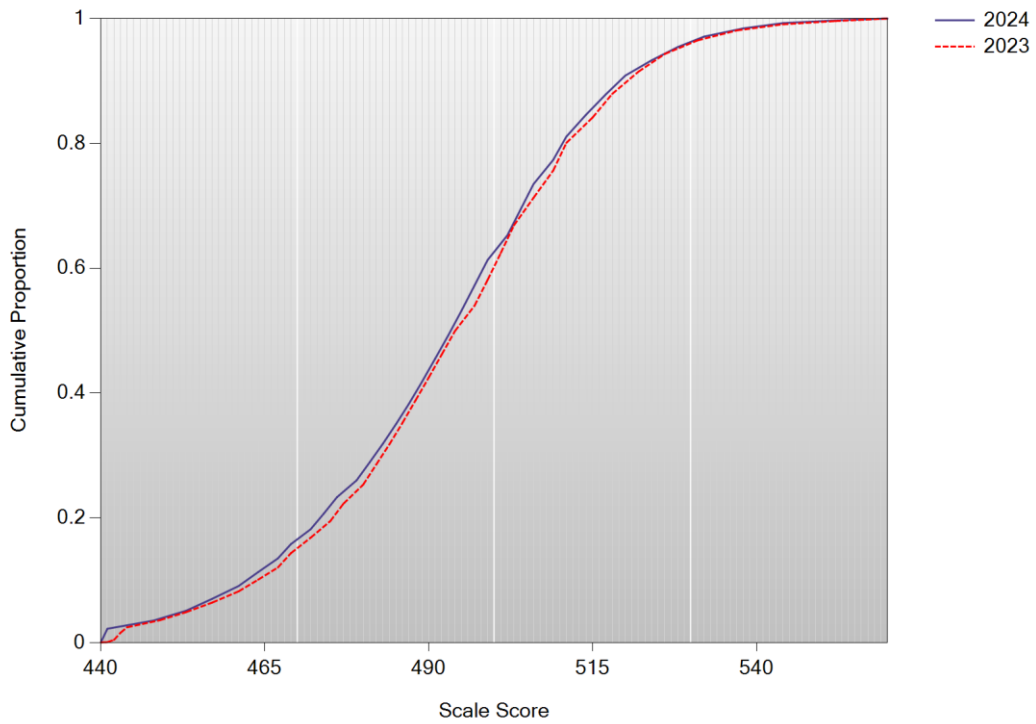
Test Characteristic Curve: English Language Arts Grade 4



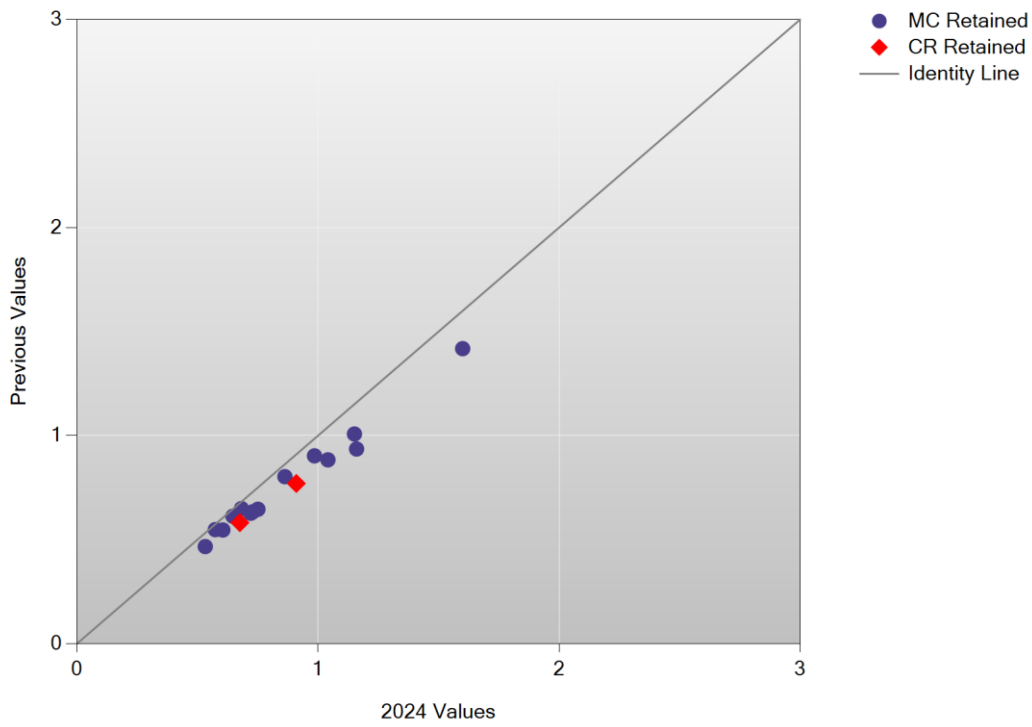
### Test Information Function: English Language Arts Grade 4



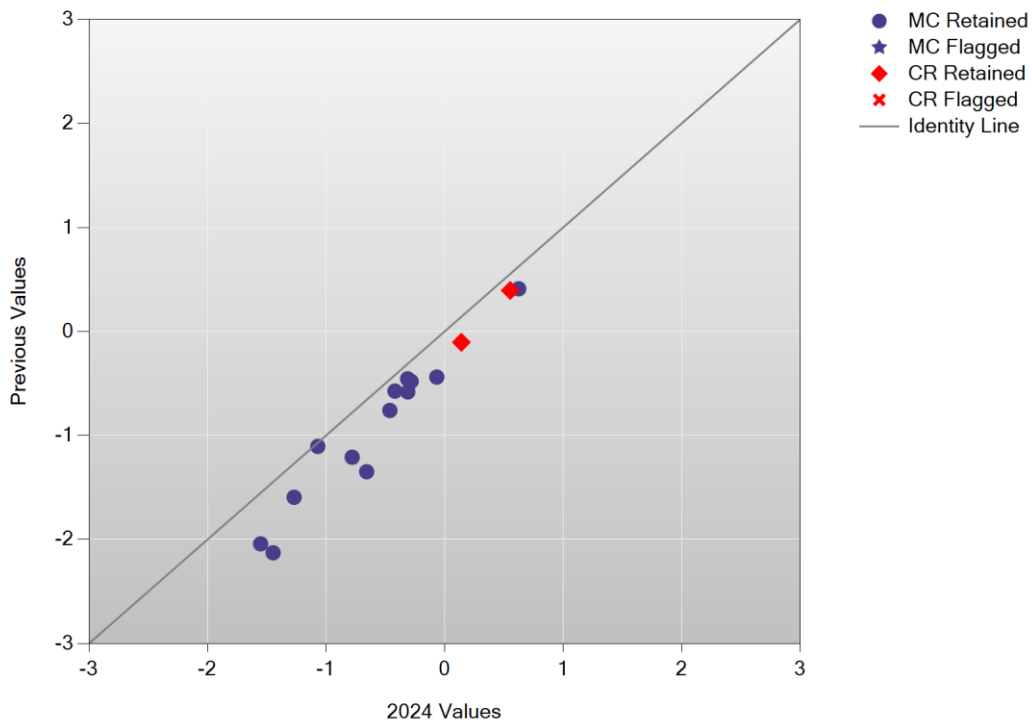
### Cumulative Scale Score Distributions: English Language Arts Grade 4



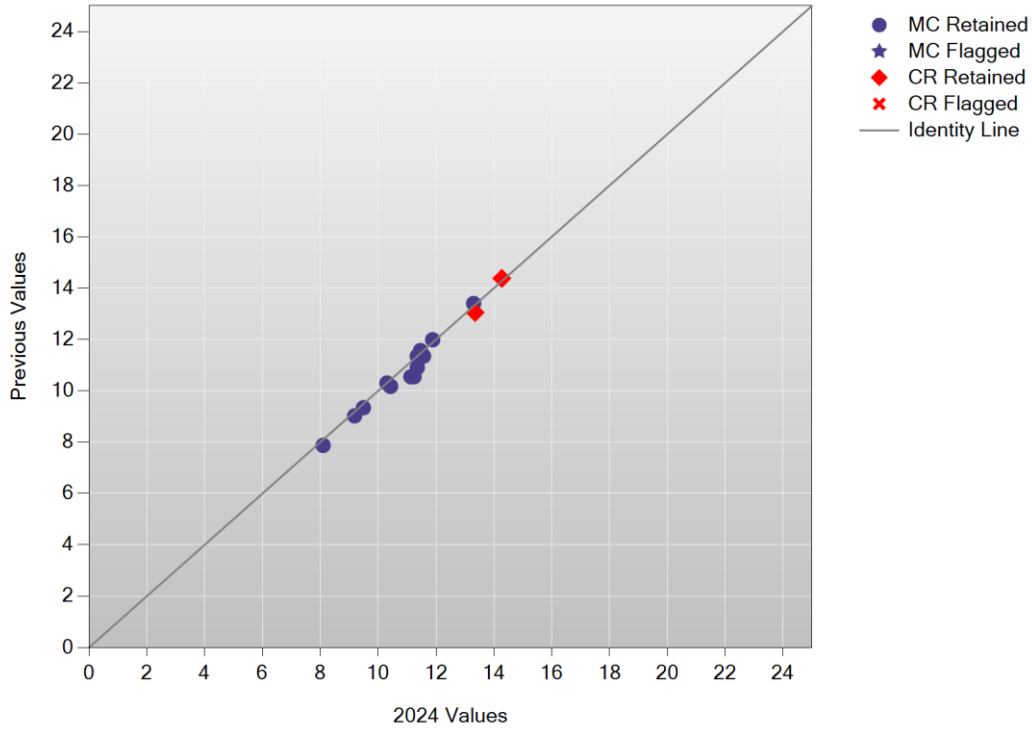
A/A Plot: English Language Arts Grade 5



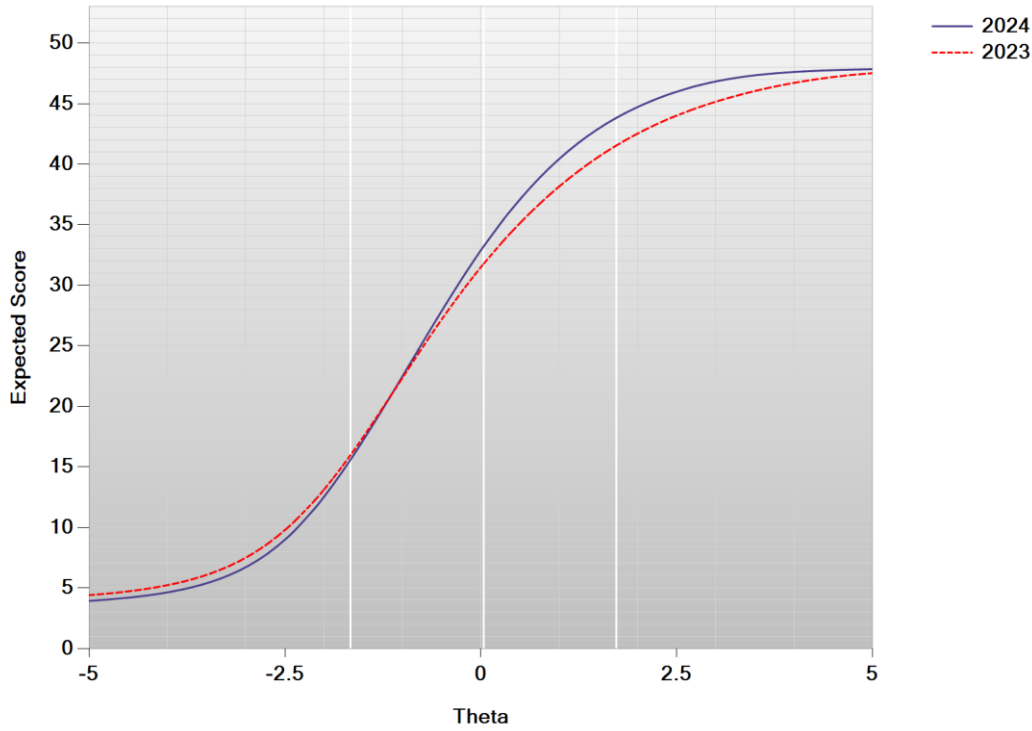
B/B Plot: English Language Arts Grade 5



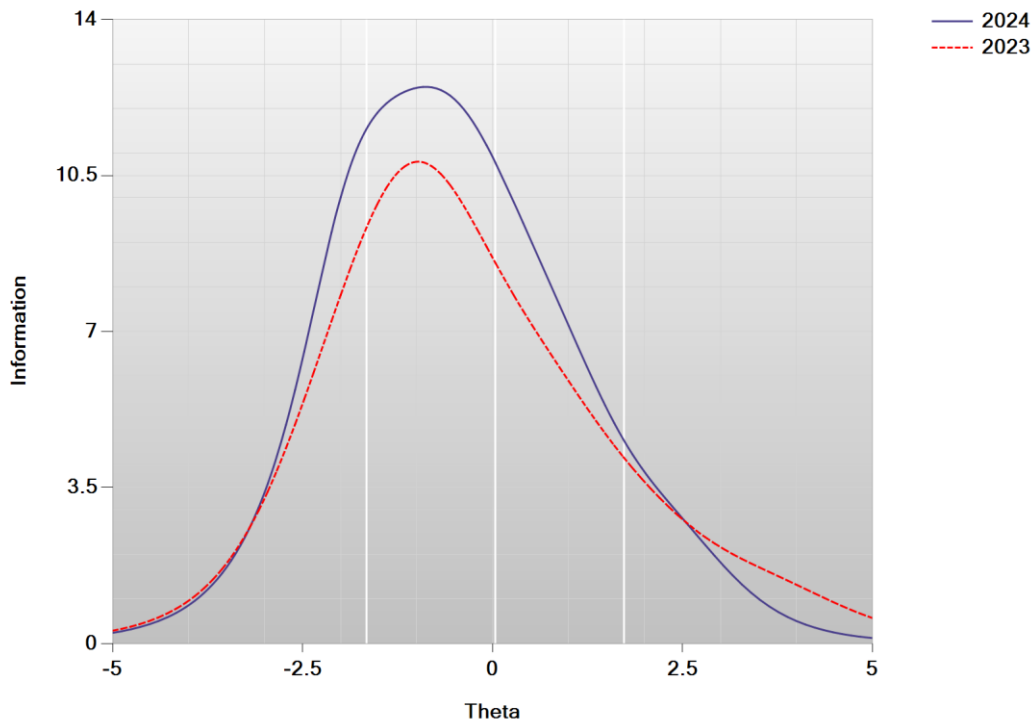
Delta Plot: English Language Arts Grade 5



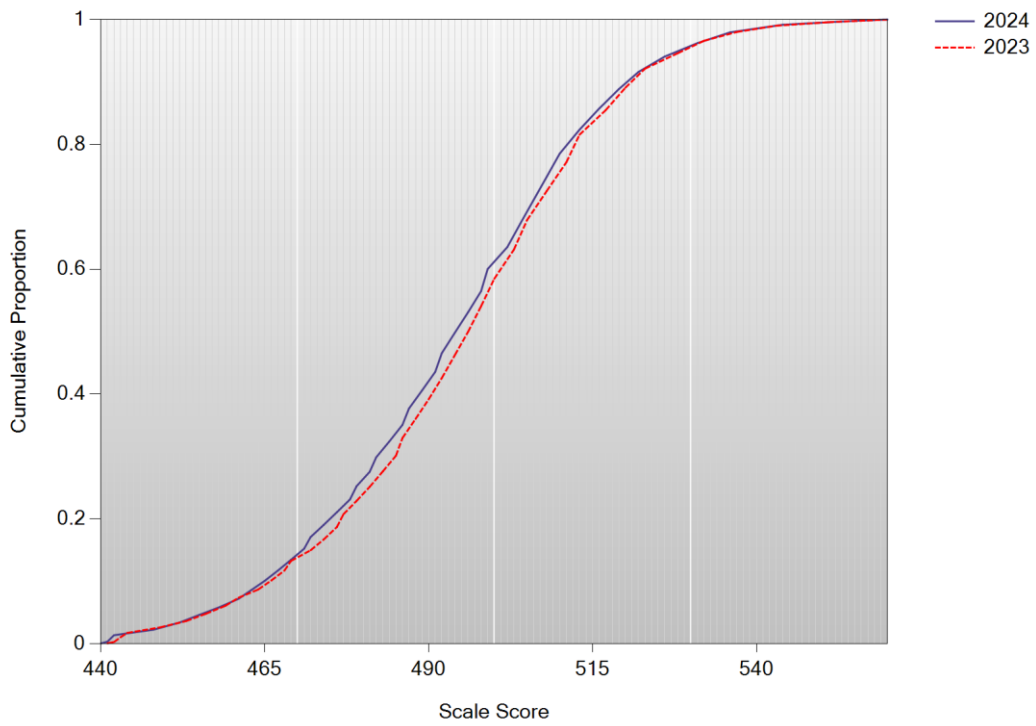
Test Characteristic Curve: English Language Arts Grade 5



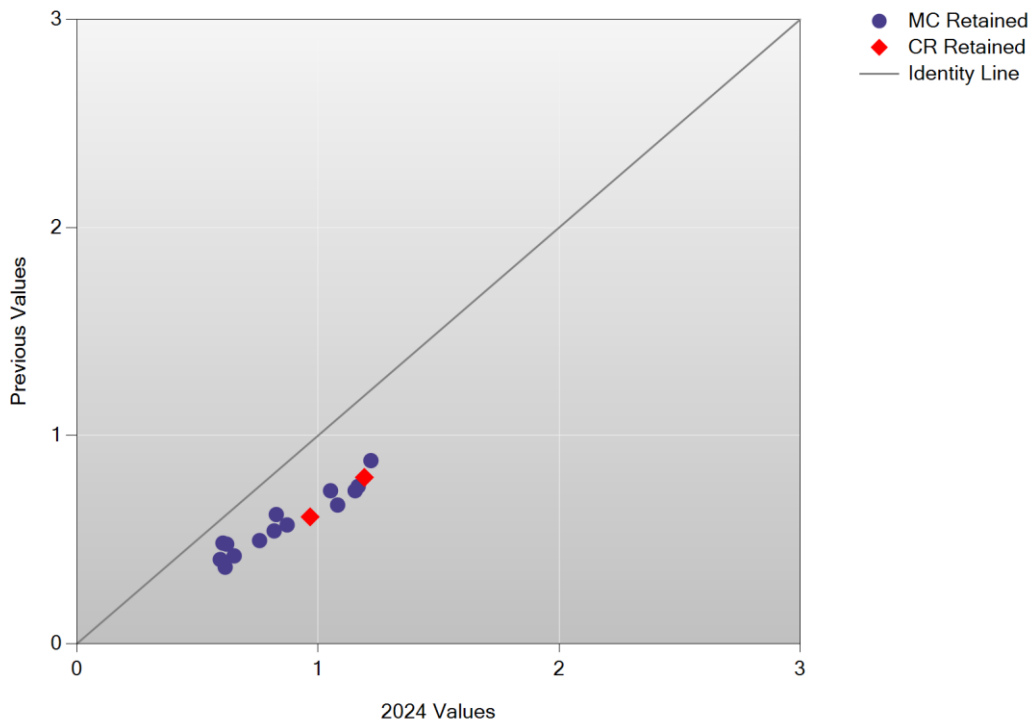
### Test Information Function: English Language Arts Grade 5



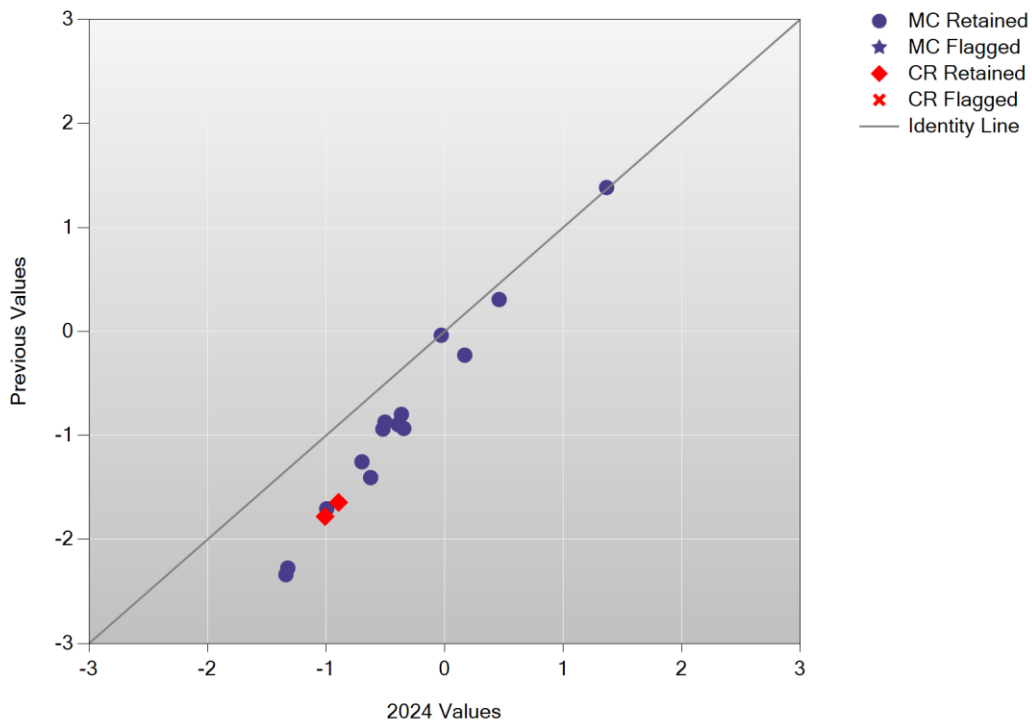
### Cumulative Scale Score Distributions: English Language Arts Grade 5



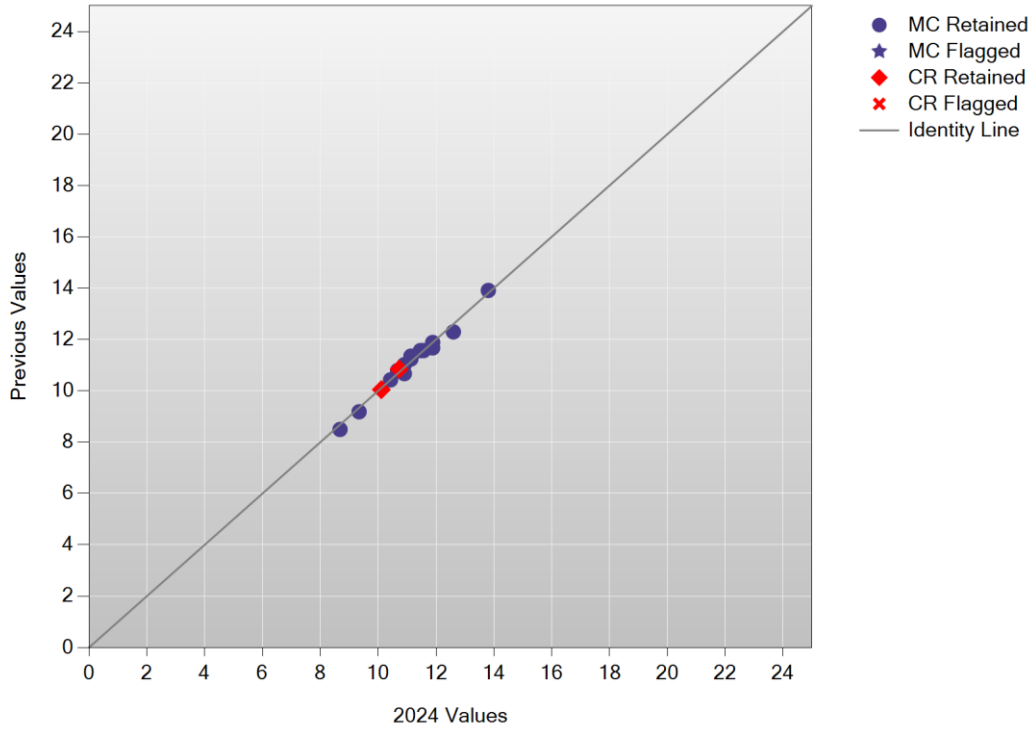
A/A Plot: English Language Arts Grade 6



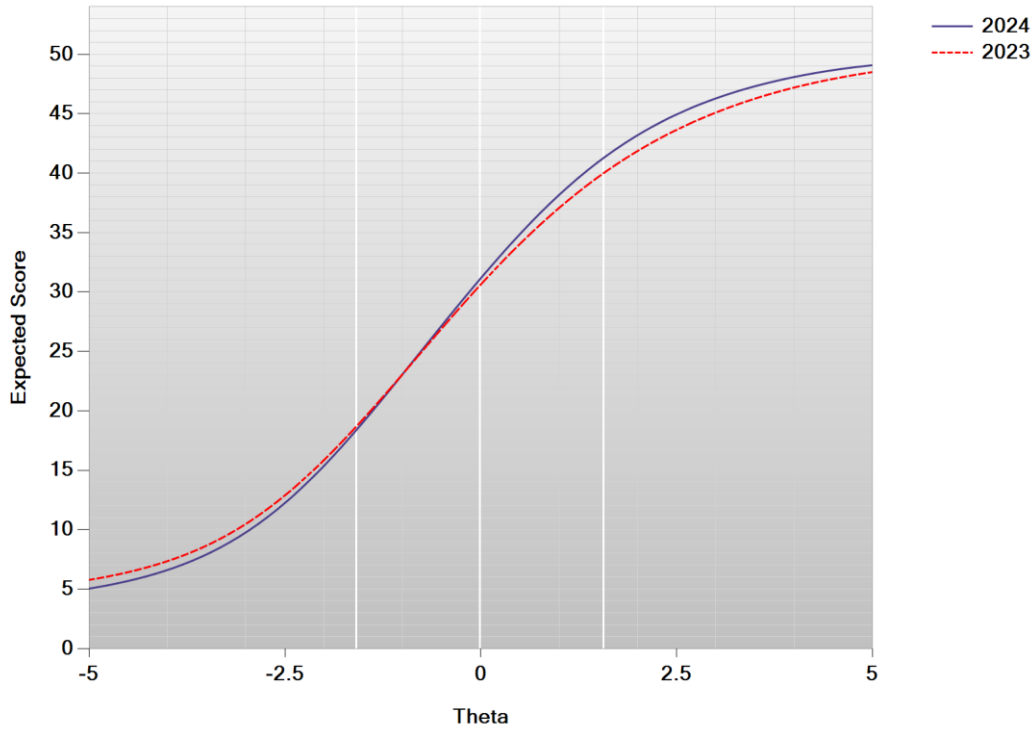
B/B Plot: English Language Arts Grade 6



Delta Plot: English Language Arts Grade 6

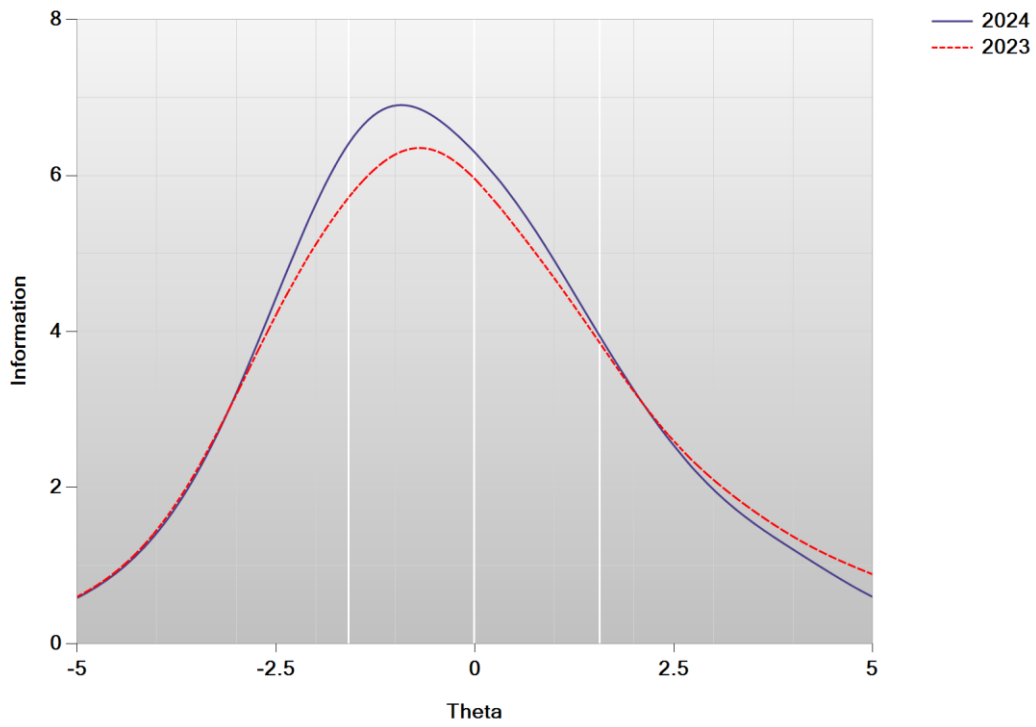


Test Characteristic Curve: English Language Arts Grade 6

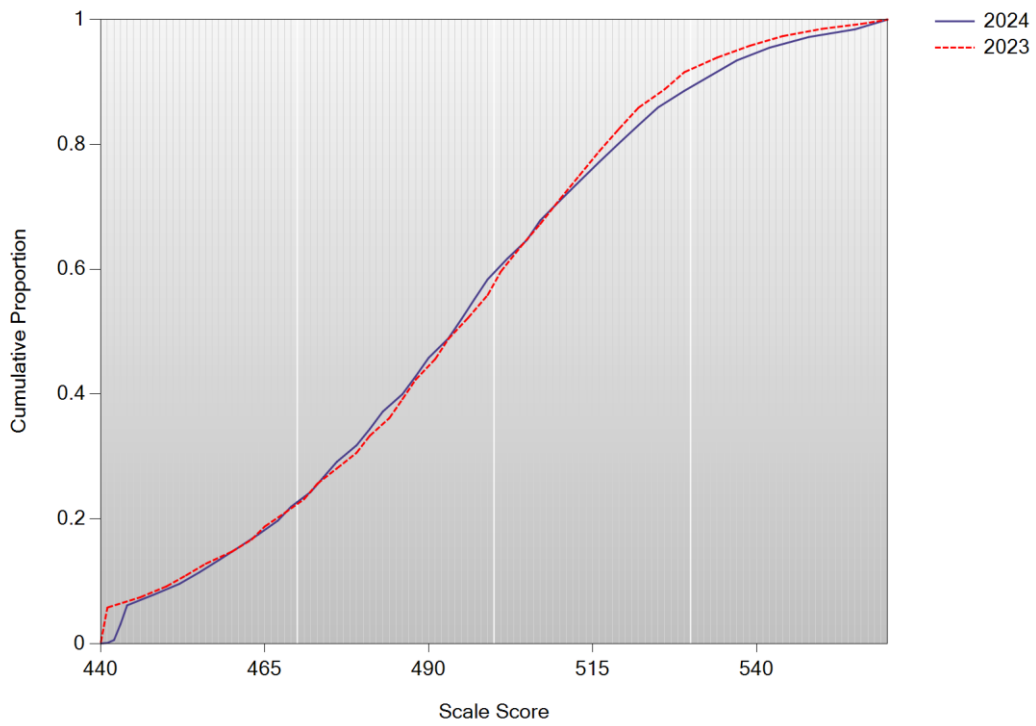




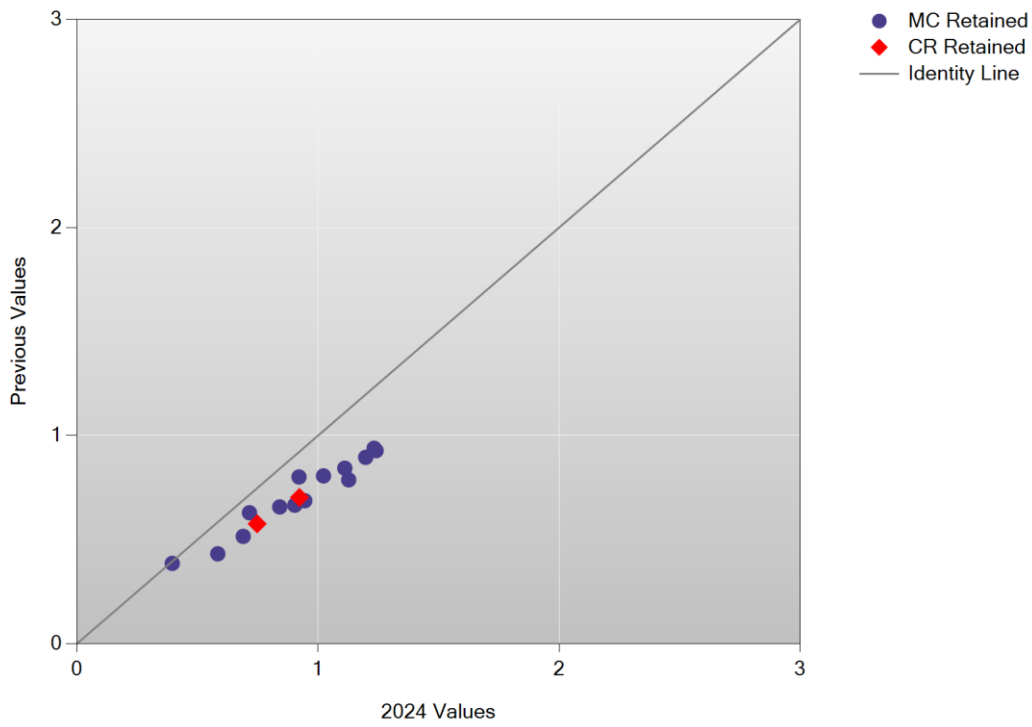
### Test Information Function: English Language Arts Grade 6



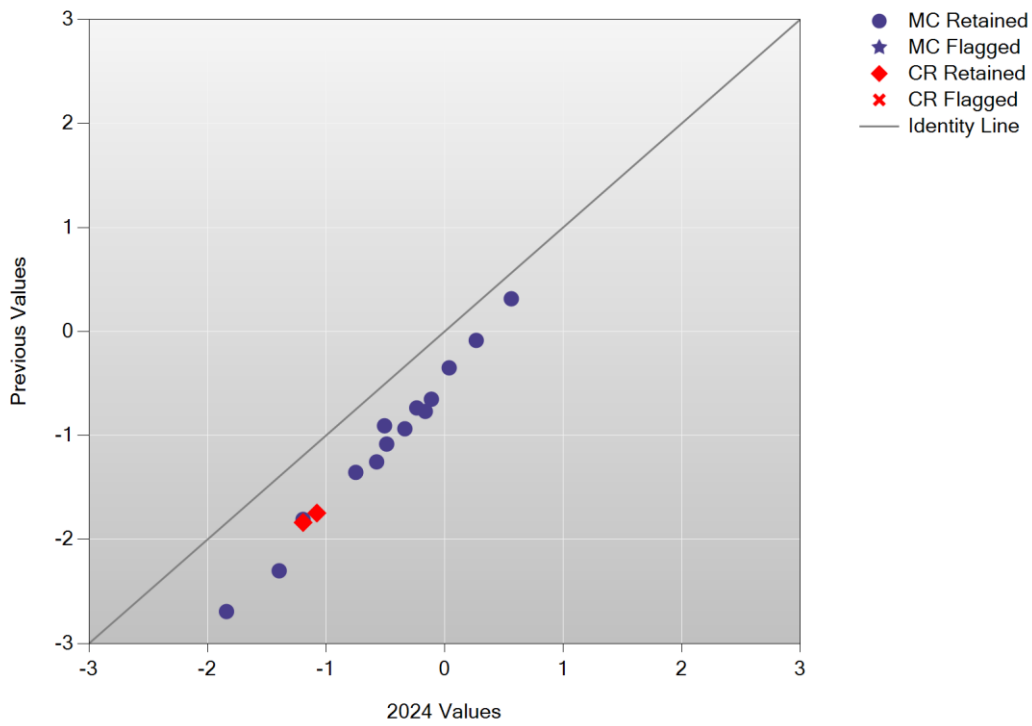
### Cumulative Scale Score Distributions: English Language Arts Grade 6



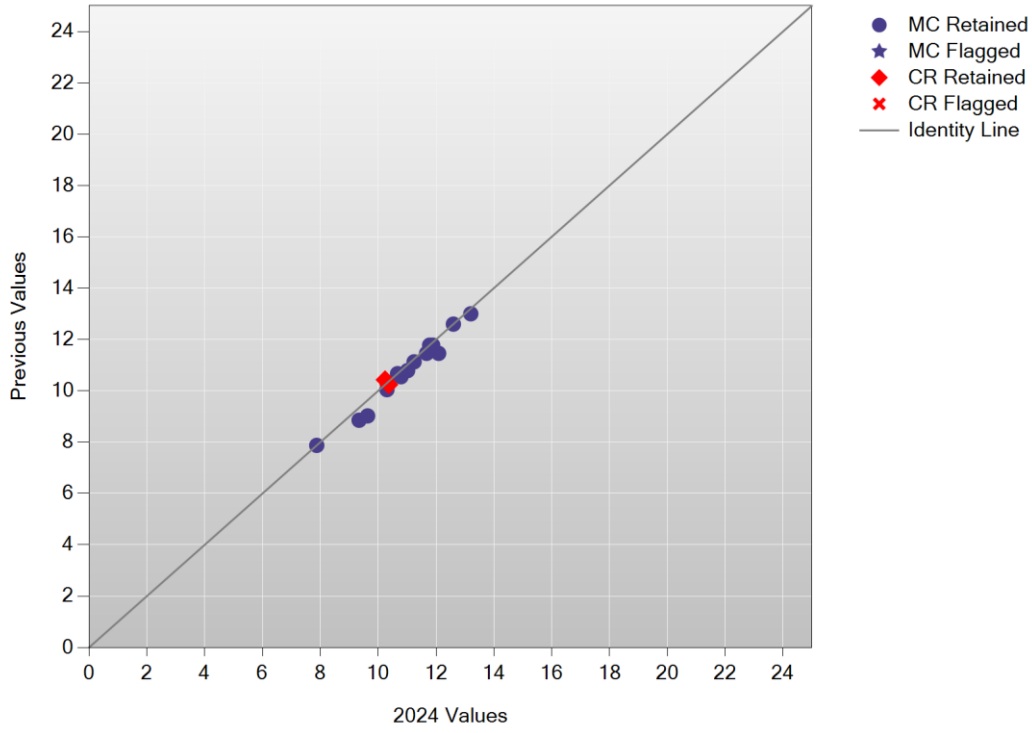
A/A Plot: English Language Arts Grade 7



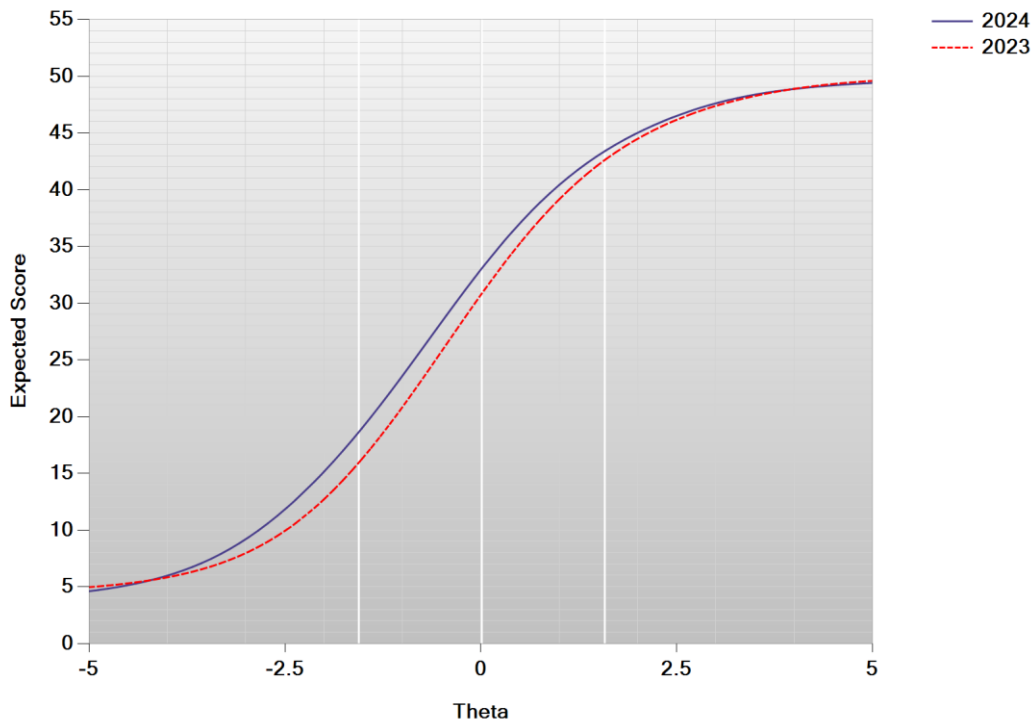
B/B Plot: English Language Arts Grade 7



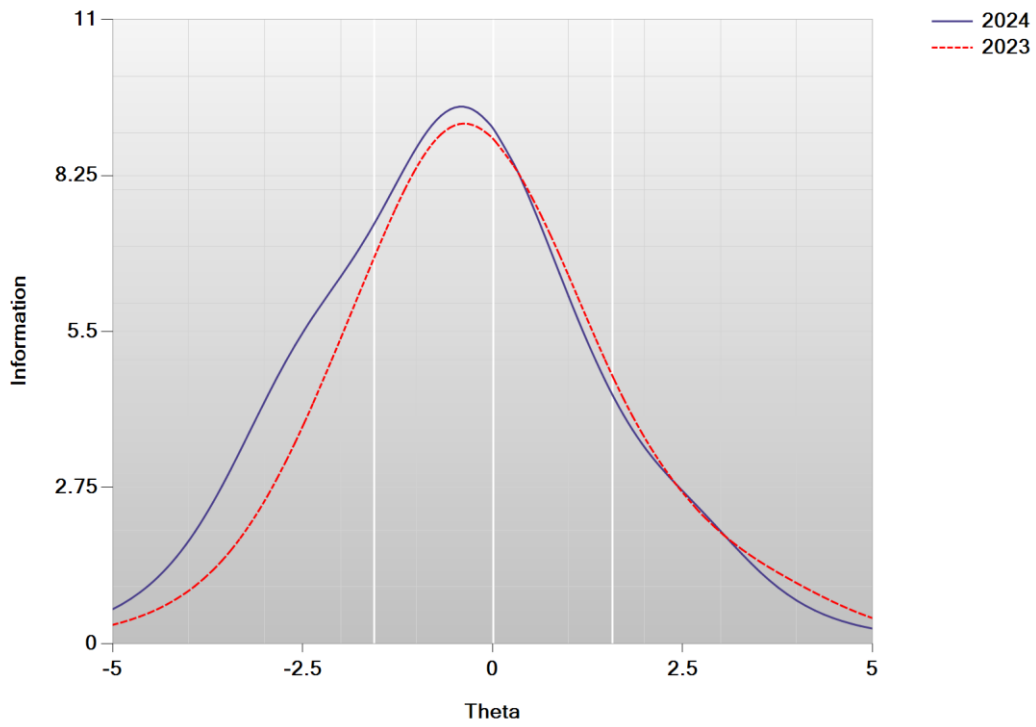
Delta Plot: English Language Arts Grade 7



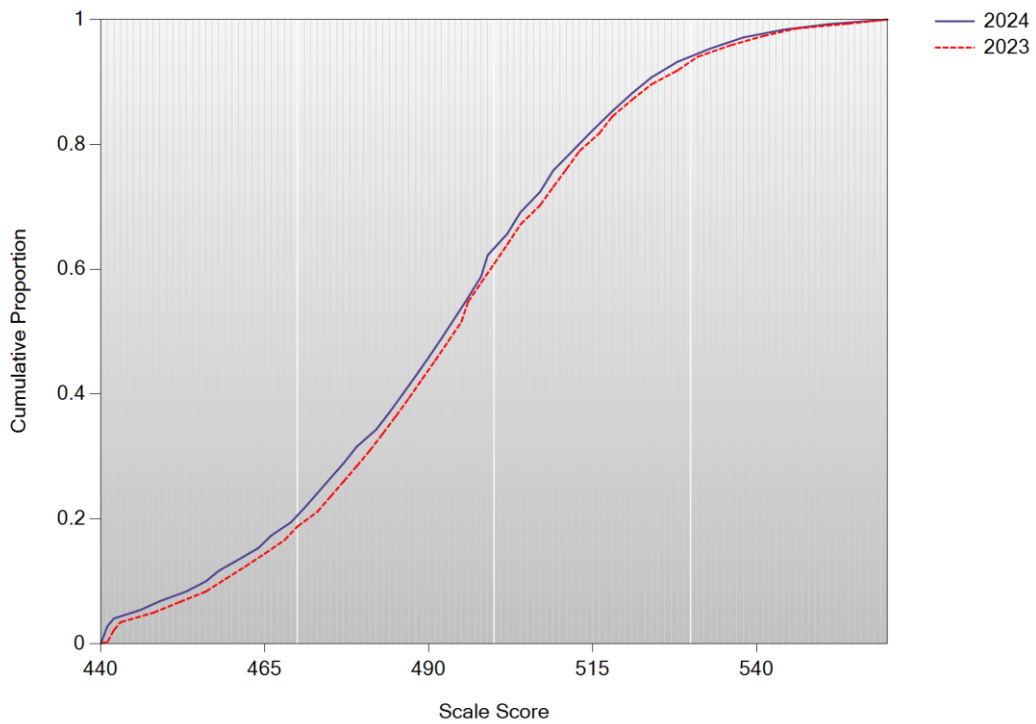
Test Characteristic Curve: English Language Arts Grade 7



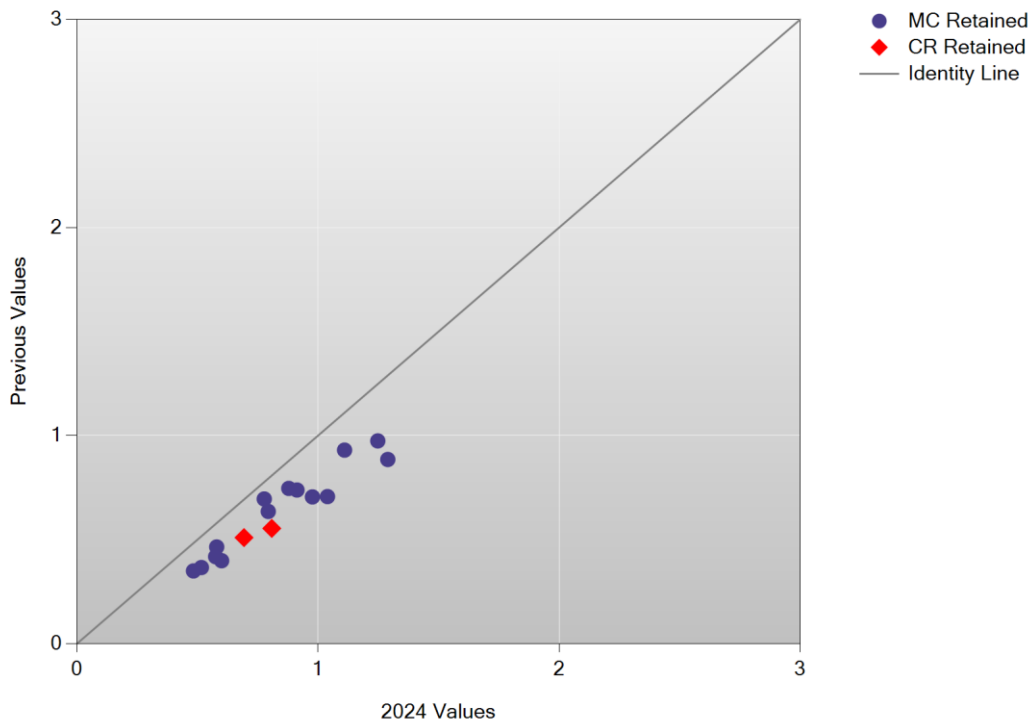
### Test Information Function: English Language Arts Grade 7



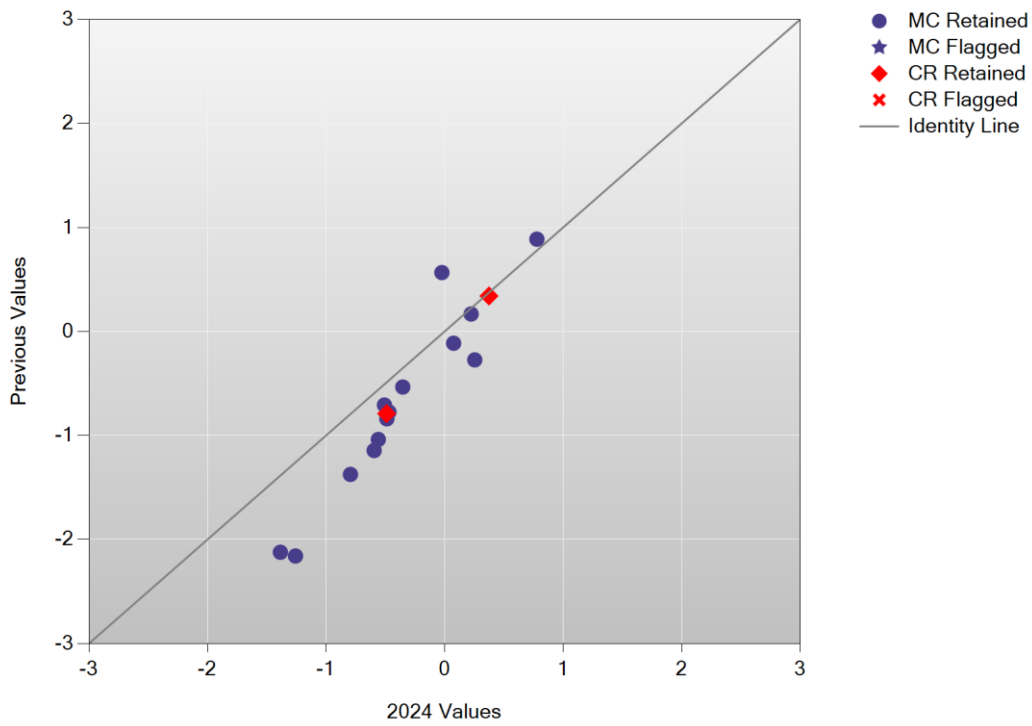
### Cumulative Scale Score Distributions: English Language Arts Grade 7



