



2023 RICAS Technical Report

Prepared by Cognia and the
Rhode Island Department of Education

3/13/2024



RIDE Rhode Island
Department
of Education

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

Table of Contents

CHAPTER 1. INTRODUCTION TO THE RHODE ISLAND COMPREHENSIVE ASSESSMENT SYSTEM	7
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 2023 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
CHAPTER 2. TEST DESIGN AND DEVELOPMENT	14
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 STANDARDS	15
2.4.2 ELA BLUEPRINTS	15
2.4.3 ELA ITEM TYPES	15
2.4.4 ELA PASSAGE TYPES	16
2.4.5 ELA COGNITIVE LEVELS	17
2.4.6 ELA FORM DESIGN	17
2.4.7 ELA REFERENCE MATERIALS	19
2.5 MATHEMATICS	19
2.5.1 MATHEMATICS STANDARDS	19
2.5.2 MATHEMATICS BLUEPRINTS	20
2.5.3 MATHEMATICS ITEM TYPES	20
2.5.4 MATHEMATICS COGNITIVE LEVELS	21
2.5.5 FORM TEST DESIGN	21
2.5.6 MATHEMATICS REFERENCE MATERIALS	22
2.6 ITEM AND TEST DEVELOPMENT PROCESS	23
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 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING TEST DESIGN AND DEVELOPMENT	29
CHAPTER 3. TEST ADMINISTRATION	32
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
CHAPTER 4. SCORING	36
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 DOUBLE-BLIND SCORING WITH THE INTELLIGENT ESSAY ASSESSOR (IEA)	43
4.5.8 MONITORING OF SCORING QUALITY	45
4.5.9 INTERRATER CONSISTENCY	46
4.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING SCORING	48
CHAPTER 5. REPORTING	49
5.1 REPORTING OF RESULTS	49
5.2 PARENT/GUARDIAN REPORT	49
5.3 REPORTING BUSINESS REQUIREMENTS	50
5.4 QUALITY ASSURANCE	50
5.5 ADDITIONAL RESOURCES	51
5.5.1 STUDENTS AND THEIR FAMILIES	51
5.5.2 EDUCATORS AND ADMINISTRATORS	52
5.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING REPORTING	52
CHAPTER 6. CLASSICAL ITEM ANALYSIS	55
6.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES	55
6.2 DIFFERENTIAL ITEM FUNCTIONING	57
6.3 DIMENSIONALITY ANALYSIS	58
6.4 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING CLASSICAL ITEM ANALYSES	60
CHAPTER 7. ITEM RESPONSE THEORY ANALYSIS	62
7.1 OVERVIEW	62
7.2 IRT	62
7.3 IRT RESULTS	65
7.4 EQUATING	65
7.5 REPORTED SCALE SCORES AND ACHIEVEMENT STANDARDS	68
7.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING ITEM RESPONSE THEORY ANALYSES	70
CHAPTER 8. RELIABILITY	72
8.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT	72
8.2 SUBGROUP RELIABILITY	73
8.3 REPORTING SUBCATEGORY RELIABILITY	73

8.4 RELIABILITY OF ACHIEVEMENT LEVEL CATEGORIZATION	74
8.5 DECISION ACCURACY AND CONSISTENCY RESULTS	75
8.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING RELIABILITY	77
CHAPTER 9. VALIDITY ARGUMENTS SUPPORTING INTENDED INTERPRETATIONS AND USES OF TEST SCORES	78
9.1 RATIONALE FOR VALIDITY ARGUMENT-CENTERED TECHNICAL REPORTING.....	78
9.2 VALIDITY ARGUMENT FOR INTERPRETATION AND USE OF RICAS TEST SCORES	79
9.2.1 CLAIMS SUPPORTING INTENDED INTERPRETATIONS OF RICAS TEST SCORES.....	80
9.2.2 CLAIMS SUPPORTING INTENDED USES OF RICAS TEST SCORES	84
9.3 VALIDATION SUMMARY.....	87
REFERENCES.....	88
APPENDICES.....	91