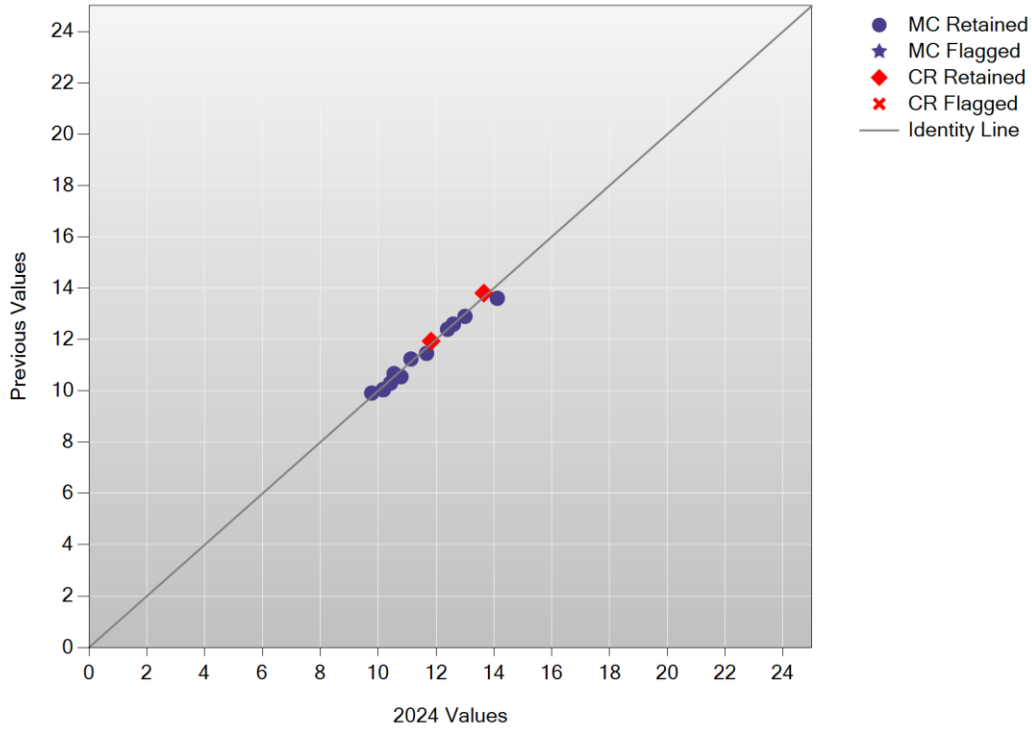
A/A Plot: English Language Arts Grade 8



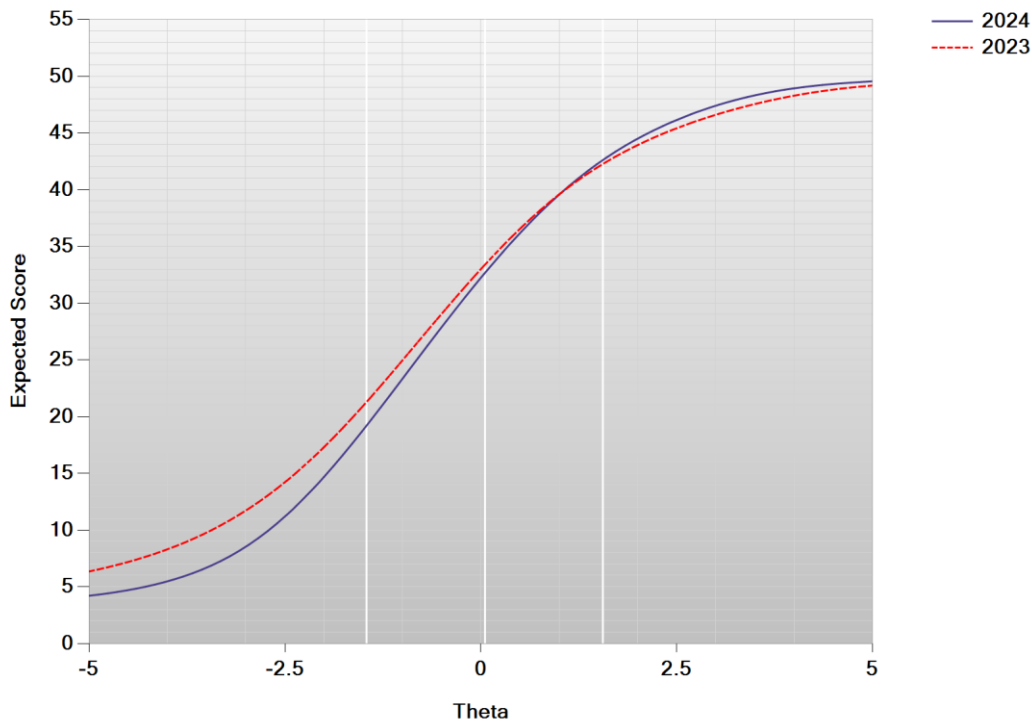
B/B Plot: English Language Arts Grade 8



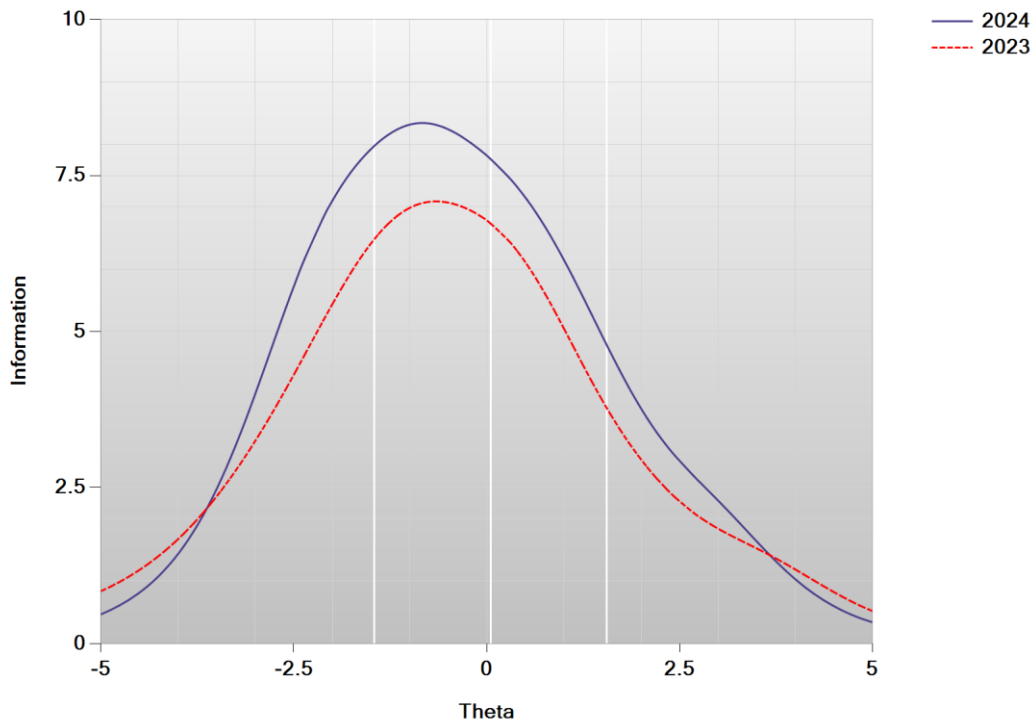
Delta Plot: English Language Arts Grade 8



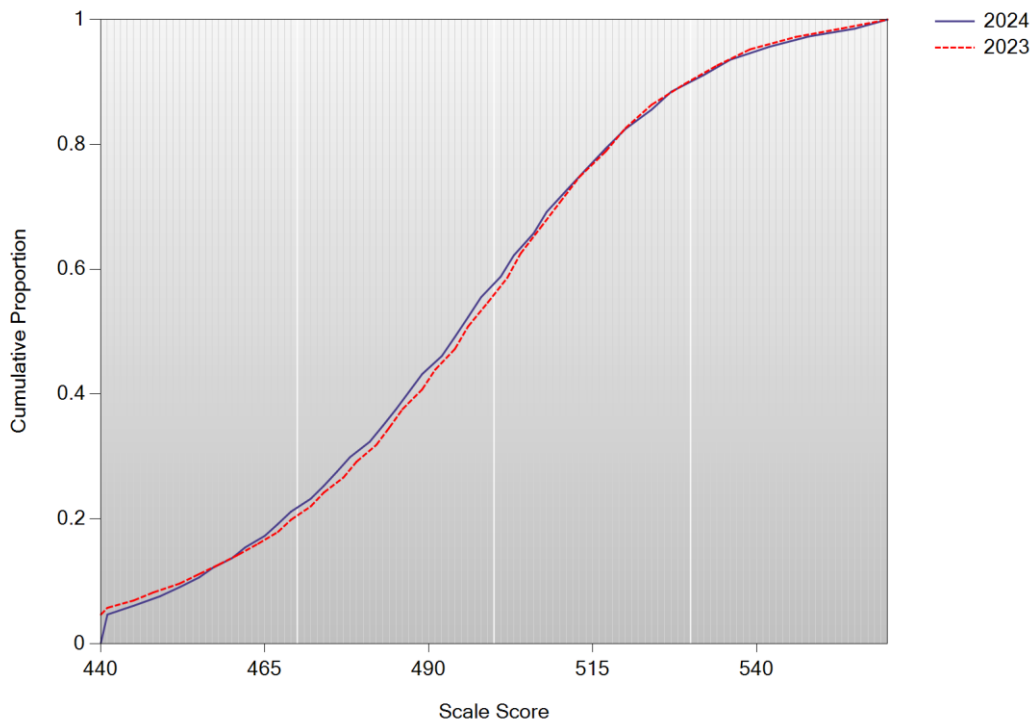
Test Characteristic Curve: English Language Arts Grade 8



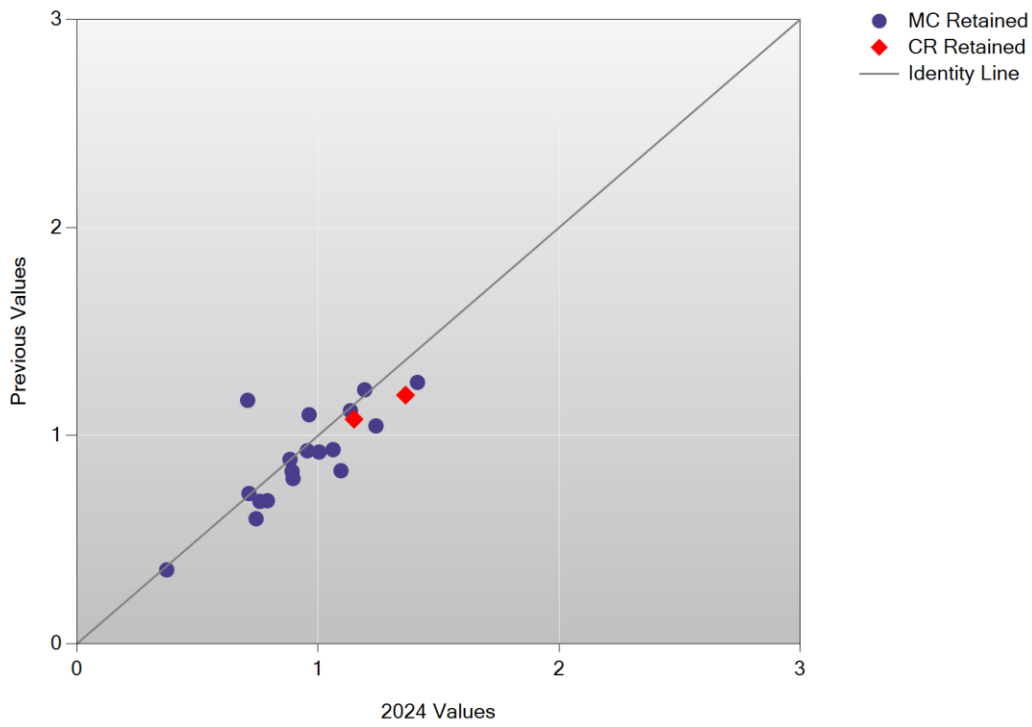
### Test Information Function: English Language Arts Grade 8



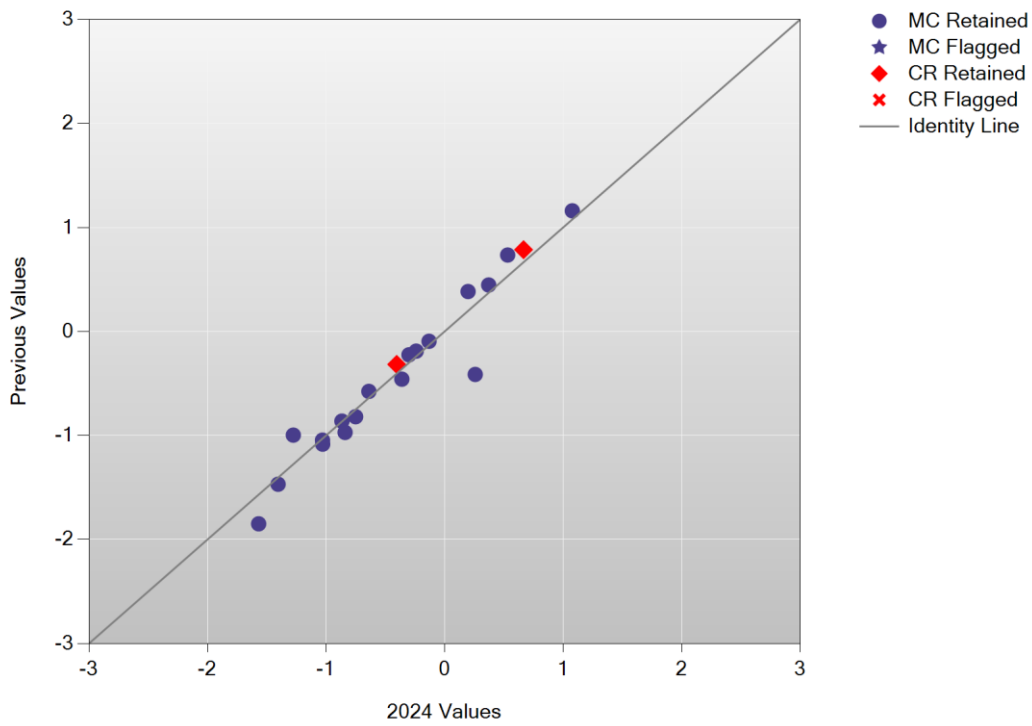
### Cumulative Scale Score Distributions: English Language Arts Grade 8



A/A Plot: Mathematics Grade 3

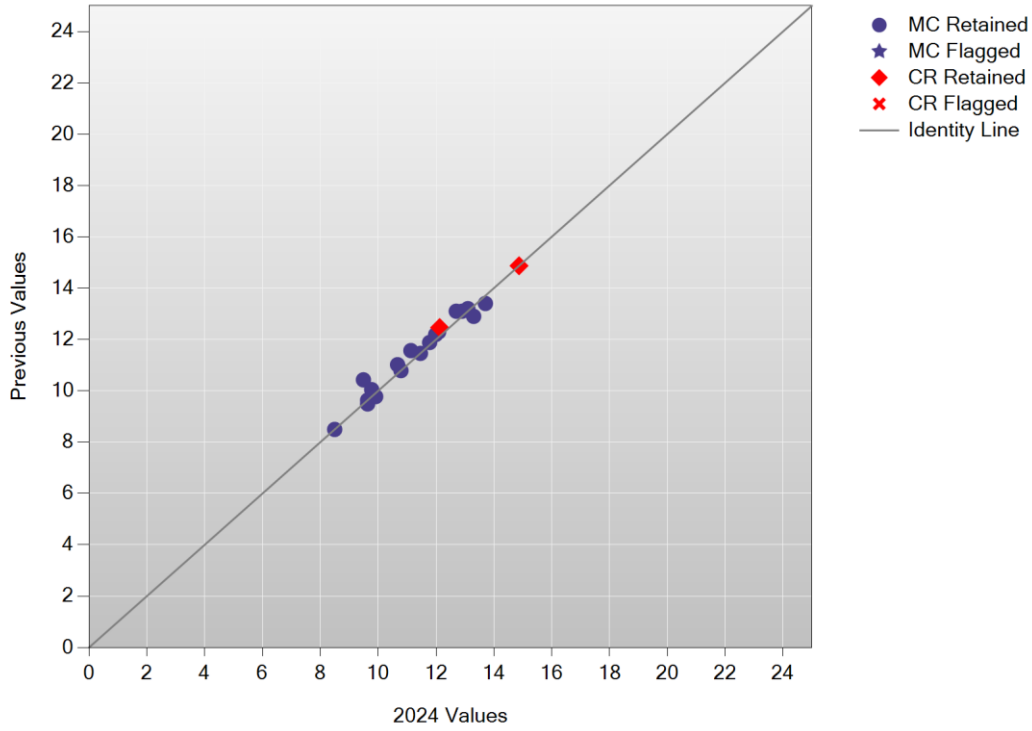


B/B Plot: Mathematics Grade 3

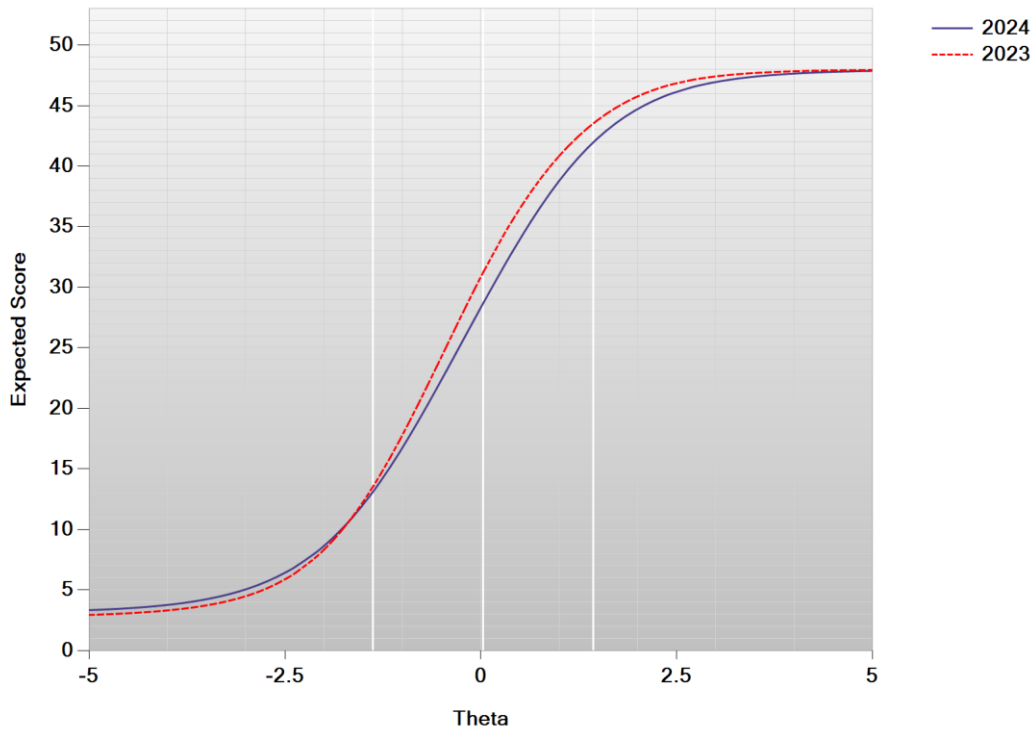




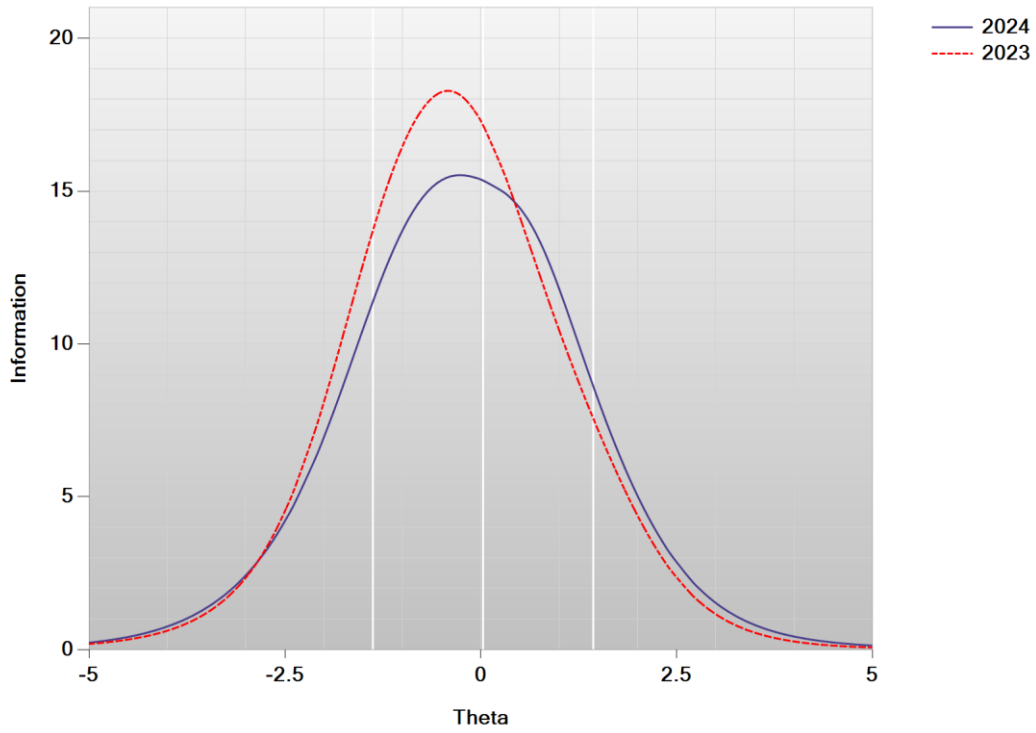
### Delta Plot: Mathematics Grade 3



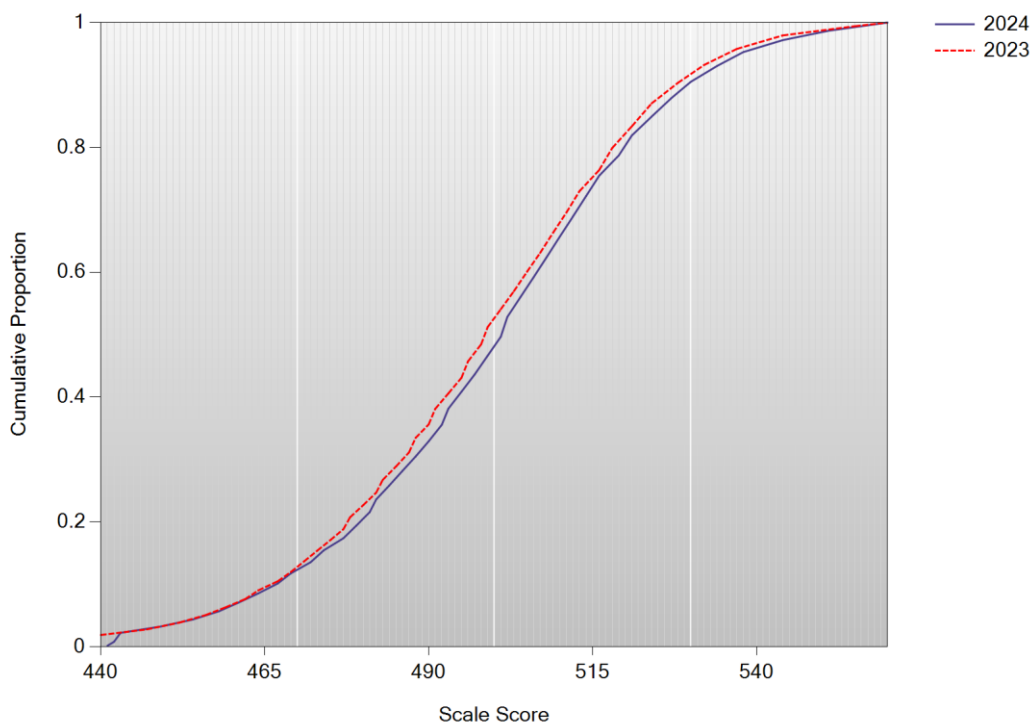
### Test Characteristic Curve: Mathematics Grade 3



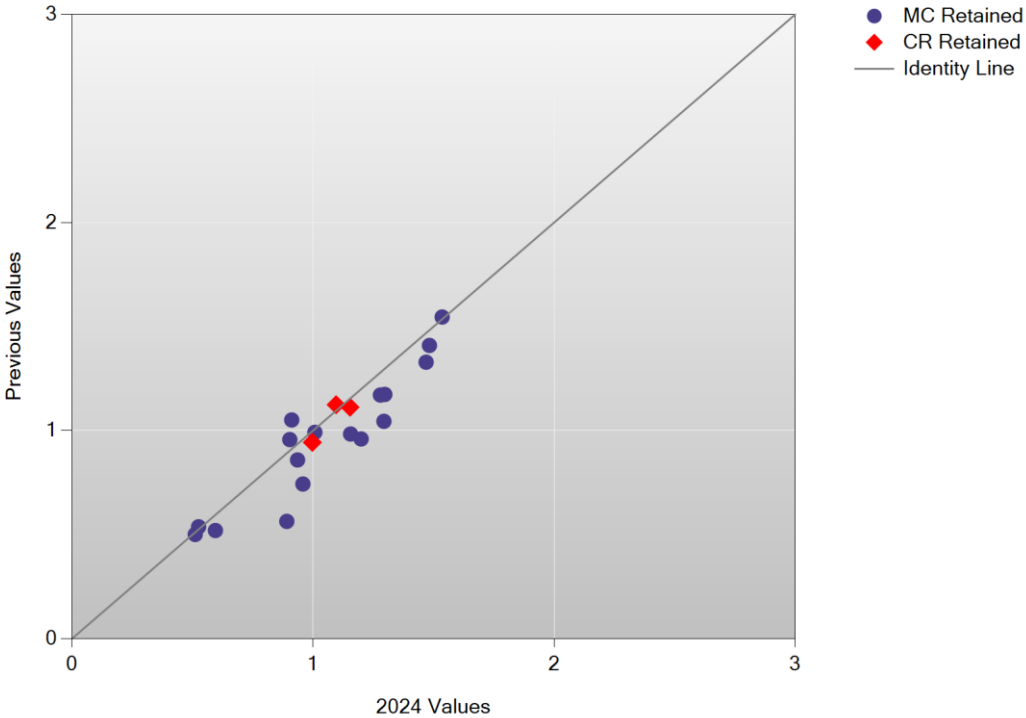
### Test Information Function: Mathematics Grade 3



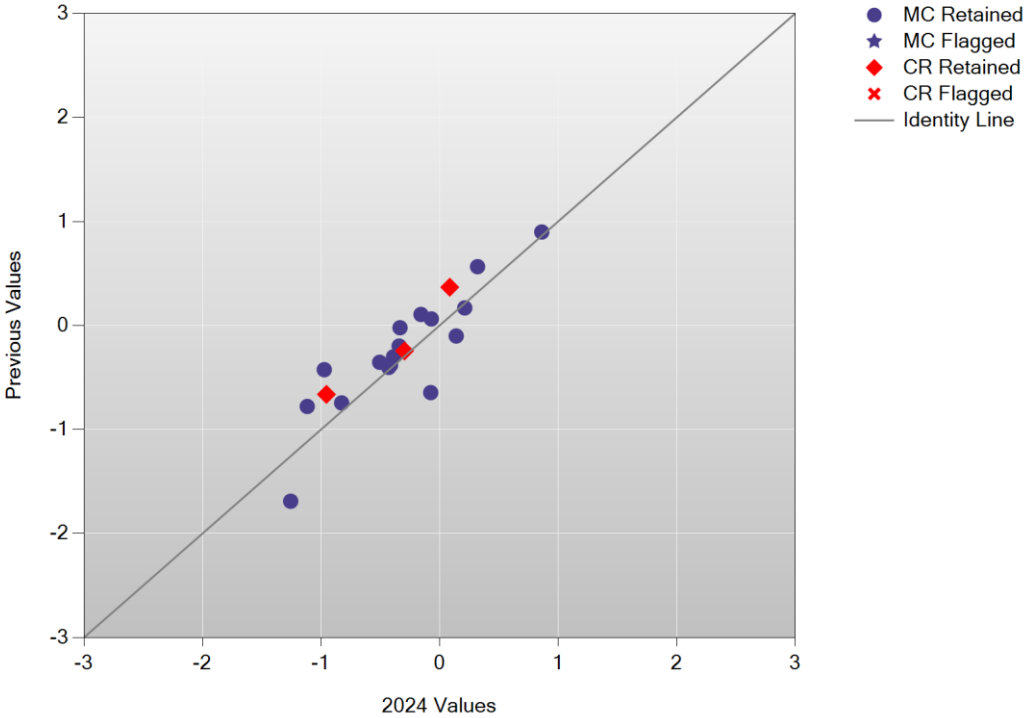
### Cumulative Scale Score Distributions: Mathematics Grade 3



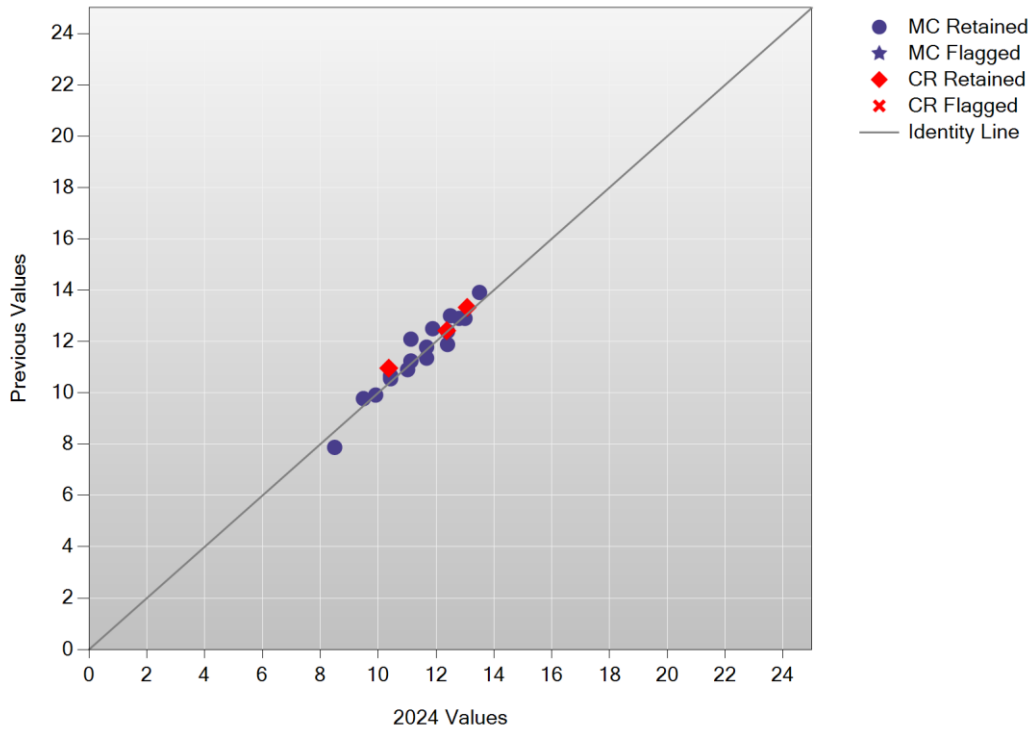
### A/A Plot: Mathematics Grade 4



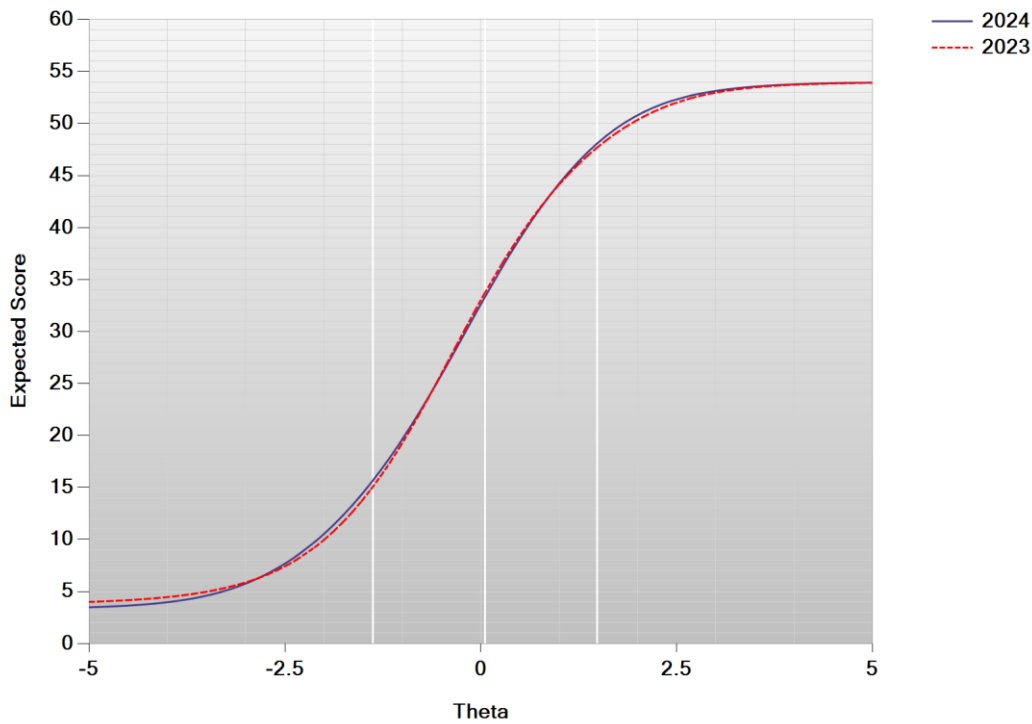
### B/B Plot: Mathematics Grade 4



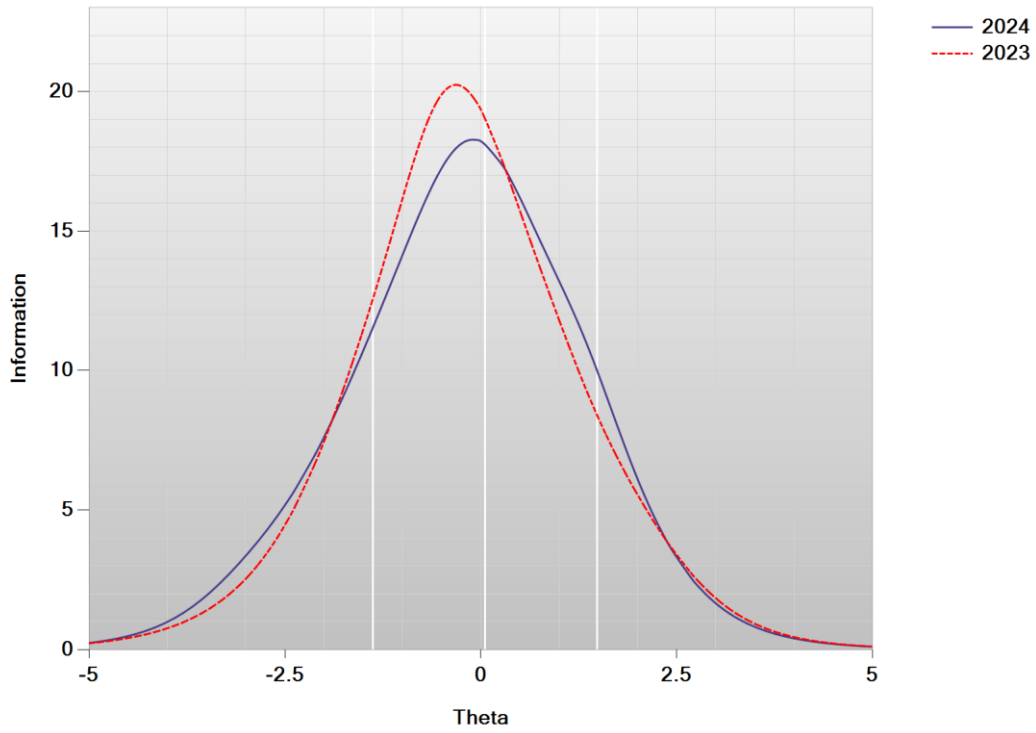
### Delta Plot: Mathematics Grade 4



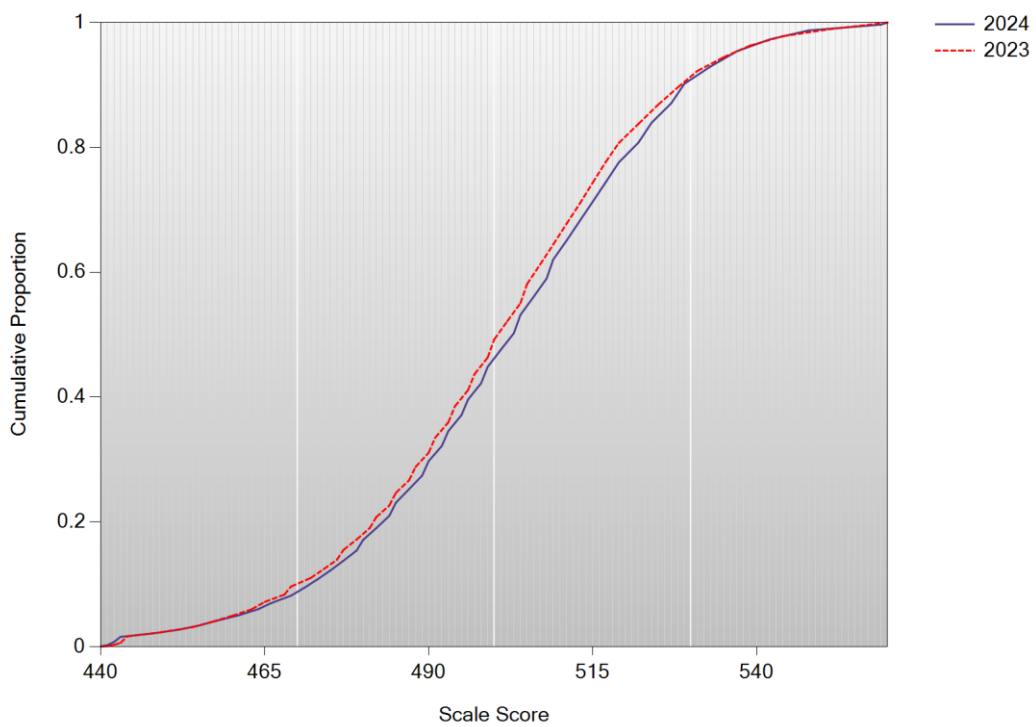
### Test Characteristic Curve: Mathematics Grade 4



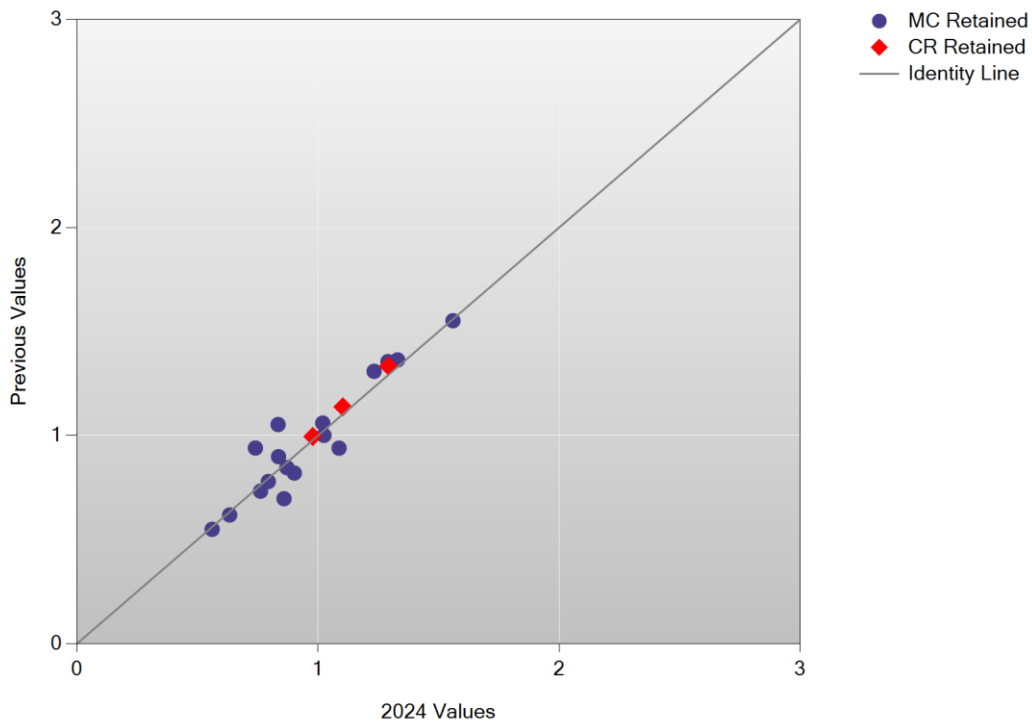
### Test Information Function: Mathematics Grade 4



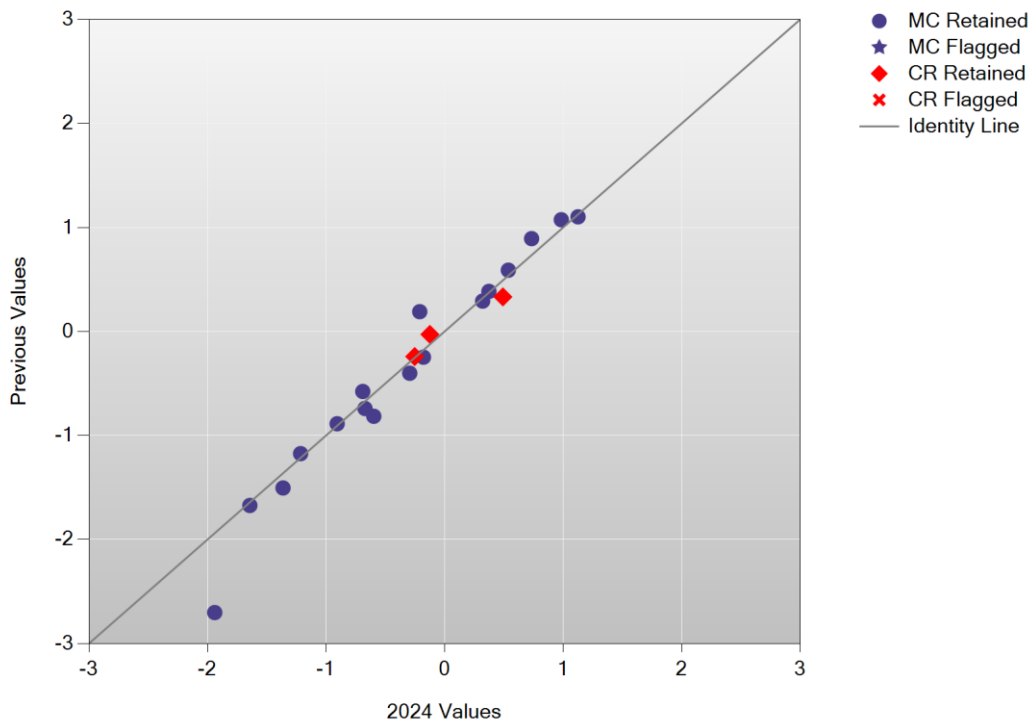
### Cumulative Scale Score Distributions: Mathematics Grade 4



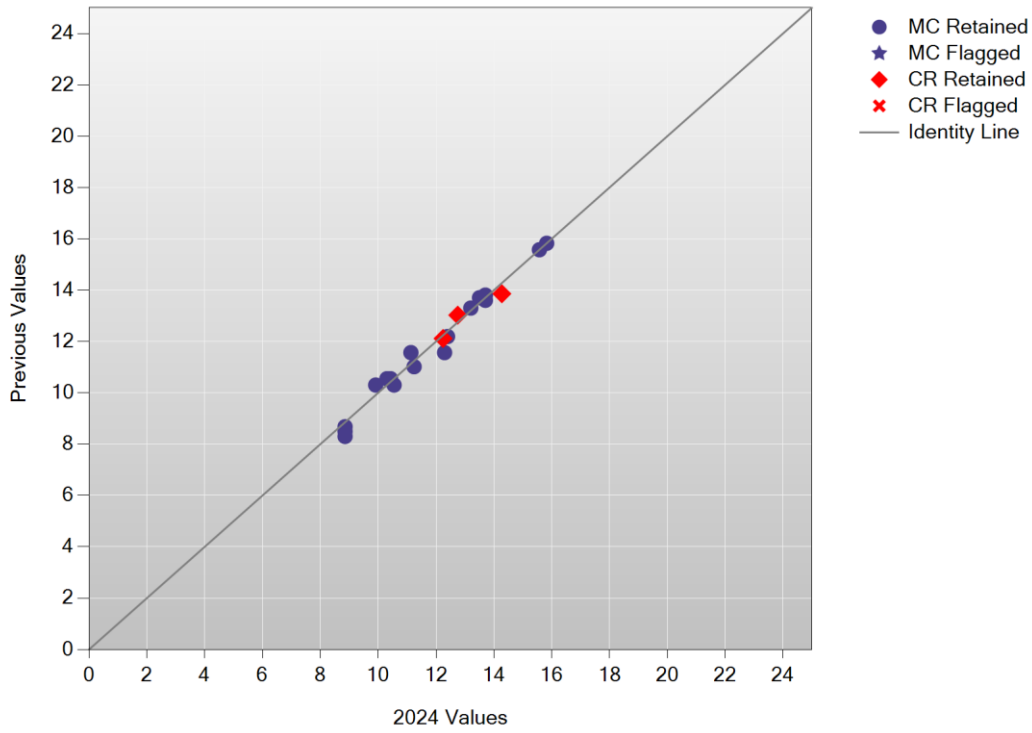
### A/A Plot: Mathematics Grade 5



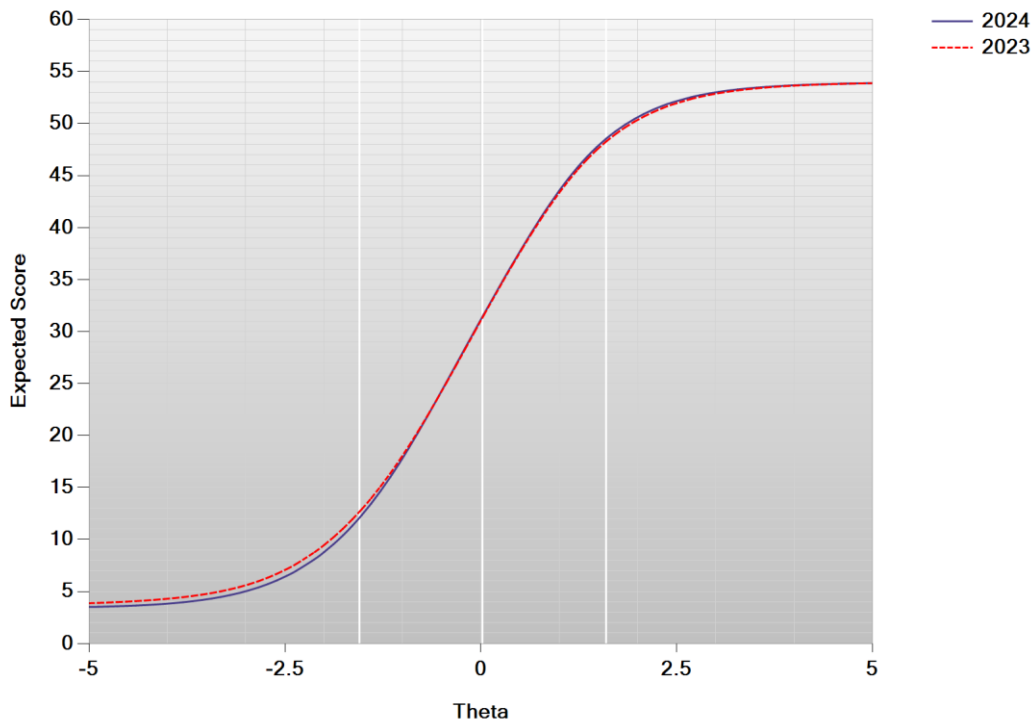
### B/B Plot: Mathematics Grade 5



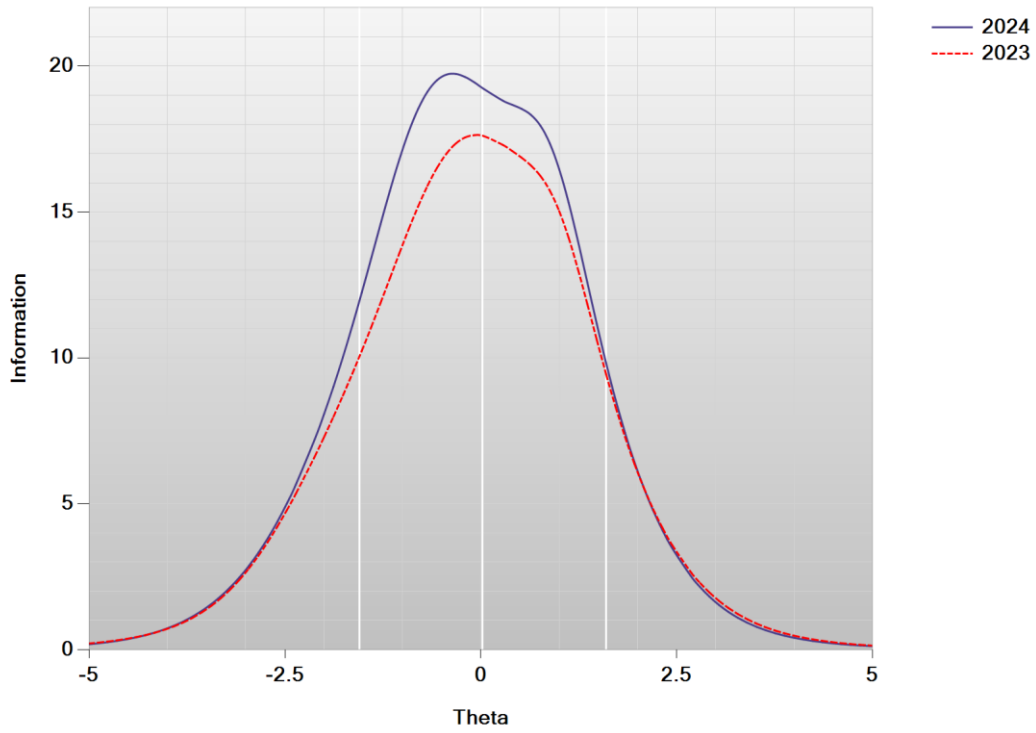
### Delta Plot: Mathematics Grade 5



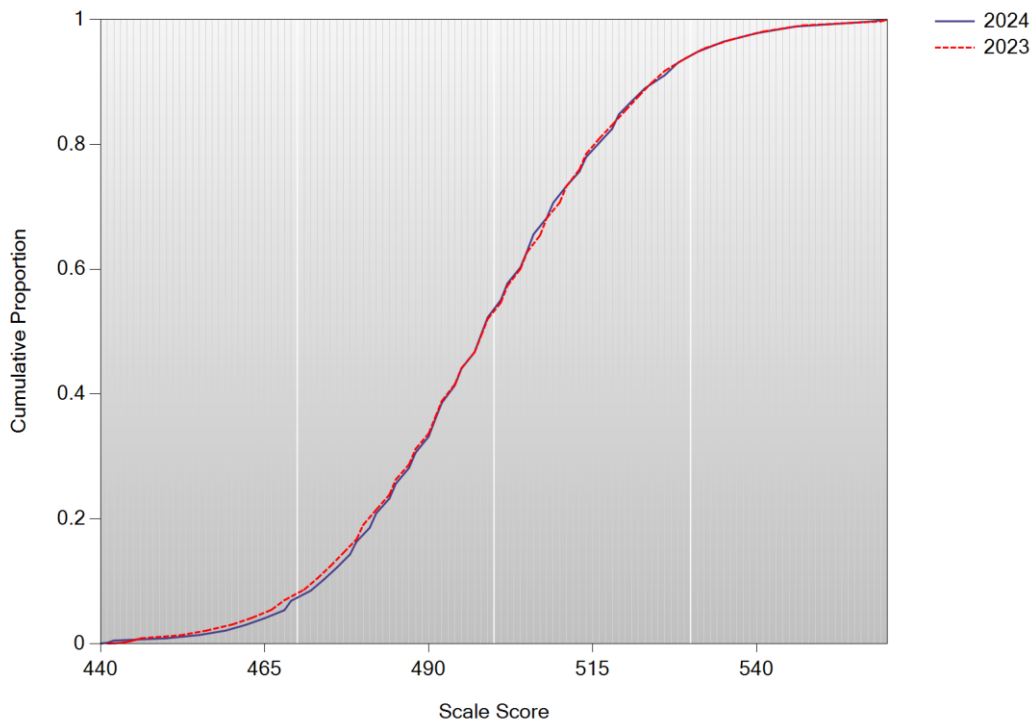
### Test Characteristic Curve: Mathematics Grade 5



### Test Information Function: Mathematics Grade 5

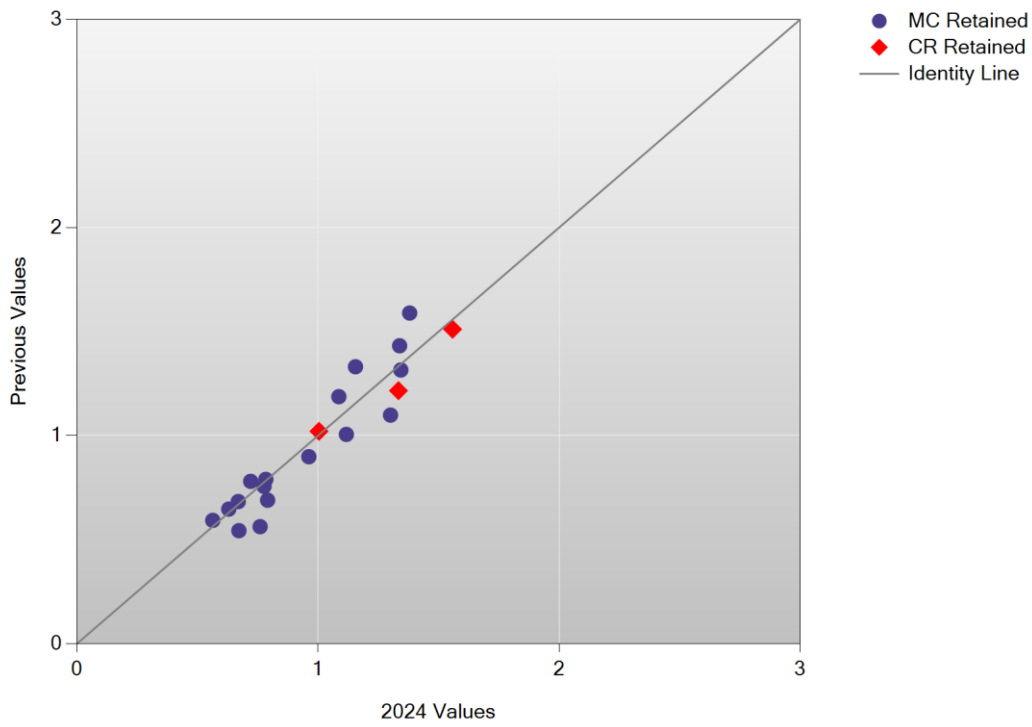


### Cumulative Scale Score Distributions: Mathematics Grade 5

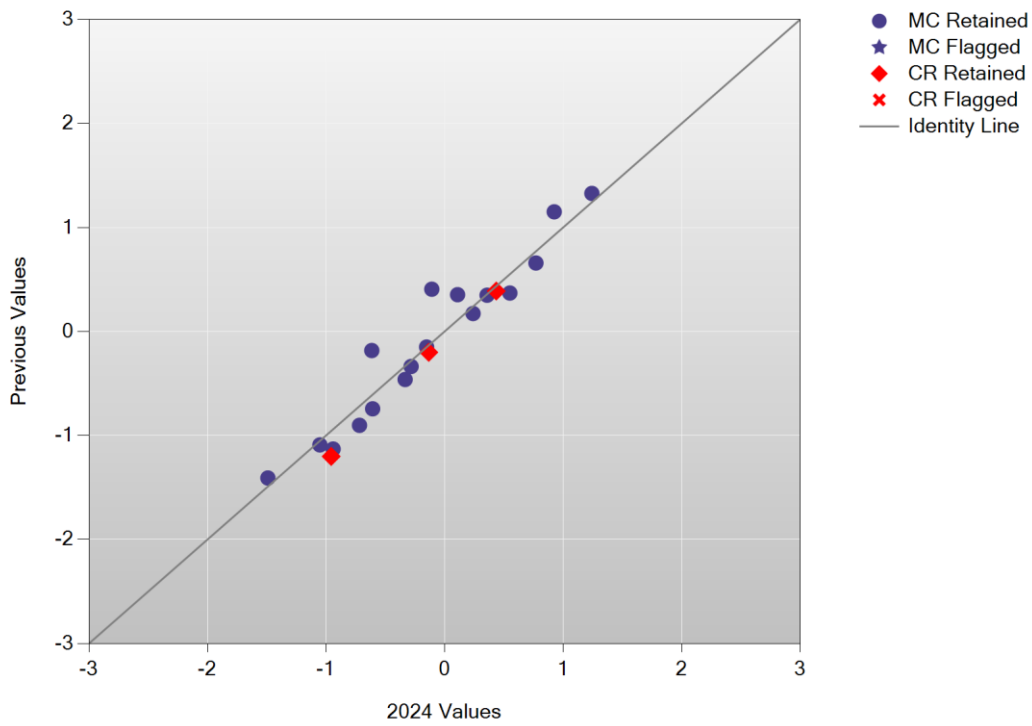




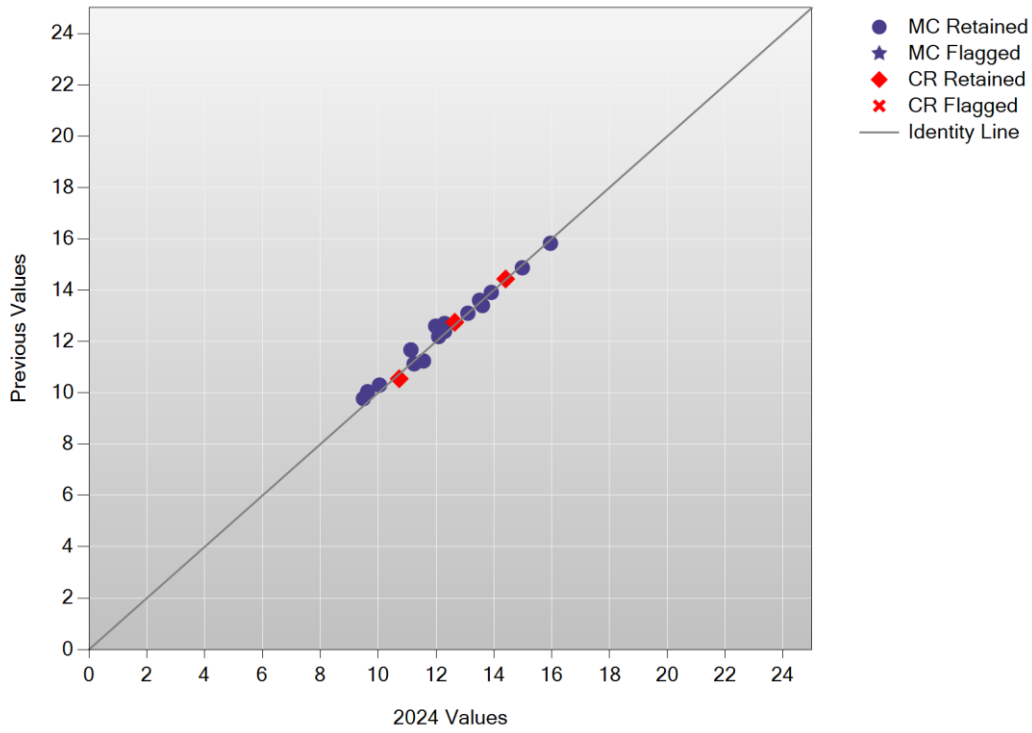
### A/A Plot: Mathematics Grade 6



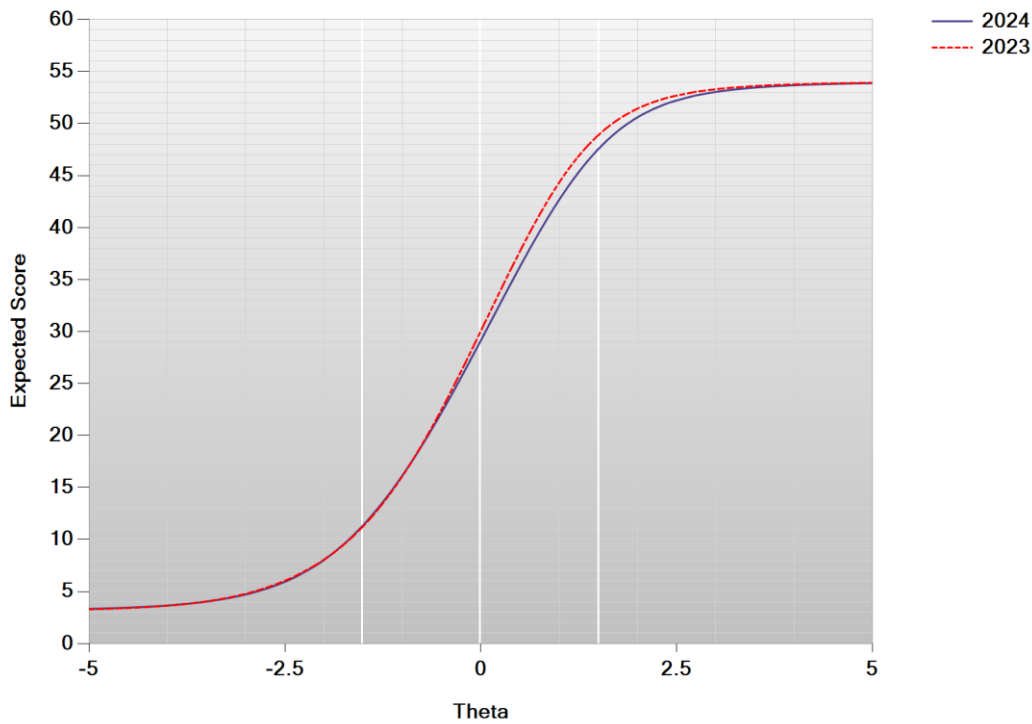
### B/B Plot: Mathematics Grade 6



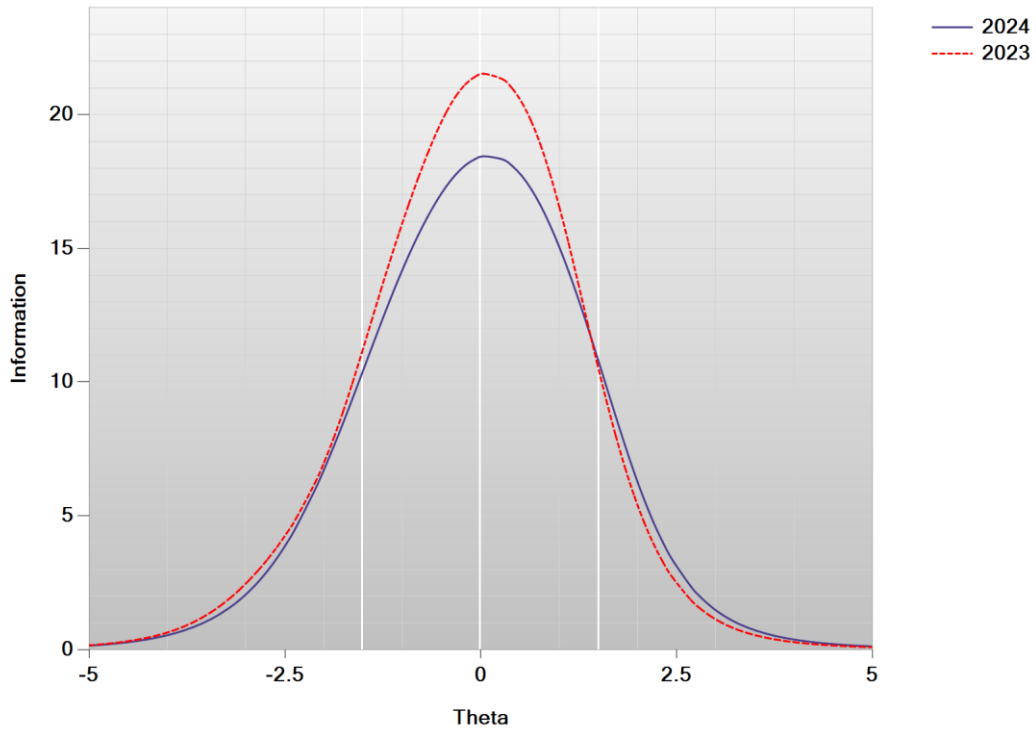
### Delta Plot: Mathematics Grade 6



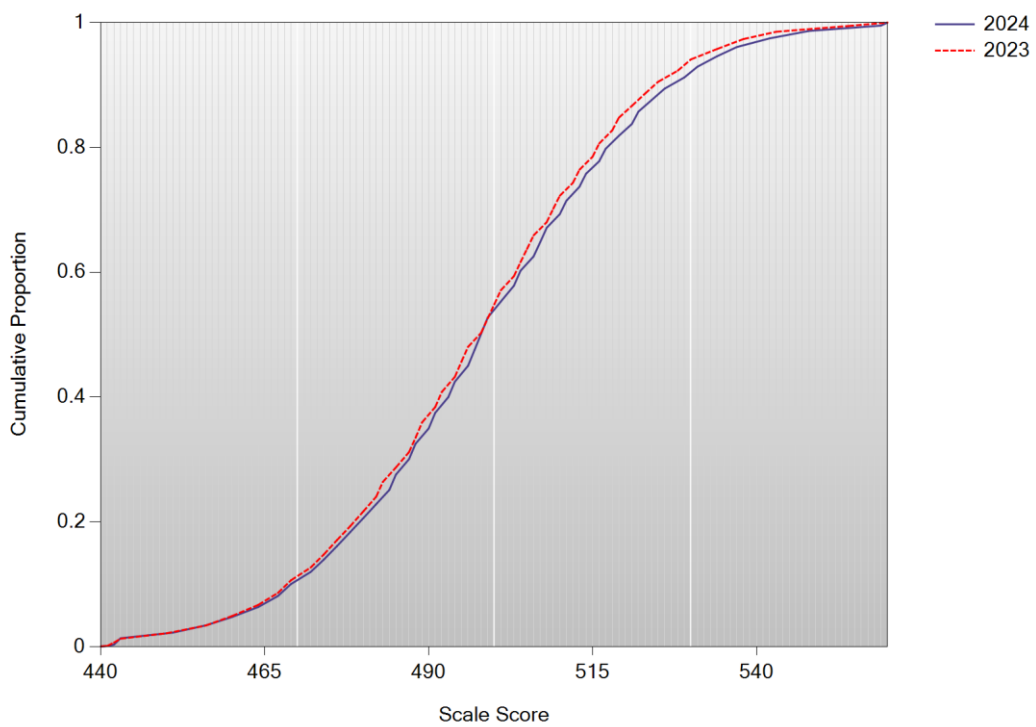
### Test Characteristic Curve: Mathematics Grade 6



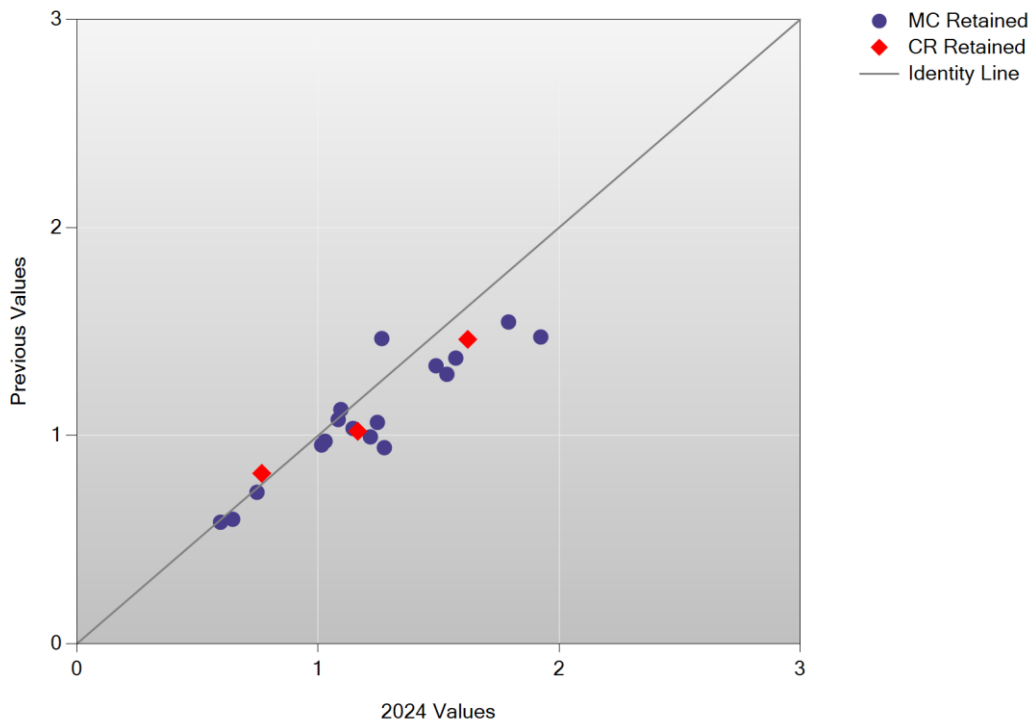
### Test Information Function: Mathematics Grade 6



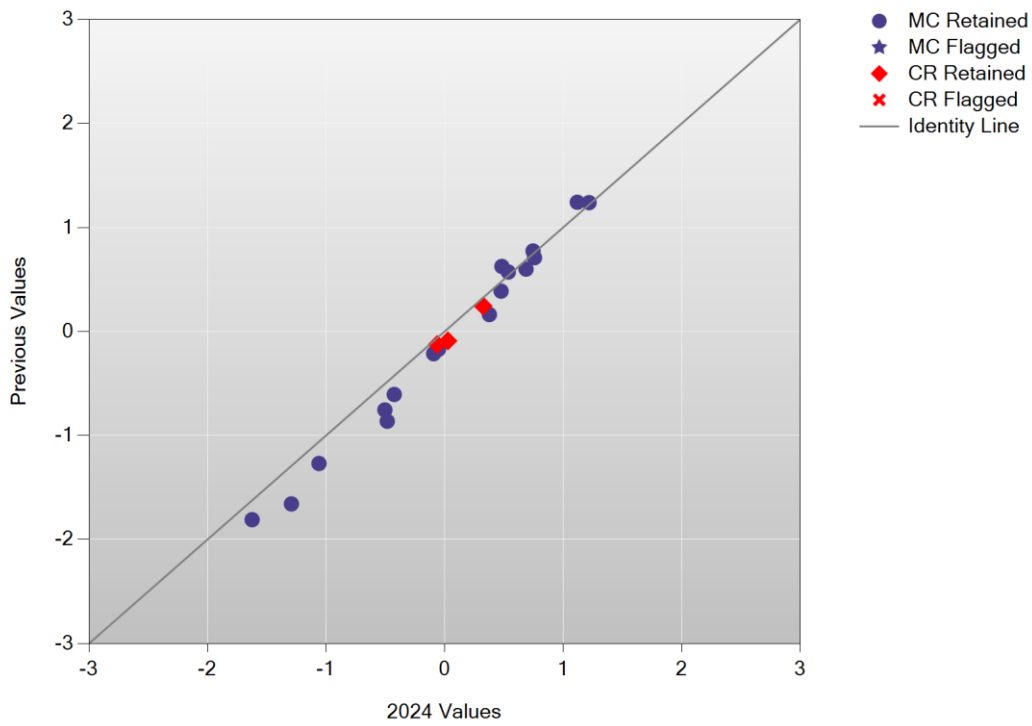
### Cumulative Scale Score Distributions: Mathematics Grade 6



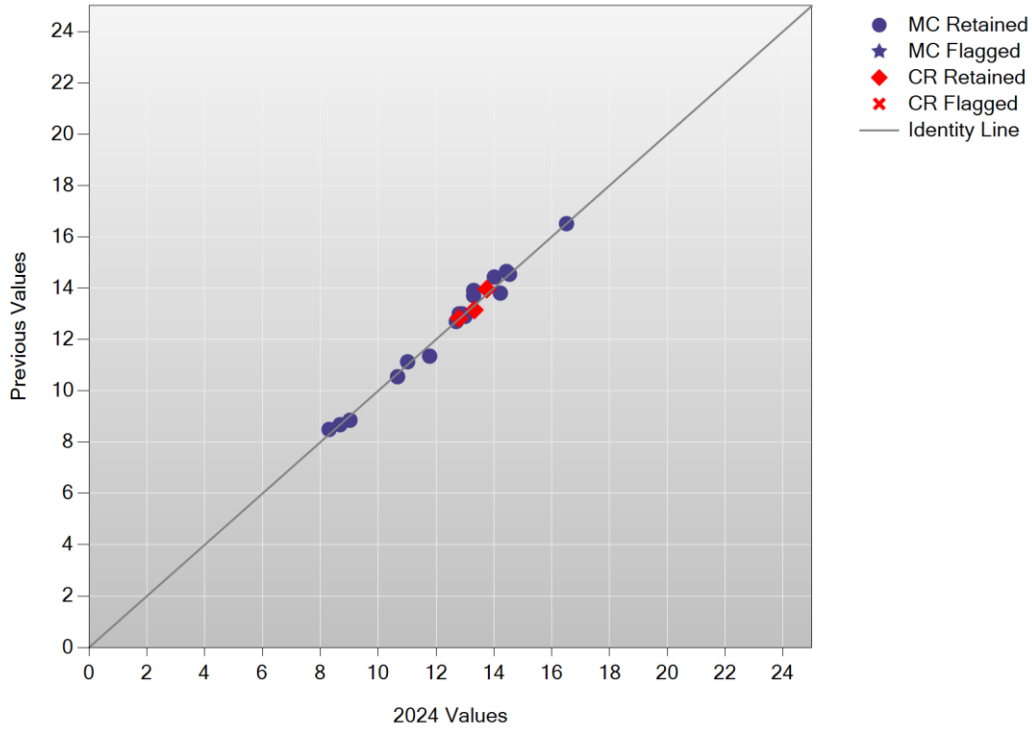
A/A Plot: Mathematics Grade 7



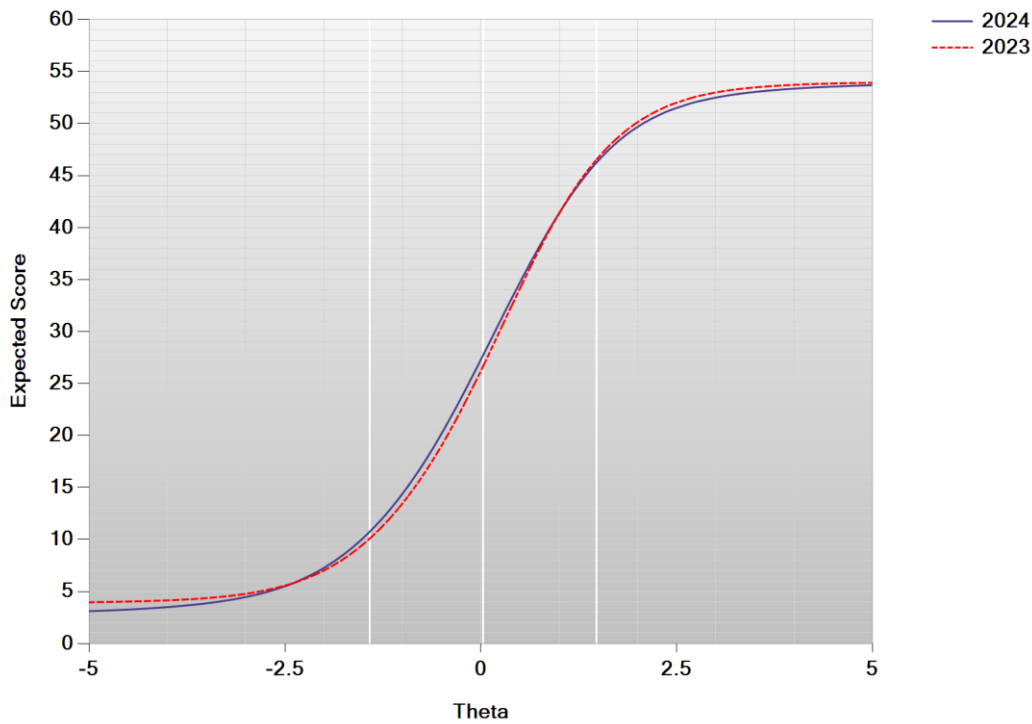
B/B Plot: Mathematics Grade 7



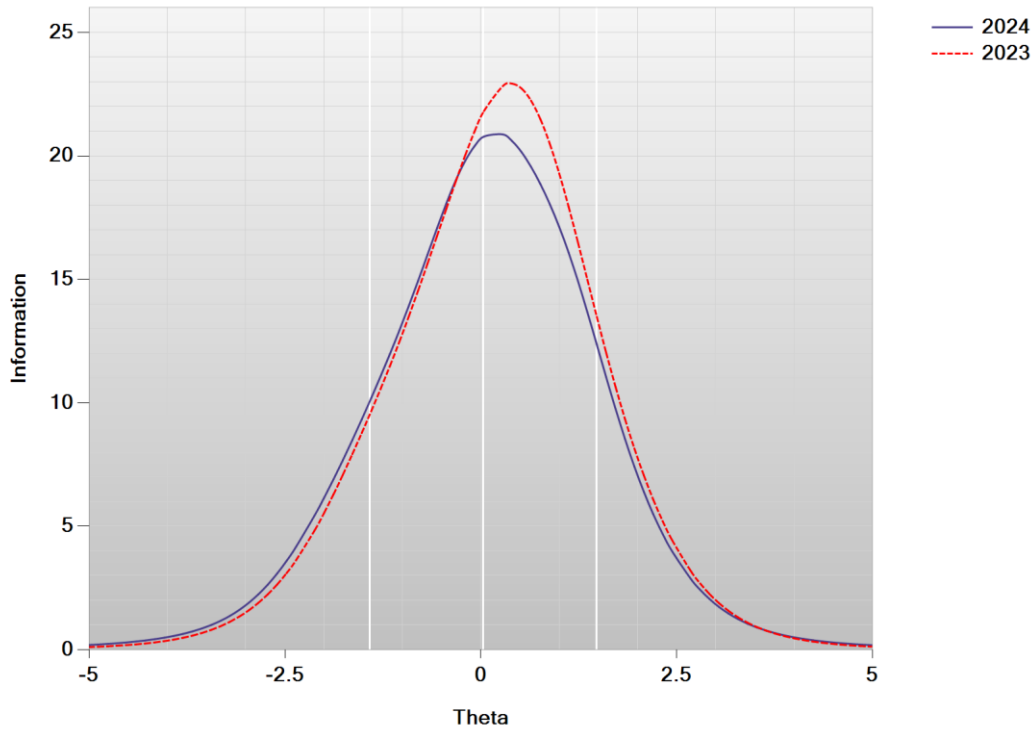
Delta Plot: Mathematics Grade 7



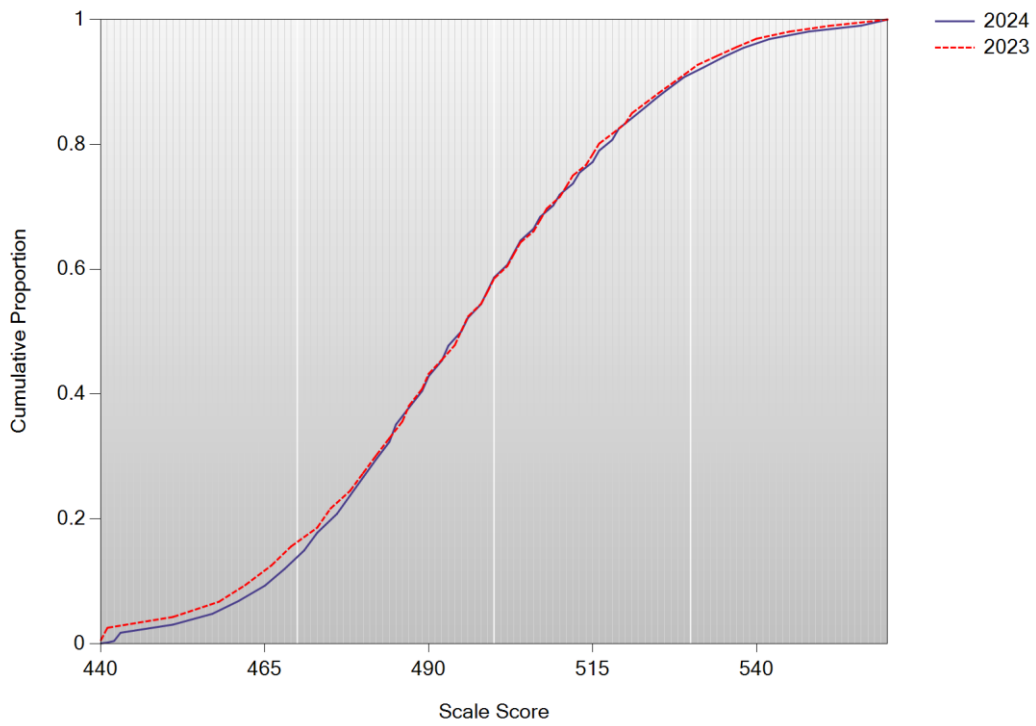
Test Characteristic Curve: Mathematics Grade 7



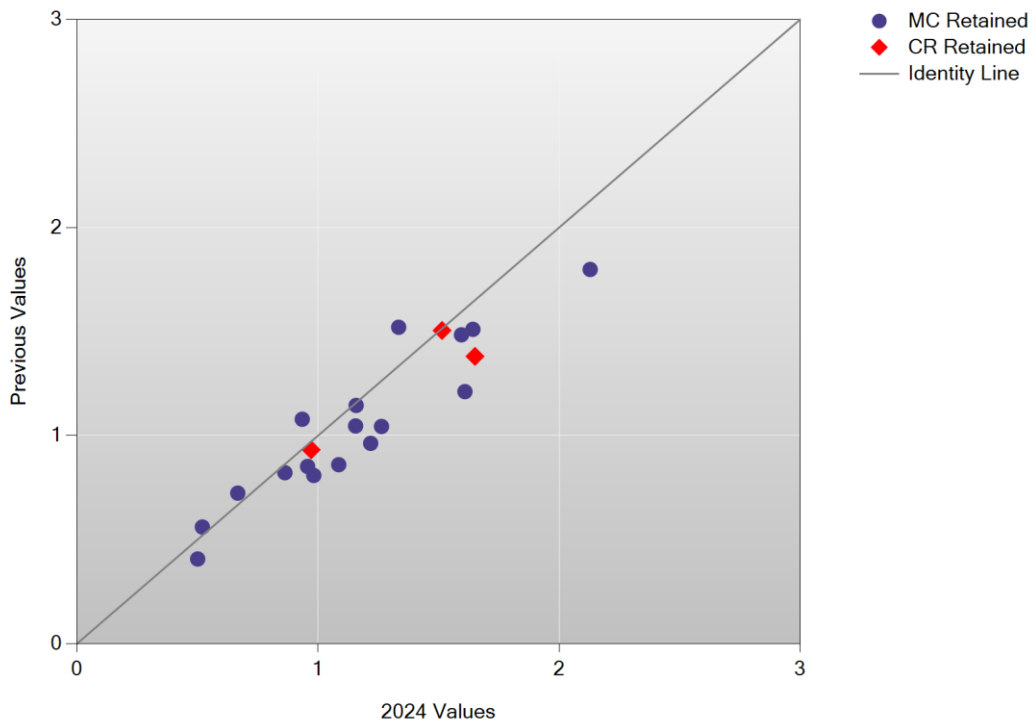
### Test Information Function: Mathematics Grade 7



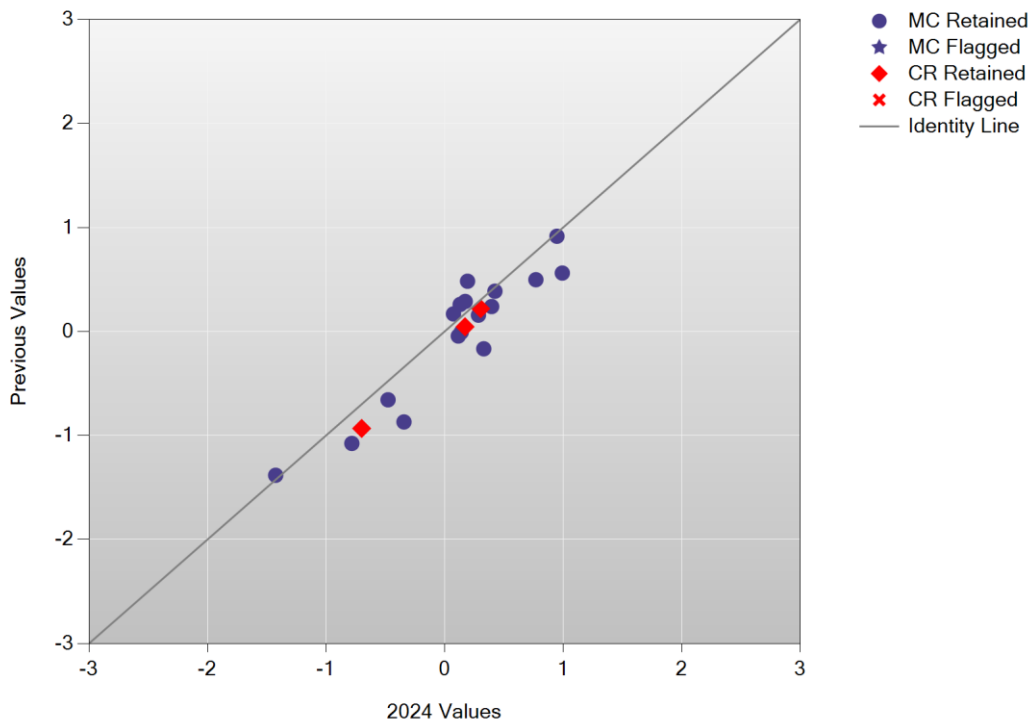
### Cumulative Scale Score Distributions: Mathematics Grade 7



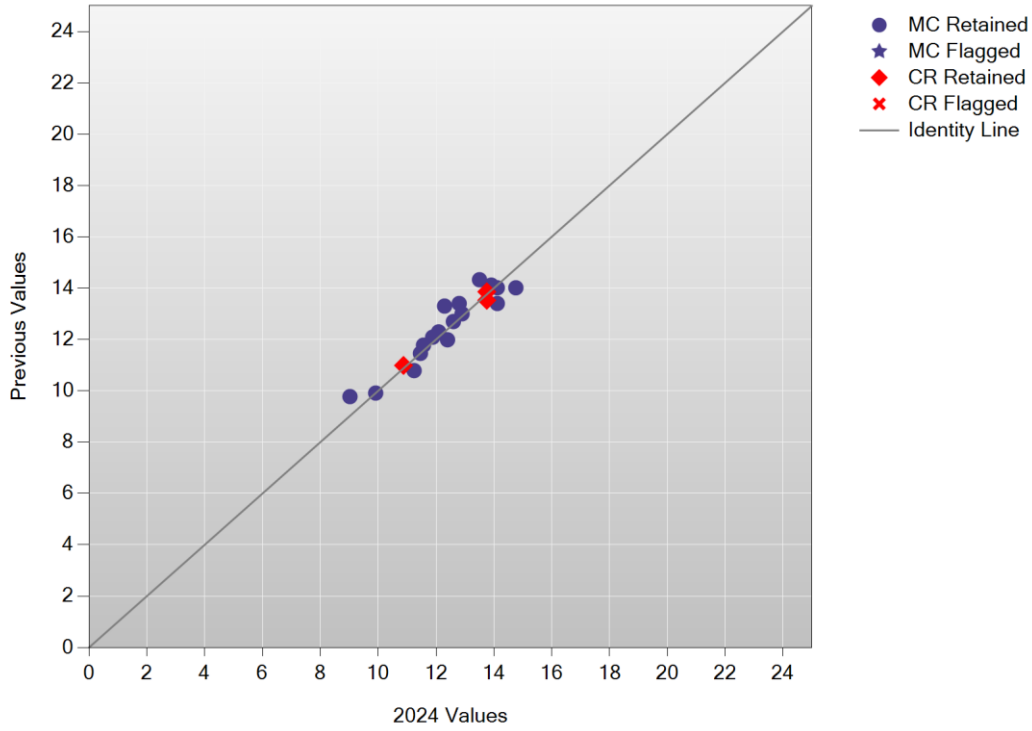
A/A Plot: Mathematics Grade 8



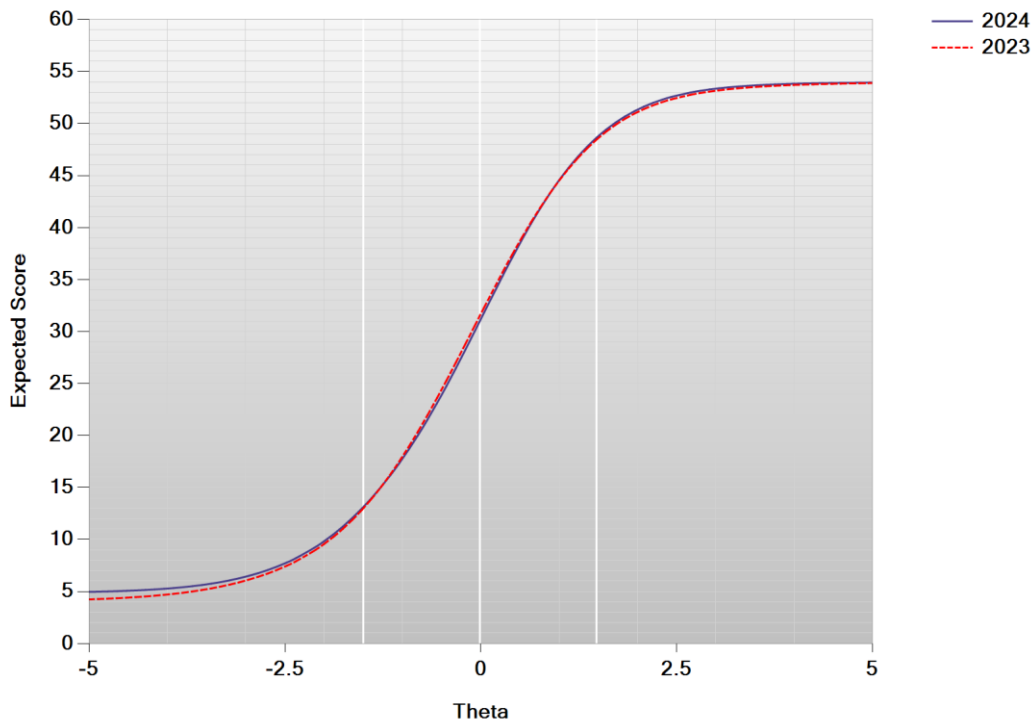
B/B Plot: Mathematics Grade 8



### Delta Plot: Mathematics Grade 8

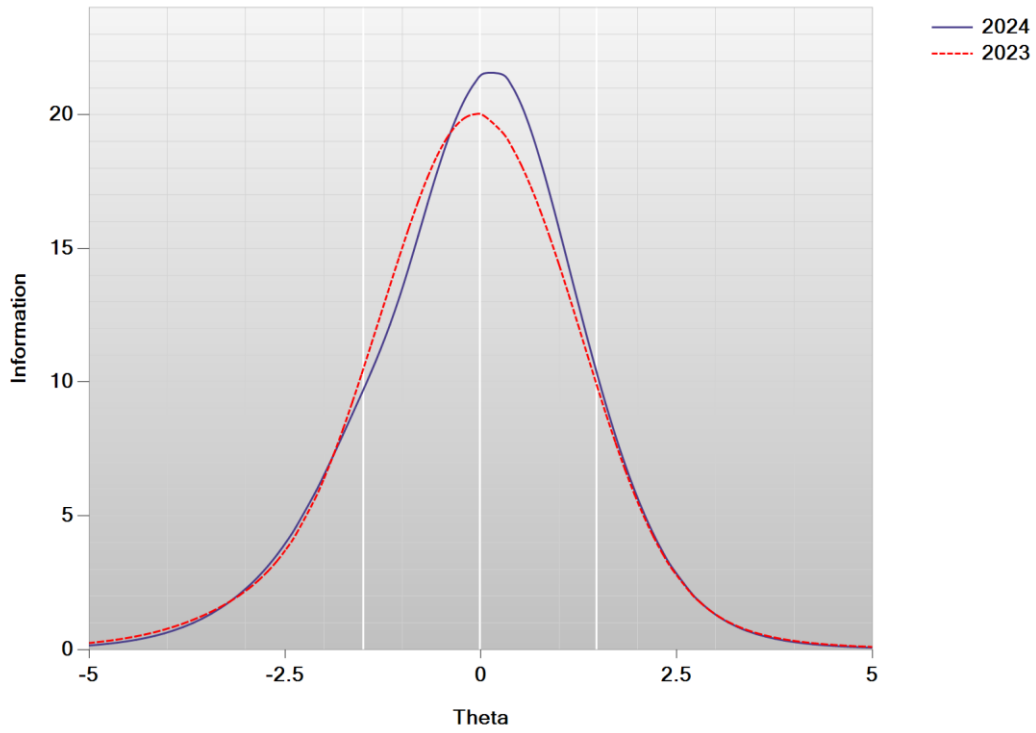


### Test Characteristic Curve: Mathematics Grade 8

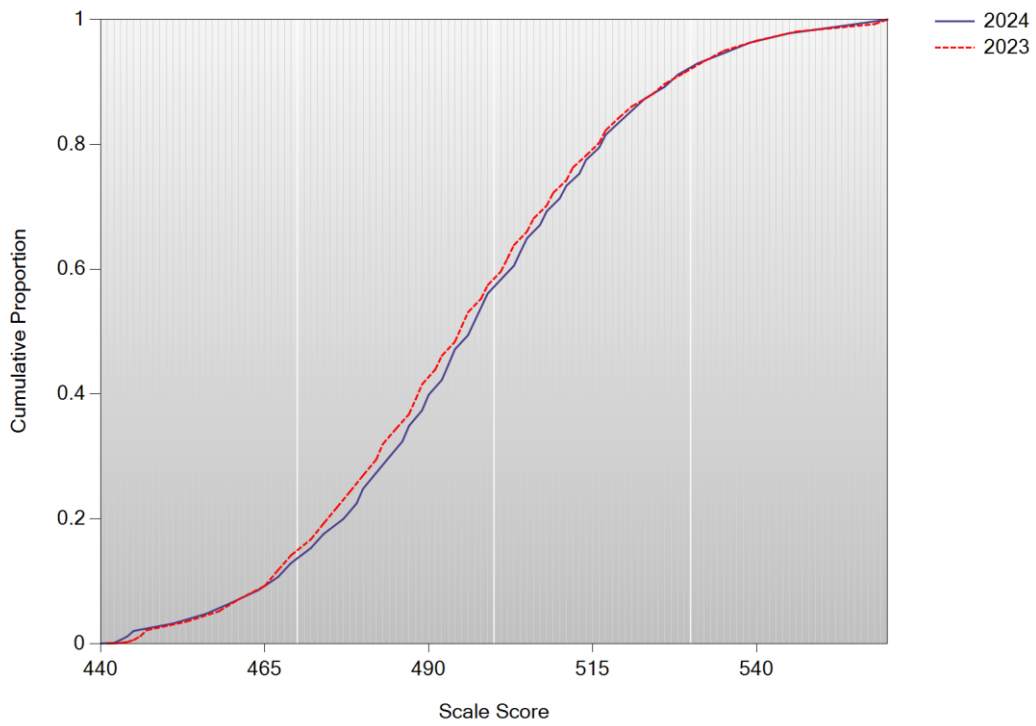




### Test Information Function: Mathematics Grade 8



### Cumulative Scale Score Distributions: Mathematics Grade 8



# Section 2.2

## Lookup Tables

Table 2.2.1  
Raw Score to Scale Score Lookup Table  
English Language Arts Grade 3

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.173	1.64	10.0	440	1	440	1
1	-3.144	1.68	10.0	441	1	440	1
2	-3.115	1.73	10.0	441	1	440	1
3	-3.086	1.78	10.0	442	1	440	1
4	-3.057	1.82	10.0	442	1	440	1
5	-3.028	1.87	10.0	443	1	440	1
6	-2.998	1.92	10.0	443	1	440	1
7	-2.969	1.98	10.0	444	1	449	1
8	-2.628	2.72	10.0	450	1	454	1
9	-2.360	3.51	10.0	455	1	459	1
10	-2.137	4.31	9.1	460	1	462	1
11	-1.945	5.11	8.3	463	1	465	1
12	-1.775	5.88	7.8	466	1	468	1
13	-1.621	6.61	7.3	469	1	471	2
14	-1.478	7.29	7.0	472	2	473	2
15	-1.345	7.91	6.7	474	2	475	2
16	-1.218	8.45	6.5	477	2	477	2
17	-1.095	8.90	6.3	479	2	479	2
18	-0.977	9.25	6.2	481	2	481	2
19	-0.861	9.51	6.1	484	2	483	2
20	-0.746	9.68	6.1	486	2	485	2
21	-0.632	9.76	6.0	488	2	487	2
22	-0.518	9.77	6.0	490	2	489	2
23	-0.404	9.73	6.0	492	2	491	2
24	-0.288	9.65	6.1	494	2	492	2
25	-0.171	9.54	6.1	497	2	494	2
26	-0.052	9.42	6.1	499	2	496	2
27	0.070	9.28	6.2	501	3	498	2
28	0.194	9.11	6.2	503	3	500	3
29	0.322	8.90	6.3	506	3	502	3
30	0.454	8.63	6.4	508	3	504	3
31	0.592	8.29	6.5	511	3	507	3
32	0.737	7.86	6.7	514	3	509	3
33	0.891	7.37	6.9	517	3	512	3
34	1.056	6.80	7.2	520	3	515	3
35	1.236	6.20	7.6	523	3	518	3
36	1.433	5.58	8.0	527	3	522	3
37	1.653	4.97	8.5	531	4	526	3
38	1.900	4.39	9.0	536	4	532	4
39	2.183	3.81	9.7	541	4	538	4
40	2.517	3.16	10.0	547	4	545	4
41	2.936	2.40	10.0	555	4	555	4
42	3.196	1.99	10.0	560	4	560	4
43	3.196	1.99	10.0	560	4	560	4
44	3.196	1.99	10.0	560	4	560	4

Table 2.2.2  
Raw Score to Scale Score Lookup Table  
English Language Arts Grade 4

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.153	1.72	10.0	440	1	440	1
1	-3.143	1.74	10.0	440	1	440	1
2	-3.133	1.76	10.0	440	1	441	1
3	-3.123	1.79	10.0	441	1	441	1
4	-3.113	1.81	10.0	441	1	442	1
5	-3.103	1.83	10.0	441	1	442	1
6	-3.093	1.85	10.0	441	1	443	1
7	-3.083	1.87	10.0	441	1	443	1
8	-2.722	2.82	10.0	448	1	444	1
9	-2.454	3.80	9.7	453	1	449	1
10	-2.237	4.77	8.6	457	1	453	1
11	-2.052	5.67	7.9	461	1	457	1
12	-1.888	6.44	7.4	464	1	461	1
13	-1.738	7.09	7.1	467	1	464	1
14	-1.599	7.62	6.8	469	1	467	1
15	-1.467	8.05	6.6	472	2	469	1
16	-1.341	8.40	6.5	474	2	472	2
17	-1.220	8.70	6.4	476	2	475	2
18	-1.103	8.96	6.3	479	2	477	2
19	-0.989	9.18	6.2	481	2	480	2
20	-0.877	9.38	6.2	483	2	482	2
21	-0.766	9.54	6.1	485	2	484	2
22	-0.657	9.67	6.1	487	2	486	2
23	-0.549	9.77	6.0	489	2	488	2
24	-0.440	9.82	6.0	491	2	490	2
25	-0.332	9.82	6.0	493	2	492	2
26	-0.223	9.78	6.0	495	2	494	2
27	-0.112	9.67	6.1	497	2	497	2
28	0.001	9.51	6.1	499	2	499	2
29	0.117	9.28	6.2	502	3	501	3
30	0.236	8.99	6.3	504	3	503	3
31	0.360	8.63	6.4	506	3	506	3
32	0.490	8.21	6.6	509	3	509	3
33	0.626	7.73	6.8	511	3	511	3
34	0.772	7.20	7.0	514	3	515	3
35	0.930	6.64	7.3	517	3	518	3
36	1.101	6.04	7.7	520	3	522	3
37	1.289	5.43	8.1	524	3	526	3
38	1.501	4.81	8.6	528	3	531	4
39	1.741	4.17	9.2	532	4	537	4
40	2.024	3.49	10.0	538	4	544	4
41	2.370	2.75	10.0	544	4	552	4
42	2.831	1.93	10.0	553	4	560	4
43	3.215	1.40	10.0	560	4	560	4
44	3.215	1.40	10.0	560	4	560	4