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	2022-23 MCAS EQUATING REPORT
APPENDIX K	RELIABILITY

List of Tables

Table 1-1 Relationship between 2023 RICAS and MCAS Tests on Critical Test Components	9
Figure 1-1 Logic of Validity Arguments for Tests	12
Figure 1-2 Chappelle (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	15
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) ¹	19
Table 2-5 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grades 3–5	20
Table 2-6 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grades 6 and 7	20
Table 2-7 Target (and Actual) Distribution of Mathematics Common Item Points by Reporting Category, Grade 8	20
Table 2-8 Mathematics Item Types and Score Points	21
Table 2-9 Distribution of Mathematics Common and Matrix Items by Grade and Item Type—Computer-based Test (CBT)	22
Table 2-10 Distribution of Mathematics Common and Matrix Items by Grade and Item Type—Paper-based Test (PBT)	22
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	44
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 ¹	48
Table 6-1 Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade	56
Table 6-2 Multidimensionality Effect Sizes by Grade and Content Area	60
Table 7-1 Number of Cycles Required for Convergence	65
Table 7-2 Stocking and Lord Constants	68
Table 7-3 Scale Score Slopes and Intercepts by Content Area and Grade	69
Table 7-4 Cut Scores on the Theta Metric and Reporting Scale by Content Area and Grade	70
Table 8-1 Raw Score Descriptive Statistics, Cronbach's Alpha, and SEMs by Content Area and Grade	73
Table 8-2 Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Achievement Level	76
Table 8-3 Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint	76
Figure 9-1 Validity Argument Logic Model	79
Figure 9-2 Chappelle (2020)'s Framework: The Arguments and the Inferential Steps	80

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 MCAS and RICAS. Also provided are updates for the 2023 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 ELA and mathematics at grades 3–8 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:

- 1) to provide information to parents/guardians and students on Rhode Island student achievement on the state’s ELA and mathematics content standards,
- 2) to provide information to support program evaluation and improvement at the school and district level, and
- (3) 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, 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 2023 RICAS Technical Report is to document the technical quality and essential design characteristics of the 2023 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.

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) and factor analysis.

1.2 ORGANIZATION OF THIS REPORT

This report provides information regarding the spring 2023 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 2023 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 2023 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
Scoring	
Machine-scored items	Identical
Hand-scored items	Identical
Psychometric Quality	Identical
Reporting	
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 – 2022-23 MCAS Equating Report
- Appendix K – Reliability

1.3 UPDATES FOR THE 2023 ADMINISTRATION

Massachusetts DESE changed in the ELA essay scoring for grades 3 through 8 to remove the dependency between the two trait scores. Until Spring 2022, students in grades 3–8 could only receive up to 1 point for Conventions if they obtained a 0 score on Idea Development. Starting in Spring 2023, students were allowed to receive full credit (up to 3 points) for Conventions regardless of the score on Idea Development. The rule was applied to all 2023 field-test and operational responses for all essays in grades 3–8. Because RIDE administers the MCAS tests for ELA grades 3–8, these changes apply to RICAS.

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:

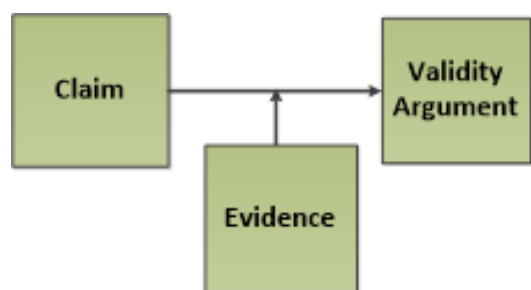
- Parents/guardians and students can use test scores and their underlying interpretations to monitor academic achievement and participate in decisions regarding student learning to support student growth.
- Educators can use test scores and their underlying interpretations to support curricular planning 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 achievement level descriptors for achievement level classifications, item design information for marker items on the scale).

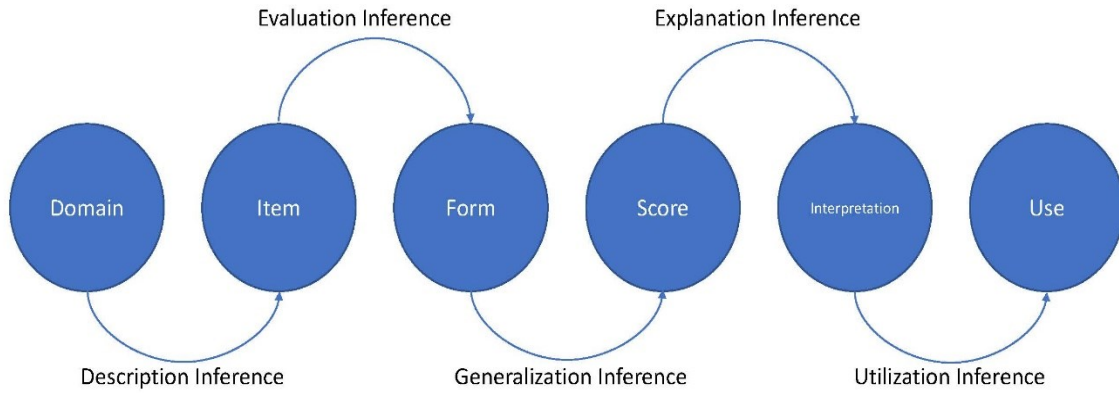
The Standards also provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five areas: test content, response processes, internal structure, relationship to other variables, and consequences 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 Chappelle’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 Chappelle (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 Chappelle (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 2023 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 Standards

The 2023 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)

2.4.2 ELA Blueprints

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

2.4.3 ELA Item Types

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

2.4.4 ELA Passage Types

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 included three passage sets, with forms in some grades containing two literary passage sets and one informational passage set. Forms in other grades contained 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 be 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).