Table 2.2.3  
Raw Score to Scale Score Lookup Table  
English Language Arts Grade 5

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.360	2.08	10.0	440	1	440	1
1	-3.339	2.14	10.0	440	1	441	1
2	-3.318	2.20	10.0	441	1	441	1
3	-3.296	2.27	10.0	441	1	442	1
4	-3.275	2.34	10.0	441	1	442	1
5	-3.254	2.40	10.0	442	1	443	1
6	-3.233	2.48	10.0	442	1	443	1
7	-2.920	3.78	9.1	448	1	444	1
8	-2.686	5.12	7.8	452	1	449	1
9	-2.498	6.42	7.0	455	1	453	1
10	-2.337	7.63	6.4	458	1	456	1
11	-2.195	8.70	6.0	461	1	459	1
12	-2.065	9.60	5.7	463	1	461	1
13	-1.946	10.33	5.4	465	1	464	1
14	-1.833	10.90	5.3	467	1	466	1
15	-1.725	11.33	5.2	469	1	468	1
16	-1.621	11.66	5.2	471	2	469	1
17	-1.521	11.90	5.1	472	2	472	2
18	-1.422	12.08	5.1	474	2	474	2
19	-1.326	12.22	5.1	476	2	476	2
20	-1.231	12.32	5.0	478	2	477	2
21	-1.138	12.40	5.0	479	2	479	2
22	-1.045	12.45	5.0	481	2	481	2
23	-0.952	12.48	5.0	482	2	483	2
24	-0.860	12.49	5.0	484	2	485	2
25	-0.768	12.47	5.0	486	2	486	2
26	-0.676	12.41	5.0	487	2	488	2
27	-0.582	12.32	5.0	489	2	490	2
28	-0.488	12.19	5.1	491	2	492	2
29	-0.392	12.01	5.1	492	2	494	2
30	-0.295	11.79	5.1	494	2	496	2
31	-0.195	11.53	5.2	496	2	498	2
32	-0.092	11.22	5.3	498	2	500	3
33	0.014	10.87	5.4	499	2	503	3
34	0.124	10.48	5.5	502	3	505	3
35	0.239	10.05	5.6	504	3	508	3
36	0.359	9.59	5.7	506	3	511	3
37	0.487	9.11	5.9	508	3	513	3
38	0.623	8.59	6.0	510	3	517	3
39	0.768	8.03	6.2	513	3	520	3
40	0.927	7.42	6.5	516	3	523	3
41	1.100	6.75	6.8	519	3	528	3
42	1.295	6.01	7.2	522	3	532	4
43	1.517	5.22	7.7	526	3	537	4
44	1.778	4.41	8.4	531	4	543	4
45	2.097	3.63	9.3	536	4	551	4
46	2.506	2.79	10.0	544	4	560	4
47	3.132	1.58	10.0	555	4	560	4
48	3.430	1.10	10.0	560	4	560	4

Table 2.2.4  
Raw Score to Scale Score Lookup Table  
English Language Arts Grade 6

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.171	2.83	10.0	440	1	440	1
1	-3.149	2.88	10.0	440	1	440	1
2	-3.126	2.93	10.0	441	1	440	1
3	-3.103	2.98	10.0	441	1	440	1
4	-3.081	3.03	10.0	442	1	440	1
5	-3.058	3.08	10.0	442	1	441	1
6	-3.036	3.13	10.0	443	1	441	1
7	-3.013	3.18	10.0	443	1	441	1
8	-2.991	3.23	10.0	443	1	441	1
9	-2.968	3.28	10.0	444	1	441	1
10	-2.945	3.34	10.0	444	1	441	1
11	-2.734	3.85	9.7	448	1	446	1
12	-2.544	4.32	9.1	452	1	450	1
13	-2.370	4.76	8.7	455	1	453	1
14	-2.208	5.15	8.4	458	1	456	1
15	-2.056	5.51	8.1	461	1	460	1
16	-1.911	5.82	7.9	464	1	463	1
17	-1.772	6.10	7.7	467	1	465	1
18	-1.638	6.33	7.5	469	1	468	1
19	-1.508	6.52	7.4	472	2	471	2
20	-1.380	6.67	7.3	474	2	473	2
21	-1.255	6.79	7.3	476	2	476	2
22	-1.131	6.86	7.2	479	2	479	2
23	-1.008	6.90	7.2	481	2	481	2
24	-0.886	6.90	7.2	483	2	484	2
25	-0.765	6.88	7.2	486	2	486	2
26	-0.643	6.83	7.3	488	2	488	2
27	-0.521	6.76	7.3	490	2	491	2
28	-0.398	6.67	7.4	493	2	493	2
29	-0.273	6.57	7.4	495	2	496	2
30	-0.148	6.45	7.5	497	2	499	2
31	-0.021	6.32	7.6	499	2	501	3
32	0.109	6.17	7.6	502	3	504	3
33	0.241	6.02	7.7	505	3	507	3
34	0.377	5.84	7.9	507	3	510	3
35	0.517	5.65	8.0	510	3	513	3
36	0.662	5.44	8.1	513	3	516	3
37	0.813	5.21	8.3	516	3	519	3
38	0.971	4.96	8.5	519	3	522	3
39	1.138	4.68	8.8	522	3	526	3
40	1.317	4.37	9.1	525	3	529	3
41	1.509	4.04	9.4	529	3	534	4
42	1.719	3.69	9.9	533	4	539	4
43	1.952	3.32	10.0	537	4	544	4
44	2.214	2.93	10.0	542	4	550	4
45	2.518	2.52	10.0	548	4	556	4
46	2.879	2.10	10.0	555	4	560	4
47	3.150	1.83	10.0	560	4	560	4
48	3.150	1.83	10.0	560	4	560	4
49	3.150	1.83	10.0	560	4	560	4
50	3.150	1.83	10.0	560	4	560	4

Table 2.2.5  
Raw Score to Scale Score Lookup Table  
English Language Arts Grade 7

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.131	3.92	9.7	440	1	440	1
1	-3.121	3.94	9.6	440	1	440	1
2	-3.111	3.97	9.6	440	1	441	1
3	-3.102	3.99	9.6	441	1	441	1
4	-3.092	4.02	9.5	441	1	441	1
5	-3.082	4.05	9.5	441	1	442	1
6	-3.073	4.07	9.5	441	1	442	1
7	-3.063	4.10	9.4	441	1	442	1
8	-3.053	4.13	9.4	441	1	443	1
9	-3.043	4.15	9.4	442	1	448	1
10	-2.831	4.70	8.8	446	1	452	1
11	-2.643	5.16	8.4	449	1	456	1
12	-2.471	5.54	8.1	453	1	459	1
13	-2.312	5.87	7.9	456	1	462	1
14	-2.162	6.16	7.7	458	1	465	1
15	-2.021	6.43	7.5	461	1	468	1
16	-1.887	6.69	7.4	464	1	470	2
17	-1.758	6.95	7.2	466	1	473	2
18	-1.634	7.23	7.1	469	1	475	2
19	-1.514	7.51	7.0	471	2	477	2
20	-1.398	7.79	6.8	473	2	479	2
21	-1.284	8.07	6.7	475	2	481	2
22	-1.173	8.34	6.6	477	2	483	2
23	-1.065	8.60	6.5	479	2	485	2
24	-0.957	8.84	6.4	482	2	487	2
25	-0.851	9.05	6.3	484	2	489	2
26	-0.746	9.22	6.3	486	2	491	2
27	-0.641	9.35	6.2	488	2	493	2
28	-0.536	9.43	6.2	490	2	495	2
29	-0.431	9.46	6.2	492	2	496	2
30	-0.324	9.45	6.2	494	2	498	2
31	-0.216	9.38	6.2	496	2	500	3
32	-0.107	9.25	6.3	498	2	502	3
33	0.006	9.08	6.3	499	2	504	3
34	0.121	8.84	6.4	502	3	507	3
35	0.240	8.56	6.5	504	3	509	3
36	0.364	8.22	6.7	507	3	511	3
37	0.493	7.82	6.8	509	3	513	3
38	0.630	7.37	7.0	512	3	516	3
39	0.775	6.88	7.3	515	3	518	3
40	0.931	6.35	7.6	518	3	521	3
41	1.100	5.79	7.9	521	3	524	3
42	1.286	5.20	8.4	524	3	528	3
43	1.493	4.60	8.9	528	3	531	4
44	1.727	4.02	9.5	533	4	536	4
45	1.997	3.47	10.0	538	4	541	4
46	2.314	2.95	10.0	544	4	546	4
47	2.701	2.41	10.0	551	4	554	4
48	3.153	1.76	10.0	560	4	560	4
49	3.153	1.76	10.0	560	4	560	4
50	3.153	1.76	10.0	560	4	560	4

Table 2.2.6  
Raw Score to Scale Score Lookup Table  
English Language Arts Grade 8

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.964	4.12	9.8	440	1	440	1
1	-2.957	4.15	9.8	440	1	440	1
2	-2.949	4.17	9.7	440	1	440	1
3	-2.942	4.20	9.7	440	1	440	1
4	-2.935	4.22	9.7	441	1	440	1
5	-2.927	4.25	9.7	441	1	440	1
6	-2.920	4.27	9.6	441	1	440	1
7	-2.913	4.30	9.6	441	1	440	1
8	-2.905	4.33	9.6	441	1	440	1
9	-2.898	4.35	9.5	441	1	440	1
10	-2.704	5.03	8.9	445	1	440	1
11	-2.531	5.62	8.4	449	1	440	1
12	-2.374	6.12	8.0	452	1	441	1
13	-2.228	6.54	7.8	455	1	445	1
14	-2.091	6.89	7.6	457	1	448	1
15	-1.961	7.18	7.4	460	1	452	1
16	-1.836	7.42	7.3	462	1	455	1
17	-1.714	7.63	7.2	465	1	458	1
18	-1.597	7.80	7.1	467	1	461	1
19	-1.481	7.95	7.1	469	1	464	1
20	-1.368	8.07	7.0	472	2	467	1
21	-1.257	8.17	7.0	474	2	469	1
22	-1.146	8.25	6.9	476	2	472	2
23	-1.037	8.30	6.9	478	2	474	2
24	-0.927	8.33	6.9	481	2	477	2
25	-0.818	8.34	6.9	483	2	479	2
26	-0.709	8.33	6.9	485	2	482	2
27	-0.598	8.29	6.9	487	2	484	2
28	-0.487	8.23	6.9	489	2	486	2
29	-0.375	8.16	7.0	492	2	489	2
30	-0.261	8.07	7.0	494	2	491	2
31	-0.145	7.97	7.1	496	2	494	2
32	-0.027	7.85	7.1	498	2	496	2
33	0.093	7.72	7.2	501	3	499	2
34	0.217	7.56	7.2	503	3	502	3
35	0.345	7.39	7.3	506	3	504	3
36	0.477	7.19	7.4	508	3	507	3
37	0.614	6.95	7.5	511	3	510	3
38	0.758	6.67	7.7	514	3	513	3
39	0.909	6.35	7.9	517	3	517	3
40	1.071	5.98	8.1	520	3	520	3
41	1.244	5.56	8.4	524	3	524	3
42	1.432	5.09	8.8	527	3	529	3
43	1.640	4.58	9.3	532	4	534	4
44	1.873	4.03	9.9	536	4	539	4
45	2.139	3.50	10.0	542	4	546	4
46	2.450	3.00	10.0	548	4	553	4
47	2.822	2.51	10.0	555	4	560	4
48	3.066	2.20	10.0	560	4	560	4
49	3.066	2.20	10.0	560	4	560	4
50	3.066	2.20	10.0	560	4	560	4



Table 2.2.7  
Raw Score to Scale Score Lookup Table  
Mathematics Grade 3

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.782	3.09	10.0	440	1	440	1
1	-2.756	3.18	10.0	441	1	440	1
2	-2.731	3.27	10.0	441	1	440	1
3	-2.705	3.37	10.0	442	1	440	1
4	-2.679	3.46	10.0	442	1	440	1
5	-2.654	3.56	10.0	443	1	440	1
6	-2.628	3.66	10.0	443	1	447	1
7	-2.346	4.94	9.6	449	1	452	1
8	-2.124	6.19	8.6	454	1	456	1
9	-1.939	7.37	7.9	458	1	459	1
10	-1.779	8.48	7.3	461	1	462	1
11	-1.637	9.51	6.9	464	1	464	1
12	-1.508	10.46	6.6	467	1	467	1
13	-1.389	11.31	6.4	469	1	469	1
14	-1.278	12.07	6.1	472	2	471	2
15	-1.173	12.74	6.0	474	2	473	2
16	-1.073	13.33	5.9	477	2	475	2
17	-0.976	13.83	5.7	479	2	477	2
18	-0.883	14.26	5.7	481	2	478	2
19	-0.793	14.62	5.6	482	2	480	2
20	-0.705	14.91	5.5	484	2	482	2
21	-0.618	15.14	5.5	486	2	483	2
22	-0.532	15.31	5.5	488	2	485	2
23	-0.447	15.43	5.4	490	2	487	2
24	-0.363	15.50	5.4	492	2	488	2
25	-0.279	15.52	5.4	493	2	490	2
26	-0.196	15.51	5.4	495	2	491	2
27	-0.112	15.47	5.4	497	2	493	2
28	-0.027	15.40	5.4	499	2	495	2
29	0.058	15.31	5.5	501	3	496	2
30	0.144	15.20	5.5	502	3	498	2
31	0.231	15.07	5.5	504	3	499	2
32	0.319	14.91	5.5	506	3	501	3
33	0.410	14.70	5.6	508	3	503	3
34	0.502	14.44	5.6	510	3	505	3
35	0.597	14.09	5.7	512	3	507	3
36	0.696	13.65	5.8	514	3	509	3
37	0.800	13.09	5.9	516	3	511	3
38	0.909	12.42	6.1	519	3	513	3
39	1.025	11.63	6.3	521	3	516	3
40	1.150	10.72	6.5	524	3	518	3
41	1.287	9.70	6.9	527	3	521	3
42	1.440	8.58	7.3	530	4	524	3
43	1.615	7.36	7.9	534	4	528	3
44	1.822	6.04	8.7	538	4	532	4
45	2.080	4.62	9.9	544	4	537	4
46	2.435	3.08	10.0	551	4	544	4
47	2.837	1.87	10.0	560	4	555	4
48	2.837	1.87	10.0	560	4	560	4

Table 2.2.8 Raw Score to Scale Score Lookup Table Mathematics Grade 4

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.812	3.99	10.0	440	1	440	1
1	-2.789	4.07	10.0	440	1	441	1
2	-2.766	4.15	10.0	441	1	441	1
3	-2.743	4.23	10.0	441	1	442	1
4	-2.721	4.31	10.0	442	1	442	1
5	-2.698	4.40	10.0	442	1	443	1
6	-2.675	4.48	9.9	443	1	444	1
7	-2.653	4.57	9.8	443	1	444	1
8	-2.434	5.45	9.0	448	1	449	1
9	-2.247	6.31	8.3	452	1	454	1
10	-2.083	7.16	7.8	455	1	457	1
11	-1.935	7.99	7.4	458	1	460	1
12	-1.800	8.80	7.1	461	1	463	1
13	-1.675	9.58	6.8	464	1	465	1
14	-1.558	10.33	6.5	466	1	468	1
15	-1.448	11.06	6.3	469	1	469	1
16	-1.344	11.76	6.1	471	2	472	2
17	-1.245	12.44	5.9	473	2	474	2
18	-1.150	13.10	5.8	475	2	476	2
19	-1.059	13.73	5.7	477	2	477	2
20	-0.971	14.34	5.5	479	2	479	2
21	-0.885	14.92	5.4	480	2	481	2
22	-0.803	15.48	5.3	482	2	482	2
23	-0.722	16.00	5.2	484	2	484	2
24	-0.643	16.48	5.2	485	2	485	2
25	-0.566	16.92	5.1	487	2	487	2
26	-0.490	17.30	5.0	489	2	488	2
27	-0.414	17.63	5.0	490	2	490	2
28	-0.340	17.90	4.9	492	2	491	2
29	-0.266	18.09	4.9	493	2	493	2
30	-0.193	18.22	4.9	495	2	494	2
31	-0.119	18.27	4.9	496	2	496	2
32	-0.045	18.25	4.9	498	2	497	2
33	0.029	18.15	4.9	499	2	499	2
34	0.104	17.98	4.9	501	3	500	3
35	0.181	17.74	5.0	503	3	502	3
36	0.258	17.43	5.0	504	3	504	3
37	0.338	17.07	5.1	506	3	505	3
38	0.419	16.64	5.1	508	3	507	3
39	0.503	16.17	5.2	509	3	509	3
40	0.589	15.66	5.3	511	3	511	3
41	0.679	15.12	5.4	513	3	513	3
42	0.773	14.55	5.5	515	3	515	3
43	0.871	13.97	5.6	517	3	517	3
44	0.975	13.35	5.7	519	3	519	3
45	1.085	12.69	5.9	522	3	522	3
46	1.203	11.95	6.1	524	3	525	3
47	1.331	11.10	6.3	527	3	528	3
48	1.472	10.09	6.6	529	3	531	4
49	1.632	8.87	7.0	533	4	535	4
50	1.819	7.42	7.7	537	4	539	4
51	2.051	5.77	8.7	542	4	544	4
52	2.364	3.93	10.0	548	4	551	4
53	2.885	1.94	10.0	559	4	560	4
54	2.920	1.85	10.0	560	4	560	4

Table 2.2.9 Raw Score to Scale Score Lookup Table Mathematics Grade 5

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.126	2.33	10.0	440	1	440	1
1	-3.103	2.40	10.0	440	1	441	1
2	-3.080	2.47	10.0	441	1	442	1
3	-3.057	2.55	10.0	441	1	443	1
4	-3.033	2.62	10.0	442	1	444	1
5	-3.010	2.70	10.0	442	1	445	1
6	-2.623	4.25	9.2	450	1	446	1
7	-2.353	5.70	8.0	455	1	452	1
8	-2.141	7.06	7.2	459	1	456	1
9	-1.966	8.34	6.6	462	1	460	1
10	-1.814	9.56	6.2	465	1	463	1
11	-1.680	10.74	5.8	468	1	466	1
12	-1.558	11.86	5.5	469	1	468	1
13	-1.447	12.94	5.3	472	2	471	2
14	-1.344	13.95	5.1	474	2	473	2
15	-1.247	14.90	4.9	476	2	475	2
16	-1.155	15.77	4.8	478	2	477	2
17	-1.068	16.56	4.7	479	2	479	2
18	-0.984	17.26	4.6	481	2	480	2
19	-0.902	17.87	4.5	482	2	482	2
20	-0.823	18.40	4.4	484	2	484	2
21	-0.746	18.82	4.4	485	2	485	2
22	-0.670	19.17	4.3	487	2	487	2
23	-0.596	19.42	4.3	488	2	488	2
24	-0.522	19.60	4.3	490	2	490	2
25	-0.448	19.70	4.3	491	2	491	2
26	-0.375	19.74	4.3	492	2	492	2
27	-0.303	19.73	4.3	494	2	494	2
28	-0.230	19.66	4.3	495	2	495	2
29	-0.157	19.57	4.3	497	2	497	2
30	-0.084	19.45	4.3	498	2	498	2
31	-0.011	19.31	4.3	499	2	499	2
32	0.063	19.17	4.3	501	3	501	3
33	0.138	19.04	4.4	502	3	502	3
34	0.213	18.91	4.4	504	3	504	3
35	0.289	18.81	4.4	505	3	505	3
36	0.366	18.71	4.4	506	3	507	3
37	0.444	18.63	4.4	508	3	508	3
38	0.523	18.53	4.4	509	3	510	3
39	0.603	18.41	4.4	511	3	511	3
40	0.686	18.21	4.5	513	3	513	3
41	0.770	17.92	4.5	514	3	514	3
42	0.858	17.49	4.6	516	3	516	3
43	0.949	16.88	4.6	518	3	518	3
44	1.045	16.07	4.8	519	3	520	3
45	1.148	15.03	4.9	521	3	522	3
46	1.259	13.78	5.1	523	3	524	3
47	1.380	12.34	5.4	526	3	526	3
48	1.517	10.73	5.8	528	3	529	3
49	1.675	9.02	6.3	531	4	532	4
50	1.864	7.23	7.1	535	4	536	4
51	2.102	5.40	8.2	540	4	541	4
52	2.431	3.53	10.0	546	4	547	4
53	2.997	1.62	10.0	557	4	559	4
54	3.176	1.26	10.0	560	4	560	4

Table 2.2.10 Raw Score to Scale Score Lookup Table Mathematics Grade 6

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.028	1.98	10.0	440	1	440	1
1	-2.993	2.07	10.0	441	1	441	1
2	-2.958	2.17	10.0	441	1	441	1
3	-2.923	2.27	10.0	442	1	442	1
4	-2.889	2.38	10.0	443	1	442	1
5	-2.854	2.49	10.0	443	1	443	1
6	-2.480	3.96	10.0	451	1	450	1
7	-2.215	5.37	8.6	456	1	456	1
8	-2.007	6.69	7.7	460	1	460	1
9	-1.833	7.91	7.1	464	1	464	1
10	-1.683	9.04	6.6	467	1	467	1
11	-1.548	10.08	6.3	469	1	469	1
12	-1.426	11.03	6.0	472	2	472	2
13	-1.313	11.90	5.8	474	2	474	2
14	-1.208	12.71	5.6	476	2	476	2
15	-1.109	13.44	5.4	478	2	478	2
16	-1.015	14.10	5.3	480	2	480	2
17	-0.925	14.71	5.2	482	2	482	2
18	-0.838	15.26	5.1	484	2	483	2
19	-0.754	15.76	5.0	485	2	485	2
20	-0.673	16.22	4.9	487	2	487	2
21	-0.594	16.63	4.9	488	2	488	2
22	-0.517	16.99	4.8	490	2	489	2
23	-0.441	17.32	4.8	491	2	491	2
24	-0.367	17.60	4.7	493	2	492	2
25	-0.294	17.85	4.7	494	2	494	2
26	-0.222	18.05	4.7	496	2	495	2
27	-0.150	18.21	4.7	497	2	496	2
28	-0.080	18.33	4.6	499	2	498	2
29	-0.009	18.41	4.6	499	2	499	2
30	0.060	18.45	4.6	501	3	500	3
31	0.130	18.46	4.6	503	3	501	3
32	0.200	18.42	4.6	504	3	503	3
33	0.270	18.34	4.6	506	3	504	3
34	0.340	18.22	4.7	507	3	505	3
35	0.411	18.06	4.7	508	3	506	3
36	0.482	17.85	4.7	510	3	508	3
37	0.555	17.60	4.7	511	3	509	3
38	0.628	17.29	4.8	513	3	510	3
39	0.704	16.93	4.8	514	3	512	3
40	0.781	16.50	4.9	516	3	513	3
41	0.860	16.02	5.0	517	3	515	3
42	0.943	15.47	5.1	519	3	516	3
43	1.029	14.85	5.2	521	3	518	3
44	1.120	14.15	5.3	522	3	519	3
45	1.216	13.35	5.4	524	3	521	3
46	1.320	12.45	5.6	526	3	523	3
47	1.433	11.43	5.9	529	3	525	3
48	1.559	10.25	6.2	531	4	528	3
49	1.703	8.89	6.7	534	4	530	4
50	1.876	7.32	7.3	537	4	534	4
51	2.095	5.53	8.4	542	4	538	4
52	2.406	3.56	10.0	548	4	543	4
53	2.964	1.54	10.0	559	4	554	4
54	3.011	1.44	10.0	560	4	560	4

Table 2.2.11 Raw Score to Scale Score Lookup Table Mathematics Grade 7

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.859	2.18	10.0	440	1	440	1
1	-2.829	2.27	10.0	441	1	440	1
2	-2.799	2.36	10.0	441	1	440	1
3	-2.768	2.46	10.0	442	1	440	1
4	-2.738	2.56	10.0	443	1	441	1
5	-2.708	2.67	10.0	443	1	441	1
6	-2.332	4.29	10.0	451	1	451	1
7	-2.064	5.77	8.6	457	1	458	1
8	-1.852	7.07	7.8	461	1	462	1
9	-1.675	8.24	7.2	465	1	466	1
10	-1.521	9.31	6.8	468	1	469	1
11	-1.384	10.29	6.5	471	2	473	2
12	-1.260	11.21	6.2	473	2	475	2
13	-1.146	12.09	6.0	476	2	478	2
14	-1.040	12.92	5.8	478	2	480	2
15	-0.941	13.74	5.6	480	2	482	2
16	-0.847	14.53	5.4	482	2	484	2
17	-0.759	15.31	5.3	484	2	486	2
18	-0.674	16.05	5.2	485	2	487	2
19	-0.593	16.77	5.1	487	2	489	2
20	-0.515	17.45	5.0	489	2	490	2
21	-0.439	18.09	4.9	490	2	492	2
22	-0.365	18.67	4.8	492	2	494	2
23	-0.294	19.19	4.7	493	2	495	2
24	-0.223	19.65	4.7	495	2	496	2
25	-0.154	20.04	4.6	496	2	498	2
26	-0.086	20.37	4.6	498	2	499	2
27	-0.019	20.62	4.6	499	2	500	3
28	0.048	20.80	4.6	500	3	502	3
29	0.114	20.91	4.5	502	3	503	3
30	0.181	20.94	4.5	503	3	504	3
31	0.247	20.91	4.5	504	3	506	3
32	0.314	20.81	4.6	506	3	507	3
33	0.382	20.65	4.6	507	3	508	3
34	0.450	20.43	4.6	509	3	510	3
35	0.519	20.16	4.6	510	3	511	3
36	0.589	19.83	4.7	512	3	512	3
37	0.661	19.45	4.7	513	3	514	3
38	0.734	19.02	4.8	515	3	515	3
39	0.810	18.54	4.8	516	3	516	3
40	0.887	17.99	4.9	518	3	518	3
41	0.968	17.38	5.0	519	3	520	3
42	1.053	16.68	5.1	521	3	521	3
43	1.142	15.89	5.2	523	3	523	3
44	1.236	14.97	5.4	525	3	525	3
45	1.337	13.92	5.6	527	3	527	3
46	1.447	12.72	5.8	529	3	529	3
47	1.569	11.37	6.2	532	4	531	4
48	1.708	9.87	6.6	535	4	534	4
49	1.870	8.24	7.2	538	4	537	4
50	2.068	6.49	8.1	542	4	540	4
51	2.328	4.64	9.6	548	4	545	4
52	2.710	2.74	10.0	556	4	551	4
53	2.922	2.04	10.0	560	4	560	4
54	2.922	2.04	10.0	560	4	560	4

Table 2.2.12 Raw Score to Scale Score Lookup Table Mathematics Grade 8

Raw Score	Theta	Information	SE (Scale Score)	2024		2023	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.983	2.32	10.0	440	1	440	1
1	-2.949	2.41	10.0	441	1	441	1
2	-2.914	2.50	10.0	441	1	442	1
3	-2.880	2.60	10.0	442	1	443	1
4	-2.846	2.71	10.0	443	1	444	1
5	-2.812	2.81	10.0	443	1	445	1
6	-2.777	2.92	10.0	444	1	446	1
7	-2.743	3.03	10.0	445	1	447	1
8	-2.415	4.32	9.7	451	1	453	1
9	-2.169	5.56	8.6	456	1	458	1
10	-1.969	6.72	7.8	460	1	461	1
11	-1.799	7.78	7.2	464	1	465	1
12	-1.650	8.74	6.8	467	1	467	1
13	-1.515	9.61	6.5	469	1	469	1
14	-1.392	10.43	6.2	472	2	472	2
15	-1.279	11.23	6.0	474	2	474	2
16	-1.172	12.05	5.8	477	2	476	2
17	-1.073	12.88	5.6	479	2	478	2
18	-0.979	13.72	5.4	480	2	480	2
19	-0.889	14.58	5.3	482	2	482	2
20	-0.804	15.43	5.1	484	2	483	2
21	-0.722	16.25	5.0	486	2	485	2
22	-0.643	17.04	4.9	487	2	487	2
23	-0.567	17.77	4.8	489	2	488	2
24	-0.493	18.45	4.7	490	2	489	2
25	-0.421	19.06	4.6	492	2	491	2
26	-0.351	19.61	4.6	493	2	492	2
27	-0.282	20.10	4.5	494	2	494	2
28	-0.213	20.52	4.5	496	2	495	2
29	-0.146	20.88	4.4	497	2	496	2
30	-0.079	21.18	4.4	499	2	498	2
31	-0.013	21.41	4.4	499	2	499	2
32	0.054	21.57	4.3	501	3	501	3
33	0.121	21.65	4.3	503	3	502	3
34	0.188	21.65	4.3	504	3	503	3
35	0.255	21.56	4.3	505	3	505	3
36	0.324	21.37	4.4	507	3	506	3
37	0.393	21.09	4.4	508	3	508	3
38	0.464	20.71	4.4	510	3	509	3
39	0.537	20.23	4.5	511	3	511	3
40	0.612	19.64	4.6	513	3	512	3
41	0.690	18.96	4.6	514	3	514	3
42	0.772	18.17	4.7	516	3	516	3
43	0.857	17.29	4.9	517	3	517	3
44	0.947	16.31	5.0	519	3	519	3
45	1.044	15.22	5.2	521	3	521	3
46	1.148	14.04	5.4	523	3	524	3
47	1.261	12.74	5.7	526	3	526	3
48	1.388	11.34	6.0	528	3	529	3
49	1.532	9.81	6.4	531	4	532	4
50	1.702	8.15	7.1	535	4	535	4
51	1.913	6.33	8.0	539	4	540	4
52	2.202	4.33	9.7	545	4	546	4
53	2.686	2.12	10.0	554	4	558	4
54	2.966	1.37	10.0	560	4	560	4

# Section 2.3

## Cumulative Scale Score Distribution Tables

Table 2.3.1  
Cumulative Scale Score Distribution  
English Language Arts Grade 3

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	1	0.00002	0.00002
441	NM	13	0.00022	0.00023
442	NM	110	0.00183	0.00207
443	NM	536	0.00894	0.01101
444	NM	573	0.00956	0.02056
450	NM	884	0.01474	0.03530
455	NM	1055	0.01759	0.05290
460	NM	1315	0.02193	0.07483
463	NM	1415	0.02360	0.09842
466	NM	1471	0.02453	0.12296
469	NM	1633	0.02723	0.15019
472	PM	1557	0.02597	0.17615
474	PM	1549	0.02583	0.20198
477	PM	1584	0.02642	0.22840
479	PM	1714	0.02858	0.25698
481	PM	1690	0.02818	0.28517
484	PM	1759	0.02933	0.31450
486	PM	1773	0.02957	0.34407
488	PM	1953	0.03257	0.37664
490	PM	2024	0.03375	0.41039
492	PM	2058	0.03432	0.44471
494	PM	2132	0.03555	0.48026
497	PM	2237	0.03731	0.51757
499	PM	2327	0.03881	0.55637
501	ME	2359	0.03934	0.59571
503	ME	2399	0.04001	0.63572
506	ME	2365	0.03944	0.67516
508	ME	2518	0.04199	0.71715
511	ME	2495	0.04161	0.75876
514	ME	2403	0.04007	0.79883
517	ME	2308	0.03849	0.83732
520	ME	2054	0.03425	0.87158
523	ME	1950	0.03252	0.90409
527	ME	1715	0.02860	0.93269
531	EE	1395	0.02326	0.95596
536	EE	1078	0.01798	0.97393
541	EE	774	0.01291	0.98684
547	EE	407	0.00679	0.99363
555	EE	238	0.00397	0.99760
560	EE	144	0.00240	1.00000



Table 2.3.2  
Cumulative Scale Score Distribution  
English Language Arts Grade 4

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	17	0.00028	0.00028
441	NM	1328	0.02173	0.02201
448	NM	793	0.01298	0.03499
453	NM	969	0.01586	0.05085
457	NM	1174	0.01921	0.07006
461	NM	1243	0.02034	0.09040
464	NM	1372	0.02245	0.11286
467	NM	1337	0.02188	0.13474
469	NM	1405	0.02299	0.15773
472	PM	1453	0.02378	0.18151
474	PM	1529	0.02502	0.20654
476	PM	1598	0.02615	0.23269
479	PM	1654	0.02707	0.25976
481	PM	1774	0.02903	0.28879
483	PM	1799	0.02944	0.31823
485	PM	1924	0.03149	0.34972
487	PM	2018	0.03303	0.38275
489	PM	2163	0.03540	0.41815
491	PM	2255	0.03690	0.45505
493	PM	2283	0.03736	0.49241
495	PM	2393	0.03916	0.53158
497	PM	2491	0.04077	0.57235
499	PM	2492	0.04078	0.61313
502	ME	2406	0.03938	0.65250
504	ME	2513	0.04113	0.69363
506	ME	2514	0.04114	0.73478
509	ME	2365	0.03871	0.77348
511	ME	2278	0.03728	0.81076
514	ME	2140	0.03502	0.84578
517	ME	1983	0.03245	0.87824
520	ME	1856	0.03037	0.90861
524	ME	1493	0.02443	0.93305
528	ME	1302	0.02131	0.95436
532	EE	1019	0.01668	0.97103
538	EE	806	0.01319	0.98422
544	EE	516	0.00844	0.99267
553	EE	296	0.00484	0.99751
560	EE	152	0.00249	1.00000

Table 2.3.3  
Cumulative Scale Score Distribution  
English Language Arts Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	4	0.00006	0.00006
441	NM	182	0.00291	0.00297
442	NM	643	0.01027	0.01324
448	NM	548	0.00875	0.02200
452	NM	722	0.01153	0.03353
455	NM	765	0.01222	0.04575
458	NM	755	0.01206	0.05781
461	NM	861	0.01375	0.07156
463	NM	902	0.01441	0.08597
465	NM	907	0.01449	0.10046
467	NM	1024	0.01636	0.11682
469	NM	1084	0.01732	0.13414
471	PM	1108	0.01770	0.15183
472	PM	1180	0.01885	0.17068
474	PM	1220	0.01949	0.19017
476	PM	1264	0.02019	0.21036
478	PM	1285	0.02053	0.23089
479	PM	1324	0.02115	0.25204
481	PM	1443	0.02305	0.27509
482	PM	1462	0.02335	0.29845
484	PM	1587	0.02535	0.32380
486	PM	1662	0.02655	0.35035
487	PM	1644	0.02626	0.37661
489	PM	1822	0.02910	0.40571
491	PM	1848	0.02952	0.43523
492	PM	1881	0.03005	0.46528
494	PM	2057	0.03286	0.49814
496	PM	2038	0.03256	0.53069
498	PM	2129	0.03401	0.56470
499	PM	2220	0.03546	0.60017
502	ME	2201	0.03516	0.63533
504	ME	2386	0.03811	0.67344
506	ME	2329	0.03720	0.71064
508	ME	2336	0.03732	0.74796
510	ME	2337	0.03733	0.78529
513	ME	2383	0.03807	0.82336
516	ME	2121	0.03388	0.85724
519	ME	1951	0.03117	0.88840
522	ME	1726	0.02757	0.91598
526	ME	1544	0.02466	0.94064
531	EE	1352	0.02160	0.96224
536	EE	1082	0.01728	0.97952
544	EE	761	0.01216	0.99168
555	EE	394	0.00629	0.99797
560	EE	127	0.00203	1.00000

Table 2.3.4  
Cumulative Scale Score Distribution  
English Language Arts Grade 6

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	3	0.00005	0.00005
441	NM	51	0.00081	0.00086
442	NM	287	0.00458	0.00544
443	NM	1608	0.02565	0.03109
444	NM	1879	0.02997	0.06106
448	NM	1055	0.01683	0.07789
452	NM	1113	0.01775	0.09565
455	NM	1141	0.01820	0.11385
458	NM	1240	0.01978	0.13363
461	NM	1264	0.02016	0.15379
464	NM	1303	0.02079	0.17458
467	NM	1397	0.02228	0.19686
469	NM	1393	0.02222	0.21908
472	PM	1486	0.02370	0.24279
474	PM	1522	0.02428	0.26707
476	PM	1513	0.02414	0.29120
479	PM	1652	0.02635	0.31756
481	PM	1627	0.02595	0.34351
483	PM	1764	0.02814	0.37165
486	PM	1752	0.02795	0.39960
488	PM	1777	0.02835	0.42794
490	PM	1876	0.02993	0.45787
493	PM	1941	0.03096	0.48883
495	PM	1945	0.03103	0.51986
497	PM	2023	0.03227	0.55213
499	PM	1976	0.03152	0.58365
502	ME	2078	0.03315	0.61680
505	ME	1882	0.03002	0.64682
507	ME	1943	0.03099	0.67782
510	ME	1986	0.03168	0.70950
513	ME	1954	0.03117	0.74067
516	ME	1932	0.03082	0.77149
519	ME	1874	0.02989	0.80138
522	ME	1836	0.02929	0.83067
525	ME	1792	0.02859	0.85926
529	ME	1662	0.02651	0.88577
533	EE	1528	0.02437	0.91014
537	EE	1524	0.02431	0.93445
542	EE	1286	0.02051	0.95497
548	EE	1068	0.01704	0.97200
555	EE	767	0.01224	0.98424
560	EE	988	0.01576	1.00000

Table 2.3.5  
Cumulative Scale Score Distribution  
English Language Arts Grade 7

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	14	0.00022	0.00022
441	NM	1754	0.02757	0.02779
442	NM	785	0.01234	0.04013
446	NM	855	0.01344	0.05356
449	NM	907	0.01426	0.06782
453	NM	977	0.01536	0.08317
456	NM	1038	0.01631	0.09949
458	NM	1089	0.01712	0.11660
461	NM	1139	0.01790	0.13450
464	NM	1161	0.01825	0.15275
466	NM	1296	0.02037	0.17312
469	NM	1345	0.02114	0.19426
471	PM	1427	0.02243	0.21669
473	PM	1517	0.02384	0.24053
475	PM	1541	0.02422	0.26475
477	PM	1514	0.02380	0.28855
479	PM	1695	0.02664	0.31519
482	PM	1760	0.02766	0.34285
484	PM	1736	0.02728	0.37013
486	PM	1841	0.02893	0.39907
488	PM	1870	0.02939	0.42846
490	PM	1911	0.03003	0.45849
492	PM	1978	0.03109	0.48958
494	PM	2047	0.03217	0.52175
496	PM	2061	0.03239	0.55414
498	PM	2161	0.03396	0.58811
499	PM	2176	0.03420	0.62231
502	ME	2191	0.03444	0.65674
504	ME	2185	0.03434	0.69109
507	ME	2089	0.03283	0.72392
509	ME	2168	0.03407	0.75799
512	ME	2041	0.03208	0.79007
515	ME	2054	0.03228	0.82235
518	ME	1945	0.03057	0.85292
521	ME	1825	0.02868	0.88161
524	ME	1642	0.02581	0.90741
528	ME	1597	0.02510	0.93251
533	EE	1343	0.02111	0.95362
538	EE	1123	0.01765	0.97127
544	EE	782	0.01229	0.98356
551	EE	579	0.00910	0.99266
560	EE	467	0.00734	1.00000

Table 2.3.6  
Cumulative Scale Score Distribution  
English Language Arts Grade 8

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	67	0.00104	0.00104
441	NM	2908	0.04534	0.04638
445	NM	904	0.01409	0.06048
449	NM	964	0.01503	0.07551
452	NM	941	0.01467	0.09018
455	NM	1008	0.01572	0.10590
457	NM	950	0.01481	0.12071
460	NM	1027	0.01601	0.13672
462	NM	1100	0.01715	0.15387
465	NM	1194	0.01862	0.17248
467	NM	1217	0.01897	0.19146
469	NM	1274	0.01986	0.21132
472	PM	1315	0.02050	0.23182
474	PM	1330	0.02074	0.25256
476	PM	1449	0.02259	0.27515
478	PM	1506	0.02348	0.29863
481	PM	1577	0.02459	0.32322
483	PM	1636	0.02551	0.34873
485	PM	1694	0.02641	0.37514
487	PM	1805	0.02814	0.40328
489	PM	1822	0.02841	0.43169
492	PM	1856	0.02894	0.46062
494	PM	1964	0.03062	0.49125
496	PM	2027	0.03160	0.52285
498	PM	2063	0.03216	0.55501
501	ME	2121	0.03307	0.58808
503	ME	2185	0.03407	0.62215
506	ME	2246	0.03502	0.65717
508	ME	2213	0.03450	0.69167
511	ME	2203	0.03435	0.72602
514	ME	2172	0.03386	0.75988
517	ME	2137	0.03332	0.79320
520	ME	2011	0.03135	0.82455
524	ME	1989	0.03101	0.85556
527	ME	1847	0.02880	0.88436
532	EE	1728	0.02694	0.91130
536	EE	1557	0.02428	0.93558
542	EE	1340	0.02089	0.95647
548	EE	1052	0.01640	0.97287
555	EE	802	0.01250	0.98538
560	EE	938	0.01462	1.00000

Table 2.3.7  
Cumulative Scale Score Distribution  
Mathematics Grade 3

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
441	NM	60	0.00120	0.00120
442	NM	315	0.00629	0.00749
443	NM	718	0.01435	0.02184
449	NM	494	0.00987	0.03171
454	NM	572	0.01143	0.04314
458	NM	662	0.01323	0.05637
461	NM	698	0.01395	0.07032
464	NM	735	0.01469	0.08500
467	NM	805	0.01609	0.10109
469	NM	813	0.01625	0.11733
472	PM	889	0.01776	0.13510
474	PM	951	0.01900	0.15410
477	PM	972	0.01942	0.17352
479	PM	1030	0.02058	0.19410
481	PM	1060	0.02118	0.21528
482	PM	1024	0.02046	0.23574
484	PM	1121	0.02240	0.25814
486	PM	1164	0.02326	0.28140
488	PM	1150	0.02298	0.30438
490	PM	1223	0.02444	0.32882
492	PM	1321	0.02640	0.35521
493	PM	1294	0.02586	0.38107
495	PM	1360	0.02717	0.40824
497	PM	1376	0.02749	0.43574
499	PM	1520	0.03037	0.46611
501	ME	1516	0.03029	0.49640
502	ME	1587	0.03171	0.52811
504	ME	1596	0.03189	0.56000
506	ME	1584	0.03165	0.59166
508	ME	1621	0.03239	0.62405
510	ME	1638	0.03273	0.65678
512	ME	1612	0.03221	0.68899
514	ME	1635	0.03267	0.72166
516	ME	1638	0.03273	0.75439
519	ME	1634	0.03265	0.78704
521	ME	1606	0.03209	0.81913
524	ME	1502	0.03001	0.84914
527	ME	1481	0.02959	0.87873
530	EE	1323	0.02644	0.90517
534	EE	1265	0.02528	0.93044
538	EE	1114	0.02226	0.95270
544	EE	954	0.01906	0.97177
551	EE	754	0.01507	0.98683
560	EE	659	0.01317	1.00000

Table 2.3.8  
Cumulative Scale Score Distribution  
Mathematics Grade 4

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	10	0.00020	0.00020
441	NM	95	0.00187	0.00206
442	NM	265	0.00521	0.00727
443	NM	413	0.00812	0.01539
448	NM	287	0.00564	0.02103
452	NM	307	0.00603	0.02706
455	NM	338	0.00664	0.03371
458	NM	430	0.00845	0.04216
461	NM	393	0.00772	0.04988
464	NM	503	0.00989	0.05977
466	NM	501	0.00985	0.06961
469	NM	600	0.01179	0.08140
471	PM	633	0.01244	0.09384
473	PM	684	0.01344	0.10729
475	PM	735	0.01445	0.12173
477	PM	800	0.01572	0.13746
479	PM	823	0.01617	0.15363
480	PM	872	0.01714	0.17077
482	PM	960	0.01887	0.18963
484	PM	1010	0.01985	0.20948
485	PM	1066	0.02095	0.23044
487	PM	1101	0.02164	0.25207
489	PM	1104	0.02170	0.27377
490	PM	1187	0.02333	0.29710
492	PM	1231	0.02419	0.32129
493	PM	1218	0.02394	0.34523
495	PM	1297	0.02549	0.37072
496	PM	1273	0.02502	0.39574
498	PM	1328	0.02610	0.42184
499	PM	1338	0.02630	0.44813
501	ME	1392	0.02736	0.47549
503	ME	1358	0.02669	0.50218
504	ME	1492	0.02932	0.53150
506	ME	1466	0.02881	0.56032
508	ME	1505	0.02958	0.58989
509	ME	1541	0.03029	0.62018
511	ME	1516	0.02979	0.64997
513	ME	1592	0.03129	0.68126
515	ME	1588	0.03121	0.71247
517	ME	1597	0.03139	0.74386
519	ME	1626	0.03196	0.77581
522	ME	1619	0.03182	0.80763
524	ME	1618	0.03180	0.83943
527	ME	1586	0.03117	0.87060
529	ME	1591	0.03127	0.90187
533	EE	1409	0.02769	0.92956
537	EE	1239	0.02435	0.95391
542	EE	968	0.01902	0.97294
548	EE	738	0.01450	0.98744
559	EE	470	0.00924	0.99668
560	EE	169	0.00332	1.00000

Table 2.3.9  
Cumulative Scale Score Distribution  
Mathematics Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	1	0.00002	0.00002
441	NM	59	0.00110	0.00112
442	NM	190	0.00356	0.00468
450	NM	191	0.00358	0.00826
455	NM	283	0.00530	0.01356
459	NM	374	0.00700	0.02056
462	NM	475	0.00890	0.02946
465	NM	590	0.01105	0.04051
468	NM	688	0.01288	0.05339
469	NM	792	0.01483	0.06822
472	PM	870	0.01629	0.08451
474	PM	952	0.01783	0.10234
476	PM	1032	0.01933	0.12167
478	PM	1129	0.02114	0.14281
479	PM	1094	0.02049	0.16330
481	PM	1179	0.02208	0.18538
482	PM	1244	0.02330	0.20867
484	PM	1254	0.02348	0.23216
485	PM	1289	0.02414	0.25630
487	PM	1331	0.02493	0.28122
488	PM	1314	0.02461	0.30583
490	PM	1376	0.02577	0.33160
491	PM	1474	0.02760	0.35920
492	PM	1410	0.02640	0.38561
494	PM	1504	0.02817	0.41377
495	PM	1438	0.02693	0.44070
497	PM	1402	0.02626	0.46696
498	PM	1495	0.02800	0.49495
499	PM	1475	0.02762	0.52258
501	ME	1488	0.02787	0.55044
502	ME	1410	0.02640	0.57685
504	ME	1390	0.02603	0.60288
505	ME	1351	0.02530	0.62818
506	ME	1481	0.02773	0.65591
508	ME	1392	0.02607	0.68198
509	ME	1297	0.02429	0.70627
511	ME	1423	0.02665	0.73292
513	ME	1252	0.02345	0.75636
514	ME	1224	0.02292	0.77928
516	ME	1225	0.02294	0.80222
518	ME	1216	0.02277	0.82500
519	ME	1221	0.02287	0.84786
521	ME	1149	0.02152	0.86938
523	ME	1100	0.02060	0.88998
526	ME	1106	0.02071	0.91069
528	ME	1067	0.01998	0.93067
531	EE	937	0.01755	0.94822
535	EE	877	0.01642	0.96464
540	EE	716	0.01341	0.97805
546	EE	577	0.01081	0.98886
557	EE	407	0.00762	0.99648
560	EE	188	0.00352	1.00000



Table 2.3.10  
Cumulative Scale Score Distribution  
Mathematics Grade 6

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	1	0.00002	0.00002
441	NM	44	0.00080	0.00082
442	NM	117	0.00213	0.00295
443	NM	566	0.01030	0.01325
451	NM	490	0.00892	0.02216
456	NM	639	0.01163	0.03379
460	NM	737	0.01341	0.04720
464	NM	878	0.01598	0.06318
467	NM	967	0.01760	0.08078
469	NM	1060	0.01929	0.10007
472	PM	1064	0.01936	0.11943
474	PM	1084	0.01973	0.13915
476	PM	1183	0.02153	0.16068
478	PM	1225	0.02229	0.18297
480	PM	1230	0.02238	0.20535
482	PM	1254	0.02282	0.22817
484	PM	1248	0.02271	0.25088
485	PM	1349	0.02455	0.27543
487	PM	1361	0.02477	0.30020
488	PM	1360	0.02475	0.32494
490	PM	1342	0.02442	0.34936
491	PM	1381	0.02513	0.37450
493	PM	1385	0.02520	0.39970
494	PM	1362	0.02478	0.42448
496	PM	1390	0.02529	0.44978
497	PM	1398	0.02544	0.47522
499	PM	2846	0.05179	0.52700
501	ME	1415	0.02575	0.55275
503	ME	1394	0.02537	0.57812
504	ME	1317	0.02397	0.60209
506	ME	1265	0.02302	0.62510
507	ME	1242	0.02260	0.64771
508	ME	1283	0.02335	0.67105
510	ME	1198	0.02180	0.69285
511	ME	1178	0.02144	0.71429
513	ME	1236	0.02249	0.73678
514	ME	1148	0.02089	0.75767
516	ME	1094	0.01991	0.77758
517	ME	1105	0.02011	0.79769
519	ME	1137	0.02069	0.81838
521	ME	1048	0.01907	0.83745
522	ME	1089	0.01982	0.85726
524	ME	1015	0.01847	0.87573
526	ME	1008	0.01834	0.89408
529	ME	988	0.01798	0.91205
531	EE	943	0.01716	0.92921
534	EE	924	0.01681	0.94603
537	EE	809	0.01472	0.96075
542	EE	761	0.01385	0.97460
548	EE	654	0.01190	0.98650
559	EE	474	0.00863	0.99512
560	EE	268	0.00488	1.00000

Table 2.3.11  
Cumulative Scale Score Distribution  
Mathematics Grade 7

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	4	0.00007	0.00007
441	NM	81	0.00143	0.00151
442	NM	132	0.00234	0.00384
443	NM	754	0.01336	0.01720
451	NM	733	0.01299	0.03019
457	NM	978	0.01733	0.04751
461	NM	1157	0.02050	0.06801
465	NM	1386	0.02455	0.09256
468	NM	1507	0.02670	0.11926
471	PM	1676	0.02969	0.14895
473	PM	1610	0.02852	0.17747
476	PM	1701	0.03013	0.20761
478	PM	1652	0.02927	0.23687
480	PM	1653	0.02928	0.26616
482	PM	1647	0.02918	0.29533
484	PM	1558	0.02760	0.32293
485	PM	1573	0.02787	0.35080
487	PM	1534	0.02718	0.37798
489	PM	1503	0.02663	0.40460
490	PM	1338	0.02370	0.42831
492	PM	1409	0.02496	0.45327
493	PM	1348	0.02388	0.47715
495	PM	1316	0.02331	0.50046
496	PM	1223	0.02167	0.52213
498	PM	1231	0.02181	0.54393
499	PM	1228	0.02175	0.56569
500	ME	1196	0.02119	0.58688
502	ME	1130	0.02002	0.60689
503	ME	1095	0.01940	0.62629
504	ME	1094	0.01938	0.64567
506	ME	1072	0.01899	0.66466
507	ME	1074	0.01903	0.68369
509	ME	1031	0.01826	0.70196
510	ME	994	0.01761	0.71956
512	ME	986	0.01747	0.73703
513	ME	989	0.01752	0.75455
515	ME	959	0.01699	0.77154
516	ME	1035	0.01834	0.78988
518	ME	969	0.01717	0.80704
519	ME	1020	0.01807	0.82511
521	ME	951	0.01685	0.84196
523	ME	966	0.01711	0.85907
525	ME	977	0.01731	0.87638
527	ME	903	0.01600	0.89238
529	ME	878	0.01555	0.90793
532	EE	889	0.01575	0.92368
535	EE	923	0.01635	0.94003
538	EE	821	0.01454	0.95458
542	EE	801	0.01419	0.96877
548	EE	679	0.01203	0.98080
556	EE	537	0.00951	0.99031
560	EE	547	0.00969	1.00000

Table 2.3.12  
Cumulative Scale Score Distribution  
Mathematics Grade 8

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	1	0.00002	0.00002
441	NM	16	0.00028	0.00030
442	NM	35	0.00061	0.00091
443	NM	272	0.00474	0.00564
444	NM	315	0.00548	0.01112
445	NM	514	0.00895	0.02007
451	NM	684	0.01191	0.03198
456	NM	883	0.01537	0.04735
460	NM	1052	0.01831	0.06567
464	NM	1134	0.01974	0.08541
467	NM	1200	0.02089	0.10630
469	NM	1294	0.02253	0.12883
472	PM	1365	0.02376	0.15259
474	PM	1330	0.02315	0.17575
477	PM	1374	0.02392	0.19967
479	PM	1414	0.02462	0.22428
480	PM	1367	0.02380	0.24808
482	PM	1461	0.02543	0.27352
484	PM	1458	0.02538	0.29890
486	PM	1441	0.02509	0.32398
487	PM	1420	0.02472	0.34871
489	PM	1440	0.02507	0.37377
490	PM	1422	0.02476	0.39853
492	PM	1361	0.02369	0.42222
493	PM	1395	0.02429	0.44651
494	PM	1430	0.02490	0.47141
496	PM	1302	0.02267	0.49407
497	PM	1281	0.02230	0.51637
499	PM	2554	0.04446	0.56084
501	ME	1288	0.02242	0.58326
503	ME	1268	0.02207	0.60533
504	ME	1294	0.02253	0.62786
505	ME	1240	0.02159	0.64945
507	ME	1237	0.02154	0.67098
508	ME	1241	0.02160	0.69259
510	ME	1199	0.02087	0.71346
511	ME	1152	0.02006	0.73352
513	ME	1121	0.01952	0.75303
514	ME	1260	0.02194	0.77497
516	ME	1131	0.01969	0.79466
517	ME	1185	0.02063	0.81529
519	ME	1124	0.01957	0.83486
521	ME	1105	0.01924	0.85409
523	ME	1114	0.01939	0.87349
526	ME	1075	0.01871	0.89220
528	ME	1106	0.01925	0.91146
531	EE	1043	0.01816	0.92961
535	EE	943	0.01642	0.94603
539	EE	975	0.01697	0.96301
545	EE	861	0.01499	0.97799
554	EE	749	0.01304	0.99103
560	EE	515	0.00897	1.00000

# Section 2.4

## Rescore Analysis Results

This section shows the results of rescore analyses. Rescore analyses are conducted on human-scored items to ensure consistency in scoring across years. To detect rater drift, 200 student responses from a previous administration are *rescored* using raters during the current administration. Then, the resulting scores from the current year are compared to the previous scores (on the same set of 200 student responses). Effect sizes (i.e., Cohen’s *d*) are calculated using the means and standard deviations of the two sets of scores. The threshold for flagging an item is 0.5.