2.4.5 ELA Cognitive Levels

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.

2.4.6 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 2023 ELA test, by grade.

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

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

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

¹ The paper form is derived from Form 1 of the CBT.

2.4.7 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 Standards

The 2023 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

2.5.2 Mathematics Blueprints

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 (31)	20 (20)	15 (15)
Number and Operations in Base Ten	15 (17)	20 (20)	30 (30)
Number and Operations – Fractions	20 (19)	30 (30)	25 (26)
Geometry	10 (8)	10 (9)	10 (9)
Measurement and Data	25 (25)	20 (20)	20 (20)
Total	100	100	100

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)
Total	100	100

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 (37)
Functions	20 (20)
Geometry	30 (30)
Statistics and Probability	10 (13)
Total	100

2.5.3 Mathematics Item Types

The 2023 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

2.5.4 Mathematics Cognitive Levels

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.

2.5.5 Form Test Design

Test 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 two 1-point selected-response or short-answer items, 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 two 1-point selected-response or short-answer items, two 2-point selected-response or short-answer items, 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		CR		Totals	SR/MS SA/TE	CR	Totals
		(1 pt.)	(2 pt.)	(3 pt.)	(4 pt.)	# (pt.)	(1 or 2 pt.)	(3 or 4 pt.)	# (pt.)
3	1	36	0	4	0	40 (48)	3	1	4 (6)
4	1	34	2	0	4	40 (54)	3	1	4 (8-9)
5	1	34	2	0	4	40 (54)	3	1	4 (8-9)
6	1	34	2	0	4	40 (54)	3	1	4 (8-9)
7	1	34	2	0	4	40 (54)	4	2	6 (13-14)
8	1	34	2	0	4	40 (54)	4	2	6 (13-14)

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		CR		Totals	SR/MS/SA	CR	Totals
		(1 pt.)	(2 pt.)	(3 pt.)	(4 pt.)	# (pt.)	(1 or 2 pt.)	(3 or 4 pt.)	# (pt.)
3	1	36	0	4	0	40 (48)	3	1	4 (6)
4	1	34	2	0	4	40 (54)	3	1	4 (8-9)
5	1	34	2	0	4	40 (54)	3	1	4 (8-9)
6	1	34	2	0	4	40 (54)	3	1	4 (8-9)
7	1	34	2	0	4	40 (54)	4	2	6 (13-14)
8	1	34	2	0	4	40 (54)	4	2	6 (13-14)

2.5.6 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.

Table 2-11 Overview of Item and Test Development Process

Phase	Development Step	Detail of the Process
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 RICAS test.
	Development of initial item versions	Cognia's test developers generate items and edit items from subcontractors that are aligned to Massachusetts standards and specifications.
Item Review and Refinement	Review of initial item versions by DESE and educators	<ol style="list-style-type: none"> 1. Cognia sends draft items to DESE test developers for review. 2. DESE test developers review and edit items prior to presenting the items to ADCs. 3. ADCs review items and make recommendations. 4. BSC reviews items and makes recommendations.
	Revision of initial items based on review	DESE test developers edit & revise items based on recommendations from ADC & BSC.
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.
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.
	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.
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).
	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.
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).

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 area 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 2.4.4. 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, the 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 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.

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 Form 1.
- **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 2023 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 3.4.1.

2.7 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 2.4.1 for ELA and 2.5.1 for Mathematics explicitly state that items across all grades within those content areas “were aligned to and measured the ... learning standards as articulated in the Rhode Island Core Standards,” detailing the specific standards addressed by items available for RICAS assessments. Subsections 2.4.3, 2.4.4, and 2.4.5 describe item types, passage types, and cognitive levels for items on the ELA assessments. Subsections 2.5.3 and 2.5.4 describe the item types and cognitive levels for items on the mathematics assessments. Subsection 2.6.1 describes item development and review procedures, and Subsection 2.6.2 describes item field testing and subsequent review, acceptance, and revision

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: Subsections 2.4.2 and 2.4.6 describe the blueprints and test design specifications for ELA, while Subsections 2.5.2 and 2.5.5 cover 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. 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.