Table 2.4.1  
Rescore Analysis  
English Language Arts Grade 3

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA00287	3	0.92500	1.02500	0.82021	0.87073	0.12192	False
IA00288	3	0.90000	1.04000	0.77004	0.89577	0.18181	False

Table 2.4.2  
Rescore Analysis  
English Language Arts Grade 4

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA00225	3	1.43000	1.25500	0.71249	0.62605	-0.24562	False
IA00226	3	1.21000	1.27000	0.83027	0.75495	0.07227	False

Table 2.4.3  
Rescore Analysis  
Mathematics Grade 3

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01081	3	0.86000	0.89500	0.96178	0.96885	0.03639	False
IA10020	3	1.61500	1.60000	1.00089	0.95106	-0.01499	False

Table 2.4.4  
Rescore Analysis  
Mathematics Grade 4

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01057	4	2.23500	2.22000	1.20292	1.24473	-0.01247	False
IA12525	4	1.98000	1.93000	1.19446	1.16701	-0.04186	False

Table 2.4.5  
Rescore Analysis  
Mathematics Grade 5

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01159	4	1.84000	1.79500	1.32027	1.35727	-0.03408	False
IA02736	4	2.05000	2.03000	1.43099	1.39601	-0.01398	False

Table 2.4.6  
Rescore Analysis  
Mathematics Grade 6

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA00972	4	2.85500	2.76500	1.12709	1.15605	-0.07985	False
IA02078	4	1.41500	1.41000	1.41520	1.42902	-0.00353	False

Table 2.4.7  
Rescore Analysis  
Mathematics Grade 7

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01069	4	2.11000	2.10500	1.14211	1.13597	-0.00438	False
IA02722	4	1.96500	1.94000	1.61767	1.60289	-0.01545	False

Table 2.4.8  
Rescore Analysis  
Mathematics Grade 8

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA00864	4	2.75000	2.73500	1.43800	1.44054	-0.01043	False
IA12109	4	1.81000	1.83000	1.58634	1.62614	0.01261	False

# Section 2.5

## Tabled Delta Analysis Results

Table 2.5.1  
Delta Analysis  
English Language Arts Grade 3

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00279 (EL308822)	0.78000	0.77000	9.91123	10.04461	1	False	-0.40934
IA00280 (EL308824)	0.66000	0.63000	11.35015	11.67259	1	False	-0.05862
IA00281 (EL308826)	0.59000	0.55000	12.08982	12.49735	1	False	1.14251
IA00282 (EL308827)	0.65000	0.60000	11.45872	11.98661	1	False	1.89333
IA00283 (EL308835)	0.57000	0.56000	12.29450	12.39612	1	False	-1.46411
IA00284 (EL308837)	0.66000	0.64000	11.35015	11.56616	1	False	-1.03791
IA00285 (EL308838)	0.77000	0.76000	10.04461	10.17479	1	False	-0.45521
IA00286 (EL308842)	0.41000	0.43000	13.91018	13.70550	1	False	0.44127
IA00287 (EL308855)	0.37333	0.39000	14.29215	14.11728	3	False	-0.04893
IA00288 (EL308857)	0.32000	0.31667	14.87080	14.90816	3	False	-0.69190
IA00443 (EL626042844)	0.66000	0.66000	11.35015	11.35015	1	False	0.00476
IA00444 (EL626043062)	0.67000	0.65000	11.24035	11.45872	1	False	-1.07831
IA00445 (EL626043435)	0.51500	0.51500	12.84957	12.84957	2	False	-0.84274
IA00446 (EL626049849)	0.53000	0.50000	12.69892	13.00000	1	False	0.50719
IA00450 (EL626050679)	0.67000	0.65000	11.24035	11.45872	1	False	-1.07831
IA00451 (EL626050927)	0.51000	0.54000	12.89972	12.59827	1	False	1.90293
IA00452 (EL626051097)	0.55000	0.52000	12.49735	12.79939	2	False	0.40201
IA00453 (EL626051328)	0.64000	0.65000	11.56616	11.45872	1	False	0.87138



Table 2.5.2  
Delta Analysis  
English Language Arts Grade 4

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00218 (EL307705)	0.81000	0.82000	9.48841	9.33854	1	False	1.24905
IA00219 (EL307709)	0.77000	0.77000	10.04461	10.04461	1	False	-0.37941
IA00220 (EL307710)	0.43000	0.45000	13.70550	13.50265	1	False	0.46328
IA00221 (EL307713)	0.52000	0.53000	12.79939	12.69892	1	False	-0.25251
IA00222 (EL307714)	0.76000	0.77000	10.17479	10.04461	1	False	0.84586
IA00223 (EL307719)	0.56000	0.57000	12.39612	12.29450	1	False	-0.11690
IA00224 (EL307724)	0.75000	0.73000	10.30204	10.54875	1	False	-0.03571
IA00225 (EL307728)	0.45333	0.44333	13.46898	13.57009	3	False	-0.47413
IA00226 (EL307729)	0.41667	0.42000	13.84171	13.80757	3	False	-1.21874
IA00289 (EL309792)	0.64000	0.62000	11.56616	11.77808	1	False	0.01597
IA00407 (EL624647403)	0.41000	0.42000	13.91018	13.80757	1	False	-0.57431
IA00408 (EL624647580)	0.58000	0.54000	12.19243	12.59827	1	False	2.09426
IA00411 (EL624652450)	0.78000	0.76000	9.91123	10.17479	1	False	0.00759
IA00412 (EL624652621)	0.90000	0.89000	7.87379	8.09389	1	False	-1.04339
IA00414 (EL624652989)	0.43000	0.44000	13.70550	13.60388	1	False	-0.52077
IA00415 (EL624653348)	0.67000	0.63000	11.24035	11.67259	1	False	2.05722
IA00416 (EL624653492)	0.72000	0.71000	10.66863	10.78646	2	False	-1.17545
IA00419 (EL624654711)	0.80000	0.79500	9.63352	9.70443	2	False	-0.94191

Table 2.5.3  
Delta Analysis  
English Language Arts Grade 5

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00495 (EL626304658)	0.75000	0.75000	10.30204	10.30204	1	False	0.20886
IA00497 (EL626304969)	0.73000	0.67000	10.54875	11.24035	1	False	2.45837
IA00500 (EL626332335)	0.66000	0.64000	11.35015	11.56616	1	False	-1.04829
IA00501 (EL626332592)	0.84000	0.83000	9.02217	9.18334	1	False	-0.59879
IA00502 (EL626333002)	0.82000	0.81000	9.33854	9.48841	1	False	-0.62778
IA00505 (EL626355215)	0.60000	0.61000	11.98661	11.88272	1	False	0.40487
IA00506 (EL626355557)	0.64000	0.65000	11.56616	11.45872	1	False	0.59214
IA00508 (EL626356291)	0.36500	0.37500	14.38050	14.27456	2	False	-0.48260
IA00638 (EL627351056)	0.66000	0.66000	11.35015	11.35015	1	False	-0.18691
IA01669 (EL711809263)	0.76000	0.74000	10.17479	10.42662	1	False	-1.20527
IA01670 (EL711809592)	0.76000	0.74000	10.17479	10.42662	1	False	-1.20527
IA01671 (EL711827203)	0.90000	0.89000	7.87379	8.09389	1	False	-0.63712
IA01672 (EL711827807)	0.73000	0.68000	10.54875	11.12920	1	False	1.56815
IA01679 (EL711868011)	0.49500	0.46500	13.05013	13.35138	2	False	0.27628
IA01680 (EL711900602)	0.70000	0.66000	10.90240	11.35015	1	False	0.63875
IA01691 (EL712167015)	0.46000	0.47000	13.40173	13.30108	1	False	-0.15538

Table 2.5.4  
Delta Analysis  
English Language Arts Grade 6

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00173 (EL303496)	0.74000	0.74000	10.42662	10.42662	1	False	-0.96345
IA00174 (EL303500)	0.64000	0.64000	11.56616	11.56616	1	False	-1.30105
IA00175 (EL303504)	0.61000	0.61000	11.88272	11.88272	1	False	-1.27260
IA00176 (EL303508)	0.71000	0.72000	10.78646	10.66863	1	False	0.24762
IA00177 (EL303510)	0.87000	0.86000	8.49444	8.67872	1	False	-0.21550
IA00178 (EL303513)	0.66000	0.68000	11.35015	11.12920	1	False	1.23378
IA00179 (EL303514)	0.63000	0.61000	11.67259	11.88272	1	False	1.01514
IA00180 (EL303518)	0.57000	0.54000	12.29450	12.59827	1	False	2.24640
IA00515 (EL626864414)	0.83000	0.82000	9.18334	9.33854	1	False	-0.33668
IA00517 (EL626864724)	0.69000	0.70000	11.01660	10.90240	1	False	0.13889
IA00518 (EL626865003)	0.67000	0.68000	11.24035	11.12920	1	False	0.03840
IA00520 (EL626865416)	0.41000	0.42000	13.91018	13.80757	1	False	-0.84803
IA00522 (EL626865773)	0.72000	0.70000	10.66863	10.90240	1	False	0.98193
IA00523 (EL626865942)	0.64000	0.65000	11.56616	11.45872	1	False	-0.09946
IA00528 (EL626867605)	0.77000	0.76500	10.04461	10.11008	2	False	-1.08498
IA00530 (EL626868748)	0.70500	0.71500	10.84466	10.72779	2	False	0.21959

Table 2.5.5  
Delta Analysis  
English Language Arts Grade 7

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00065 (EL292160)	0.73000	0.71000	10.54875	10.78646	1	False	-0.87428
IA00066 (EL292163)	0.65000	0.59000	11.45872	12.08982	1	False	2.06383
IA00067 (EL292168)	0.50000	0.48000	13.00000	13.20061	1	False	-0.78093
IA00068 (EL292170)	0.71000	0.69000	10.78646	11.01660	1	False	-0.89359
IA00069 (EL292172)	0.62000	0.62000	11.77808	11.77808	1	False	-0.00403
IA00070 (EL292176)	0.62000	0.61000	11.77808	11.88272	1	False	-0.75026
IA00081 (EL293802)	0.68000	0.67000	11.12920	11.24035	1	False	-0.70184
IA00082 (EL293804)	0.54000	0.54000	12.59827	12.59827	1	False	-0.12378
IA00257 (EL308358)	0.85000	0.82000	8.85427	9.33854	1	False	0.63653
IA00258 (EL308360)	0.75500	0.74500	10.23876	10.36465	2	False	-0.67695
IA00262 (EL308382)	0.65000	0.63000	11.45872	11.67259	1	False	-0.91145
IA00265 (EL308389)	0.90000	0.90000	7.87379	7.87379	1	False	0.56603
IA00269 (EL308397)	0.84000	0.80000	9.02217	9.63352	1	False	1.56720
IA00655 (EL628647210)	0.72000	0.72000	10.66863	10.66863	1	False	0.15796
IA00657 (EL628647689)	0.77000	0.75000	10.04461	10.30204	1	False	-0.80730
IA00658 (EL628653398)	0.74000	0.75500	10.42662	10.23876	2	False	1.53288

Table 2.5.6  
Delta Analysis  
English Language Arts Grade 8

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00056 (EL290795)	0.77000	0.76000	10.04461	10.17479	1	False	-0.52986
IA00057 (EL290798)	0.78000	0.79000	9.91123	9.77432	1	False	0.28064
IA00058 (EL290799)	0.75000	0.74000	10.30204	10.42662	1	False	-0.63407
IA00059 (EL290800)	0.67000	0.68000	11.24035	11.12920	1	False	0.23733
IA00060 (EL290801)	0.77000	0.76000	10.04461	10.17479	1	False	-0.52986
IA00061 (EL290805)	0.56000	0.56000	12.39612	12.39612	1	False	-0.74367
IA00062 (EL290808)	0.54000	0.54000	12.59827	12.59827	1	False	-0.70859
IA00063 (EL290814)	0.44000	0.39000	13.60388	14.11728	1	False	2.92653
IA00368 (EL623873883)	0.72000	0.73000	10.66863	10.54875	1	False	0.23107
IA00371 (EL623951471)	0.60500	0.61500	11.93476	11.83050	2	False	0.28463
IA00373 (EL623952377)	0.42000	0.43500	13.80757	13.65463	2	False	1.12718
IA00374 (EL623952612)	0.73000	0.71000	10.54875	10.78646	1	False	0.52587
IA00378 (EL623955555)	0.51000	0.50000	12.89972	13.00000	1	False	-1.34321
IA00379 (EL623955757)	0.54000	0.54000	12.59827	12.59827	1	False	-0.70859
IA00383 (EL623959265)	0.65000	0.63000	11.45872	11.67259	1	False	0.11447
IA00699 (EL632808123)	0.77000	0.76000	10.04461	10.17479	1	False	-0.52986

Table 2.5.7  
Delta Analysis  
Mathematics Grade 3

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00799 (MA260559)	0.49000	0.53000	13.10028	12.69892	1	False	0.49128
IA00834 (MA293457)	0.81000	0.80000	9.48841	9.63352	1	False	0.58860
IA00838 (MA293524)	0.77000	0.79000	10.04461	9.77432	1	False	-0.70179
IA00850 (MA297405)	0.71000	0.71000	10.78646	10.78646	1	False	-0.40684
IA00852 (MA297438)	0.65000	0.65000	11.45872	11.45872	1	False	-0.51178
IA00924 (MA306310)	0.48000	0.49000	13.20061	13.10028	1	False	-1.13778
IA00925 (MA306315)	0.79000	0.78000	9.77432	9.91123	1	False	0.49923
IA00930 (MA306359)	0.64000	0.68000	11.56616	11.12920	1	False	0.44635
IA00932 (MA306375)	0.49000	0.51000	13.10028	12.89972	1	False	-0.60589
IA01019 (MA311277)	0.74000	0.81000	10.42662	9.48841	1	False	3.00720
IA01081 (MA623656013)	0.32000	0.32000	14.87080	14.87080	3	False	-1.04440
IA02323 (MA301611A)	0.87000	0.87000	8.49444	8.49444	1	False	-0.04906
IA04659 (MA307303)	0.46000	0.43000	13.40173	13.70550	1	False	0.84463
IA04760 (MA713752330)	0.80000	0.80000	9.63352	9.63352	1	False	-0.22687
IA04813 (MA735572247)	0.69000	0.72000	11.01660	10.66863	1	False	-0.12570
IA04828 (MA735653938)	0.57000	0.59000	12.29450	12.08982	1	False	-0.70909
IA04844 (MA735735757)	0.61000	0.62000	11.88272	11.77808	1	False	-1.14975
IA07855 (MA900579464)	0.51000	0.47000	12.89972	13.30108	1	False	1.45624
IA10020 (MA001639117)	0.55333	0.58667	12.46365	12.12409	3	False	0.05428
IA12584 (MA900373094)	0.58000	0.60000	12.19243	11.98661	1	False	-0.71885

Table 2.5.8  
Delta Analysis  
Mathematics Grade 4

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00828 (MA287237)	0.79000	0.81000	9.77432	9.48841	1	False	0.13031
IA00841 (MA293718)	0.73000	0.74000	10.54875	10.42662	1	False	-0.85876
IA00861 (MA297629)	0.90000	0.87000	7.87379	8.49444	1	False	0.92295
IA00906 (MA301811)	0.72000	0.74000	10.66863	10.42662	1	False	-0.36651
IA00958 (MA307055)	0.51000	0.50000	12.89972	13.00000	1	False	0.30691
IA00961 (MA307081)	0.61000	0.56000	11.88272	12.39612	1	False	1.79939
IA00963 (MA307085)	0.70000	0.69000	10.90240	11.01660	1	False	-0.30529
IA01049 (MA311537)	0.67000	0.68000	11.24035	11.12920	1	False	-1.14099
IA01055 (MA311572)	0.55000	0.61000	12.49735	11.88272	1	False	0.67202
IA01057 (MA311581)	0.55750	0.56250	12.42146	12.37076	4	False	-0.52539
IA01093 (MA623879088)	0.69500	0.74500	10.95971	10.36465	2	False	1.10398
IA02175 (MA286769)	0.78000	0.78000	9.91123	9.91123	1	False	-1.14724
IA02819 (MA713583365)	0.59000	0.68000	12.08982	11.12920	1	False	2.34693
IA02841 (MA713774890)	0.50000	0.55000	13.00000	12.49735	1	False	0.00478
IA02902 (MA714251321)	0.41000	0.45000	13.91018	13.50265	1	False	-0.72501
IA04945 (MA800744502)	0.66000	0.63000	11.35015	11.67259	1	False	0.77112
IA04965 (MA800867144)	0.62000	0.63000	11.77808	11.67259	1	False	-0.98595
IA12273 (MA200267672)	0.56000	0.56000	12.39612	12.39612	1	False	-0.30862
IA12278 (MA200334941)	0.51000	0.52000	12.89972	12.79939	1	False	-0.58451
IA12525 (MA713956124)	0.46750	0.49250	13.32622	13.07520	4	False	-1.11012

Table 2.5.9  
Delta Analysis  
Mathematics Grade 5

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00771 (MA204911)	0.75000	0.78000	10.30204	9.91123	1	False	1.46004
IA00776 (MA221207)	0.69000	0.67000	11.01660	11.24035	1	False	-0.54219
IA00803 (MA262207)	0.75000	0.73000	10.30204	10.54875	1	False	-0.52220
IA00806 (MA272292)	0.47000	0.48000	13.30108	13.20061	1	False	-0.60012
IA00880 (MA298106)	0.24000	0.24000	15.82521	15.82521	1	False	-1.06067
IA00885 (MA299556)	0.73000	0.75000	10.54875	10.30204	1	False	0.62400
IA00943 (MA306466)	0.64000	0.68000	11.56616	11.12920	1	False	1.52607
IA00989 (MA307638)	0.87000	0.85000	8.49444	8.85427	1	False	-0.16642
IA01020 (MA311280)	0.43000	0.45000	13.70550	13.50265	1	False	-0.09318
IA01029 (MA311337)	0.86000	0.85000	8.67872	8.85427	1	False	-1.16052
IA01149 (MA624347774)	0.42000	0.43000	13.80757	13.70550	1	False	-0.66724
IA01155 (MA624357395)	0.41500	0.37500	13.85881	14.27456	2	False	0.94915
IA01159 (MA624377498)	0.58750	0.57500	12.11553	12.24353	4	False	-0.90803
IA02552 (MA311324)	0.44000	0.43000	13.60388	13.70550	1	False	-0.83081
IA02725 (MA704359215)	0.64000	0.57000	11.56616	12.29450	1	False	2.33799
IA02736 (MA704359678)	0.49750	0.52500	13.02507	12.74917	4	False	0.41397
IA02917 (MA715102107)	0.88000	0.85000	8.30005	8.85427	1	False	0.88213
IA04970 (MA800974344)	0.73000	0.74000	10.54875	10.42662	1	False	-0.06671
IA05002 (MA801652356)	0.58000	0.56000	12.19243	12.39612	1	False	-0.47680
IA05015 (MA801763240)	0.26000	0.26000	15.57338	15.57338	1	False	-1.09848



Table 2.5.10  
Delta Analysis  
Mathematics Grade 6

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00778 (MA221669)	0.79000	0.81000	9.77432	9.48841	1	False	-0.85988
IA00804 (MA264305)	0.75000	0.77000	10.30204	10.04461	1	False	-0.84835
IA00817 (MA280989)	0.49000	0.49000	13.10028	13.10028	1	False	-0.57194
IA00818 (MA282268)	0.41000	0.41000	13.91018	13.91018	1	False	-0.84091
IA00819 (MA282277)	0.44000	0.45000	13.60388	13.50265	1	False	-0.65000
IA00827 (MA287186)	0.55000	0.58000	12.49735	12.19243	1	False	0.15387
IA00845 (MA296349)	0.53000	0.57000	12.69892	12.29450	1	False	0.79291
IA00972 (MA307339)	0.73000	0.71500	10.54875	10.72779	4	False	1.30502
IA01058 (MA311658)	0.46000	0.44000	13.40173	13.60388	1	False	0.49035
IA02037 (MA217493)	0.63000	0.68000	11.67259	11.12920	1	False	1.25117
IA02078 (MA251350)	0.36000	0.36250	14.43384	14.40714	4	False	-0.80297
IA02597 (MA311693)	0.77000	0.80000	10.04461	9.63352	1	False	-0.05018
IA02698 (MA703179529)	0.24000	0.23000	15.82521	15.95539	1	False	-0.72832
IA02906 (MA714281467)	0.67000	0.64000	11.24035	11.56616	1	False	1.91934
IA04745 (MA703231515)	0.52500	0.53500	12.74917	12.64862	2	False	-0.93775
IA04884 (MA736365836)	0.68000	0.67000	11.12920	11.24035	1	False	0.72178
IA04899 (MA736511626)	0.54000	0.60000	12.59827	11.98661	1	False	1.95118
IA05126 (MA805103779)	0.56000	0.57000	12.39612	12.29450	1	False	-0.92244
IA05134 (MA805167086)	0.32000	0.31000	14.87080	14.98340	1	False	-0.51240
IA05135 (MA805171807)	0.58000	0.59000	12.19243	12.08982	1	False	-0.86047

Table 2.5.11  
Delta Analysis  
Mathematics Grade 7

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00831 (MA288414)	0.73000	0.72000	10.54875	10.66863	1	False	-0.45377
IA00847 (MA296358)	0.53000	0.53000	12.69892	12.69892	1	False	-0.90568
IA00909 (MA301846)	0.87000	0.88000	8.49444	8.30005	1	False	0.27788
IA00948 (MA306600)	0.85000	0.84000	8.85427	9.02217	1	False	-0.40131
IA00949 (MA306605)	0.50000	0.51000	13.00000	12.89972	1	False	-1.03213
IA01006 (MA311093)	0.86000	0.86000	8.67872	8.67872	1	False	-1.02099
IA01011 (MA311109)	0.41000	0.47000	13.91018	13.30108	1	False	2.15307
IA01018 (MA311140)	0.42000	0.38000	13.80757	14.22192	1	False	1.97344
IA01069 (MA316886)	0.51750	0.52000	12.82448	12.79939	4	False	-1.05032
IA01097 (MA623950280)	0.35000	0.35000	14.54128	14.54128	1	False	-0.62151
IA02722 (MA703943185)	0.48500	0.46750	13.15043	13.32622	4	False	0.31290
IA04486 (MA227988)	0.68000	0.69000	11.12920	11.01660	1	False	-0.66299
IA04538 (MA282218)	0.66000	0.62000	11.35015	11.77808	1	False	1.68316
IA04591 (MA298183)	0.50000	0.52000	13.00000	12.79939	1	False	-0.37634
IA04593 (MA298192)	0.19000	0.19000	16.51159	16.51159	1	False	-0.31760
IA04629 (MA303730)	0.51000	0.50000	12.89972	13.00000	1	False	-0.21932
IA07614 (MA306636)	0.43000	0.47000	13.70550	13.30108	1	False	0.84686
IA08208 (MA904265644)	0.40500	0.42500	13.96170	13.75647	2	False	-0.49450
IA10259 (MA005177715)	0.36000	0.40000	14.43384	14.01339	1	False	0.83929
IA10268 (MA005205470)	0.34000	0.36000	14.64985	14.43384	1	False	-0.53014

Table 2.5.12  
Delta Analysis  
Mathematics Grade 8

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00849 (MA296757)	0.62000	0.64000	11.77808	11.56616	1	False	-0.90473
IA00858 (MA297513)	0.78000	0.78000	9.91123	9.91123	1	False	-0.28617
IA00864 (MA297652)	0.69250	0.70250	10.98820	10.87358	4	False	-0.88851
IA00865 (MA297656)	0.60000	0.56000	11.98661	12.39612	1	False	0.58953
IA00905 (MA301702)	0.46000	0.52000	13.40173	12.79939	1	False	0.70643
IA00979 (MA307472)	0.65000	0.65000	11.45872	11.45872	1	False	-0.61963
IA00985 (MA307570)	0.53000	0.54000	12.69892	12.59827	1	False	-1.06571
IA01033 (MA311384)	0.57000	0.59000	12.29450	12.08982	1	False	-0.81680
IA01037 (MA311414)	0.39000	0.41000	14.11728	13.91018	1	False	-0.41622
IA01042 (MA311448)	0.50000	0.51000	13.00000	12.89972	1	False	-1.00206
IA01125 (MA624247061)	0.41500	0.42500	13.85881	13.75647	2	False	-0.81035
IA02495 (MA309741)	0.37000	0.45000	14.32741	13.50265	1	False	1.62441
IA04665 (MA307399)	0.47000	0.57000	13.30108	12.29450	1	False	1.99057
IA04678 (MA309738)	0.40000	0.39000	14.01339	14.11728	1	False	-0.83451
IA04719 (MA311462)	0.59000	0.61000	12.08982	11.88272	1	False	-0.85311
IA04957 (MA800770988)	0.46000	0.39000	13.40173	14.11728	1	False	1.27320
IA05059 (MA803856627)	0.71000	0.67000	10.78646	11.24035	1	False	0.99149
IA05070 (MA804042487)	0.40000	0.33000	14.01339	14.75965	1	False	1.24065
IA10142 (MA002243189)	0.79000	0.84000	9.77432	9.02217	1	False	0.40869
IA12109 (MA002234564)	0.44750	0.42500	13.52792	13.75647	4	False	-0.32718

# Section 2.6

## Tabled B/B Analysis Results

Table 2.6.1 b/b Analysis English Language Arts Grade 3

Item Id	Old b	New b	Std Dist	Flag
IA00279 (EL308822)	-1.19830	-0.81510	-0.85348	False
IA00280 (EL308824)	-0.58890	-0.37460	0.02315	False
IA00281 (EL308826)	0.22670	0.50940	0.32834	False
IA00282 (EL308827)	-0.47740	-0.12230	-0.32774	False
IA00283 (EL308835)	-0.18710	0.10960	-0.44542	False
IA00284 (EL308837)	-0.44530	-0.18760	-0.90151	False
IA00285 (EL308838)	-0.88420	-0.58840	-0.39696	False
IA00286 (EL308842)	0.34210	0.72160	1.91337	False
IA00287 (EL308855)	0.78583	0.59937	2.26922	False
IA00288 (EL308857)	1.15683	0.97997	1.27430	False
IA00443 (EL626042844)	-0.77730	-0.59000	0.82922	False
IA00444 (EL626043062)	-0.48490	-0.20930	-1.05263	False
IA00445 (EL626043435)	-0.37490	-0.06235	-0.66746	False
IA00446 (EL626049849)	0.06220	0.35790	0.12183	False
IA00450 (EL626050679)	-0.45320	-0.13060	-0.71325	False
IA00451 (EL626050927)	0.13620	0.42840	0.24665	False
IA00452 (EL626051097)	-0.30750	-0.05210	-1.19130	False
IA00453 (EL626051328)	-0.57110	-0.20940	-0.45631	False

Table 2.6.2 b/b Analysis English Language Arts Grade 4

Item Id	Old b	New b	Std Dist	Flag
IA00218 (EL307705)	-1.73240	-1.36900	-1.12068	False
IA00219 (EL307709)	-1.07490	-1.21080	2.86074	False
IA00220 (EL307710)	0.18730	0.58590	0.72737	False
IA00221 (EL307713)	0.09620	0.35710	-0.75064	False
IA00222 (EL307714)	-1.13520	-1.05810	0.73049	False
IA00223 (EL307719)	-0.12200	0.19600	-0.33621	False
IA00224 (EL307724)	-0.87340	-0.65880	-0.87782	False
IA00225 (EL307728)	-0.10186	0.26470	0.17572	False
IA00226 (EL307729)	0.56296	0.47350	1.11631	False
IA00289 (EL309792)	-0.38780	-0.36250	0.68048	False
IA00407 (EL624647403)	0.79560	0.82520	-0.28092	False
IA00408 (EL624647580)	-0.43560	-0.05950	0.01449	False
IA00411 (EL624652450)	-1.13410	-0.81450	-1.10450	False
IA00412 (EL624652621)	-2.04980	-1.57240	-0.20159	False
IA00414 (EL624652989)	0.70380	0.84750	-1.37592	False
IA00415 (EL624653348)	-0.53070	-0.15510	-0.06435	False
IA00416 (EL624653492)	-1.31845	-0.87005	0.06900	False
IA00419 (EL624654711)	-1.74605	-1.29760	-0.26199	False

Table 2.6.3 b/b Analysis English Language Arts Grade 5

Item Id	Old b	New b	Std Dist	Flag
IA00495 (EL626304658)	-1.10510	-1.06970	2.29381	False
IA00497 (EL626304969)	-0.43840	-0.06570	0.17725	False
IA00500 (EL626332335)	-0.47970	-0.28240	-0.39485	False
IA00501 (EL626332592)	-2.04200	-1.55250	-1.03306	False
IA00502 (EL626333002)	-1.59450	-1.26970	0.00493	False
IA00505 (EL626355215)	-0.47950	-0.29450	-0.26531	False
IA00506 (EL626355557)	-0.58130	-0.31070	-1.00960	False
IA00508 (EL626356291)	0.39485	0.55395	-0.77289	False
IA00638 (EL627351056)	-0.45490	-0.31220	0.14274	False
IA01669 (EL711809263)	-1.20930	-0.77950	-0.42720	False
IA01670 (EL711809592)	-0.75930	-0.46190	-1.01378	False
IA01671 (EL711827203)	-2.12800	-1.44710	0.78502	False
IA01672 (EL711827807)	-1.34860	-0.65750	2.11329	False
IA01679 (EL711868011)	-0.10355	0.14180	-0.64284	False
IA01680 (EL711900602)	-0.57280	-0.41880	0.20809	False
IA01691 (EL712167015)	0.41050	0.62480	-0.16561	False

Table 2.6.4 b/b Analysis English Language Arts Grade 6

Item Id	Old b	New b	Std Dist	Flag
IA00173 (EL303496)	-1.40490	-0.62430	0.85354	False
IA00174 (EL303500)	-0.87130	-0.50300	0.83465	False
IA00175 (EL303504)	-0.22770	0.17010	0.21529	False
IA00176 (EL303508)	-1.70470	-0.99480	-0.71791	False
IA00177 (EL303510)	-2.33850	-1.33880	0.04244	False
IA00178 (EL303513)	-0.89460	-0.39260	-1.14445	False
IA00179 (EL303514)	-0.03620	-0.02820	2.67108	False
IA00180 (EL303518)	0.30740	0.46010	-1.14081	False
IA00515 (EL626864414)	-2.27480	-1.32320	-0.41910	False
IA00517 (EL626864724)	-1.25290	-0.69710	-0.36094	False
IA00518 (EL626865003)	-0.93850	-0.52060	0.36558	False
IA00520 (EL626865416)	1.38450	1.36660	1.03370	False
IA00522 (EL626865773)	-0.79740	-0.36450	-0.50718	False
IA00523 (EL626865942)	-0.93230	-0.34460	-0.01421	False
IA00528 (EL626867605)	-1.77925	-1.00805	-0.98803	False
IA00530 (EL626868748)	-1.64355	-0.89480	-0.72363	False

Table 2.6.5 b/b Analysis English Language Arts Grade 7

Item Id	Old b	New b	Std Dist	Flag
IA00065 (EL292160)	-0.90660	-0.50700	1.95384	False
IA00066 (EL292163)	-0.34950	0.03860	-0.56151	False
IA00067 (EL292168)	0.31490	0.56270	-0.24896	False
IA00068 (EL292170)	-1.08250	-0.48880	-0.68794	False
IA00069 (EL292172)	-0.73500	-0.23520	-1.38834	False
IA00070 (EL292176)	-0.65150	-0.11150	0.08569	False
IA00081 (EL293802)	-0.93560	-0.33550	0.22214	False
IA00082 (EL293804)	-0.08550	0.26700	-0.96454	False
IA00257 (EL308358)	-1.80700	-1.19450	0.93459	False
IA00258 (EL308360)	-1.74560	-1.07750	-0.83069	False
IA00262 (EL308382)	-0.76810	-0.16280	1.20496	False
IA00265 (EL308389)	-2.69280	-1.84020	-0.87021	False
IA00269 (EL308397)	-2.30110	-1.39560	1.31066	False
IA00655 (EL628647210)	-1.25380	-0.57340	0.71479	False
IA00657 (EL628647689)	-1.35460	-0.74960	-1.15614	False
IA00658 (EL628653398)	-1.83675	-1.19355	0.28166	False

Table 2.6.6 b/b Analysis English Language Arts Grade 8

Item Id	Old b	New b	Std Dist	Flag
IA00056 (EL290795)	-2.15880	-1.25830	-0.25329	False
IA00057 (EL290798)	-2.12280	-1.38570	-0.26965	False
IA00058 (EL290799)	-0.83980	-0.48740	-0.73069	False
IA00059 (EL290800)	-0.70730	-0.50750	0.12571	False
IA00060 (EL290801)	-0.77460	-0.46950	-0.53186	False
IA00061 (EL290805)	-0.11290	0.07600	-0.30422	False
IA00062 (EL290808)	0.16820	0.22400	-0.59418	False
IA00063 (EL290814)	0.88790	0.77860	0.07050	False
IA00368 (EL623873883)	-1.03830	-0.55920	-0.53484	False
IA00371 (EL623951471)	-0.78990	-0.48915	-0.45477	False
IA00373 (EL623952377)	0.34170	0.37460	-0.29858	False
IA00374 (EL623952612)	-1.14430	-0.59460	-0.26204	False
IA00378 (EL623955555)	-0.27310	0.25450	1.96623	False
IA00379 (EL623955757)	0.56660	-0.02300	2.93117	False
IA00383 (EL623959265)	-0.53400	-0.35280	-0.20377	False
IA00699 (EL632808123)	-1.37450	-0.79460	-0.65571	False

Table 2.6.7 b/b Analysis Mathematics Grade 3

Item Id	Old b	New b	Std Dist	Flag
IA00799 (MA260559)	0.38435	0.19810	0.18536	False
IA00834 (MA293457)	-1.04475	-1.03030	-0.43591	False
IA00838 (MA293524)	-1.08336	-1.02900	-0.68155	False
IA00850 (MA297405)	-0.81918	-0.75080	-0.49742	False
IA00852 (MA297438)	-0.45849	-0.36060	-0.12228	False
IA00924 (MA306310)	1.16083	1.07750	-0.51972	False
IA00925 (MA306315)	-0.97068	-0.83970	-0.15735	False
IA00930 (MA306359)	-0.57583	-0.63890	-0.15465	False
IA00932 (MA306375)	0.73550	0.53190	0.12591	False
IA01019 (MA311277)	-0.99639	-1.27660	1.49531	False
IA01081 (MA623656013)	0.78644	0.66650	-0.45456	False
IA02323 (MA301611A)	-1.84954	-1.56910	0.39747	False
IA04659 (MA307303)	0.44720	0.37200	-0.58278	False
IA04760 (MA713752330)	-1.46878	-1.40480	-0.55376	False
IA04813 (MA735572247)	-0.86180	-0.86640	-0.40046	False
IA04828 (MA735653938)	-0.22403	-0.30070	-0.23931	False
IA04844 (MA735735757)	-0.18889	-0.23990	-0.42705	False
IA07855 (MA900579464)	-0.41290	0.25850	3.70600	False
IA10020 (MA001639117)	-0.31623	-0.40417	-0.11871	False
IA12584 (MA900373094)	-0.09400	-0.13140	-0.56454	False

Table 2.6.8 b/b Analysis Mathematics Grade 4

Item Id	Old b	New b	Std Dist	Flag
IA00828 (MA287237)	-0.77841	-1.11660	0.58398	False
IA00841 (MA293718)	-0.19751	-0.34140	-0.73093	False
IA00861 (MA297629)	-1.68976	-1.25660	1.48479	False
IA00906 (MA301811)	-0.64521	-0.07480	2.62247	False
IA00958 (MA307055)	0.16895	0.21210	-0.15852	False
IA00961 (MA307081)	-0.10105	0.14030	0.89874	False
IA00963 (MA307085)	-0.40294	-0.42960	-0.75055	False
IA01049 (MA311537)	-0.30079	-0.38860	-1.02123	False
IA01055 (MA311572)	-0.02167	-0.33400	0.18524	False
IA01057 (MA311581)	-0.24431	-0.29603	-0.84320	False
IA01093 (MA623879088)	-0.66234	-0.95475	0.28102	False
IA02175 (MA286769)	-0.74467	-0.82670	-0.90848	False
IA02819 (MA713583365)	-0.42498	-0.97330	1.68276	False
IA02841 (MA713774890)	0.10634	-0.15690	-0.14080	False
IA02902 (MA714251321)	0.56595	0.32020	-0.39330	False
IA04945 (MA800744502)	-0.38040	-0.41370	-0.78153	False
IA04965 (MA800867144)	-0.35378	-0.50550	-0.63419	False
IA12273 (MA200267672)	0.89900	0.86220	-0.38045	False
IA12278 (MA200334941)	0.06330	-0.06940	-0.88153	False
IA12525 (MA713956124)	0.36842	0.08568	-0.11431	False

Table 2.6.9 b/b Analysis Mathematics Grade 5

Item Id	Old b	New b	Std Dist	Flag
IA00771 (MA204911)	-1.17443	-1.21510	0.48657	False
IA00776 (MA221207)	-0.74023	-0.67190	-0.95537	False
IA00803 (MA262207)	-0.81500	-0.59750	-0.00097	False
IA00806 (MA272292)	0.29223	0.32100	-0.45593	False
IA00880 (MA298106)	1.10301	1.12600	0.38335	False
IA00885 (MA299556)	-0.57658	-0.69030	0.47675	False
IA00943 (MA306466)	0.19086	-0.20960	2.17266	False
IA00989 (MA307638)	-1.50445	-1.36320	-0.76168	False
IA01020 (MA311280)	0.58970	0.53780	-0.84347	False
IA01029 (MA311337)	-1.67303	-1.64330	0.41085	False
IA01149 (MA624347774)	0.38574	0.37410	-0.71101	False
IA01155 (MA624357395)	0.33304	0.49225	0.74380	False
IA01159 (MA624377498)	-0.24085	-0.25208	-0.79946	False
IA02552 (MA311324)	0.89321	0.73400	-0.73471	False
IA02725 (MA704359215)	-0.40130	-0.29410	-0.52320	False
IA02736 (MA704359678)	-0.02780	-0.12408	-0.28042	False
IA02917 (MA715102107)	-2.70247	-1.93960	2.75473	False
IA04970 (MA800974344)	-0.88694	-0.90660	-0.01519	False
IA05002 (MA801652356)	-0.24707	-0.17860	-0.69672	False
IA05015 (MA801763240)	1.07490	0.98460	-0.65058	False

Table 2.6.10 b/b Analysis Mathematics Grade 6

Item Id	Old b	New b	Std Dist	Flag
IA00778 (MA221669)	-1.13035	-0.94060	-0.18431	False
IA00804 (MA264305)	-1.40884	-1.49090	0.25797	False
IA00817 (MA280989)	0.34905	0.35850	-0.73992	False
IA00818 (MA282268)	0.65855	0.77060	0.30487	False
IA00819 (MA282277)	0.17311	0.24020	-0.37434	False
IA00827 (MA287186)	-0.14937	-0.15190	-1.15446	False
IA00845 (MA296349)	0.35420	0.10960	0.48713	False
IA00972 (MA307339)	-1.20129	-0.95668	0.22475	False
IA01058 (MA311658)	0.36997	0.55050	0.68885	False
IA02037 (MA217493)	-0.18307	-0.61450	2.37298	False
IA02078 (MA251350)	0.38907	0.43613	-0.40344	False
IA02597 (MA311693)	-1.09057	-1.05230	-0.93906	False
IA02698 (MA703179529)	1.32797	1.24270	-0.90414	False
IA02906 (MA714281467)	-0.74360	-0.60740	-0.38264	False
IA04745 (MA703231515)	-0.20044	-0.13350	-0.61198	False
IA04884 (MA736365836)	-0.90196	-0.71810	-0.08852	False
IA04899 (MA736511626)	0.40690	-0.10820	2.69191	False
IA05126 (MA805103779)	-0.33613	-0.28270	-0.80963	False
IA05134 (MA805167086)	1.15120	0.92480	-0.16783	False
IA05135 (MA805171807)	-0.46156	-0.33310	-0.26820	False



Table 2.6.11 b/b Analysis Mathematics Grade 7

Item Id	Old b	New b	Std Dist	Flag
IA00831 (MA288414)	-0.75473	-0.50390	-0.09032	False
IA00847 (MA296358)	-0.21338	-0.09260	-0.97601	False
IA00909 (MA301846)	-1.65805	-1.29260	-0.06325	False
IA00948 (MA306600)	-1.26915	-1.05990	-0.14772	False
IA00949 (MA306605)	-0.17121	-0.05210	-1.04432	False
IA01006 (MA311093)	-1.80972	-1.62400	1.57737	False
IA01011 (MA311109)	0.62617	0.48330	2.01907	False
IA01018 (MA311140)	0.16283	0.37780	1.40506	False
IA01069 (MA316886)	-0.12564	-0.06280	-0.08864	False
IA01097 (MA623950280)	0.38816	0.47650	-0.45482	False
IA02722 (MA703943185)	-0.08998	0.02773	-1.03222	False
IA04486 (MA227988)	-0.60540	-0.42510	-1.06980	False
IA04538 (MA282218)	-0.86266	-0.48400	2.07030	False
IA04591 (MA298183)	0.60011	0.68700	0.01977	False
IA04593 (MA298192)	1.23918	1.21870	-0.49613	False
IA04629 (MA303730)	0.77432	0.74620	-0.50083	False
IA07614 (MA306636)	0.57159	0.53780	0.08650	False
IA08208 (MA904265644)	0.24535	0.32455	-0.96582	False
IA10259 (MA005177715)	1.24300	1.11930	0.19577	False
IA10268 (MA005205470)	0.70990	0.75850	-0.44395	False

Table 2.6.12 b/b Analysis Mathematics Grade 8

Item Id	Old b	New b	Std Dist	Flag
IA00849 (MA296757)	-0.65691	-0.47720	-0.95939	False
IA00858 (MA297513)	-1.07637	-0.78270	-0.26696	False
IA00864 (MA297652)	-0.93219	-0.69943	-0.66511	False
IA00865 (MA297656)	-0.16632	0.33040	1.54866	False
IA00905 (MA301702)	0.25997	0.13160	0.71198	False
IA00979 (MA307472)	-0.04356	0.11560	-0.89548	False
IA00985 (MA307570)	0.38769	0.42470	-0.55142	False
IA01033 (MA311384)	-0.00738	0.14030	-0.96737	False
IA01037 (MA311414)	0.15717	0.28550	-1.05218	False
IA01042 (MA311448)	0.16909	0.07470	0.49351	False
IA01125 (MA624247061)	0.21933	0.30455	-0.84754	False
IA02495 (MA309741)	0.28898	0.17400	0.60313	False
IA04665 (MA307399)	0.48343	0.19370	1.82248	False
IA04678 (MA309738)	0.91601	0.94780	-0.69844	False
IA04719 (MA311462)	0.49800	0.77020	0.12755	False
IA04957 (MA800770988)	0.24010	0.39700	-0.81255	False
IA05059 (MA803856627)	-0.86979	-0.34280	1.52473	False
IA05070 (MA804042487)	0.56218	0.99320	1.32031	False
IA10142 (MA002243189)	-1.38240	-1.42530	0.65879	False
IA12109 (MA002234564)	0.04537	0.17243	-1.09472	False

# Section 2.7

## Tabled Beta Analysis Results

Table 2.7.1 Beta Analysis English Language Arts Grade 3

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00279 (EL308822)	2	0.78000	0.77000	ON03	ON03	-0.00234	False
IA00280 (EL308824)	2	0.66000	0.63000	ON03	ON03	-0.02647	False
IA00281 (EL308826)	2	0.59000	0.55000	ON03	ON03	-0.00984	False
IA00282 (EL308827)	2	0.65000	0.60000	ON03	ON03	-0.00363	False
IA00283 (EL308835)	2	0.57000	0.56000	ON03	ON03	-0.01414	False
IA00284 (EL308837)	2	0.66000	0.64000	ON03	ON03	-0.00915	False
IA00285 (EL308838)	2	0.77000	0.76000	ON03	ON03	-0.00697	False
IA00286 (EL308842)	2	0.41000	0.43000	ON03	ON03	-0.05944	True
IA00287 (EL308855)	4	1.12000	1.17000	ON03	ON03	0.05302	True
IA00288 (EL308857)	4	0.96000	0.95000	ON03	ON03	0.02587	False
IA00443 (EL626042844)	2	0.66000	0.66000	ON04	ON04	0.00110	False
IA00444 (EL626043062)	2	0.67000	0.65000	ON04	ON04	-0.01853	False
IA00445 (EL626043435)	3	1.03000	1.03000	ON04	ON04	-0.01783	False
IA00446 (EL626049849)	2	0.53000	0.50000	ON04	ON04	-0.00995	False
IA00450 (EL626050679)	2	0.67000	0.65000	ON04	ON04	-0.01899	False
IA00451 (EL626050927)	2	0.51000	0.54000	ON04	ON04	-0.00047	False
IA00452 (EL626051097)	3	1.10000	1.04000	ON04	ON04	-0.00422	False
IA00453 (EL626051328)	2	0.64000	0.65000	ON04	ON04	-0.00166	False