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 2022–2023, 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 AND 2022–2023 administrations. For more information, consult Appendix B of the 2021 and 2022 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 2023, 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	4/28	4/17	3/27-4/28	5/1
Mathematics	5/26	4/17	4/24-5/26	5/30

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	Session 2	Total
	Recommended Testing Time (min)	Recommended Testing Time (min)	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 2023 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 2023 Test Coordinator’s Manual (TCM), Grades 3–8* and the *2023 Test Administrator’s Manuals (TAMs)*. In spring 2023, 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, 2023* for details on how ELL students participate in spring 2023 RICAS.
- Students with significant cognitive disabilities who are eligible for the alternate assessment, the Dynamic Learning Maps (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 math assessment as referenced in chapter 2, Subsection 2.5.6.
- ensuring that all students who require supplemental supports (math 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

There were no changes to scoring procedures in 2023 versus the previous administrations. 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 2023 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

Once responses to hand-scored items for a student were sorted into item-specific groups, student responses were scored. In short, all like items (asset ids) are grouped together and scored as a group opposed to the entire student scored as a collective unit, the items are scored as collective units. Scorers within each item group scored one response at a time. However, if there was a need to see a student's responses across all the hand-scored items, scoring leadership had access to the student's entire answer booklet. 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 the same stringent quality control measures that were 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 2023 RICAS responses:

- Cognia Staff
 - The *Scoring Director for Content and Quality* provided guidance, direction, and leadership to RICAS scoring.
 - 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 Cognia.
 - *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 Cognia 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	Scorers		Leadership	
	Number	Percent	Number	Percent
Bachelor's degree or higher	1,853	100%	110	100%
Master's degree or higher	793	43%	54	49%
Teaching Experience				
1-2 years	308	16.6%	15	14%
11 years or more	358	19.3%	21	19%
3-5 years	361	19.5%	18	16%
6-10 years	265	14.3%	9	8%
I have no teaching experience	450	24.3%	41	37%
Less than a year	156	8.4%	12	11%
Scoring Experience				
1 year scoring	891	48%	46	42%
2 years scoring	384	21%	17	15%
3 years scoring	117	6%	8	7%
4 years scoring	80	4%	10	9%
5 years scoring	91	5%	6	5%
6 years scoring	57	3%	5	5%
7 years scoring	30	2%	3	3%
8 years scoring	31	2%	2	2%
9 years scoring	46	2%	4	4%
10 years scoring	33	2%	2	2%
11 years scoring	30	2%	7	6%
12 years scoring	54	3%	--	--
13 years scoring	9	0%	--	--

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 agreement¹ (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 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, a 10% sample of the human scored ELA essay responses were scored via automated scoring using Pearson’s Intelligent Essay Assessor (IEA). All other 3–8 ELA and mathematics responses were done by human scorers at a rate of 10% human-to-human double-blind scoring rate. 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 2023 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:

- 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

¹ “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.

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 student responses received at least one human score. This was the only human response unless the response was independently read and scored by two human scorers (i.e., it was slotted for double-blind scoring).

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 treated as a human scorer 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 views responses recently scored by a particular scorer and assigns a score to that same response. Scoring leadership then compared scores and advised/counseled the scorer as necessary.

4.5.7 Double-Blind Scoring with the Intelligent Essay Assessor (IEA)

The Intelligent Essay Assessor (IEA) is used to score student responses to essay prompts.² 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 used operationally for the fourth consecutive year as the second double-blind score. 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. 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.

² 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 N Counts by Prompt

Grade	Prompt	Training Set Size	Evaluation Set Size
3	EL909882556	1263	631
4	EL007459900	1305	652
5	EL030400392	1283	639
	EL624182427	1195	598
6	EL007051004	1251	624
	EL807016586	1312	656
7	EL006653237	1275	637
	EL713375305	1186	594
8	EL007062902	1278	640
	EL007253494	1268	636

Table 4-3 Metrics for Evaluating Automated Scoring³

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.