Table 2.7.2 Beta Analysis English Language Arts Grade 4

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00218 (EL307705)	2	0.81000	0.82000	ON03	ON03	-0.00096	False
IA00219 (EL307709)	2	0.77000	0.77000	ON03	ON03	0.01441	False
IA00220 (EL307710)	2	0.43000	0.45000	ON03	ON03	-0.02233	False
IA00221 (EL307713)	2	0.52000	0.53000	ON03	ON03	0.00565	False
IA00222 (EL307714)	2	0.76000	0.77000	ON03	ON03	0.01944	False
IA00223 (EL307719)	2	0.56000	0.57000	ON03	ON03	-0.00306	False
IA00224 (EL307724)	2	0.75000	0.73000	ON03	ON03	0.01072	False
IA00225 (EL307728)	4	1.36000	1.33000	ON03	ON03	-0.02506	False
IA00226 (EL307729)	4	1.25000	1.26000	ON03	ON03	0.02004	False
IA00289 (EL309792)	2	0.64000	0.62000	ON03	ON03	0.03852	False
IA00407 (EL624647403)	2	0.41000	0.42000	ON04	ON04	0.01909	False
IA00408 (EL624647580)	2	0.58000	0.54000	ON04	ON04	-0.02332	False
IA00411 (EL624652450)	2	0.78000	0.76000	ON04	ON04	0.00624	False
IA00412 (EL624652621)	2	0.90000	0.89000	ON04	ON04	-0.01074	False
IA00414 (EL624652989)	2	0.43000	0.44000	ON04	ON04	0.01915	False
IA00415 (EL624653348)	2	0.67000	0.63000	ON04	ON04	-0.02012	False
IA00416 (EL624653492)	3	1.44000	1.42000	ON04	ON04	-0.01194	False
IA00419 (EL624654711)	3	1.60000	1.59000	ON04	ON04	-0.00453	False

Table 2.7.3 Beta Analysis English Language Arts Grade 5

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00495 (EL626304658)	2	0.75000	0.75000	ON03	ON03	0.02914	False
IA00497 (EL626304969)	2	0.73000	0.67000	ON03	ON03	-0.03390	False
IA00500 (EL626332335)	2	0.66000	0.64000	ON03	ON03	-0.00521	False
IA00501 (EL626332592)	2	0.84000	0.83000	ON03	ON03	0.00343	False
IA00502 (EL626333002)	2	0.82000	0.81000	ON03	ON03	-0.00007	False
IA00505 (EL626355215)	2	0.60000	0.61000	ON03	ON03	0.00674	False
IA00506 (EL626355557)	2	0.64000	0.65000	ON03	ON03	0.01090	False
IA00508 (EL626356291)	3	0.73000	0.75000	ON03	ON03	0.00729	False
IA00638 (EL627351056)	2	0.66000	0.66000	ON03	ON03	0.00357	False
IA01669 (EL711809263)	2	0.76000	0.74000	ON04	ON04	-0.01665	False
IA01670 (EL711809592)	2	0.76000	0.74000	ON04	ON04	0.01458	False
IA01671 (EL711827203)	2	0.90000	0.89000	ON04	ON04	-0.00079	False
IA01672 (EL711827807)	2	0.73000	0.68000	ON04	ON04	-0.03029	False
IA01679 (EL711868011)	3	0.99000	0.93000	ON04	ON04	0.00261	False
IA01680 (EL711900602)	2	0.70000	0.66000	ON04	ON04	-0.00451	False
IA01691 (EL712167015)	2	0.46000	0.47000	ON04	ON04	-0.00261	False

Table 2.7.4 Beta Analysis English Language Arts Grade 6

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00173 (EL303496)	2	0.74000	0.74000	ON04	ON04	-0.00390	False
IA00174 (EL303500)	2	0.64000	0.64000	ON04	ON04	0.01126	False
IA00175 (EL303504)	2	0.61000	0.61000	ON04	ON04	-0.02877	False
IA00176 (EL303508)	2	0.71000	0.72000	ON04	ON04	-0.01843	False
IA00177 (EL303510)	2	0.87000	0.86000	ON04	ON04	-0.00286	False
IA00178 (EL303513)	2	0.66000	0.68000	ON04	ON04	0.02574	False
IA00179 (EL303514)	2	0.63000	0.61000	ON04	ON04	0.00463	False
IA00180 (EL303518)	2	0.57000	0.54000	ON04	ON04	-0.00137	False
IA00515 (EL626864414)	2	0.83000	0.82000	ON03	ON03	-0.00111	False
IA00517 (EL626864724)	2	0.69000	0.70000	ON03	ON03	-0.01741	False
IA00518 (EL626865003)	2	0.67000	0.68000	ON03	ON03	0.01210	False
IA00520 (EL626865416)	2	0.41000	0.42000	ON03	ON03	-0.01384	False
IA00522 (EL626865773)	2	0.72000	0.70000	ON03	ON03	-0.00034	False
IA00523 (EL626865942)	2	0.64000	0.65000	ON03	ON03	-0.00110	False
IA00528 (EL626867605)	3	1.54000	1.53000	ON03	ON03	0.00502	False
IA00530 (EL626868748)	3	1.41000	1.43000	ON03	ON03	0.01086	False

Table 2.7.5 Beta Analysis English Language Arts Grade 7

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00065 (EL292160)	2	0.73000	0.71000	ON04	ON04	0.01645	False
IA00066 (EL292163)	2	0.65000	0.59000	ON04	ON04	-0.01702	False
IA00067 (EL292168)	2	0.50000	0.48000	ON04	ON04	0.00518	False
IA00068 (EL292170)	2	0.71000	0.69000	ON04	ON04	0.00013	False
IA00069 (EL292172)	2	0.62000	0.62000	ON04	ON04	-0.01004	False
IA00070 (EL292176)	2	0.62000	0.61000	ON04	ON04	0.01622	False
IA00081 (EL293802)	2	0.68000	0.67000	ON04	ON04	-0.01859	False
IA00082 (EL293804)	2	0.54000	0.54000	ON04	ON04	-0.02578	False
IA00257 (EL308358)	2	0.85000	0.82000	ON03	ON03	-0.01203	False
IA00258 (EL308360)	3	1.51000	1.49000	ON03	ON03	0.01036	False
IA00262 (EL308382)	2	0.65000	0.63000	ON03	ON03	-0.00855	False
IA00265 (EL308389)	2	0.90000	0.90000	ON03	ON03	0.00953	False
IA00269 (EL308397)	2	0.84000	0.80000	ON03	ON03	-0.00594	False
IA00655 (EL628647210)	2	0.72000	0.72000	ON03	ON03	0.00039	False
IA00657 (EL628647689)	2	0.77000	0.75000	ON03	ON03	0.00003	False
IA00658 (EL628653398)	3	1.48000	1.51000	ON03	ON03	0.01254	False

Table 2.7.6 Beta Analysis English Language Arts Grade 8

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00056 (EL290795)	2	0.77000	0.76000	ON04	ON04	-0.03654	False
IA00057 (EL290798)	2	0.78000	0.79000	ON04	ON04	-0.00494	False
IA00058 (EL290799)	2	0.75000	0.74000	ON04	ON04	-0.01275	False
IA00059 (EL290800)	2	0.67000	0.68000	ON04	ON04	0.01366	False
IA00060 (EL290801)	2	0.77000	0.76000	ON04	ON04	0.00675	False
IA00061 (EL290805)	2	0.56000	0.56000	ON04	ON04	-0.00376	False
IA00062 (EL290808)	2	0.54000	0.54000	ON04	ON04	0.00533	False
IA00063 (EL290814)	2	0.44000	0.39000	ON04	ON04	0.00397	False
IA00368 (EL623873883)	2	0.72000	0.73000	ON03	ON03	0.00027	False
IA00371 (EL623951471)	3	1.21000	1.23000	ON03	ON03	0.02581	False
IA00373 (EL623952377)	3	0.84000	0.87000	ON03	ON03	0.00700	False
IA00374 (EL623952612)	2	0.73000	0.71000	ON03	ON03	-0.00966	False
IA00378 (EL623955555)	2	0.51000	0.50000	ON03	ON03	-0.05383	True
IA00379 (EL623955757)	2	0.54000	0.54000	ON03	ON03	0.03611	False
IA00383 (EL623959265)	2	0.65000	0.63000	ON03	ON03	-0.00545	False
IA00699 (EL632808123)	2	0.77000	0.76000	ON03	ON03	-0.00304	False

Table 2.7.7 Beta Analysis Mathematics Grade 3

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00799 (MA260559)	2	0.49000	0.53000	ON05	ON16	0.00914	False
IA00834 (MA293457)	2	0.81000	0.80000	ON03	ON09	-0.00918	False
IA00838 (MA293524)	2	0.77000	0.79000	ON07	ON14	0.00626	False
IA00850 (MA297405)	2	0.71000	0.71000	ON09	ON17	-0.00471	False
IA00852 (MA297438)	2	0.65000	0.65000	ON06	ON13	-0.02739	False
IA00924 (MA306310)	2	0.48000	0.49000	ON14	ON05	0.00299	False
IA00925 (MA306315)	2	0.79000	0.78000	ON02	ON08	-0.03302	False
IA00930 (MA306359)	2	0.64000	0.68000	ON04	ON03	0.02171	False
IA00932 (MA306375)	2	0.49000	0.51000	ON13	ON12	0.00539	False
IA01019 (MA311277)	2	0.74000	0.81000	ON17	ON14	0.06082	True
IA01081 (MA623656013)	4	0.96000	0.96000	ON08	ON02	0.00155	False
IA02323 (MA301611A)	2	0.87000	0.87000	ON15	ON12	-0.00097	False
IA04659 (MA307303)	2	0.46000	0.43000	ON01	ON03	-0.01005	False
IA04760 (MA713752330)	2	0.80000	0.80000	ON11	ON07	-0.01223	False
IA04813 (MA735572247)	2	0.69000	0.72000	ON12	ON10	0.00638	False
IA04828 (MA735653938)	2	0.57000	0.59000	ON10	ON06	0.01815	False
IA04844 (MA735735757)	2	0.61000	0.62000	ON20	ON15	-0.01278	False
IA07855 (MA900579464)	2	0.51000	0.47000		ON04	-0.10121	True
IA10020 (MA001639117)	4	1.66000	1.76000	ON04	ON11	0.01626	False
IA12584 (MA900373094)	2	0.58000	0.60000	ON19	ON10	0.01051	False

Table 2.7.8 Beta Analysis Mathematics Grade 4

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00828 (MA287237)	2	0.79000	0.81000	ON06	ON11	0.02432	False
IA00841 (MA293718)	2	0.73000	0.74000	ON08	ON09	-0.00867	False
IA00861 (MA297629)	2	0.90000	0.87000	ON21	ON04	-0.03357	False
IA00906 (MA301811)	2	0.72000	0.74000	ON09	ON08	0.00739	False
IA00958 (MA307055)	2	0.51000	0.50000	ON10	ON05	-0.03013	False
IA00961 (MA307081)	2	0.61000	0.56000	ON13	ON09	-0.08024	True
IA00963 (MA307085)	2	0.70000	0.69000	ON15	ON12	-0.03371	False
IA01049 (MA311537)	2	0.67000	0.68000	ON18	ON13	-0.00773	False
IA01055 (MA311572)	2	0.55000	0.61000	ON19	ON14	0.05179	True
IA01057 (MA311581)	5	2.23000	2.25000	ON11	ON02	-0.02112	False
IA01093 (MA623879088)	3	1.39000	1.49000	ON14	ON03	0.02958	False
IA02175 (MA286769)	2	0.78000	0.78000	ON03	ON06	-0.01693	False
IA02819 (MA713583365)	2	0.59000	0.68000	ON17	ON06	0.07308	True
IA02841 (MA713774890)	2	0.50000	0.55000	ON05	ON17	0.03527	False
IA02902 (MA714251321)	2	0.41000	0.45000	ON04	ON16	0.01798	False

continued

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA04945 (MA800744502)	2	0.66000	0.63000	ON16	ON04	-0.02417	False
IA04965 (MA800867144)	2	0.62000	0.63000	ON02	ON15	-0.00076	False
IA12273 (MA200267672)	2	0.56000	0.56000	ON23	ON05	-0.00316	False
IA12278 (MA200334941)	2	0.51000	0.52000	ON21	ON07	-0.00551	False
IA12525 (MA713956124)	5	1.87000	1.97000	ON06	ON10	0.01946	False

Table 2.7.9  
Beta Analysis  
Mathematics Grade 5

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00771 (MA204911)	2	0.75000	0.78000	ON08	ON05	0.01051	False
IA00776 (MA221207)	2	0.69000	0.67000	ON20	ON05	-0.02113	False
IA00803 (MA262207)	2	0.75000	0.73000	ON05	ON06	-0.01838	False
IA00806 (MA272292)	2	0.47000	0.48000	ON23	ON09	-0.00323	False
IA00880 (MA298106)	2	0.24000	0.24000	ON04	ON14	0.00491	False
IA00885 (MA299556)	2	0.73000	0.75000	ON18	ON04	0.00818	False
IA00943 (MA306466)	2	0.64000	0.68000	ON11	ON06	0.04106	False
IA00989 (MA307638)	2	0.87000	0.85000	ON12	ON07	-0.01728	False
IA01020 (MA311280)	2	0.43000	0.45000	ON13	ON17	0.02720	False
IA01029 (MA311337)	2	0.86000	0.85000	ON17	ON07	-0.00029	False
IA01149 (MA624347774)	2	0.42000	0.43000	ON16	ON08	0.00314	False
IA01155 (MA624357395)	3	0.83000	0.75000	ON03	ON03	-0.03374	False
IA01159 (MA624377498)	5	2.35000	2.30000	ON05	ON02	0.00487	False
IA02552 (MA311324)	2	0.44000	0.43000	ON14	ON13	-0.00304	False
IA02725 (MA704359215)	2	0.64000	0.57000	12	ON12	-0.02158	False
IA02736 (MA704359678)	5	1.99000	2.10000	ON02	ON10	0.02580	False
IA02917 (MA715102107)	2	0.88000	0.85000	AT01	ON15	-0.06821	True
IA04970 (MA800974344)	2	0.73000	0.74000	ON10	ON08	0.00162	False
IA05002 (MA801652356)	2	0.58000	0.56000	ON15	ON11	-0.01971	False
IA05015 (MA801763240)	2	0.26000	0.26000	ON24	ON16	0.01046	False

Table 2.7.10  
Beta Analysis  
Mathematics Grade 6

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00778 (MA221669)	2	0.79000	0.81000	ON11	ON11	0.00222	False
IA00804 (MA264305)	2	0.75000	0.77000	ON17	ON17	0.01413	False
IA00817 (MA280989)	2	0.49000	0.49000	ON14	ON14	-0.00530	False
IA00818 (MA282268)	2	0.41000	0.41000	ON05	ON05	-0.01181	False
IA00819 (MA282277)	2	0.44000	0.45000	ON03	ON11	-0.00553	False
IA00827 (MA287186)	2	0.55000	0.58000	ON09	ON09	0.01689	False
IA00845 (MA296349)	2	0.53000	0.57000	ON16	ON16	0.02896	False
IA00972 (MA307339)	5	2.92000	2.86000	ON02	ON02	-0.01558	False
IA01058 (MA311658)	2	0.46000	0.44000	ON13	ON13	-0.03493	False
IA02037 (MA217493)	2	0.63000	0.68000	ON03	ON14	0.04206	False
IA02078 (MA251350)	5	1.44000	1.45000	ON07	ON03	-0.00698	False
IA02597 (MA311693)	2	0.77000	0.80000	ON04	ON04	0.01191	False
IA02698 (MA703179529)	2	0.24000	0.23000	ON15	ON15	-0.02655	False
IA02906 (MA714281467)	2	0.67000	0.64000	05	ON10	-0.01864	False
IA04745 (MA703231515)	3	1.05000	1.07000	ON03	ON08	-0.00513	False
IA04884 (MA736365836)	2	0.68000	0.67000	ON12	ON12	-0.02324	False
IA04899 (MA736511626)	2	0.54000	0.60000	ON24	ON10	0.10425	True
IA05126 (MA805103779)	2	0.56000	0.57000	ON07	ON07	-0.00464	False
IA05134 (MA805167086)	2	0.32000	0.31000	ON03	ON06	0.01770	False
IA05135 (MA805171807)	2	0.58000	0.59000	ON08	ON08	-0.01561	False



Table 2.7.11  
Beta Analysis  
Mathematics Grade 7

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00831 (MA288414)	2	0.73000	0.72000	ON18	ON07	-0.01428	False
IA00847 (MA296358)	2	0.53000	0.53000	ON17	ON09	0.00196	False
IA00909 (MA301846)	2	0.87000	0.88000	ON13	ON13	0.00225	False
IA00948 (MA306600)	2	0.85000	0.84000	ON11	ON11	-0.00411	False
IA00949 (MA306605)	2	0.50000	0.51000	ON05	ON05	-0.00157	False
IA01006 (MA311093)	2	0.86000	0.86000	ON20	ON05	-0.00194	False
IA01011 (MA311109)	2	0.41000	0.47000	ON12	ON12	0.04981	False
IA01018 (MA311140)	2	0.42000	0.38000	ON10	ON10	-0.03872	False
IA01069 (MA316886)	5	2.07000	2.08000	ON03	ON03	0.00738	False
IA01097 (MA623950280)	2	0.35000	0.35000	ON06	ON06	-0.01005	False
IA02722 (MA703943185)	5	1.94000	1.87000	ON02	ON02	-0.00381	False
IA04486 (MA227988)	2	0.68000	0.69000	ON09	ON09	-0.01035	False
IA04538 (MA282218)	2	0.66000	0.62000	ON16	ON08	-0.05730	True
IA04591 (MA298183)	2	0.50000	0.52000	AT01	ON03	-0.01323	False
IA04593 (MA298192)	2	0.19000	0.19000	ON15	ON10	0.00463	False
IA04629 (MA303730)	2	0.51000	0.50000	102	ON07	0.00782	False
IA07614 (MA306636)	2	0.43000	0.47000	ON13	ON08	0.01133	False
IA08208 (MA904265644)	3	0.81000	0.85000	ON07	ON05	0.00558	False
IA10259 (MA005177715)	2	0.36000	0.40000	ON01	ON04	0.03175	False
IA10268 (MA005205470)	2	0.34000	0.36000	ON02	ON14	0.01460	False

Table 2.7.12  
Beta Analysis  
Mathematics Grade 8

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00849 (MA296757)	2	0.62000	0.64000	ON17	ON06	0.00725	False
IA00858 (MA297513)	2	0.78000	0.78000	ON04	ON04	-0.01339	False
IA00864 (MA297652)	5	2.77000	2.81000	ON03	ON03	-0.01109	False
IA00865 (MA297656)	2	0.60000	0.56000	ON11	ON13	-0.05188	True
IA00905 (MA301702)	2	0.46000	0.52000	ON07	ON07	0.05845	True
IA00979 (MA307472)	2	0.65000	0.65000	ON13	ON10	-0.01066	False
IA00985 (MA307570)	2	0.53000	0.54000	ON18	ON05	0.01287	False
IA01033 (MA311384)	2	0.57000	0.59000	ON08	ON08	0.00084	False
IA01037 (MA311414)	2	0.39000	0.41000	ON10	ON10	-0.00568	False
IA01042 (MA311448)	2	0.50000	0.51000	ON05	ON13	0.00860	False
IA01125 (MA624247061)	3	0.83000	0.85000	ON14	ON14	0.00120	False
IA02495 (MA309741)	2	0.37000	0.45000	ON15	ON08	0.05828	True
IA04665 (MA307399)	2	0.47000	0.57000	ON16	ON06	0.10198	True
IA04678 (MA309738)	2	0.40000	0.39000	ON04	ON04	-0.02082	False
IA04719 (MA311462)	2	0.59000	0.61000	ON15	ON14	0.05473	True
IA04957 (MA800770988)	2	0.46000	0.39000	ON16	ON11	-0.01440	False
IA05059 (MA803856627)	2	0.71000	0.67000	ON09	ON09	-0.04759	False
IA05070 (MA804042487)	2	0.40000	0.33000	ON04	ON07	-0.08373	True
IA10142 (MA002243189)	2	0.79000	0.84000	ON14	ON12	0.02010	False
IA12109 (MA002234564)	5	1.79000	1.70000	ON05	ON02	-0.00517	False

# Section 2.8

## Final Item Parameters

Table 2.8.1  
IRT Parameters for Dichotomous Items  
English Language Arts Grade 3

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA08981 (EL006852708)	0.53228	0.01523	-0.58040	0.06492	0.14660	0.02070
IA08983 (EL006853420)	1.03798	0.02001	0.25930	0.01354	0.16180	0.00480
IA08984 (EL006853709)	0.87225	0.02348	0.78450	0.01724	0.24590	0.00510
IA09111 (EL007171237)	1.26520	0.02915	0.26390	0.01354	0.27590	0.00470
IA09112 (EL007171625)	1.04968	0.02826	0.78960	0.01489	0.26480	0.00430
IA09128 (EL007253254)	0.93269	0.01723	-0.87790	0.02720	0.16180	0.01230
IA11432 (EL206247438)	0.75285	0.02479	1.23900	0.02015	0.20070	0.00510
IA11436 (EL206252440)	0.75385	0.01970	0.53850	0.02048	0.19730	0.00670
IA11437 (EL206252891)	0.86473	0.02022	0.26290	0.01936	0.23050	0.00660
IA11441 (EL206440619)	0.76696	0.01822	-0.69940	0.03985	0.24440	0.01450
IA11445 (EL206538101)	0.87496	0.01623	-1.32390	0.03727	0.10320	0.01930
IA11446 (EL206538400)	0.85426	0.02069	-0.24550	0.02787	0.28850	0.00940
IA11450 (EL206539824)	0.69935	0.01681	0.71580	0.01780	0.10560	0.00580
IA11454 (EL206632601)	1.26661	0.02232	-0.75070	0.01746	0.19090	0.00830
IA11458 (EL206633507)	0.90247	0.01486	-0.92280	0.02496	0.08140	0.01190
IA11462 (EL206635137)	0.67737	0.01455	-1.37170	0.05709	0.08300	0.02580
IA11463 (EL206635666)	0.82516	0.01707	-0.56330	0.02933	0.18930	0.01150
IA11464 (EL206635928)	1.07219	0.02085	-0.96600	0.02619	0.21580	0.01230
IA11466 (EL206636183)	0.85291	0.01639	-1.38140	0.04086	0.10170	0.02120
IA11467 (EL206636821)	0.84274	0.01471	-1.36920	0.03716	0.07100	0.01950
IA11468 (EL206636985)	1.17678	0.02180	-0.79590	0.02048	0.21130	0.00950
IA11470 (EL206638372)	0.58325	0.01455	-0.70310	0.05317	0.08990	0.01930
IA11493 (EL206658239)	0.57196	0.01250	-1.16580	0.06044	0.04520	0.02410
IA11494 (EL206658538)	0.50811	0.01754	0.04510	0.05731	0.18740	0.01640

Table 2.8.2  
IRT Parameters for Polytomous Items  
English Language Arts Grade 3

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA09201 (EL008648215)	0.75326	0.00678	-0.24445	0.00756	0.87825	0.01314	-0.87825	0.01305	0.00000	0.00000
IA11433 (EL206248298)	0.51170	0.00572	-0.44860	0.00880	0.46210	0.01587	-0.46210	0.01459	0.00000	0.00000
IA11442 (EL206455922)	0.71969	0.00662	0.94663	0.01201	2.50153	0.01744	-0.42317	0.01957	-2.07837	0.03231
IA11465 (EL206636010)	0.40676	0.00530	-1.01395	0.01340	0.87985	0.02667	-0.87985	0.01913	0.00000	0.00000
IA11473 (EL206639826)	0.80000	0.00814	-2.19575	0.01383	0.90665	0.02883	-0.90665	0.01779	0.00000	0.00000
IA11475 (EL206640503)	0.42634	0.00520	-0.29825	0.01132	0.83535	0.01997	-0.83535	0.01924	0.00000	0.00000
IA11481A (EL206642581#SCORE_TRAIT_Conv)	1.05062	0.01123	0.56220	0.00890	1.59300	0.01384	0.13510	0.01280	-1.72810	0.02438
IA11481D (EL206642581#SCORE_TRAIT_Ideadev)	0.92210	0.00999	1.38555	0.02193	2.84245	0.02575	1.01075	0.02402	-0.77385	0.03120

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA11442 (EL206455922)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11481A (EL206642581#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11481D (EL206642581#SCORE_TRAIT_Ideadev)	-3.07935	0.08604	0.00000	0.00000	n/a	n/a

Table 2.8.3  
IRT Parameters for Dichotomous Items  
English Language Arts Grade 4

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA06411 (EL297747)	0.87478	0.02267	-0.01150	0.02360	0.31530	0.00780
IA06412 (EL297749)	1.04868	0.01908	-0.92000	0.02296	0.17420	0.01080
IA06416 (EL297754)	0.68930	0.01913	-0.13450	0.03359	0.22070	0.01130
IA06817 (EL909139013)	0.53786	0.01797	-1.50550	0.11417	0.19180	0.03910
IA06819 (EL909146543)	0.94250	0.02356	0.47270	0.01563	0.22450	0.00530
IA06820 (EL909147325)	0.73580	0.02135	0.30550	0.02519	0.26300	0.00810
IA06824 (EL909150609)	0.89924	0.02129	-0.59940	0.02977	0.30110	0.01090
IA06825 (EL909151025)	0.77443	0.02041	0.59020	0.01818	0.18820	0.00600
IA06832 (EL909155188)	0.26549	0.00979	-2.02590	0.23525	0.06310	0.05050
IA06833 (EL909156595)	0.85791	0.02190	0.05560	0.02201	0.26090	0.00770
IA06944 (EL909952647)	0.89636	0.02372	0.44540	0.01754	0.25500	0.00580
IA06946 (EL909966060)	0.68095	0.01792	-1.06460	0.05698	0.19700	0.02160
IA07151 (EL911905412)	1.43292	0.02854	-1.78470	0.02977	0.16410	0.02200
IA07157 (EL911909187)	0.72640	0.03174	1.16510	0.02530	0.31780	0.00630
IA07159 (EL911910431)	0.37972	0.01211	-0.52400	0.08887	0.02360	0.02500
IA07160 (EL911910990)	0.54497	0.00785	-1.81680	0.04050	0.00620	0.01660
IA07162 (EL911919086)	1.09347	0.02002	-1.33660	0.02828	0.16540	0.01540
IA07264 (EL913042685)	0.58454	0.01681	-0.31370	0.04709	0.17520	0.01520
IA07282 (EL913366686)	1.03086	0.02074	-1.39710	0.03434	0.18690	0.01820
IA07305 (EL915149134)	0.84603	0.01637	-1.02150	0.03168	0.11580	0.01450
IA07307 (EL915167319)	0.89782	0.01759	-0.48050	0.02190	0.15120	0.00910
IA07312 (EL915543104)	1.19224	0.02278	-0.06190	0.01318	0.20550	0.00520
IA07313 (EL915543753)	1.19306	0.02041	-0.61040	0.01520	0.13940	0.00700
IA07315 (EL915545933)	0.73839	0.02079	0.35040	0.02307	0.22100	0.00780

Table 2.8.4  
IRT Parameters for Polytomous Items  
English Language Arts Grade 4

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA06413 (EL297750)	0.83539	0.00763	-0.77450	0.00682	0.68670	0.01272	-0.68670	0.01082	0.00000	0.00000
IA06815A (EL909132428#SCORE_TRAIT_Conv)	0.94533	0.01011	-0.02497	0.00836	1.80053	0.01682	-0.11337	0.01228	-1.68717	0.02012
IA06815D (EL909132428#SCORE_TRAIT_Ideadev)	0.73939	0.00882	0.92535	0.01196	2.05275	0.01766	0.72555	0.01666	-0.65255	0.02382
IA06827 (EL909152457)	0.50723	0.00575	-0.83135	0.00997	0.64335	0.01917	-0.64335	0.01512	0.00000	0.00000
IA06829 (EL909153399)	0.54145	0.00597	-0.66300	0.00862	0.50570	0.01593	-0.50570	0.01385	0.00000	0.00000
IA07144 (EL911875859)	0.86085	0.00724	0.29783	0.00792	1.65523	0.01300	0.12583	0.01178	-1.78107	0.02110
IA07153 (EL911907107)	0.98154	0.00924	-1.39055	0.00814	0.77205	0.01640	-0.77205	0.01132	0.00000	0.00000
IA07308 (EL915169315)	0.67637	0.00697	-0.28615	0.00677	0.29885	0.01163	-0.29885	0.01180	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA06815A (EL909132428#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA06815D (EL909132428#SCORE_TRAIT_Ideadev)	-2.12575	0.04127	0.00000	0.00000	n/a	n/a
IA07144 (EL911875859)	0.00000	0.00000	n/a	n/a	n/a	n/a

Table 2.8.5  
IRT Parameters for Dichotomous Items  
English Language Arts Grade 5

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA11380 (EL206053008)	1.30941	0.02493	-1.68960	0.02750	0.19430	0.01730
IA11388 (EL206060317)	1.01046	0.01857	-0.05040	0.01479	0.15910	0.00550
IA11390 (EL206068188)	0.53039	0.01491	-1.43680	0.09312	0.15790	0.03210
IA11397 (EL206085599)	1.55461	0.03119	-1.99180	0.02738	0.12450	0.02180
IA11398 (EL206100724)	1.01699	0.01913	-0.98910	0.02542	0.22860	0.01120
IA11400 (EL206127119)	0.58148	0.01242	-1.63800	0.06955	0.06570	0.02920
IA11401 (EL206130727)	1.00782	0.01842	-1.92920	0.03917	0.08490	0.02540
IA11420 (EL206142184)	0.78701	0.01588	-1.13310	0.03801	0.16390	0.01620
IA11574 (EL207033428)	0.83081	0.01516	-2.08170	0.05130	0.05500	0.03140
IA11576 (EL207035725)	0.38366	0.00779	-1.47710	0.07649	0.01310	0.02290
IA11578 (EL207037283)	0.80071	0.01771	-0.53170	0.03039	0.25470	0.01070
IA11579 (EL207039028)	0.63263	0.01364	-0.58820	0.03801	0.12400	0.01350
IA11581 (EL207041852)	0.50647	0.01618	0.33670	0.04252	0.16700	0.01240
IA11585 (EL207128330)	0.71393	0.01465	-0.36650	0.02750	0.12610	0.01010
IA11587 (EL207129544)	1.06914	0.01979	-0.66710	0.01953	0.20810	0.00830
IA11654 (EL207984109)	1.00964	0.01755	-0.79280	0.02080	0.16410	0.00920
IA11671 (EL208106818)	0.33586	0.00799	-1.55850	0.10340	0.01850	0.02770
IA11741 (EL208451664)	0.95708	0.02030	0.22960	0.01606	0.22510	0.00540
IA11744 (EL208452118)	0.97684	0.01465	-1.23390	0.02195	0.04480	0.01130
IA11748 (EL208452922)	0.95885	0.01913	-0.45520	0.02230	0.27580	0.00820
IA11760 (EL208571752)	0.81587	0.01308	-1.00940	0.02588	0.05190	0.01180
IA11761 (EL208575656)	0.73616	0.01348	-0.68690	0.02842	0.10020	0.01120
IA11762 (EL208577435)	1.11581	0.02712	0.51510	0.01398	0.26510	0.00430
IA11911 (EL209442821)	0.22416	0.00600	-1.99270	0.14904	0.01600	0.02750



Table 2.8.6  
IRT Parameters for Polytomous Items  
English Language Arts Grade 5

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA11386A (EL206057533#SCORE_TRAIT_Conv)	0.99641	0.01035	-0.27443	0.00739	1.51887	0.01592	0.02327	0.01157	-1.54213	0.01636
IA11386D (EL206057533#SCORE_TRAIT_Ideadev)	1.08471	0.01105	0.52630	0.00738	1.75240	0.01284	0.64210	0.01126	-0.59500	0.01452
IA11583 (EL207043067)	0.73010	0.00677	-0.65900	0.00713	0.42710	0.01299	-0.42710	0.01167	0.00000	0.00000
IA11586 (EL207128708)	0.74203	0.00667	-1.09670	0.00848	0.81440	0.01675	-0.81440	0.01229	0.00000	0.00000
IA11589A (EL207135604#SCORE_TRAIT_Conv)	0.96185	0.00988	-0.52477	0.00737	1.46903	0.01716	-0.01967	0.01179	-1.44937	0.01477
IA11589D (EL207135604#SCORE_TRAIT_Ideadev)	1.01499	0.00999	0.13243	0.00824	2.26713	0.01800	0.71633	0.01206	-0.66418	0.01398
IA11715 (EL208382001)	0.79153	0.00723	-0.03640	0.00634	0.34050	0.01064	-0.34050	0.01130	0.00000	0.00000
IA11904 (EL209368842)	0.85456	0.00824	-1.29295	0.00780	0.27715	0.01442	-0.27715	0.01254	0.00000	0.00000
IA11915 (EL209449008)	1.01611	0.00896	-1.20315	0.00696	0.52595	0.01329	-0.52595	0.01067	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA11386A (EL206057533#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11386D (EL206057533#SCORE_TRAIT_Ideadev)	-1.79950	0.02425	0.00000	0.00000	n/a	n/a
IA11589A (EL207135604#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11589D (EL207135604#SCORE_TRAIT_Ideadev)	-2.31928	0.02632	0.00000	0.00000	n/a	n/a

Table 2.8.7  
IRT Parameters for Dichotomous Items  
English Language Arts Grade 6

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA09533 (EL027853702)	0.78912	0.01581	-1.69330	0.04228	0.21000	0.01630
IA09535 (EL027973437)	0.79982	0.01645	0.37640	0.01780	0.17970	0.00480
IA09537 (EL028044247)	0.85744	0.01653	-1.07320	0.02922	0.24590	0.01000
IA09544 (EL028171110)	0.80112	0.01668	-1.24570	0.03768	0.31430	0.01200
IA09551 (EL028261263)	0.38789	0.00828	-1.96310	0.09568	0.03030	0.02820
IA09552 (EL028278272)	0.73668	0.01415	-0.57450	0.02789	0.21730	0.00840
IA09553 (EL028283260)	0.99641	0.01649	-1.45270	0.02329	0.11750	0.00990
IA09577 (EL028464862)	0.42975	0.01201	0.06380	0.05236	0.13900	0.01280
IA09579 (EL028474800)	0.37190	0.01050	1.17260	0.03308	0.02990	0.00830
IA11383 (EL206054907)	0.70276	0.01490	-0.89220	0.03694	0.25850	0.01110
IA11385 (EL206055338)	0.46402	0.01015	-2.56000	0.10577	0.06470	0.03820
IA11391 (EL206069631)	0.85267	0.01561	-2.05220	0.03886	0.07510	0.01880
IA11392 (EL206071655)	0.69289	0.01621	0.85000	0.01928	0.17390	0.00470
IA11393 (EL206072413)	0.44103	0.00709	-2.94850	0.07565	0.01810	0.02760
IA11394 (EL206073117)	0.50106	0.01181	-0.45820	0.04836	0.17780	0.01260
IA11402 (EL206131013)	0.54903	0.01248	-0.89550	0.05014	0.16830	0.01450
IA11430 (EL206156401)	0.57989	0.01090	-1.15720	0.04317	0.10960	0.01400
IA11698 (EL208346712)	0.42275	0.01133	-0.69550	0.07239	0.13980	0.01810
IA11710 (EL208352214)	0.49483	0.01066	-2.34200	0.09108	0.06570	0.03340
IA11717 (EL208431592)	0.43292	0.01264	-0.16200	0.06112	0.17730	0.01450
IA11894 (EL208956821)	0.70529	0.01946	0.71920	0.02344	0.28890	0.00560
IA11976 (EL211482852)	0.37731	0.00820	-1.64390	0.08945	0.02570	0.02490
IA12010 (EL215337440)	0.57537	0.00919	-1.78180	0.04465	0.01390	0.01700
IA12029 (EL216269224)	0.70600	0.01288	-1.18170	0.03456	0.15450	0.01200

Table 2.8.8  
IRT Parameters for Polytomous Items  
English Language Arts Grade 6

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA09555 (EL028337921)	0.63786	0.00591	-1.57610	0.01070	0.66650	0.02077	-0.66650	0.01598	0.00000	0.00000
IA09556 (EL028340179)	0.53710	0.00519	-1.19620	0.01028	0.58920	0.01941	-0.58920	0.01604	0.00000	0.00000
IA09581A (EL028556260#SCORE_TRAIT_Conv)	0.83557	0.00852	-0.10617	0.00779	1.64883	0.01611	-0.30177	0.01364	-1.34707	0.01681
IA09581D (EL028556260#SCORE_TRAIT_Ideadev)	0.82934	0.00905	0.93488	0.00725	1.27408	0.01268	0.86338	0.01318	0.19978	0.01445
IA11426A (EL206144523#SCORE_TRAIT_Conv)	0.84891	0.00870	-0.94967	0.00826	1.61293	0.02081	-0.20407	0.01360	-1.40887	0.01416
IA11426D (EL206144523#SCORE_TRAIT_Ideadev)	0.83768	0.00794	0.71774	0.01062	2.95334	0.02005	1.39864	0.01529	0.19444	0.01558
IA11429 (EL206147428)	0.22616	0.00380	-0.80205	0.01755	0.32565	0.03158	-0.32565	0.02916	0.00000	0.00000
IA11919 (EL209683892)	0.39659	0.00420	0.72805	0.01539	1.62725	0.02190	-1.62725	0.03069	0.00000	0.00000
IA12027 (EL216263506)	0.34815	0.00400	-0.68500	0.01700	1.78070	0.03166	-1.78070	0.02707	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA09581A (EL028556260#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09581D (EL028556260#SCORE_TRAIT_Ideadev)	-0.61542	0.01793	-1.72182	0.02670	0.00000	0.00000
IA11426A (EL206144523#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11426D (EL206144523#SCORE_TRAIT_Ideadev)	-1.30486	0.02159	-3.24156	0.04517	0.00000	0.00000

Table 2.8.9  
IRT Parameters for Dichotomous Items  
English Language Arts Grade 7

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA11370 (EL206038392)	0.40194	0.01019	-2.41840	0.13545	0.06840	0.04360
IA11373 (EL206040879)	0.58260	0.01165	-1.23610	0.04759	0.06210	0.01740
IA11374 (EL206046704)	0.66725	0.01426	-0.55520	0.03280	0.16780	0.01080
IA11376 (EL206048853)	0.85291	0.01380	-2.79770	0.04862	0.03120	0.03210
IA11377 (EL206049453)	0.67607	0.01650	-0.83600	0.04373	0.28250	0.01330
IA11381 (EL206054058)	0.35379	0.01554	0.49810	0.08232	0.17860	0.01830
IA11382 (EL206054673)	0.77831	0.01549	-0.74530	0.02881	0.20190	0.01020
IA11395 (EL206073551)	0.76608	0.01202	-2.41460	0.04489	0.02210	0.02520
IA11501 (EL206735123)	0.85532	0.01499	-1.13490	0.02676	0.13110	0.01110
IA11633 (EL207653069)	0.67443	0.01938	0.95300	0.02110	0.19680	0.00550
IA11661 (EL208035502)	0.70329	0.01554	-0.66820	0.03499	0.24200	0.01110
IA11665 (EL208037687)	0.66214	0.01453	-0.61900	0.03447	0.15910	0.01160
IA11666 (EL208038191)	0.45079	0.01586	0.50090	0.04798	0.18980	0.01230
IA11668 (EL208038966)	0.92446	0.02166	0.26950	0.01737	0.25120	0.00520
IA11733 (EL208449230)	1.22528	0.02111	-0.55170	0.01389	0.16020	0.00550
IA11734 (EL208449595)	0.52469	0.01202	-0.97540	0.05441	0.09140	0.01760
IA11740 (EL208451648)	0.78783	0.01632	0.24120	0.01788	0.16020	0.00550
IA11747 (EL208452611)	0.63510	0.01357	-0.39850	0.03113	0.12690	0.01020
IA11751 (EL208475824)	0.54727	0.00823	-2.70270	0.05853	0.01690	0.02630
IA11752 (EL208476152)	0.79436	0.01540	-0.44080	0.02341	0.18350	0.00800
IA11763 (EL208579240)	0.74374	0.01650	0.08230	0.02212	0.20060	0.00680
IA11831 (EL208736539)	1.05808	0.01805	-0.72200	0.01749	0.17570	0.00690
IA11988 (EL213027859)	0.79359	0.01618	-0.31330	0.02315	0.19840	0.00770
IA11993 (EL213654472)	0.50441	0.01079	-1.60540	0.07075	0.06530	0.02490

Table 2.8.10  
IRT Parameters for Polytomous Items  
English Language Arts Grade 7

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA11419 (EL206141630)	0.31082	0.00439	1.71835	0.02856	1.83165	0.03323	-1.83165	0.06157	0.00000	0.00000
IA11424A (EL206144041#SCORE_TRAIT_Conv)	0.96179	0.01017	-1.37690	0.00845	1.41720	0.02199	-0.03300	0.01435	-1.38420	0.01294
IA11424D (EL206144041#SCORE_TRAIT_Ideadev)	0.91517	0.00888	-0.14630	0.00742	2.50040	0.02077	1.24260	0.01419	-0.04560	0.01213
IA11427 (EL206145148)	0.44668	0.00466	-1.63865	0.01746	1.87385	0.03648	-1.87385	0.02231	0.00000	0.00000
IA11764 (EL208580556)	0.54256	0.00553	-0.60770	0.00842	0.41450	0.01495	-0.41450	0.01421	0.00000	0.00000
IA11828 (EL208734693)	0.55614	0.00535	-0.96480	0.00934	0.86600	0.01771	-0.86600	0.01448	0.00000	0.00000
IA12035 (EL217355629)	0.57601	0.00553	-0.73800	0.00943	0.98310	0.01754	-0.98310	0.01503	0.00000	0.00000
IA12036A (EL217467654#SCORE_TRAIT_Conv)	0.99483	0.01029	-1.00917	0.00719	1.17083	0.01745	0.07983	0.01281	-1.25067	0.01232
IA12036D (EL217467654#SCORE_TRAIT_Ideadev)	1.01617	0.00988	0.24658	0.00732	2.24198	0.01615	1.34008	0.01280	-0.06082	0.01215

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA11424A (EL206144041#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11424D (EL206144041#SCORE_TRAIT_Ideadev)	-1.16800	0.01532	-2.52940	0.02522	0.00000	0.00000
IA12036A (EL217467654#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA12036D (EL217467654#SCORE_TRAIT_Ideadev)	-1.10652	0.01606	-2.41472	0.02788	0.00000	0.00000

Table 2.8.11  
IRT Parameters for Dichotomous Items  
English Language Arts Grade 8

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA04155 (EL809734614)	0.67413	0.01314	-1.10900	0.04058	0.15350	0.01440
IA04156 (EL809863460)	0.70529	0.01101	-1.41850	0.03328	0.02850	0.01390
IA04179 (EL810133273)	0.56479	0.01026	-1.35890	0.04971	0.05030	0.01810
IA04182 (EL810180368)	0.61446	0.01256	-1.23040	0.04985	0.13550	0.01750
IA04198 (EL810439521)	0.61975	0.01009	-1.14860	0.03595	0.03470	0.01330
IA04199 (EL810456981)	0.58072	0.01561	0.55890	0.03033	0.23600	0.00780
IA04387 (EL812838757)	0.67178	0.01277	-0.37960	0.02823	0.13910	0.00900
IA04388 (EL812849329)	0.65103	0.01306	-1.44850	0.05112	0.14710	0.01910
IA04389 (EL812849633)	0.70218	0.01340	-1.67570	0.04999	0.13640	0.02080
IA09118 (EL007242472)	0.59782	0.01553	0.12280	0.03595	0.27500	0.00910
IA09119 (EL007242520)	0.74221	0.01235	-1.00200	0.02809	0.06760	0.01080
IA09120 (EL007242572)	0.62693	0.01344	0.26060	0.02458	0.13430	0.00720
IA11510 (EL206773007)	0.38307	0.01097	-1.69870	0.13552	0.10670	0.03710
IA11511 (EL206773521)	0.59095	0.01147	-1.84830	0.06263	0.05820	0.02570
IA11538 (EL206837355)	0.25491	0.00645	-2.19440	0.16220	0.02730	0.03400
IA11539 (EL206837795)	0.26108	0.01469	0.65760	0.16079	0.14960	0.02840
IA11540 (EL206839116)	0.57807	0.01101	-1.43150	0.05238	0.04920	0.01970
IA11541 (EL206840059)	0.63151	0.01139	-0.81480	0.03398	0.07940	0.01170
IA11542 (EL206840381)	0.53586	0.01361	-0.22780	0.04662	0.22640	0.01210
IA11545 (EL206843141)	0.37113	0.00946	-1.58260	0.11417	0.04700	0.03170
IA11546 (EL206844860)	0.66620	0.01578	0.31250	0.02668	0.25330	0.00710
IA11876 (EL208942592)	1.00794	0.02692	0.94550	0.01615	0.29040	0.00380
IA11877 (EL208943016)	0.60347	0.01545	0.01070	0.03651	0.26040	0.00960
IA11881 (EL208944699)	0.37549	0.00762	-1.66570	0.08594	0.01820	0.02470

Table 2.8.12  
IRT Parameters for Polytomous Items  
English Language Arts Grade 8

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA04154 (EL809713456)	0.80970	0.00687	-1.01065	0.00824	0.71555	0.01568	-0.71555	0.01272	0.00000	0.00000
IA04184 (EL810222585)	0.47813	0.00473	-0.13545	0.01468	2.05535	0.02431	-2.05535	0.02649	0.00000	0.00000
IA04200A (EL810463548#SCORE_TRAIT_Conv)	1.00376	0.01005	-0.73327	0.00725	1.42653	0.01748	0.02143	0.01244	-1.44797	0.01309
IA04200D (EL810463548#SCORE_TRAIT_Ideadev)	0.96537	0.00917	0.40920	0.00724	2.45750	0.01738	1.14490	0.01267	-0.01100	0.01243
IA09183 (EL007954597)	0.47349	0.00456	0.14865	0.01011	0.84215	0.01652	-0.84215	0.01846	0.00000	0.00000
IA11508A (EL206761648#SCORE_TRAIT_Conv)	0.91476	0.00952	-1.14133	0.00804	1.24107	0.01999	0.04597	0.01428	-1.28703	0.01315
IA11508D (EL206761648#SCORE_TRAIT_Ideadev)	0.92522	0.00876	0.23432	0.00767	2.54102	0.01892	1.41942	0.01399	0.01892	0.01268
IA11549 (EL206847632)	0.61640	0.00578	-1.10405	0.01011	0.85855	0.01977	-0.85855	0.01491	0.00000	0.00000
IA11882 (EL208945360)	0.64245	0.00569	-1.09200	0.00937	0.97300	0.01819	-0.97300	0.01399	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA04200A (EL810463548#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA04200D (EL810463548#SCORE_TRAIT_Ideadev)	-1.19020	0.01639	-2.40120	0.02622	0.00000	0.00000
IA11508A (EL206761648#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11508D (EL206761648#SCORE_TRAIT_Ideadev)	-1.26358	0.01631	-2.71578	0.02797	0.00000	0.00000

Table 2.8.13 IRT Parameters for Dichotomous Items Mathematics Grade 3

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA02179 (MA287143)	0.72096	0.01365	-0.27101	0.02680	0.05030	0.01080
IA04680 (MA310868)	1.24574	0.03275	-0.23931	0.02484	0.49710	0.00810
IA04808 (MA734752687)	0.80810	0.02595	0.61339	0.02800	0.34570	0.00830
IA07758 (MA834632167)	1.47166	0.03826	0.59073	0.01514	0.34930	0.00480
IA07784 (MA900360943)	0.91132	0.01845	-0.16020	0.02364	0.21240	0.00940
IA07790 (MA900368129)	0.70257	0.02083	0.12080	0.04010	0.33010	0.01190
IA07795 (MA900371163)	0.96210	0.01678	-0.65497	0.02419	0.13070	0.01140
IA07805 (MA900375580)	1.15321	0.01910	-0.90601	0.02092	0.10980	0.01130
IA07850 (MA900571833)	1.03256	0.02201	-0.57762	0.02757	0.28120	0.01170
IA08098 (MA902576979)	0.98012	0.02390	0.83697	0.01580	0.20430	0.00530
IA09901 (MA001044551)	0.83303	0.01522	0.70546	0.01362	0.05160	0.00440
IA12102 (MA002150122)	1.50943	0.03933	0.54442	0.01471	0.32670	0.00480
IA12170 (MA134776400)	0.99706	0.01770	-0.79477	0.02430	0.08520	0.01220
IA12472 (MA306330)	0.75722	0.01462	-0.69649	0.03356	0.06830	0.01480
IA12540 (MA735660441)	0.64408	0.01403	-2.00048	0.08204	0.06590	0.04220
IA02686 (MA703080328)	0.64645	0.00766	1.09628	0.01427	0.00000	0.00000
IA04759 (MA713744192)	0.93706	0.01219	-1.29771	0.01351	0.00000	0.00000
IA05017 (MA802238054)	0.68552	0.00782	-0.16206	0.01035	0.00000	0.00000
IA07823 (MA900438114)	0.79586	0.00982	-1.29008	0.01449	0.00000	0.00000
IA07824 (MA900439380)	0.78533	0.00933	-1.10714	0.01318	0.00000	0.00000
IA07852 (MA900577259)	0.68293	0.00831	-0.92693	0.01340	0.00000	0.00000
IA07892 (MA900734747)	1.03272	0.01122	-0.85763	0.01046	0.00000	0.00000
IA09891 (MA001034993)	0.88256	0.01020	-0.96365	0.01177	0.00000	0.00000
IA09915 (MA001061117)	0.58872	0.00896	-2.02587	0.02637	0.00000	0.00000
IA09960 (MA001330803)	0.97688	0.01058	-0.25914	0.00904	0.00000	0.00000
IA09982 (MA001444510)	0.59012	0.00728	-0.50145	0.01253	0.00000	0.00000
IA10178 (MA002740850)	0.58359	0.00772	0.76113	0.01329	0.00000	0.00000
IA10179 (MA002746011)	0.37100	0.00599	0.16394	0.01591	0.00000	0.00000
IA10181 (MA002753131)	0.71578	0.00826	0.30853	0.01024	0.00000	0.00000
IA12167 (MA134765618)	0.55953	0.00734	0.89809	0.01438	0.00000	0.00000
IA12168 (MA134772227)	0.85575	0.00977	-0.99318	0.01231	0.00000	0.00000
IA12169 (MA134775263)	0.58613	0.00723	-0.10235	0.01144	0.00000	0.00000
IA12172 (MA134826027)	1.00407	0.01014	0.18388	0.00850	0.00000	0.00000
IA12174 (MA134834397)	0.73559	0.00842	-0.24148	0.01035	0.00000	0.00000
IA12188 (MA134935661)	1.04195	0.01128	0.40408	0.00861	0.00000	0.00000
IA12650 (MA935135473)	0.80470	0.01214	-2.31002	0.02571	0.00000	0.00000