³ 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	EL909882556	Idea Development	IEA	136	90%	79%	97%	89%	83%	100%	
			Human		84%	91%	93%	79%	65%	77%	
		Conventions	IEA		87%	100%	91%	74%	88%		
			Human		87%	97%	92%	76%	79%		
4	EL007459900	Idea Development	IEA	79	92%	100%	97%	93%	71%	75%	
			Human		89%	71%	95%	83%	62%	55%	
		Conventions	IEA		98%	50%	100%	100%	91%		
			Human		91%	50%	96%	84%	87%		
5	EL030400392	Idea Development	IEA	104	78%	84%	80%	72%	67%	94%	
			Human		78%	91%	87%	63%	58%	46%	
		Conventions	IEA		83%	91%	50%	85%	91%		
				Human		80%	87%	80%	64%	78%	
	EL624182427	Idea Development	IEA	43	70%	100%	71%	61%	75%	67%	
			Human		76%	80%	84%	72%	72%	37%	
Conventions		IEA	77%		83%	85%	61%	100%			
			Human		77%	75%	81%	74%	74%		
6	EL007051004	Idea Development	IEA	110	86%	86%	85%	86%	92%	63%	0%
			Human		75%	84%	84%	72%	49%	35%	0%
		Conventions	IEA		91%	100%	90%	90%	88%		
				Human		75%	78%	76%	75%	66%	
	EL807016586	Idea Development	IEA	55	87%	100%	75%	100%	50%	67%	100%
			Human		89%	97%	92%	74%	65%	48%	71%
Conventions		IEA	95%		100%	82%	80%	100%			
			Human		90%	96%	87%	64%	84%		
7	EL006653237	Idea Development	IEA	130	85%	95%	96%	85%	93%	31%	100%
			Human		85%	93%	91%	88%	77%	51%	49%
		Conventions	IEA		91%	100%	80%	91%	93%		
				Human		88%	97%	85%	86%	85%	
	EL713375305	Idea Development	IEA	76	92%	100%	100%	87%	92%	90%	85%
			Human		85%	99%	91%	79%	81%	70%	72%
Conventions		IEA	96%		90%	94%	93%	100%			
			Human		90%	90%	89%	79%	96%		
8	EL007062902	Idea Development	IEA	132	96%	100%	90%	96%	95%	96%	100%
			Human		78%	95%	86%	75%	72%	50%	67%
		Conventions	IEA		97%	100%	91%	96%	98%		
				Human		88%	95%	85%	78%	91%	
	EL007253494	Idea Development	IEA	114	87%	100%	100%	76%	89%	71%	0%
			Human		74%	85%	84%	73%	69%	52%	0%
Conventions		IEA	85%		94%	83%	85%	83%			
			Human		79%	91%	80%	73%	79%		

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. Scorers knew in advance that some of the responses they scored were going to be scored by others, but they had no way of knowing what responses would be scored by another scorer, or whether they were the first, second, or only 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

Content Area	Grade	Number of		Percent		Correlation	% of Third Scores	LW Kappa
		Score Categories	Included Scores	Exact	Adjacent			
ELA	3	4	1,859	73.16	25.44	0.79	1.67	0.70
		5	927	72.82	26.32	0.80	1.83	0.69
	4	4	1,890	80.53	18.31	0.79	1.32	0.72
		5	939	77.32	22.26	0.79	0.64	0.70
	5	4	1,894	68.64	30.46	0.79	1.69	0.64
		5	1,894	67.00	31.73	0.78	1.69	0.63
	6	4	1,901	70.23	28.98	0.84	2.52	0.72
		6	1,901	65.70	32.40	0.84	2.52	0.70
	7	4	1,907	73.62	25.64	0.88	2.20	0.77
		6	1,907	76.61	21.50	0.92	2.20	0.81
	8	4	1,907	76.66	22.86	0.89	1.68	0.80
		6	1,907	69.27	29.37	0.89	1.68	0.77

Note. LW = linearly-weighted

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

Content Area	Grade	Number of		Percent		Correlation	% of Third Scores	LW Kappa
		Score Categories	Included Scores	Exact	Adjacent			
Mathematics	3	4	3,903	89.60	10.15	0.95	0.26	0.91
	4	5	3,906	75.40	22.50	0.89	2.10	0.79
	5	5	3,922	81.23	17.54	0.93	1.22	0.85
	6	5	3,881	87.89	11.16	0.96	0.95	0.90
	7	5	3,896	86.47	12.76	0.95	0.77	0.89
	8	5	3,924	78.90	18.83	0.93	2.27	0.82

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 Statistics¹

Subject	Grade	Number of Score Categories ²	Number of Validity Responses ³	Exact Agreement	Agreement by Score Point					
					0	1	2	3	4	5
ELA	3	4 (SR)	3,584	82.5%	93.3%	83.1%	78.5%	31.3%		
		4 (Conv)	3,282	86.6%	97.1%	91.7%	75.5%	78.8%		
		5 (ID)	3,282	84.5%	91.1%	92.6%	79.3%	64.7%	77.0%	
	4	4 (SR)	3,765	83.9%	84.9%	83.9%	81.9%	84.7%		
		4 (Conv)	3,214	90.5%	50.3%	95.6%	83.8%	86.6%		
		5 (ID)	3,214	88.7%	71.2%	95.4%	83.4%	62.2%	55.1%	
	5	4 (Conv)	6,464	78.7%	84.9%	80.7%	71.3%	76.3%		
		5 (ID)	6,464	77.2%	90.1%	84.7%	69.5%	63.3%	42.9%	
	6	4 (Conv)	6,712	82.5%	89.0%	81.7%	74.2%	75.9%		
		6 (ID)	6,712	82.1%	91.6%	88.2%	72.4%	55.3%	41.9%	70.9%
	7	4 (Conv)	6,598	88.7%	93.8%	86.9%	84.0%	91.6%		
		6 (ID)	6,598	85.1%	96.2%	91.1%	85.6%	78.4%	63.4%	69.1%
8	4 (Conv)	7,023	83.5%	93.6%	82.9%	74.5%	84.9%			
	6 (ID)	7,023	76.1%	90.3%	84.9%	73.7%	70.0%	51.2%	66.5%	
Mathematics	3	4	6,901	94.3%	95.6%	93.2%	92.5%	96.6%		
	4	5	7,208	91.4%	91.6%	91.3%	91.1%	87.5%	94.0%	
	5	5	7,100	94.8%	96.1%	94.0%	94.0%	95.5%	94.5%	
	6	5	7,107	94.9%	97.0%	92.8%	95.2%	93.2%	97.1%	
	7	5	7,150	94.1%	98.2%	93.2%	90.8%	91.9%	95.9%	
	8	5	7,327	93.2%	98.4%	92.3%	91.7%	88.6%	93.7%	