Table 2.8.14  
IRT Parameters for Polytomous Items  
Mathematics Grade 3

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA02227 (MA297399A)	1.06888	0.00874	-0.20651	0.00615	1.37699	0.01396	-0.00381	0.01023	-1.37318	0.01242
IA04639 (MA306360A)	1.01621	0.00879	0.70044	0.00640	1.10700	0.01090	0.01275	0.01117	-1.11975	0.01574
IA07856 (MA900579739)	1.19896	0.00966	-0.26524	0.00529	1.00175	0.01191	-0.01079	0.00946	-0.99096	0.01019
IA12616 (MA901132847)	1.13913	0.00944	0.24192	0.00548	1.02710	0.01076	0.03243	0.00966	-1.05953	0.01229

Item ID	Parameters and Measures of Standard Error			
	d3	SE(d3)	d4	SE(d4)
IA02227 (MA297399A)	0.00000	0.00000	n/a	n/a
IA04639 (MA306360A)	0.00000	0.00000	n/a	n/a
IA07856 (MA900579739)	0.00000	0.00000	n/a	n/a
IA12616 (MA901132847)	0.00000	0.00000	n/a	n/a

Table 2.8.15 IRT Parameters for Dichotomous Items Mathematics Grade 4

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA01047 (MA311529)	0.80888	0.01879	0.66134	0.02033	0.18120	0.00730
IA02574 (MA311538)	0.90010	0.01913	-0.79144	0.03731	0.24620	0.01670
IA04890 (MA736379879)	0.85870	0.01913	-0.92936	0.04349	0.23420	0.01960
IA04953 (MA800765452)	1.61109	0.03876	1.32743	0.01111	0.14570	0.00260
IA07872 (MA900669607)	1.00994	0.02025	-0.04320	0.02075	0.19650	0.00900
IA07952 (MA900782440)	0.83093	0.02581	1.44648	0.01907	0.20640	0.00520
IA08164 (MA903678056)	0.61091	0.01515	0.24950	0.03511	0.10760	0.01240
IA10108 (MA002143728)	0.94436	0.01599	0.25725	0.01509	0.08180	0.00620
IA10223 (MA003746997)	0.69826	0.01599	-1.73943	0.07661	0.11070	0.03900
IA12265 (MA200156568)	0.88961	0.01683	0.53905	0.01530	0.09170	0.00590
IA12271 (MA200257075)	1.15148	0.02272	0.15245	0.01719	0.24990	0.00720
IA12451 (MA301825)	1.05128	0.02334	-0.58531	0.03060	0.34090	0.01290
IA12499 (MA311557)	0.87009	0.01919	-0.41219	0.03343	0.26690	0.01340
IA12551 (MA800778188)	1.07776	0.02351	-0.42958	0.02693	0.31680	0.01150
IA12636 (MA903767905)	0.92804	0.01616	-0.63027	0.02473	0.07510	0.01230
IA14539 (MA258255)	0.79390	0.01733	0.13799	0.02620	0.18800	0.00990
IA14548 (MA286781)	1.05487	0.02491	0.79758	0.01593	0.25380	0.00550
IA02046 (MA221898)	0.79475	0.00898	-0.52275	0.01090	0.00000	0.00000
IA02825 (MA713673616)	0.96596	0.00965	0.30902	0.00838	0.00000	0.00000
IA04898 (MA736459765)	0.74611	0.00903	0.95320	0.01153	0.00000	0.00000
IA05067 (MA803975569)	0.92192	0.00999	0.16052	0.00870	0.00000	0.00000
IA07878 (MA900682939)	0.54483	0.00746	-1.01560	0.01782	0.00000	0.00000
IA07962 (MA900845882)	1.00124	0.01105	-0.55408	0.00964	0.00000	0.00000
IA08106 (MA903054837)	0.87328	0.00903	0.02062	0.00880	0.00000	0.00000
IA08116 (MA903352782)	0.55470	0.00712	0.47345	0.01216	0.00000	0.00000
IA10046 (MA001748629)	1.01314	0.01223	-1.19418	0.01279	0.00000	0.00000
IA10149 (MA002333692)	0.81780	0.01128	-1.60770	0.01855	0.00000	0.00000
IA10215 (MA003543807)	0.58645	0.00735	-0.13867	0.01153	0.00000	0.00000
IA10423 (MA803740205A)	0.53187	0.00746	-1.35231	0.02148	0.00000	0.00000
IA12108 (MA002231198)	1.03232	0.01100	-0.27459	0.00859	0.00000	0.00000
IA12114 (MA002334822)	0.75867	0.00864	-0.34753	0.01048	0.00000	0.00000
IA12261 (MA200150657)	0.85432	0.00926	-0.18111	0.00933	0.00000	0.00000
IA12526 (MA713963514)	1.06632	0.01100	0.09461	0.00807	0.00000	0.00000
IA12548 (MA800635117)	1.35494	0.01386	-0.04006	0.00744	0.00000	0.00000

Table 2.8.16  
IRT Parameters for Polytomous Items  
Mathematics Grade 4

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA02898 (MA714226023)	0.76771	0.00740	-0.48361	0.01029	1.26819	0.02048	-1.26819	0.01469	0.00000	0.00000
IA07590 (MA302496A)	1.16567	0.00909	0.14163	0.00538	1.72586	0.01426	0.42503	0.00970	-0.47832	0.00961
IA07912 (MA900750814)	0.88714	0.00757	-1.14157	0.00722	1.39232	0.02207	0.59514	0.01635	-0.34897	0.01247
IA12471 (MA304983A)	1.28538	0.01015	-0.00804	0.00426	0.99562	0.01073	0.31077	0.00885	-0.26864	0.00857
IA12527 (MA714257493)	0.89550	0.00841	0.12248	0.00901	1.31802	0.01599	-1.31802	0.01522	0.00000	0.00000
IA12570 (MA803833224)	1.21700	0.01100	-1.75363	0.00803	1.04110	0.02431	0.39021	0.01850	-0.18187	0.01501

Item ID	Parameters and Measures of Standard Error			
	d3	SE(d3)	d4	SE(d4)
IA07590 (MA302496A)	-1.67257	0.01378	0.00000	0.00000
IA07912 (MA900750814)	-1.63849	0.01155	0.00000	0.00000
IA12471 (MA304983A)	-1.03774	0.00978	0.00000	0.00000
IA12570 (MA803833224)	-1.24943	0.01146	0.00000	0.00000

Table 2.8.17 IRT Parameters for Dichotomous Items Mathematics Grade 5

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA02306 (MA301160)	1.79422	0.03802	0.97675	0.00960	0.17050	0.00270
IA04603 (MA301147)	1.31024	0.01800	-0.83520	0.01445	0.02440	0.00870
IA04613 (MA301831)	1.35005	0.02715	0.85406	0.01059	0.13210	0.00330
IA07865 (MA900660286)	0.70066	0.02175	0.66764	0.02909	0.23650	0.00970
IA08175 (MA903732126)	0.76251	0.01271	-0.47404	0.02563	0.01340	0.01210
IA08194 (MA904134306)	1.15601	0.02050	0.01843	0.01474	0.15670	0.00680
IA08204 (MA904247425)	1.01098	0.02026	-0.80463	0.03166	0.22230	0.01640
IA09871 (MA000929280)	1.21370	0.03179	0.20465	0.02098	0.47500	0.00660
IA12054 (MA001040409)	1.07485	0.03066	0.67733	0.01850	0.35980	0.00590
IA12297 (MA201546463)	1.26396	0.02394	0.20238	0.01346	0.21250	0.00580
IA12318 (MA201757235)	1.76428	0.03880	0.72166	0.01049	0.27190	0.00350
IA12475 (MA306459)	0.83387	0.01788	-0.42150	0.03127	0.12300	0.01460
IA12493 (MA311286)	1.15904	0.02240	-0.46454	0.02177	0.24840	0.01070
IA12494 (MA311299)	0.85466	0.01776	-1.34548	0.05225	0.14920	0.03010
IA12556 (MA801221646)	0.70797	0.01557	-0.15018	0.03156	0.05330	0.01360
IA12581 (MA900353683)	1.26342	0.02365	-0.66639	0.02197	0.26600	0.01180
IA12596 (MA900643667)	1.29117	0.02567	-0.78009	0.02355	0.26760	0.01320
IA02253 (MA298010)	0.89732	0.01509	-2.01438	0.02315	0.00000	0.00000
IA02913 (MA715102062)	0.73144	0.00832	-0.22924	0.00940	0.00000	0.00000
IA02923 (MA715102279)	1.06599	0.01063	0.75125	0.00930	0.00000	0.00000
IA02927 (MA715102381)	0.76613	0.00909	-1.04082	0.01267	0.00000	0.00000
IA05080 (MA804341327)	0.71510	0.00820	0.29638	0.00970	0.00000	0.00000
IA07785 (MA900362696)	1.17264	0.01200	-0.37232	0.00762	0.00000	0.00000
IA07800 (MA900371480)	1.24816	0.01349	-0.87557	0.00891	0.00000	0.00000
IA07860 (MA900647188)	0.95596	0.01063	-0.65393	0.00910	0.00000	0.00000
IA07965 (MA900885042)	0.40051	0.00642	1.16070	0.02216	0.00000	0.00000
IA08210 (MA904338797)	1.23116	0.01319	-0.64641	0.00821	0.00000	0.00000
IA09861 (MA000827038)	1.20051	0.01200	0.37277	0.00752	0.00000	0.00000
IA09927 (MA001143716)	0.92833	0.00945	0.51001	0.00900	0.00000	0.00000
IA10279 (MA006260011)	0.75693	0.00945	-1.06734	0.01306	0.00000	0.00000
IA10348 (MA306437A)	0.60982	0.00749	0.73403	0.01257	0.00000	0.00000
IA12051 (MA000950721)	1.09903	0.01212	-0.83194	0.00920	0.00000	0.00000
IA12322 (MA201858644)	0.59039	0.00766	-0.57912	0.01207	0.00000	0.00000
IA12324 (MA201936073)	0.73459	0.00962	-1.33193	0.01554	0.00000	0.00000

Table 2.8.18  
IRT Parameters for Polytomous Items  
Mathematics Grade 5

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA01153 (MA624355953)	0.66210	0.00689	-0.33660	0.01107	1.36401	0.02097	-1.36401	0.01720	0.00000	0.00000
IA05028 (MA802358485)	1.08584	0.00939	0.13524	0.00609	0.69071	0.01030	-0.69071	0.01079	0.00000	0.00000
IA10187 (MA002829660)	1.29337	0.00968	-0.00396	0.00442	1.31996	0.01093	0.44653	0.00847	-0.38246	0.00838
IA12530 (MA715102421)	0.98519	0.00802	-0.01143	0.00701	2.53451	0.02249	0.03545	0.01050	-0.86459	0.01165
IA12635 (MA903746888)	1.05572	0.00844	0.29878	0.00524	1.50722	0.01208	0.43204	0.00917	-0.34620	0.00974
IA12638 (MA903948538)	1.31399	0.00998	-0.59849	0.00496	1.30061	0.01453	0.61984	0.01081	-0.46454	0.00852

Item ID	Parameters and Measures of Standard Error			
	d3	SE(d3)	d4	SE(d4)
IA10187 (MA002829660)	-1.38403	0.01138	0.00000	0.00000
IA12530 (MA715102421)	-1.70536	0.01517	0.00000	0.00000
IA12635 (MA903746888)	-1.59306	0.01500	0.00000	0.00000
IA12638 (MA903948538)	-1.45591	0.00960	0.00000	0.00000

Table 2.8.19 IRT Parameters for Dichotomous Items Mathematics Grade 6

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA04511 (MA272299)	0.90827	0.01376	-1.43237	0.03064	0.01480	0.01870
IA04725 (MA311660)	1.09673	0.02271	0.64771	0.01341	0.19530	0.00460
IA05139 (MA805276878)	0.42676	0.02426	1.42669	0.05916	0.28620	0.01460
IA07688 (MA736359148)	0.90004	0.02365	0.85221	0.01777	0.24030	0.00570
IA07690 (MA736361568)	1.05043	0.01928	-0.43048	0.02107	0.22100	0.00950
IA07774 (MA900284082)	1.37835	0.03111	1.10077	0.01192	0.20580	0.00310
IA07831 (MA900470149)	0.54610	0.01547	-1.37502	0.10108	0.13150	0.03900
IA07833 (MA900540139)	1.20668	0.02177	0.64515	0.01107	0.10890	0.00360
IA10156 (MA002527442)	0.70340	0.01939	0.90435	0.02298	0.20990	0.00710
IA12064 (MA001528147)	1.08137	0.01945	0.33584	0.01277	0.10010	0.00500
IA12074 (MA001545119)	0.81379	0.01879	-0.06637	0.02841	0.24480	0.01040
IA12079 (MA001551849)	0.94485	0.02210	-0.23140	0.02830	0.36250	0.00990
IA12080 (MA001555259)	0.76158	0.01542	-1.53261	0.05544	0.07940	0.03020
IA12122 (MA002536026)	1.09121	0.02083	0.47299	0.01362	0.17070	0.00480
IA12136 (MA003476915)	0.89534	0.01950	-1.25000	0.04565	0.20950	0.02290
IA12216 (MA136077359)	0.64550	0.01961	1.18409	0.02309	0.16590	0.00700
IA14631 (MA311690)	1.21436	0.02298	-0.11638	0.01564	0.22620	0.00690
IA02696 (MA703178216)	0.51240	0.00652	0.24231	0.01213	0.00000	0.00000
IA04744 (MA703176747)	0.70876	0.00768	0.86434	0.01224	0.00000	0.00000
IA07699 (MA800365824)	0.65406	0.00735	-0.27481	0.01011	0.00000	0.00000
IA07747 (MA805276792)	1.22723	0.01216	0.31935	0.00755	0.00000	0.00000
IA07821 (MA900437517)	0.65240	0.00740	0.46235	0.01085	0.00000	0.00000
IA07830 (MA900462230)	1.07132	0.01083	-0.58881	0.00841	0.00000	0.00000
IA08099 (MA902635513)	0.97159	0.01166	-1.39386	0.01319	0.00000	0.00000
IA10005 (MA001557959)	0.71257	0.00834	-1.13125	0.01319	0.00000	0.00000
IA12081 (MA001555876)	0.92341	0.01011	-0.91621	0.01043	0.00000	0.00000
IA12083 (MA001574310)	0.97485	0.01011	-0.96728	0.01021	0.00000	0.00000
IA12119 (MA002532873)	1.13231	0.01111	-0.16553	0.00755	0.00000	0.00000
IA12211 (MA136042342)	0.59643	0.00713	0.07100	0.01064	0.00000	0.00000
IA12218 (MA136079412)	0.60185	0.00812	-1.86820	0.02203	0.00000	0.00000
IA12230 (MA136282011)	0.76512	0.00895	-1.24968	0.01351	0.00000	0.00000
IA12247 (MA136500757)	0.72434	0.00812	-0.64914	0.01107	0.00000	0.00000
IA12404 (MA219859831)	0.67660	0.00729	-0.29301	0.00990	0.00000	0.00000
IA14572 (MA298162)	1.02037	0.01050	-0.81832	0.00936	0.00000	0.00000

Table 2.8.20  
IRT Parameters for Polytomous Items  
Mathematics Grade 6

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA01134 (MA624253312)	1.06629	0.00917	-0.20299	0.00599	0.61533	0.01076	-0.61533	0.00998	0.00000	0.00000
IA02454 (MA307295)	1.23823	0.00934	0.36606	0.00424	1.21843	0.00951	0.13013	0.00839	-0.44413	0.00923
IA04892 (MA736443588)	1.00844	0.00873	-0.36547	0.00639	0.64991	0.01178	-0.64991	0.01031	0.00000	0.00000
IA07663 (MA311694)	1.05303	0.00807	-0.42093	0.00472	1.19031	0.01302	0.40165	0.01001	-0.53843	0.00891
IA12138 (MA003843559)	1.17624	0.00945	0.50874	0.00451	0.95753	0.00908	0.31974	0.00880	-0.24835	0.00964
IA12246 (MA136463214)	1.49763	0.01177	0.70681	0.00408	0.85979	0.00794	0.15264	0.00830	-0.26840	0.00915

Item ID	Parameters and Measures of Standard Error			
	d3	SE(d3)	d4	SE(d4)
IA02454 (MA307295)	-0.90443	0.01067	0.00000	0.00000
IA07663 (MA311694)	-1.05353	0.00983	0.00000	0.00000
IA12138 (MA003843559)	-1.02892	0.01245	0.00000	0.00000
IA12246 (MA136463214)	-0.74403	0.01080	0.00000	0.00000

Table 2.8.21 IRT Parameters for Dichotomous Items Mathematics Grade 7

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA01009 (MA311103)	1.13031	0.02314	0.29788	0.01632	0.30210	0.00540
IA02058 (MA228036)	0.84119	0.02108	0.26713	0.02586	0.32860	0.00790
IA04471 (MA205602)	1.11924	0.02251	1.01140	0.01254	0.14390	0.00320
IA07549 (MA289832)	0.89410	0.01552	-0.31264	0.02142	0.13320	0.00910
IA07723 (MA801600966)	0.91147	0.02076	0.99986	0.01587	0.19130	0.00440
IA07844 (MA900559852)	1.25626	0.02998	0.96356	0.01365	0.26970	0.00360
IA08196 (MA904158907)	0.48736	0.01149	-0.87264	0.07060	0.03260	0.02530
IA10255 (MA005169543)	0.65276	0.01255	-0.31608	0.03097	0.03180	0.01240
IA12288 (MA201085794)	1.37553	0.01949	-0.03402	0.01032	0.09690	0.00410
IA12427 (MA288411)	1.27866	0.02733	-0.00416	0.01676	0.35540	0.00610
IA12491 (MA311086)	1.63594	0.04428	1.13384	0.01243	0.29470	0.00290
IA12642 (MA904183531)	0.71970	0.01509	-0.44073	0.03497	0.13330	0.01400
IA14543 (MA272872)	0.32000	0.00789	-1.68607	0.12299	0.02470	0.03270
IA14549 (MA287437)	1.38766	0.02314	0.11284	0.01166	0.19200	0.00450
IA14601 (MA306621)	1.18460	0.01917	-0.01126	0.01332	0.15040	0.00540
IA02528 (MA311075)	1.35959	0.01298	-0.49734	0.00766	0.00000	0.00000
IA04765 (MA713848042)	1.32580	0.01245	-0.26590	0.00744	0.00000	0.00000
IA04766 (MA713848086)	1.01480	0.00974	0.66807	0.00944	0.00000	0.00000
IA05090 (MA804458974)	0.80301	0.00810	0.08331	0.00921	0.00000	0.00000
IA07714 (MA801082691)	0.48439	0.00604	0.50467	0.01421	0.00000	0.00000
IA07737 (MA804653792)	0.72701	0.00826	-0.65896	0.01032	0.00000	0.00000
IA07857 (MA900635073)	0.51024	0.00625	0.89429	0.01554	0.00000	0.00000
IA07861 (MA900647693)	1.01713	0.01022	0.56495	0.00899	0.00000	0.00000
IA10251 (MA005076023)	0.96269	0.01033	-1.21142	0.01121	0.00000	0.00000
IA10427 (MA804637729)	1.04388	0.00985	-0.10961	0.00810	0.00000	0.00000
IA10441 (MA900736982)	0.85041	0.00900	-0.46160	0.00932	0.00000	0.00000
IA10476 (MA907259763)	0.69094	0.00805	1.17546	0.01432	0.00000	0.00000
IA12062 (MA001385352)	1.18158	0.01245	0.62434	0.00844	0.00000	0.00000
IA12264 (MA200155755)	1.00924	0.00948	0.49158	0.00910	0.00000	0.00000
IA12286 (MA201077581)	0.72367	0.00869	-1.36338	0.01410	0.00000	0.00000
IA12559 (MA801371719)	0.54275	0.00651	-0.57760	0.01221	0.00000	0.00000
IA12560 (MA801379630)	0.75121	0.00763	0.34627	0.00988	0.00000	0.00000
IA12572 (MA804440567)	0.72304	0.00747	0.58904	0.01077	0.00000	0.00000
IA12644 (MA904260053)	0.52083	0.00673	-0.71857	0.01310	0.00000	0.00000



Table 2.8.22  
IRT Parameters for Polytomous Items  
Mathematics Grade 7

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA04575 (MA295758)	1.30016	0.00922	-0.18246	0.00482	1.79066	0.01396	0.12463	0.00887	-0.61687	0.00905
IA04749 (MA703930934)	0.35845	0.00487	0.15807	0.02340	2.28004	0.03759	-2.28004	0.04328	0.00000	0.00000
IA07677 (MA715009356)	1.26256	0.00953	-0.22719	0.00406	0.87733	0.01017	0.15049	0.00857	-0.29607	0.00847
IA10287 (MA007839825)	1.29365	0.00943	0.35479	0.00461	1.60362	0.01080	0.11220	0.00875	-0.68580	0.01020
IA12284 (MA200915673)	1.05765	0.00890	-0.11794	0.00616	0.63216	0.01081	-0.63216	0.01054	0.00000	0.00000
IA12428 (MA289853)	1.16702	0.00853	0.21726	0.00506	1.44933	0.01120	0.71383	0.00937	-0.47534	0.00993

Item ID	Parameters and Measures of Standard Error			
	d3	SE(d3)	d4	SE(d4)
IA04575 (MA295758)	-1.29842	0.01050	0.00000	0.00000
IA07677 (MA715009356)	-0.73175	0.00896	0.00000	0.00000
IA10287 (MA007839825)	-1.03002	0.01131	0.00000	0.00000
IA12428 (MA289853)	-1.68782	0.01414	0.00000	0.00000

Table 2.8.23 IRT Parameters for Dichotomous Items Mathematics Grade 8

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA00877 (MA298049)	0.69948	0.01242	-2.07975	0.05791	0.03140	0.03310
IA02165 (MA284198)	0.64559	0.01311	-1.19831	0.05129	0.03750	0.02340
IA04487 (MA228246)	1.06323	0.01961	-0.23956	0.01787	0.20200	0.00750
IA04515 (MA272928)	1.48311	0.02292	-0.59517	0.01268	0.13610	0.00650
IA04713 (MA311424)	0.85793	0.01967	0.63559	0.01831	0.23050	0.00580
IA07836 (MA900545813)	0.78827	0.02041	0.42061	0.02427	0.26820	0.00760
IA08002 (MA901142533)	1.08295	0.02302	0.21909	0.01666	0.28070	0.00590
IA08084 (MA902303951)	1.04937	0.01716	-0.55513	0.01831	0.13900	0.00850
IA10119 (MA002159901)	1.18390	0.02105	-0.38549	0.01743	0.25910	0.00750
IA10225 (MA003777138)	1.09835	0.02462	-0.67790	0.02747	0.38040	0.01080
IA10240 (MA004058494)	0.96405	0.01951	0.00521	0.02041	0.28560	0.00720
IA12090 (MA001734156)	1.05156	0.02223	0.34130	0.01666	0.29100	0.00550
IA12112 (MA002275463)	1.35360	0.03304	1.16792	0.01302	0.21000	0.00290
IA12139 (MA003955575)	1.17809	0.01983	0.43693	0.01081	0.09710	0.00370
IA12309 (MA201656350)	1.39090	0.03267	0.21600	0.01555	0.43240	0.00490
IA12347 (MA202724888)	0.95621	0.01913	-0.89542	0.03066	0.22100	0.01430
IA12366 (MA202931621)	1.51823	0.02398	-0.39641	0.01191	0.17020	0.00570
IA12547 (MA800571638)	1.39746	0.02851	0.24743	0.01357	0.33950	0.00460
IA12552 (MA800977322)	1.57670	0.02755	0.45006	0.00993	0.18650	0.00330
IA12618 (MA901139952)	1.45795	0.02473	0.16647	0.01125	0.22620	0.00420
IA14609 (MA307537)	1.01878	0.02281	0.74126	0.01500	0.24190	0.00450
IA14610 (MA307560)	0.74233	0.01572	-0.57587	0.03552	0.16270	0.01420
IA02933 (MA715919523)	1.31043	0.01258	-0.01101	0.00728	0.00000	0.00000
IA04711 (MA311419)	1.04293	0.01119	1.03743	0.01059	0.00000	0.00000
IA04969 (MA800974248)	1.12388	0.01045	0.10845	0.00783	0.00000	0.00000
IA07703 (MA800475031)	0.73018	0.00863	-1.07180	0.01191	0.00000	0.00000
IA08083 (MA902303258)	0.71690	0.00757	-0.35140	0.00982	0.00000	0.00000
IA08235 (MA905906652)	0.71877	0.00767	-0.11227	0.00949	0.00000	0.00000
IA10065 (MA001865862)	0.98009	0.01178	-1.54180	0.01302	0.00000	0.00000
IA12093 (MA001738779)	0.83160	0.00826	-0.56749	0.00938	0.00000	0.00000
IA12110 (MA002239631)	0.42292	0.00581	0.20453	0.01401	0.00000	0.00000
IA12304 (MA201642904)	0.70331	0.00757	-0.11392	0.00982	0.00000	0.00000
IA12575 (MA804577289)	0.81220	0.00815	-0.49105	0.00916	0.00000	0.00000
IA12625 (MA902264555)	0.91959	0.00927	0.30391	0.00871	0.00000	0.00000

Table 2.8.24  
IRT Parameters for Polytomous Items  
Mathematics Grade 8

Item ID	Parameters and Measures of Standard Error									
	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
IA00822 (MA284774)	1.46963	0.01045	-0.25042	0.00406	1.17931	0.01063	0.27284	0.00834	-0.37045	0.00805
IA04718 (MA311452)	0.96703	0.00741	-1.18530	0.00605	1.40714	0.01855	0.41342	0.01297	-0.37382	0.01061
IA05073 (MA804152353)	0.85574	0.00751	-0.79626	0.00779	0.78304	0.01507	-0.78304	0.01169	0.00000	0.00000
IA08071 (MA902268353)	0.83144	0.00698	0.59274	0.00812	0.97359	0.01207	-0.97359	0.01581	0.00000	0.00000
IA12382 (MA203081760)	1.18502	0.00885	0.38278	0.00427	0.96659	0.00902	0.22767	0.00863	-0.36863	0.00951
IA12627 (MA902364058)	1.25377	0.00922	0.24586	0.00497	1.88230	0.01240	0.22598	0.00882	-0.66604	0.00994

Item ID	Parameters and Measures of Standard Error			
	d3	SE(d3)	d4	SE(d4)
IA00822 (MA284774)	-1.08169	0.00912	0.00000	0.00000
IA04718 (MA311452)	-1.44674	0.01034	0.00000	0.00000
IA12382 (MA203081760)	-0.82562	0.01091	0.00000	0.00000
IA12627 (MA902364058)	-1.44224	0.01281	0.00000	0.00000

# Section 2.9

## Decision Accuracy and Consistency (DAC)

Table 2.9.1  
 DAC Results  
 English Language Arts Grade 3

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
59965	0.90	0.60	Overall	0.81	0.73	0.10	0.09
			Cut 1	0.95	0.92	0.02	0.03
			Cut 2	0.91	0.87	0.05	0.05
			Cut 3	0.96	0.94	0.03	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.85	0.75		
			Perf 2	0.81	0.74		
			Perf 3	0.81	0.75		
			Perf 4	0.75	0.57		

Table 2.9.2  
 DAC Results  
 English Language Arts Grade 4

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
61103	0.90	0.60	Overall	0.81	0.74	0.10	0.09
			Cut 1	0.94	0.92	0.02	0.03
			Cut 2	0.91	0.87	0.04	0.05
			Cut 3	0.97	0.95	0.03	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.84	0.73		
			Perf 2	0.83	0.77		
			Perf 3	0.79	0.73		
			Perf 4	0.70	0.48		

Table 2.9.3  
 DAC Results  
 English Language Arts Grade 5

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
62601	0.92	0.65	Overall	0.84	0.77	0.08	0.08
			Cut 1	0.95	0.93	0.02	0.03
			Cut 2	0.92	0.88	0.04	0.04
			Cut 3	0.97	0.96	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.84	0.73		
			Perf 2	0.86	0.81		
			Perf 3	0.82	0.76		
			Perf 4	0.81	0.66		

Table 2.9.4  
 DAC Results  
 English Language Arts Grade 6

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
62688	0.92	0.59	Overall	0.79	0.71	0.10	0.10
			Cut 1	0.93	0.90	0.03	0.04
			Cut 2	0.91	0.88	0.04	0.04
			Cut 3	0.95	0.93	0.03	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.85	0.77		
			Perf 2	0.78	0.70		
			Perf 3	0.77	0.68		
			Perf 4	0.80	0.68		

Table 2.9.5  
 DAC Results  
 English Language Arts Grade 7

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
63626	0.92	0.61	Overall	0.81	0.73	0.10	0.09
			Cut 1	0.94	0.91	0.03	0.03
			Cut 2	0.91	0.88	0.04	0.05
			Cut 3	0.96	0.94	0.03	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.84	0.75		
			Perf 2	0.83	0.77		
			Perf 3	0.78	0.70		
			Perf 4	0.74	0.57		

Table 2.9.6  
 DAC Results  
 English Language Arts Grade 8

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
64139	0.92	0.61	Overall	0.80	0.72	0.10	0.10
			Cut 1	0.94	0.91	0.03	0.03
			Cut 2	0.91	0.88	0.04	0.04
			Cut 3	0.95	0.93	0.03	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.87	0.79		
			Perf 2	0.78	0.70		
			Perf 3	0.79	0.71		
			Perf 4	0.79	0.67		

Table 2.9.7  
 DAC Results  
 Mathematics Grade 3

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
50046	0.93	0.64	Overall	0.83	0.76	0.09	0.08
			Cut 1	0.96	0.94	0.02	0.02
			Cut 2	0.92	0.89	0.05	0.04
			Cut 3	0.95	0.93	0.03	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.83	0.75		
			Perf 2	0.82	0.75		
			Perf 3	0.84	0.78		
			Perf 4	0.82	0.70		

Table 2.9.8  
 DAC Results  
 Mathematics Grade 4

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
50882	0.93	0.65	Overall	0.84	0.78	0.08	0.08
			Cut 1	0.97	0.95	0.01	0.02
			Cut 2	0.92	0.89	0.04	0.04
			Cut 3	0.96	0.93	0.02	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.83	0.71		
			Perf 2	0.85	0.79		
			Perf 3	0.86	0.81		
			Perf 4	0.78	0.65		

Table 2.9.9  
 DAC Results  
 Mathematics Grade 5

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
53399	0.94	0.69	Overall	0.86	0.81	0.07	0.07
			Cut 1	0.97	0.95	0.01	0.02
			Cut 2	0.92	0.89	0.04	0.04
			Cut 3	0.97	0.96	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.80	0.65		
			Perf 2	0.87	0.83		
			Perf 3	0.87	0.82		
			Perf 4	0.83	0.72		

Table 2.9.10  
 DAC Results  
 Mathematics Grade 6

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
54954	0.94	0.67	Overall	0.85	0.79	0.07	0.08
			Cut 1	0.96	0.94	0.02	0.02
			Cut 2	0.92	0.89	0.04	0.04
			Cut 3	0.97	0.95	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.83	0.72		
			Perf 2	0.86	0.81		
			Perf 3	0.85	0.80		
			Perf 4	0.84	0.73		

Table 2.9.11  
 DAC Results  
 Mathematics Grade 7

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
56448	0.94	0.67	Overall	0.84	0.78	0.08	0.08
			Cut 1	0.95	0.93	0.02	0.03
			Cut 2	0.93	0.89	0.04	0.04
			Cut 3	0.97	0.95	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.80	0.69		
			Perf 2	0.86	0.81		
			Perf 3	0.84	0.78		
			Perf 4	0.84	0.74		

Table 2.9.12  
 DAC Results  
 Mathematics Grade 8

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
57441	0.94	0.67	Overall	0.85	0.78	0.08	0.08
			Cut 1	0.95	0.93	0.02	0.03
			Cut 2	0.92	0.89	0.04	0.04
			Cut 3	0.97	0.95	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.84	0.74		
			Perf 2	0.85	0.81		
			Perf 3	0.84	0.78		
			Perf 4	0.84	0.74		

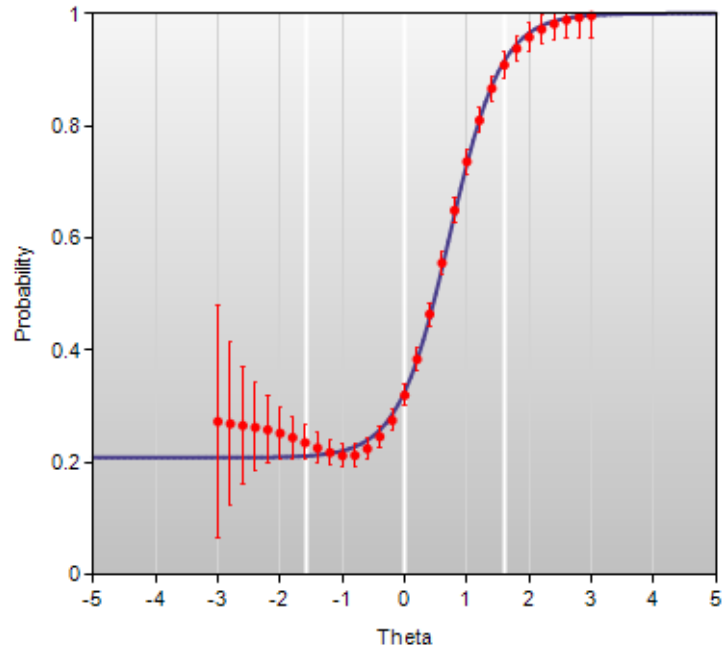
# Section 2.10

## Fit Plots of Watchlist Items



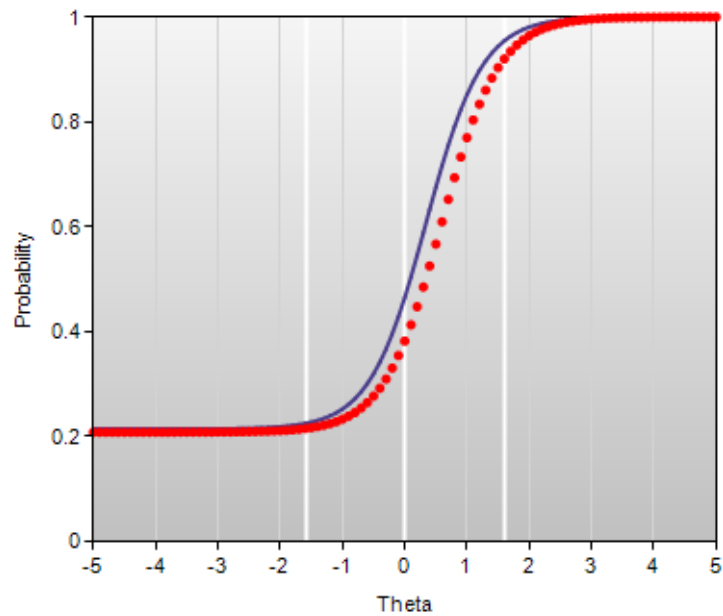
# Initial Calibration

English Language Arts Grade 3: IA00286



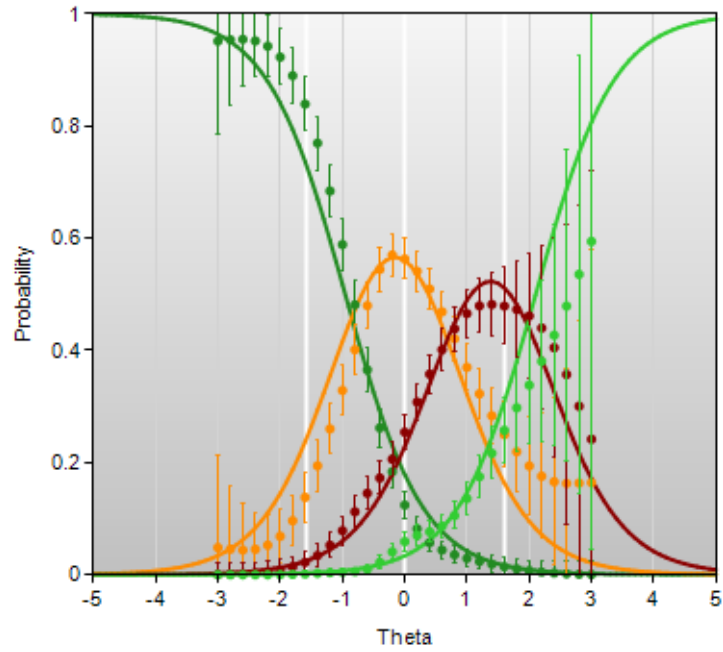
# Beta Chart

English Language Arts Grade 3: IA00286  
(EL308842)



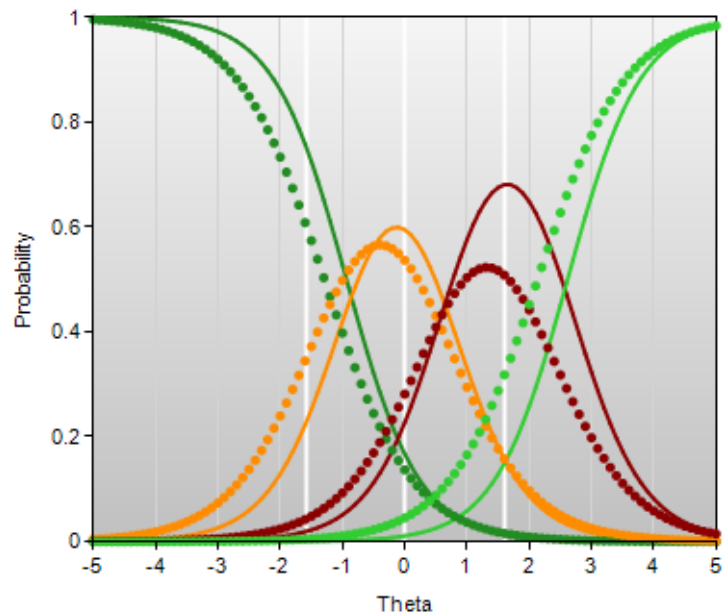
# Initial Calibration

English Language Arts Grade 3: IA00287



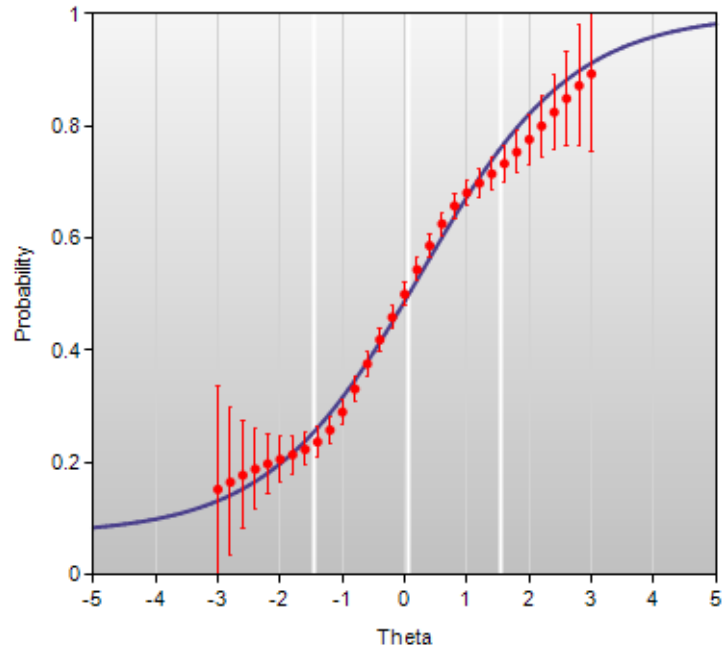
# Beta Chart

English Language Arts Grade 3: IA00287  
(EL308855)



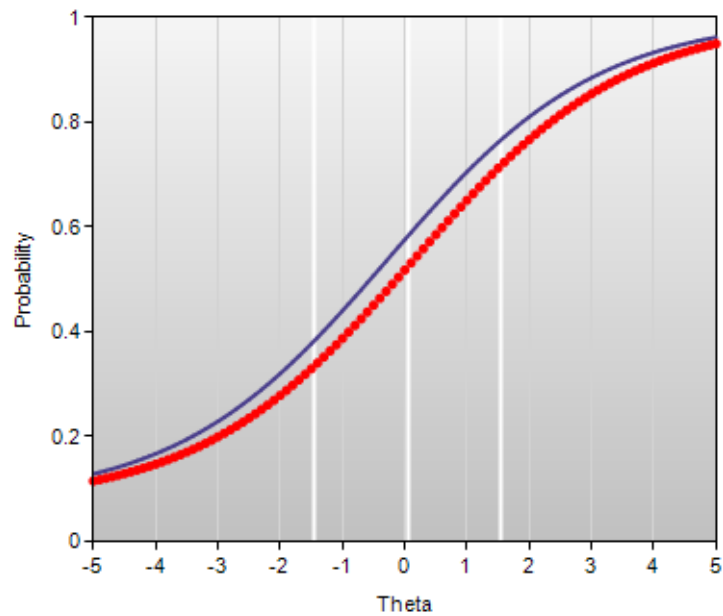
# Initial Calibration

English Language Arts Grade 8: IA00378



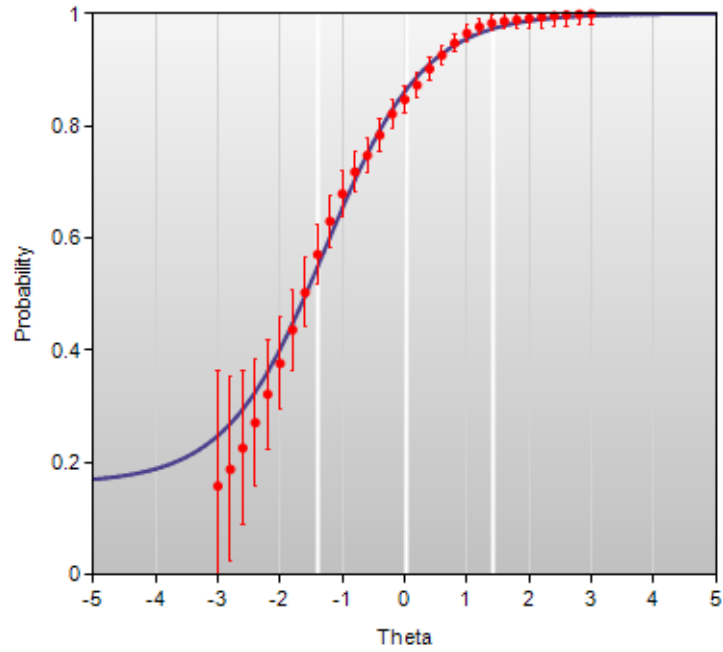
# Beta Chart

English Language Arts Grade 8: IA00378  
(EL62395555)



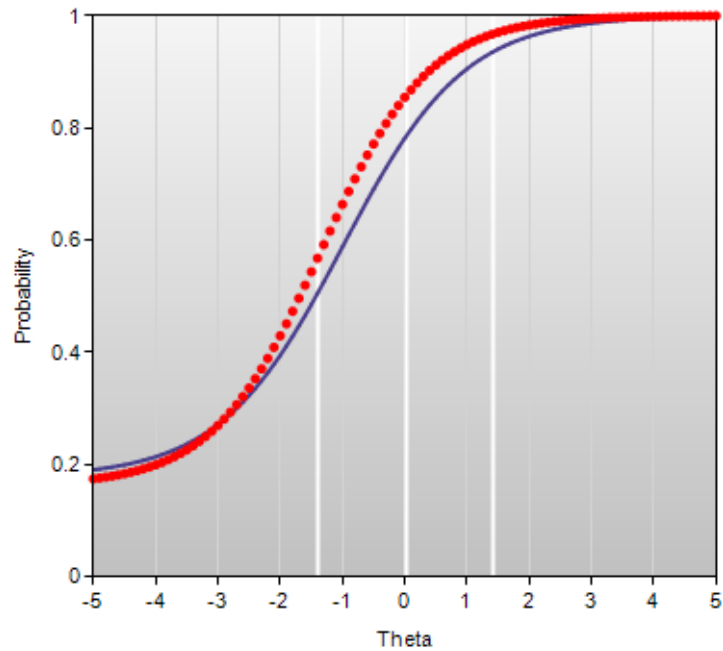
# Initial Calibration

Mathematics Grade 3: IA01019



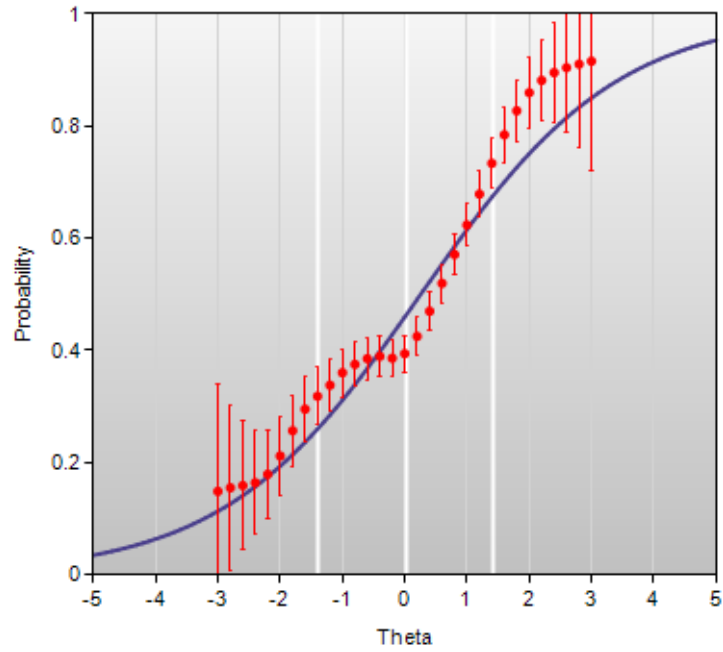
# Beta Chart

Mathematics Grade 3: IA01019 (MA311277)



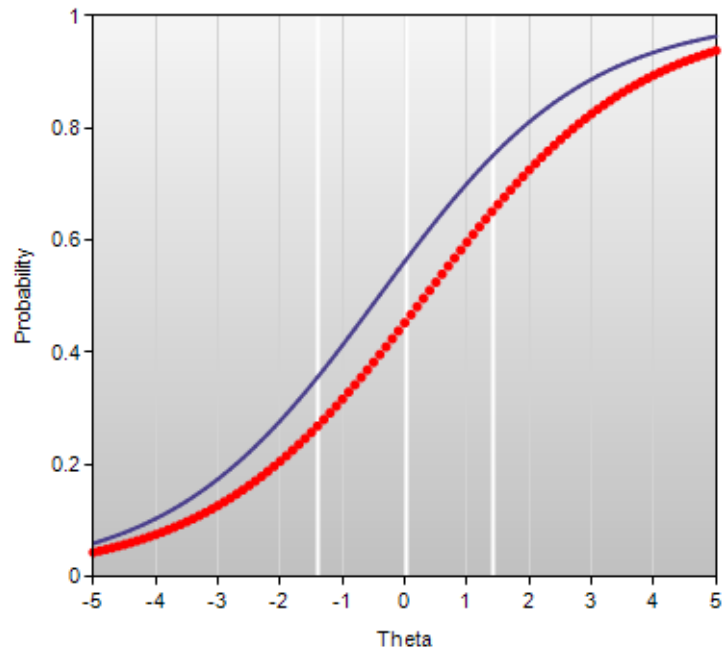
# Initial Calibration

Mathematics Grade 3: IA07855



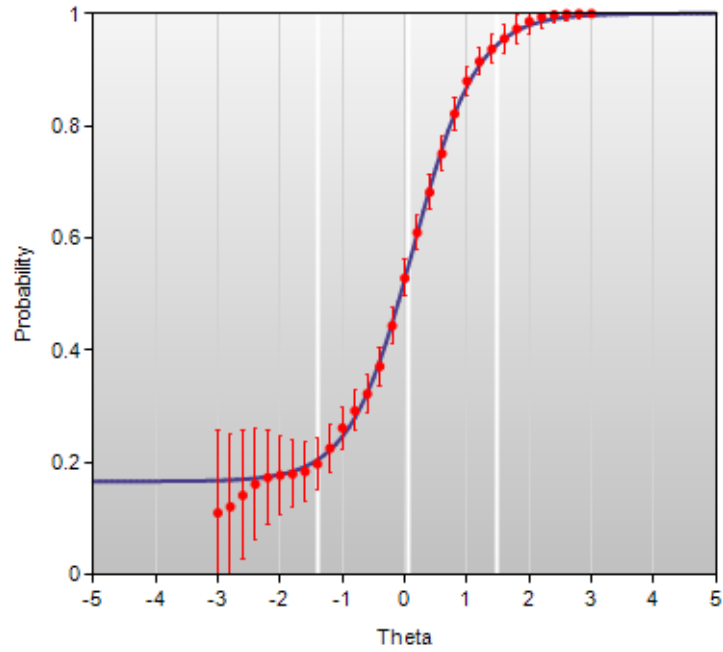
# Beta Chart

Mathematics Grade 3: IA07855 (MA900579464)



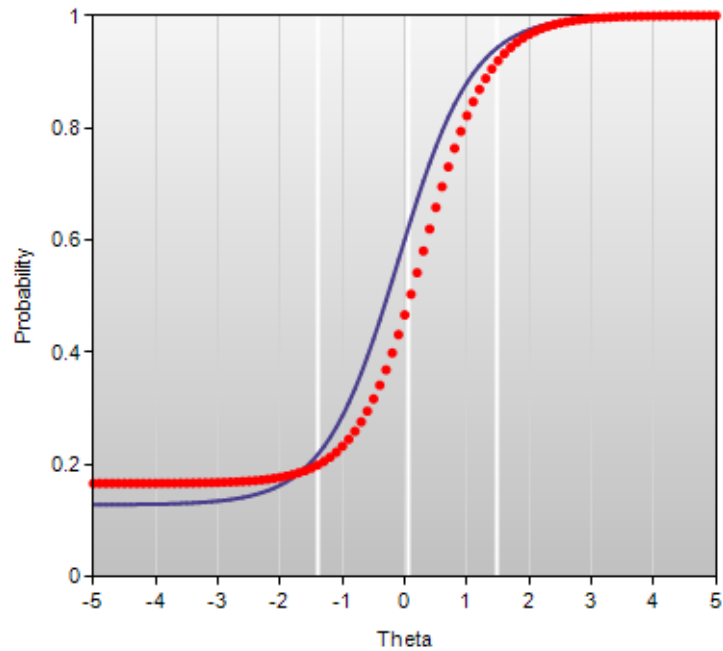
# Initial Calibration

Mathematics Grade 4: IA00961



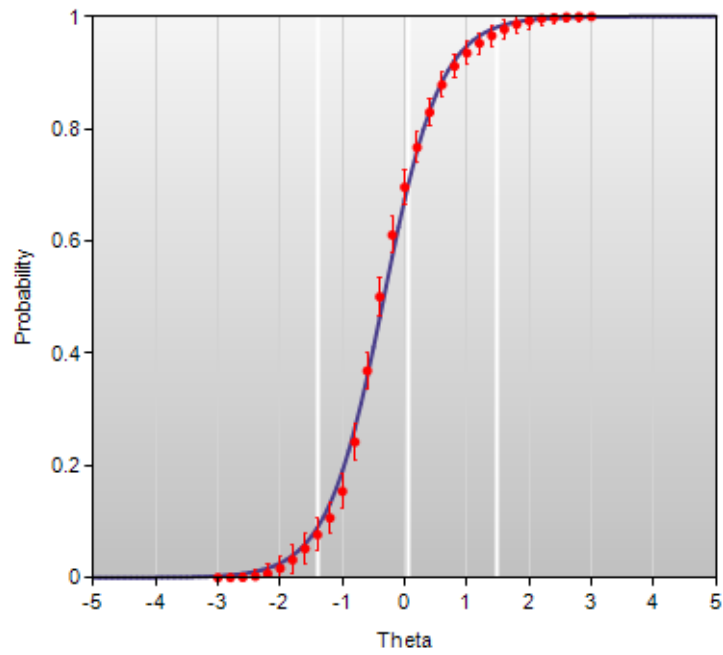
# Beta Chart

Mathematics Grade 4: IA00961 (MA307081)



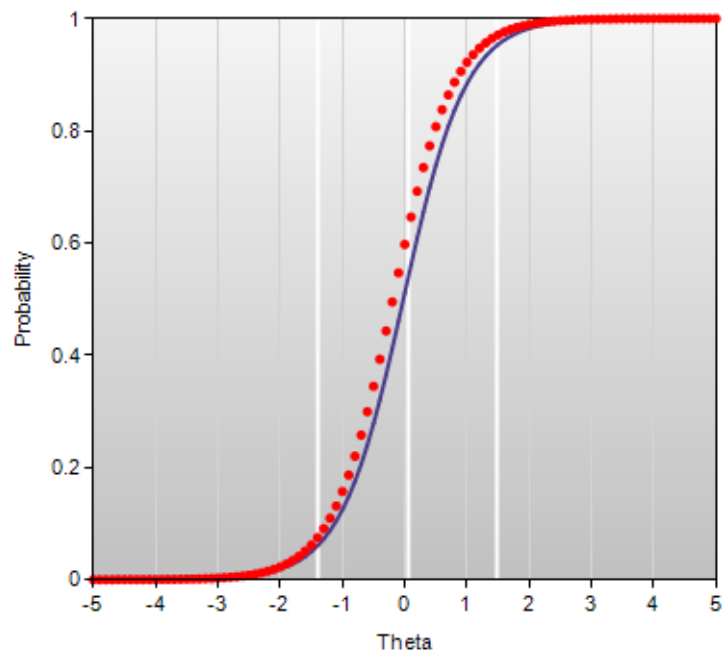
# Initial Calibration

Mathematics Grade 4: IA01055



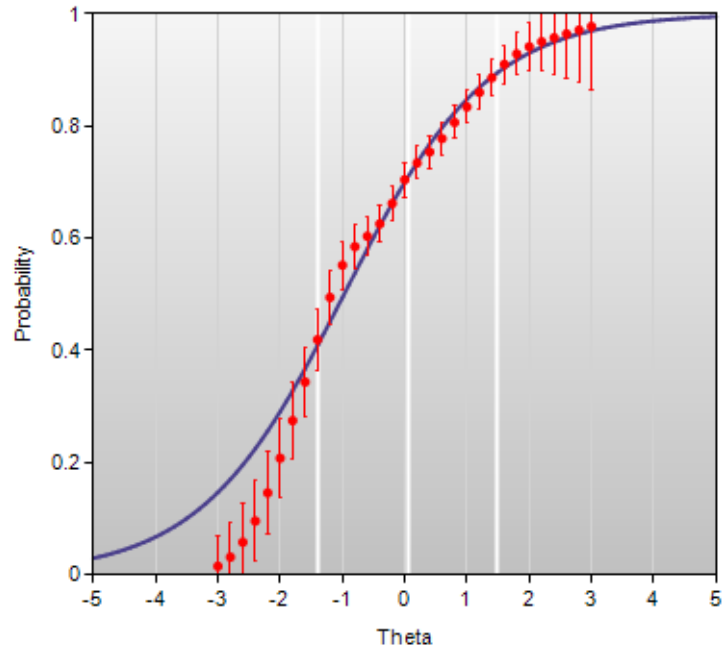
# Beta Chart

Mathematics Grade 4: IA01055 (MA311572)



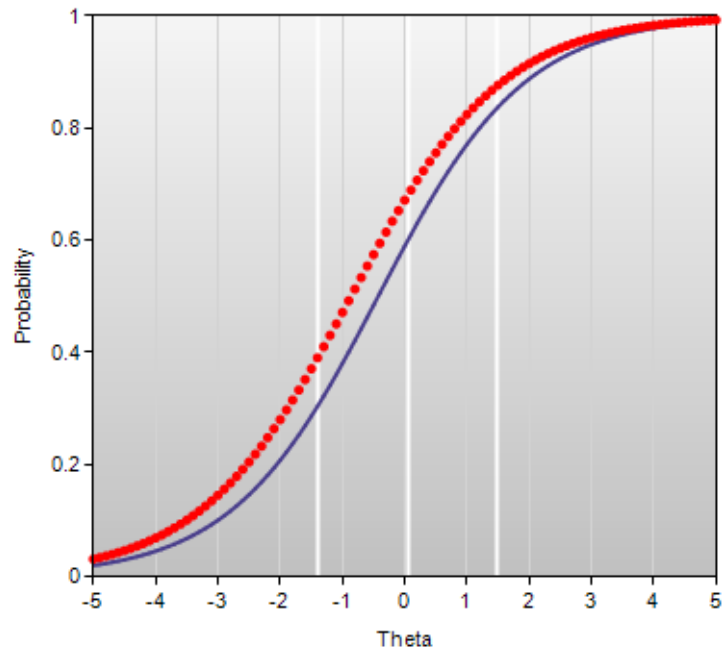
# Initial Calibration

Mathematics Grade 4: IA02819



# Beta Chart

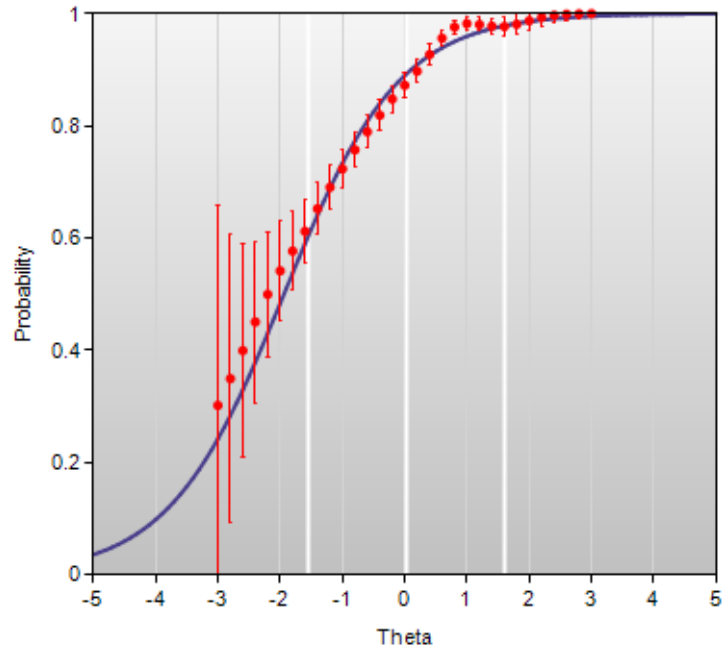
Mathematics Grade 4: IA02819 (MA713583365)





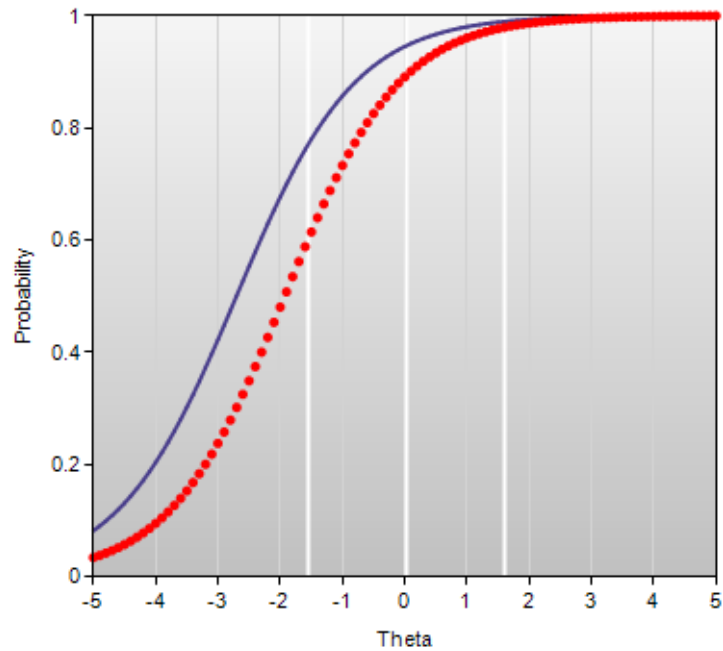
# Initial Calibration

Mathematics Grade 5: IA02917



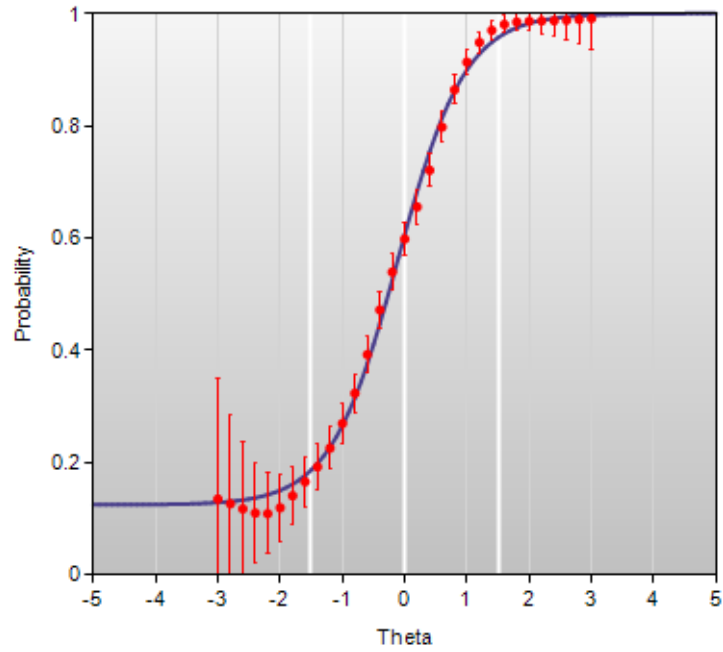
# Beta Chart

Mathematics Grade 5: IA02917 (MA715102107)



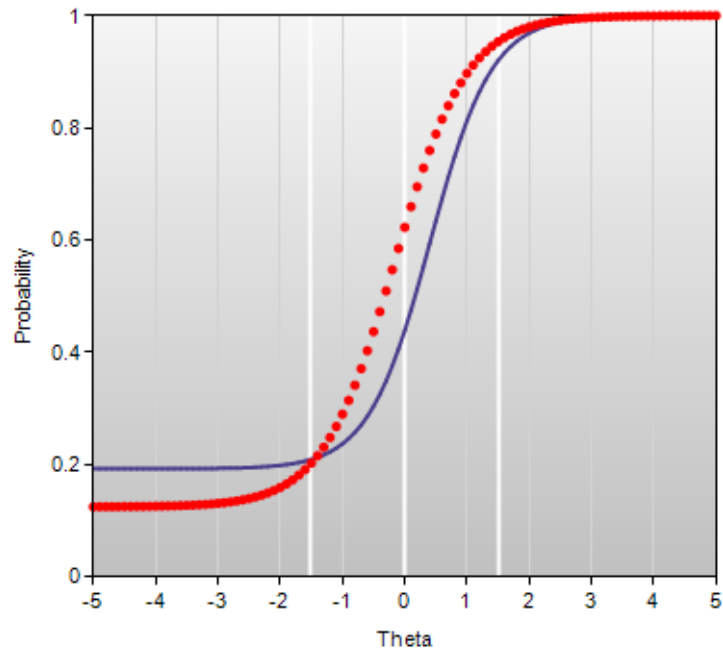
# Initial Calibration

Mathematics Grade 6: IA04899



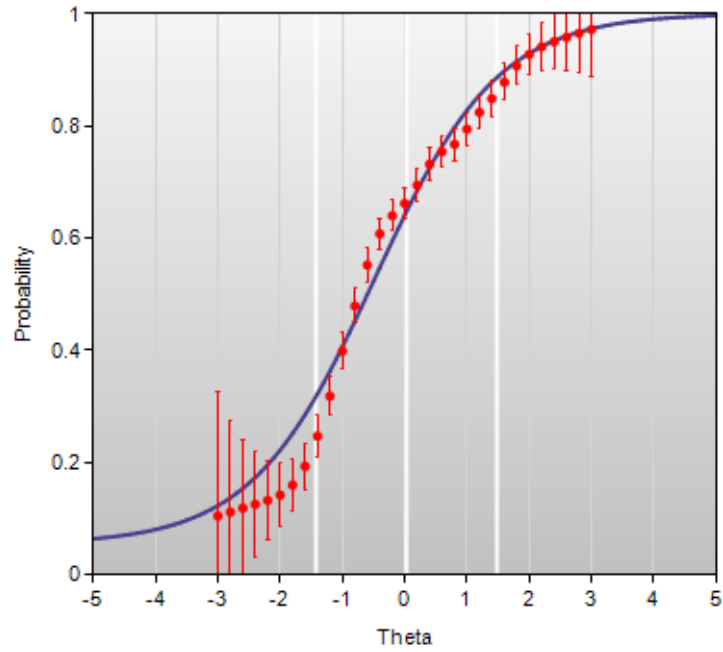
# Beta Chart

Mathematics Grade 6: IA04899 (MA736511626)



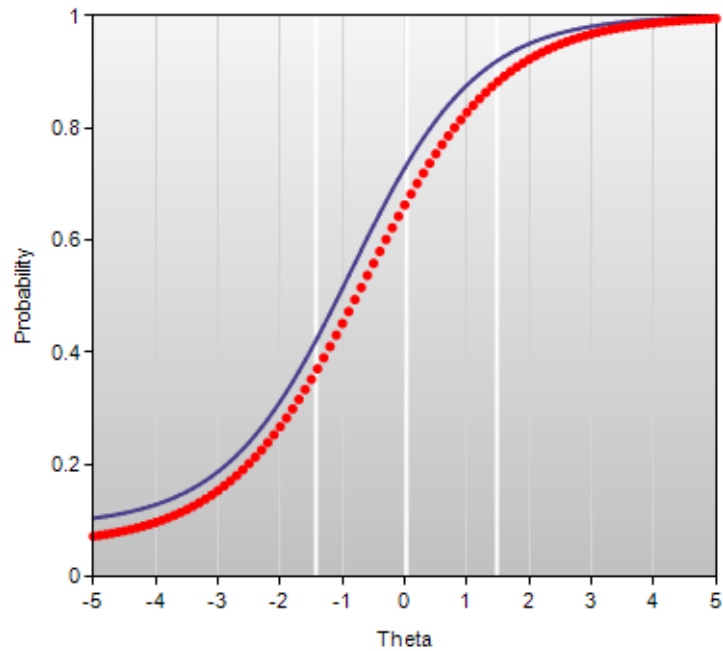
# Initial Calibration

Mathematics Grade 7: IA04538



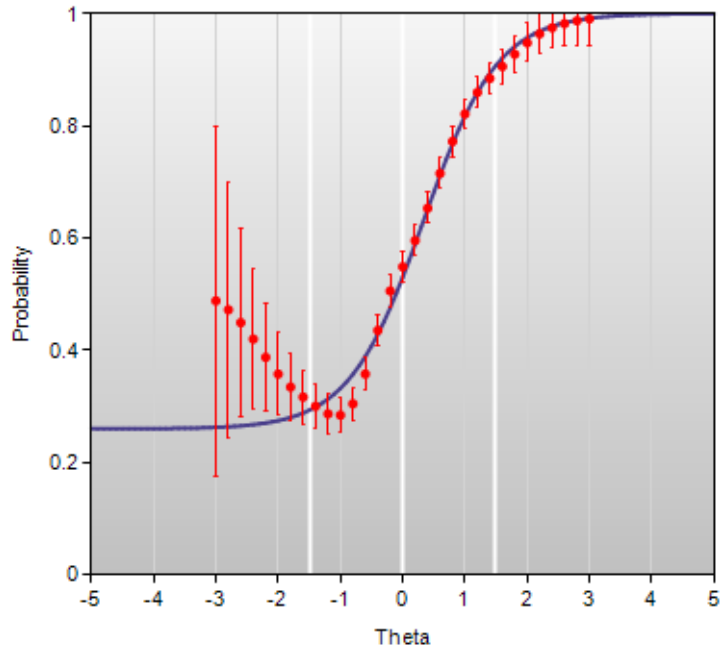
# Beta Chart

Mathematics Grade 7: IA04538 (MA282218)



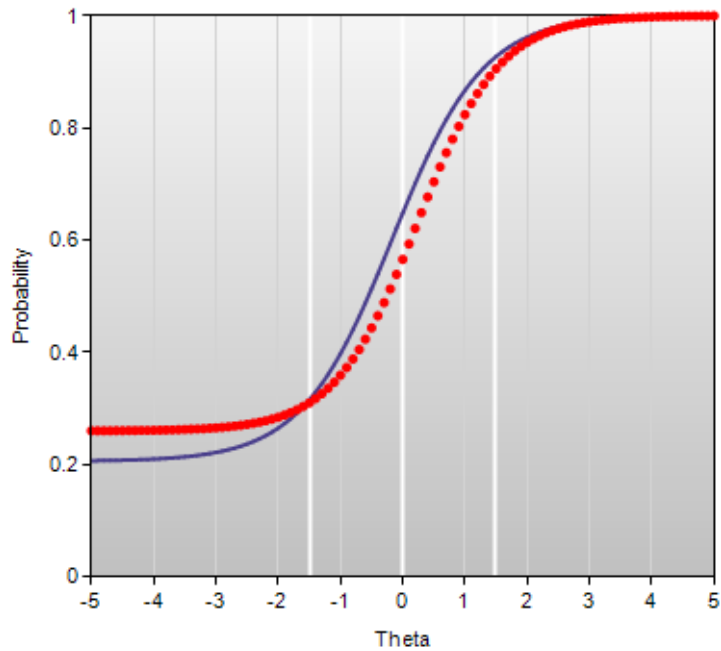
# Initial Calibration

Mathematics Grade 8: IA00865



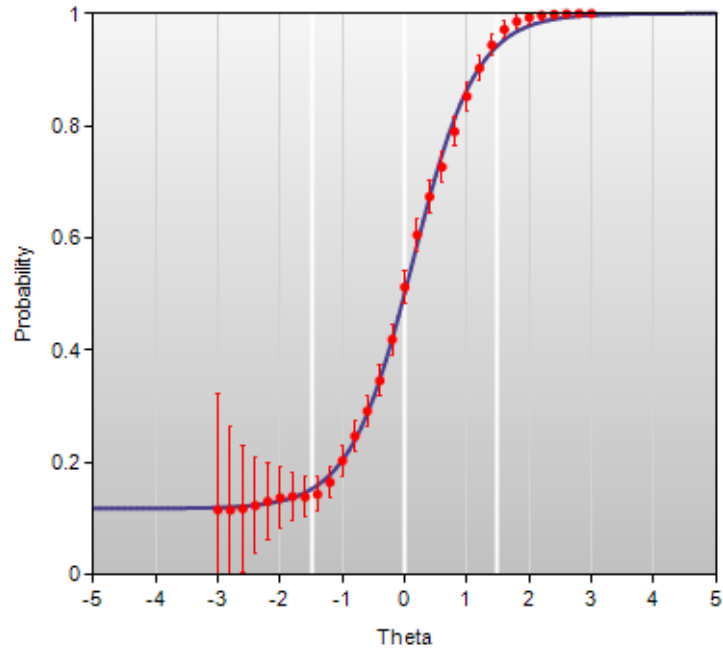
# Beta Chart

Mathematics Grade 8: IA00865 (MA297656)



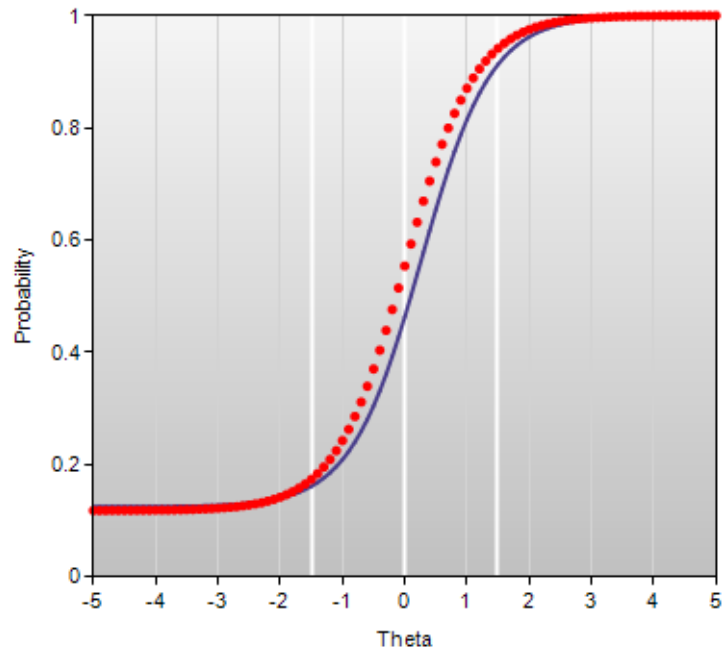
# Initial Calibration

Mathematics Grade 8: IA00905



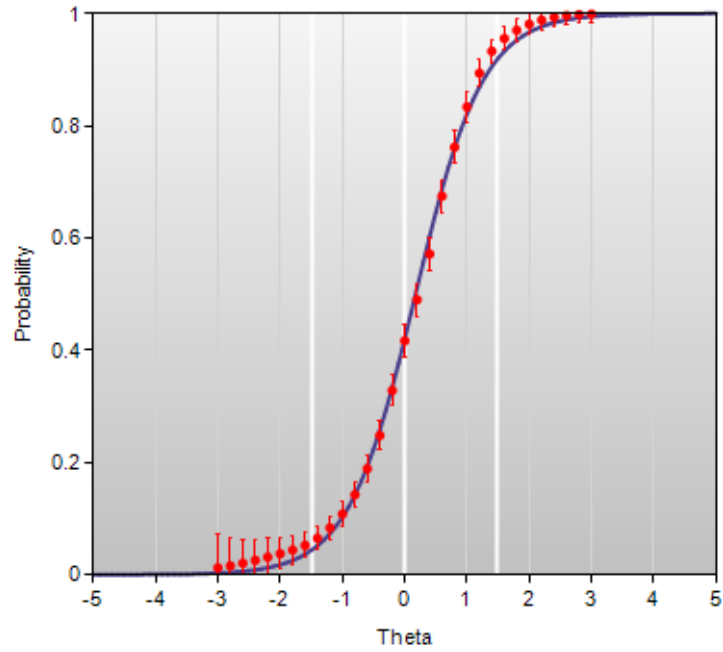
# Beta Chart

Mathematics Grade 8: IA00905 (MA301702)



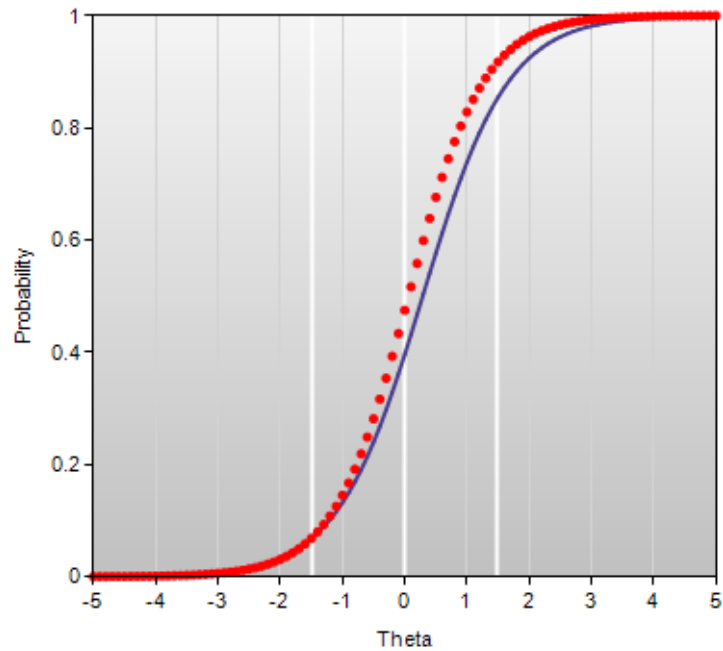
# Initial Calibration

Mathematics Grade 8: IA02495



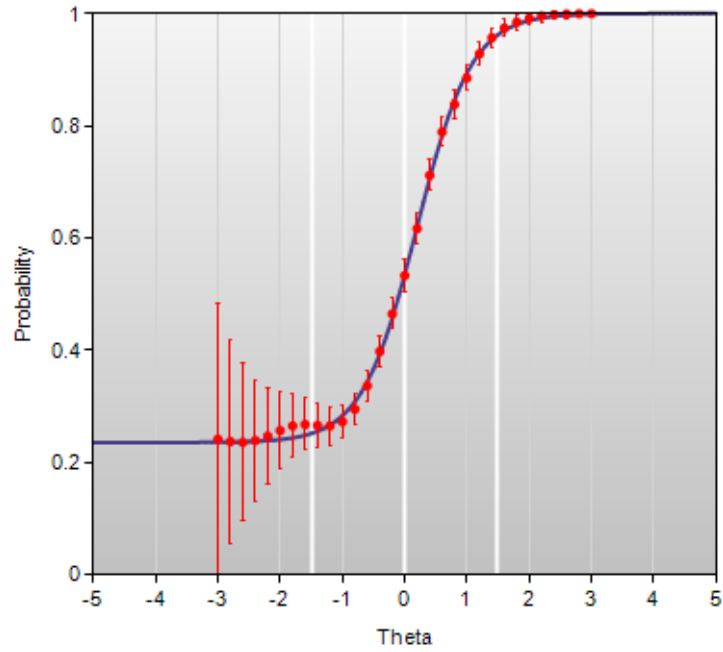
# Beta Chart

Mathematics Grade 8: IA02495 (MA309741)



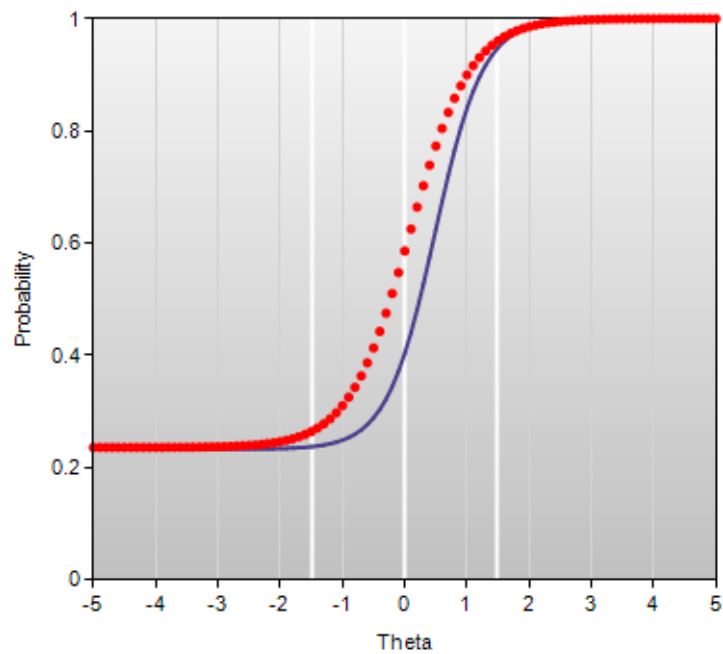
# Initial Calibration

Mathematics Grade 8: IA04665



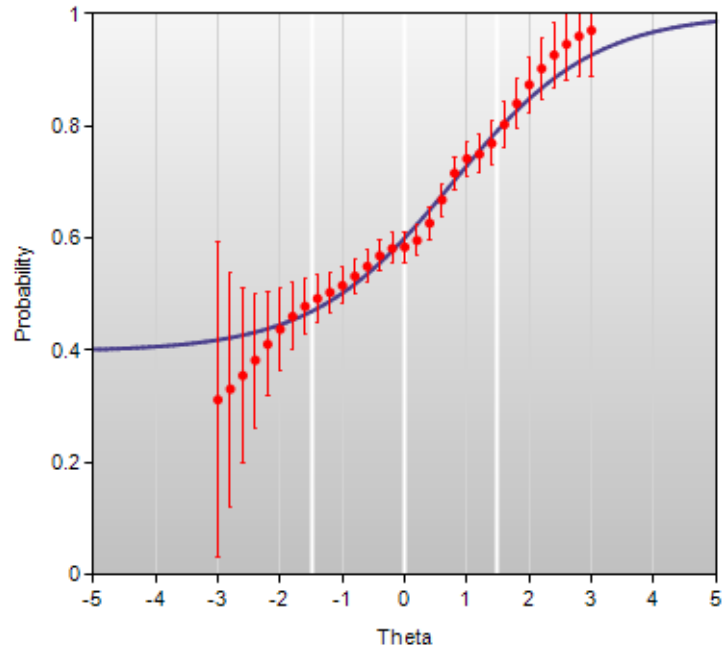
# Beta Chart

Mathematics Grade 8: IA04665 (MA307399)



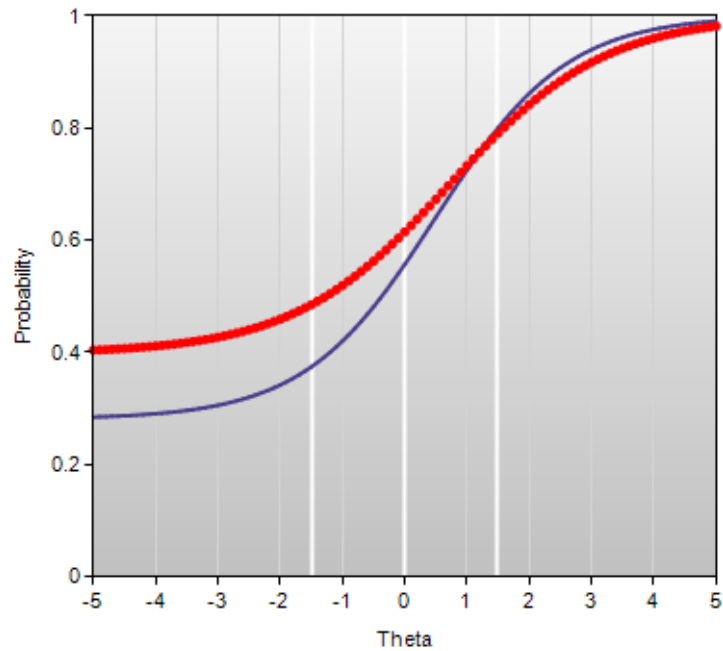
# Initial Calibration

Mathematics Grade 8: IA04719



# Beta Chart

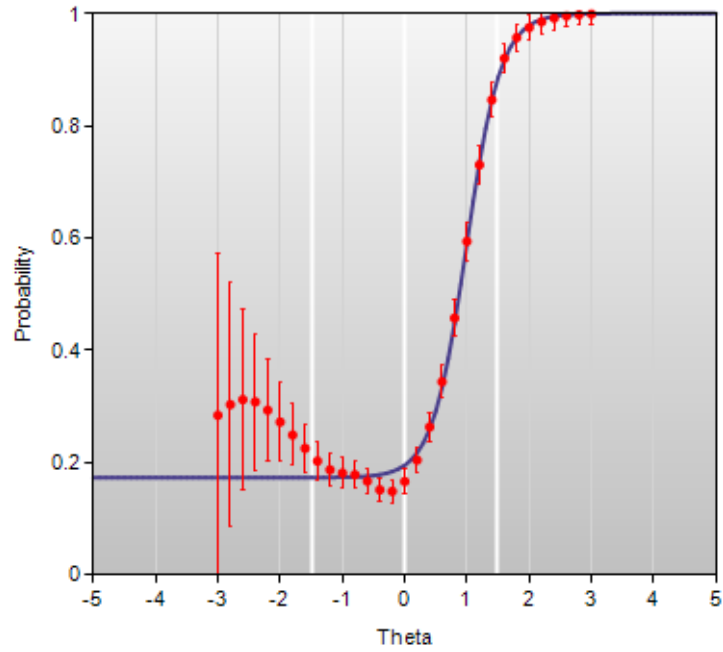
Mathematics Grade 8: IA04719 (MA311462)





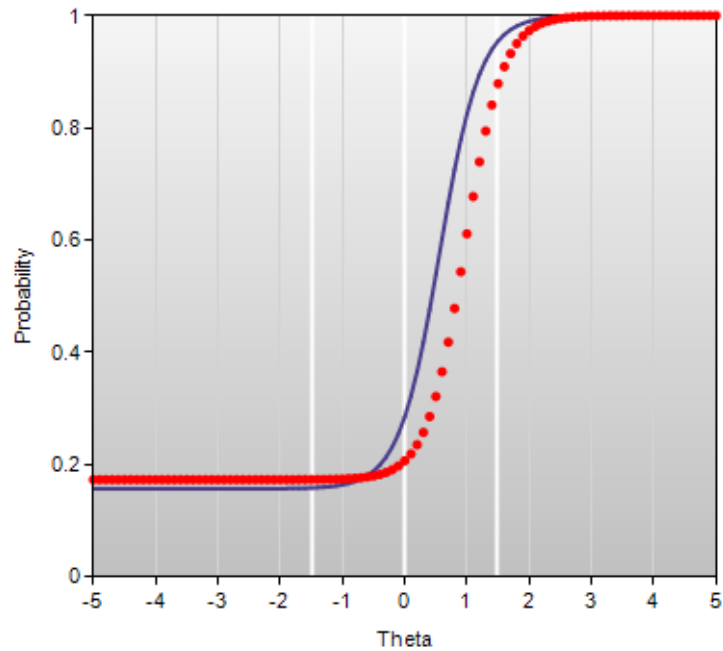
# Initial Calibration

Mathematics Grade 8: IA05070



# Beta Chart

Mathematics Grade 8: IA05070 (MA804042487)



**APPENDIX K**  
**RELIABILITY**

**Table K-1: Subgroup Reliabilities—ELA**

Grade	Subgroup	Number of Students	Maximum	Raw Mean	Score Standard Deviation	Alpha	SEM
3	All Students	9,400	44	22.74	8.89	0.89	2.90
	ELL	1,529	44	17.95	7.84	0.86	2.97
	Economically Disadvantaged	4,328	44	19.66	8.21	0.87	2.94
	African American	816	44	19.79	8.45	0.88	2.93
	Asian	321	44	25.83	8.43	0.88	2.90
	Hispanic	2,801	44	19.65	8.17	0.87	2.95
	Native American/Alaska Native	70	44	18.24	8.61	0.89	2.91
	White	4,818	44	24.87	8.71	0.89	2.86
	Pacific Islander/Hawaiian	17	44	17.59	6.47	0.77	3.07
	Multiracial	557	44	23.17	8.87	0.90	2.87
	Male	4,853	44	22.16	8.97	0.90	2.90
	Female	4,546	44	23.37	8.77	0.89	2.90
	Special Education	1,816	44	15.52	7.38	0.84	2.91
4	All Students	9,690	44	22.78	8.93	0.88	3.05
	ELL	1,685	44	17.58	8.03	0.86	3.05
	Economically Disadvantaged	4,590	44	19.41	8.19	0.86	3.07
	African American	831	44	20.03	8.18	0.86	3.08
	Asian	297	44	26.06	8.68	0.88	3.00
	Hispanic	2,985	44	19.69	8.47	0.87	3.06
	Native American/Alaska Native	85	44	18.65	7.46	0.83	3.07
	White	4,925	44	25.04	8.64	0.88	3.02
	Pacific Islander/Hawaiian	18	44	19.50	10.09	0.90	3.21
	Multiracial	549	44	22.45	8.71	0.88	3.06
	Male	4,982	44	22.24	8.87	0.88	3.03
	Female	4,707	44	23.35	8.96	0.88	3.05
	Special Education	1,869	44	14.96	7.22	0.84	2.93
5	All Students	9,716	48	26.22	10.70	0.92	3.09
	ELL	1,544	48	19.89	10.24	0.91	3.12
	Economically Disadvantaged	4,561	48	22.06	10.03	0.90	3.10
	African American	850	48	23.21	10.24	0.91	3.08
	Asian	334	48	30.28	10.32	0.91	3.06
	Hispanic	2,964	48	22.32	10.32	0.91	3.12
	Native American/Alaska Native	73	48	19.79	10.12	0.91	3.11
	White	4,933	48	28.84	10.13	0.91	3.04
	Pacific Islander/Hawaiian	13	48	28.54	10.46	0.91	3.12
	Multiracial	549	48	26.62	10.69	0.92	3.06
	Male	4,862	48	25.04	10.82	0.92	3.09
	Female	4,848	48	27.39	10.45	0.91	3.06
	Special Education	1,669	48	15.49	8.27	0.87	2.96
6	All Students	9,611	50	24.36	10.70	0.90	3.36
	ELL	1,614	50	18.22	9.33	0.88	3.22
	Economically Disadvantaged	4,448	50	20.09	9.57	0.89	3.24
	African American	877	50	20.54	9.66	0.89	3.26
	Asian	326	50	28.62	11.35	0.90	3.50
	Hispanic	2,886	50	20.05	9.51	0.88	3.25
	Native American/Alaska Native	73	50	16.84	9.01	0.89	3.05
	White	4,933	50	27.48	10.38	0.89	3.38
	Pacific Islander/Hawaiian	14	50	21.14	12.02	0.94	3.03
	Multiracial	502	50	23.63	10.28	0.90	3.31
	Male	4,936	50	22.79	10.39	0.90	3.28
	Female	4,674	50	26.02	10.78	0.90	3.41
	Special Education	1,641	50	14.44	7.17	0.83	2.98
7	All Students	9,737	50	25.58	10.77	0.90	3.33
	ELL	1,685	50	19.01	9.53	0.88	3.24
	Economically Disadvantaged	4,379	50	21.26	9.75	0.89	3.27

continued

Grade	Subgroup	Number of		Raw		Alpha	SEM
		Students	Maximum	Mean	Score Standard Deviation		
7	African American	982	50	21.38	10.28	0.90	3.24
	Asian	318	50	29.37	10.63	0.91	3.27
	Hispanic	2,791	50	21.20	9.80	0.89	3.26
	Native American/Alaska Native	73	50	20.60	9.48	0.89	3.19
	White	5,078	50	28.62	10.24	0.90	3.30
	Pacific Islander/Hawaiian	20	50	26.75	9.26	0.87	3.31
	Multiracial	475	50	25.66	10.91	0.91	3.33
	Male	5,004	50	24.25	10.65	0.90	3.33
	Female	4,727	50	26.97	10.71	0.91	3.29
8	Special Education	1,623	50	15.84	7.84	0.84	3.11
	All Students	9,845	50	25.56	11.54	0.92	3.33
	ELL	1,846	50	19.08	10.41	0.90	3.24
	Economically Disadvantaged	4,362	50	20.93	10.39	0.90	3.26
	African American	844	50	20.94	10.38	0.90	3.24
	Asian	318	50	29.72	11.44	0.92	3.24
	Hispanic	2,933	50	21.27	10.72	0.91	3.26
	Native American/Alaska Native	68	50	17.54	10.65	0.91	3.21
	White	5,172	50	28.66	11.07	0.91	3.31
	Pacific Islander/Hawaiian	18	50	25.56	12.37	0.93	3.26
	Multiracial	492	50	24.92	11.69	0.92	3.37
	Male	5,088	50	23.77	11.39	0.91	3.33
	Female	4,756	50	27.47	11.40	0.92	3.30
	Special Education	1,658	50	15.34	7.93	0.85	3.05

**Table K-2. Subgroup Reliabilities—Mathematics**

Grade	Subgroup	Number of		Raw		Alpha	SEM
		Students	Maximum	Mean	Score Standard Deviation		
3	All Students	9,579	48	24.21	11.54	0.93	3.02
	ELL	1,702	48	19.26	10.47	0.92	2.94
	Economically Disadvantaged	4,441	48	20.15	10.62	0.92	2.97
	African American	828	48	19.90	10.50	0.92	2.95
	Asian	332	48	29.20	11.72	0.94	2.96
	Hispanic	2,936	48	20.36	10.58	0.92	2.97
	Native American/Alaska Native	74	48	19.68	10.70	0.93	2.93
	White	4,835	48	26.98	11.27	0.93	3.02
	Pacific Islander/Hawaiian	18	48	15.33	7.66	0.86	2.89
	Multiracial	556	48	24.77	11.99	0.94	3.00
	Male	4,941	48	24.86	11.86	0.94	3.02
	Female	4,637	48	23.52	11.16	0.93	3.02
	Special Education	1,827	48	15.77	10.14	0.92	2.87
4	All Students	9,860	54	27.74	12.68	0.93	3.40
	ELL	1,857	54	21.38	11.83	0.92	3.42
	Economically Disadvantaged	4,695	54	22.93	11.48	0.91	3.43
	African American	861	54	23.28	11.49	0.91	3.45
	Asian	304	54	33.64	12.94	0.94	3.25
	Hispanic	3,099	54	23.20	11.71	0.91	3.44
	Native American/Alaska Native	89	54	21.47	11.14	0.91	3.38
	White	4,933	54	31.20	12.25	0.93	3.32
	Pacific Islander/Hawaiian	19	54	22.47	13.25	0.93	3.61
	Multiracial	555	54	27.19	12.52	0.93	3.39
	Male	5,073	54	28.94	12.91	0.93	3.39
	Female	4,786	54	26.47	12.31	0.92	3.41
	Special Education	1,873	54	17.37	10.60	0.90	3.33
5	All Students	9,904	54	24.94	11.96	0.92	3.31
	ELL	1,717	54	18.92	10.27	0.90	3.19

continued

Grade	Subgroup	Number of		Raw Score		Alpha	SEM
		Students	Maximum	Mean	Standard Deviation		
5	Economically Disadvantaged	4,684	54	20.08	10.13	0.90	3.23
	African American	872	54	21.24	10.46	0.90	3.25
	Asian	343	54	31.64	12.40	0.93	3.28
	Hispanic	3,080	54	20.16	10.11	0.90	3.23
	Native American/Alaska Native	80	54	17.49	10.06	0.90	3.10
	White	4,962	54	28.18	11.95	0.92	3.32
	Pacific Islander/Hawaiian	14	54	26.64	13.80	0.94	3.28
	Multiracial	553	54	25.26	12.01	0.92	3.34
	Male	4,967	54	25.29	12.50	0.93	3.29
	Female	4,931	54	24.60	11.39	0.91	3.33
	Special Education	1,681	54	14.92	8.58	0.88	3.02
6	All Students	9,763	54	21.71	12.19	0.92	3.43
	ELL	1,787	54	15.93	10.13	0.90	3.15
	Economically Disadvantaged	4,541	54	16.84	9.86	0.90	3.19
	African American	902	54	16.65	10.00	0.90	3.16
	Asian	335	54	29.55	13.85	0.93	3.63
	Hispanic	3,001	54	16.83	9.76	0.89	3.19
	Native American/Alaska Native	77	54	14.31	9.10	0.89	3.00
	White	4,932	54	25.34	12.29	0.92	3.54
	Pacific Islander/Hawaiian	15	54	17.33	12.92	0.94	3.22
	Multiracial	501	54	20.35	12.06	0.92	3.39
	Male	5,019	54	21.60	12.28	0.92	3.42
	Female	4,743	54	21.83	12.10	0.92	3.44
	Special Education	1,629	54	11.71	7.47	0.86	2.84
7	All Students	9,895	54	19.80	12.13	0.92	3.37
	ELL	1,882	54	12.94	8.77	0.88	3.01
	Economically Disadvantaged	4,466	54	14.75	9.33	0.89	3.12
	African American	1,011	54	14.55	9.21	0.89	3.10
	Asian	323	54	26.18	14.02	0.94	3.48
	Hispanic	2,899	54	14.41	9.11	0.88	3.10
	Native American/Alaska Native	75	54	13.81	8.45	0.87	3.04
	White	5,091	54	23.60	12.45	0.92	3.46
	Pacific Islander/Hawaiian	23	54	21.04	14.11	0.94	3.53
	Multiracial	473	54	19.70	11.94	0.92	3.37
	Male	5,085	54	20.23	12.47	0.93	3.38
	Female	4,804	54	19.33	11.74	0.92	3.35
	Special Education	1,609	54	10.80	6.59	0.82	2.79
8	All Students	9,963	54	22.57	12.48	0.92	3.42
	ELL	2,023	54	15.71	9.35	0.88	3.20
	Economically Disadvantaged	4,439	54	17.28	9.77	0.89	3.27
	African American	865	54	17.17	10.28	0.90	3.26
	Asian	323	54	28.12	14.33	0.94	3.48
	Hispanic	3,050	54	17.26	9.97	0.89	3.27
	Native American/Alaska Native	67	54	15.27	8.71	0.87	3.19
	White	5,148	54	26.44	12.49	0.92	3.43
	Pacific Islander/Hawaiian	21	54	17.81	11.71	0.92	3.22
	Multiracial	489	54	22.10	12.65	0.93	3.41
	Male	5,161	54	22.58	12.88	0.93	3.43
	Female	4,802	54	22.56	12.03	0.92	3.40
	Special Education	1,639	54	13.07	6.98	0.81	3.05

**Table K-3. Reliabilities by Reporting Categories, Grade, and Content Area—ELA**

Grade	Item Reporting Category	Label	Number		Raw Score		Alpha	SEM
			of Items	Maximum	Mean	Standard Deviation		
3	1	Reading	23	28	15.49	5.96	0.85	2.29
	2	Language	8	12	6.15	2.85	0.67	1.64
	3	Writing	1	4	1.11	0.84	--	--
4	1	Reading	24	30	16.53	6.39	0.83	2.60
	2	Language	7	10	5.29	2.37	0.68	1.35
	3	Writing	1	4	0.96	0.98	--	--
5	1	Reading	24	29	17.56	6.76	0.87	2.47
	2	Language	7	11	6.03	2.87	0.76	1.40
	3	Writing	2	8	2.63	1.90	0.84	0.76
6	1	Reading	24	29	16.78	6.32	0.84	2.51
	2	Language	7	11	5.14	2.88	0.72	1.53
	3	Writing	2	10	2.44	2.41	0.78	1.13
7	1	Reading	24	28	15.86	6.06	0.84	2.41
	2	Language	7	12	6.12	3.17	0.73	1.64
	3	Writing	2	10	3.6	2.49	0.90	0.80
8	1	Reading	24	29	16	6.69	0.87	2.44
	2	Language	7	11	6.17	3.14	0.74	1.60
	3	Writing	2	10	3.39	2.59	0.91	0.76

**Table K-4. Reliabilities by Reporting Categories, Grade, and Content Area—Mathematics**

Grade	Item Reporting Category	Label	Number		Raw Score		Alpha	SEM
			of Items	Maximum	Mean	Standard Deviation		
3	1	Operations and Algebraic Thinking	12	14	7.56	3.92	0.84	1.55
	2	Number and Operations in Base Ten	5	7	3.23	2.19	0.71	1.18
	3	Number and Operations-Fractions	8	10	5.03	2.63	0.73	1.36
	4	Measurement and Data	10	12	5.55	3.05	0.73	1.58
	5	Geometry	5	5	2.84	1.30	0.51	0.91
4	1	Operations and Algebraic Thinking	7	10	6.02	2.41	0.64	1.45
	2	Number and Operations in Base Ten	8	11	6.43	3.17	0.75	1.59
	3	Number and Operations-Fractions	12	16	7.69	4.14	0.83	1.70
	4	Measurement and Data	8	11	4.82	2.90	0.64	1.74
	5	Geometry	5	6	2.79	1.73	0.65	1.02
5	1	Operations and Algebraic Thinking	7	8	3.77	2.13	0.68	1.21
	2	Number and Operations in Base Ten	12	16	8.5	4.25	0.82	1.80
	3	Number and Operations-Fractions	10	13	4.28	2.96	0.69	1.64
	4	Measurement and Data	7	10	5.2	2.65	0.69	1.47
	5	Geometry	4	7	3.19	1.77	0.59	1.14
6	1	Ratios and Proportional Relationships	10	11	5.9	3.02	0.78	1.42
	2	The Number System	8	11	4.41	2.63	0.63	1.60
	3	Expressions and Equations	12	16	6.53	4.05	0.76	1.98
	4	Geometry	5	8	2.93	2.19	0.67	1.26
	5	Statistics and Probability	5	8	1.94	1.98	0.55	1.32
7	1	Ratios and Proportional Relationships	8	11	4.16	2.61	0.65	1.54
	2	The Number System	10	11	3.9	2.94	0.79	1.35
	3	Expressions and Equations	10	13	4.77	3.30	0.76	1.61
	4	Geometry	4	8	2.67	2.24	0.52	1.55
	5	Statistics and Probability	8	11	4.29	2.58	0.66	1.51
8	1	Number System & Expressions/Equations	17	21	8.08	5.11	0.87	1.86
	2	Functions	8	11	4.54	2.78	0.62	1.72
	3	Geometry	10	14	5.63	3.66	0.77	1.76
	4	Statistics and Probability	5	8	4.33	2.25	0.56	1.49