¹Includes all operational and equating items for ELA and mathematics.

²SR= Short response; Conv= Conventions; ID=Idea Development

³This 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 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). Parent/Guardian Report PDFs were run 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 a stand-alone single page (11" x 17") color report that is folded; 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 and to enable comparisons with other students at the school, district, and state levels. Three of the four sections are developed by Cognia/Pearson while one section is developed by RIDE.

Outside Portion

The outside portion of the Parent/Guardian Report has two pages, a front page and a back page. The front page 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, range of scores, and growth percentile.

In 2023, QR codes were added to the front page. The QR code leads parents/guardians to a video specific to their student. The video explained the results of the RICAS tests in context. The videos were produced by Pearson’s Spotlight team and were produced in English and in the student’s home language when the home language was in the top 10 languages in Rhode Island. See section 5.4 and an accompanying document on reporting business requirements for more information.

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

Inside Portion

The inside portion of the report also has two pages, one dedicated to ELA results and one dedicated to mathematics results.

Each page contains the achievement level, scaled score, and standard error of the scaled score for each content area tested. 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 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 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.

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 demonstration data to test the application of the decision rules (see Appendix F). The populated reports were then 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>), provides access to the Rhode Island Assessment Data Portal and guides its use. Additionally, it contains links to the following guides for interpretation and use of RICAS results and reports:

A “Family Guide to Understanding RICA (<https://ride.ri.gov/media/28271/download>), which provides not only guidance for properly interpreting RICAS results, but also on appropriate use of the results.

A RICAS Individual Student Report (ISR) Guide (<https://ride.ri.gov/media/28631/download>), explaining 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 2023 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 2023. 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: multiple choice (MC), open response (OR), and writing prompt (WP). 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	33	0.56	0.13	0.47	0.11
		MC	26	0.59	0.10	0.44	0.09
		OR	5	0.55	0.15	0.58	0.07
		WP	2	0.27	0.05	0.64	0.01
	4	ALL	33	0.58	0.14	0.41	0.11
		MC	26	0.61	0.13	0.38	0.09
		OR	5	0.52	0.10	0.47	0.10
		WP	2	0.34	0.09	0.63	0.02
	5	ALL	33	0.60	0.16	0.45	0.13
		MC	24	0.63	0.14	0.40	0.10
		OR	5	0.64	0.11	0.56	0.07
		WP	4	0.35	0.09	0.67	0.02
	6	ALL	33	0.55	0.13	0.44	0.14
		MC	24	0.59	0.11	0.38	0.08
		OR	5	0.53	0.10	0.46	0.14
		WP	4	0.35	0.11	0.73	0.01
	7	ALL	33	0.54	0.14	0.44	0.13
		MC	24	0.58	0.11	0.39	0.06
		OR	5	0.50	0.12	0.47	0.08
		WP	4	0.34	0.12	0.73	0.06
	8	ALL	33	0.59	0.11	0.44	0.15
		MC	24	0.62	0.10	0.39	0.10
		OR	5	0.61	0.09	0.43	0.08
		WP	4	0.43	0.11	0.77	0.02
Mathematics	3	ALL	40	0.58	0.13	0.50	0.11
		MC	16	0.58	0.13	0.44	0.09
		OR	24	0.59	0.14	0.55	0.11
		ALL	40	0.58	0.13	0.48	0.11
	4	MC	20	0.61	0.14	0.42	0.09
		OR	20	0.56	0.12	0.55	0.07
		ALL	40	0.50	0.13	0.45	0.12
		MC	18	0.49	0.15	0.38	0.11
	5	OR	22	0.52	0.11	0.51	0.10
		ALL	40	0.48	0.13	0.46	0.15
		MC	16	0.51	0.13	0.36	0.12
		OR	24	0.46	0.12	0.53	0.12
	6	ALL	40	0.39	0.11	0.48	0.16
		MC	17	0.41	0.11	0.36	0.13
		OR	23	0.37	0.11	0.57	0.12
		ALL	40	0.44	0.14	0.46	0.15
	7	MC	16	0.50	0.12	0.37	0.11
		OR	24	0.40	0.14	0.53	0.14

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 statistics 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 multiple-choice items tend to be higher (indicating that students performed better on these items) than the difficulty indices for open-response items because multiple-choice 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 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 for open-response items 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 2023 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 2023 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 2023 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 multiple-choice 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 2023 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 ELL⁴
- 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 2023 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. 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 2023 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 2023 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 as their basic statistical building block the estimated average conditional covariances for item pairs. 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 strictly unidimensional, all conditional covariances are expected to take on values within random noise of

⁴ ELL = English Language Learner (includes current and former English Language Learners).

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 2023. The data for each grade were split into a training sample and a cross-validation sample. For all grades, there were over 9,600 student examinees per test form in both ELA and mathematics, so every training sample and cross-validation sample had at least 4,800 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 2023 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		
		2021	2022	2023
ELA	3	0.15	0.15	0.14
	4	0.13	0.21	0.17
	5	0.21	0.21	0.17
	6	0.30	0.25	0.22
	7	0.30	0.30	0.34
	8	0.19	0.30	0.32
	Average	0.21	0.24	0.23
Mathematics	3	0.21	0.20	0.17
	4	0.16	0.18	0.14
	5	0.16	0.18	0.10
	6	0.13	0.13	0.15
	7	0.10	0.13	0.13
	8	0.18	0.18	0.14
	Average	0.16	0.17	0.14

The DETECT values indicate weak ($0.2 < \text{DETECT} < 0.4$) or very weak ($\text{DETECT} < 0.2$) multidimensionality for all the 2023 RICAS test forms, which are consistent with 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 multiple-choice and open-response item types. Inspection of the DETECT clusters indicated that multiple-choice/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 multiple-choice 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 multiple-choice and open-response items.

In summary, for the 2023 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 multiple-choice 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 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 SY 22-23 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 (θ), 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 θ). Another way to think of θ is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between θ and $P(\theta)$ (van der Linden, 2016; Hambleton & van der Linden, 1997; Hambleton & Swaminathan, 1985). The process of determining the mathematical relationship between θ 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 θ and $P(\theta)$. Once the item parameters are known, an estimate of θ 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 2023 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:

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

where

U represents the scored response on an item,
 i indexes the items,
 j indexes students,
 α represents item discrimination,
 b represents item difficulty,
 c is the pseudo guessing parameter,
 θ 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:

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

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 θ . This implies that a polytomous item with $(k + 1)$ categories can be characterized by k item category threshold curves (ICTCs) of the 2PL form:

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

where

U indexes the scored response on an item,
 i indexes the items,
 j indexes students,
 k indexes threshold,
 θ is the student's latent person parameter,
 α 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 θ , 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,

θ 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:

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

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, θ_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 θ_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 θ_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, θ_j runs from -4 to 4), and

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

The expected raw score monotonically increases with θ_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 θ_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 θ_j is approximately equal to the inverse of the square root of the statistical information at θ_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 θ 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 2023 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	30	8	
	4	27	7	
	5	43	8	
	6	26	12	
	7	21	9	
	8	119	12	
	Mathematics	3	66	--
		4	71	--
5		50	--	
6		43	--	
7		80	--	
8		33	--	

7.4 EQUATING

Section 7.4 summarizes the equating procedure and results to place the 2023 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 2023 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 2023 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 2023 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]”⁵) 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 β , 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 θ 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.

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.

⁵ 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>

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 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 2023 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 2023 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	1.14	-0.21
	4	1.07	-0.26
	5	1.14	-0.21
	6	1.44	-0.36
	7	1.22	-0.30
	8	1.41	-0.21
Mathematics	3	1.08	-0.03
	4	1.03	0.10
	5	1.01	-0.02
	6	1.03	-0.11
	7	1.11	-0.14
	8	1.10	-0.19

7.5 REPORTED SCALE SCORES AND ACHIEVEMENT STANDARDS

Because the θ 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 θ scale. As the three θ cutpoints from the standard setting have equal intervals (see Section 2.3 for more detail on cuts), one single linear transformation was sufficient to transform the θ 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 2023 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 2023 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 θ . Since the θ 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 θ 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			Min	Scale Score			Max
		Cut 1	Cut 2	Cut 3		Cut 1	Cut 2	Cut 3	
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
	3	-1.377	0.027	1.432	440	470	500	530	560
	4	-1.379	0.054	1.487	440	470	500	530	560
Mathematics	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 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 the SY 22-23 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 2023 RICAS tests is the α coefficient of Cronbach (1951). This approach is most easily understood as an extension of a related procedure, the 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 α 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 α 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 α 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
 σ_x^2 represents the total test variance.

Table 8-1 presents descriptive statistics, Cronbach's α 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 0.94, 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 (α)	SEM
			Maximum	Mean	Standard Deviation		
ELA	3	9,647	44	22.70	9.82	0.91	2.87
	4	9,728	44	23.71	8.59	0.88	2.93
	5	9,707	48	26.14	9.85	0.91	2.96
	6	9,728	50	24.50	10.16	0.90	3.18
	7	9,868	50	23.75	10.89	0.90	3.40
	8	9,958	50	27.22	10.61	0.90	3.29
Mathematics	3	9,792	48	24.54	12.45	0.94	3.03
	4	9,845	54	27.46	12.86	0.93	3.33
	5	9,817	54	24.68	12.05	0.92	3.39
	6	9,806	54	22.44	12.67	0.93	3.45
	7	9,948	54	19.09	12.46	0.93	3.38
	8	10,071	54	22.08	12.24	0.92	3.44

Because of the dependency of the α 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 2023 RICAS online forms. Appendix K presents reliabilities for various subgroups of interest for ELA and mathematics, respectively. Cronbach’s α 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 0.82 to 0.96 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, α , 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 2023 RICAS content areas. Results and reporting category descriptions are presented in Appendix K. The reliability coefficients for the reporting subcategories range from 0.41 to 0.88, with a median of 0.74 and a standard deviation of 0.11. 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 2023 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 estimated 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 2023 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 2023 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 2023 RICAS tests using Cohen’s (1960) coefficient κ (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 κ 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 2023 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 (0.79–0.85), consistency (0.71–0.79), and kappa (0.58–0.67) 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 0.86 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, 86% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.78 indicates that 78% 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 2023 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 0.91–0.94 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	0.82 (0.75)	0.62	0.86 (0.78)	0.81 (0.75)	0.81 (0.74)	0.76 (0.59)
	4	0.81 (0.74)	0.60	0.83 (0.76)	0.83 (0.77)	0.76 (0.70)	0.69 (0.39)
	5	0.83 (0.76)	0.63	0.85 (0.79)	0.83 (0.77)	0.81 (0.75)	0.67 (0.38)
	6	0.79 (0.71)	0.58	0.87 (0.81)	0.79 (0.71)	0.74 (0.66)	0.56 (0.36)
	7	0.81 (0.73)	0.60	0.85 (0.78)	0.81 (0.75)	0.76 (0.67)	0.74 (0.55)
	8	0.79 (0.71)	0.58	0.87 (0.81)	0.79 (0.70)	0.72 (0.63)	0.66 (0.49)
Mathematics	3	0.83 (0.76)	0.66	0.88 (0.82)	0.84 (0.78)	0.80 (0.73)	0.72 (0.55)
	4	0.84 (0.77)	0.66	0.86 (0.81)	0.85 (0.79)	0.81 (0.75)	0.74 (0.55)
	5	0.85 (0.79)	0.66	0.82 (0.76)	0.87 (0.83)	0.84 (0.76)	0.77 (0.57)
	6	0.85 (0.79)	0.67	0.88 (0.81)	0.86 (0.81)	0.80 (0.74)	0.75 (0.51)
	7	0.84 (0.77)	0.66	0.89 (0.83)	0.83 (0.77)	0.79 (0.71)	0.75 (0.56)
	8	0.84 (0.77)	0.65	0.85 (0.80)	0.84 (0.78)	0.81 (0.72)	0.82 (0.66)

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	False		Accuracy	False		Accuracy	False	
		(consistency)	Positive	Negative	(consistency)	Positive	Negative	(consistency)	Positive	Negative
ELA	3	0.94 (0.91)	0.03	0.04	0.91 (0.88)	0.05	0.04	0.97 (0.96)	0.02	0.01
	4	0.93 (0.91)	0.03	0.03	0.91 (0.87)	0.05	0.05	0.97 (0.96)	0.03	0.00
	5	0.94 (0.91)	0.03	0.03	0.91 (0.88)	0.05	0.04	0.97 (0.97)	0.03	0.00
	6	0.92 (0.89)	0.04	0.04	0.91 (0.88)	0.04	0.05	0.96 (0.94)	0.03	0.01
	7	0.92 (0.89)	0.04	0.04	0.92 (0.88)	0.04	0.04	0.97 (0.96)	0.02	0.01
	8	0.92 (0.89)	0.04	0.04	0.92 (0.88)	0.04	0.05	0.95 (0.93)	0.03	0.02
Mathematics	3	0.94 (0.91)	0.03	0.03	0.93 (0.90)	0.03	0.04	0.96 (0.95)	0.02	0.01
	4	0.94 (0.92)	0.03	0.02	0.93 (0.90)	0.04	0.03	0.96 (0.95)	0.03	0.01
	5	0.94 (0.92)	0.03	0.03	0.93 (0.89)	0.04	0.03	0.99 (0.98)	0.01	0.00
	6	0.94 (0.91)	0.03	0.03	0.93 (0.90)	0.03	0.04	0.98 (0.97)	0.02	0.00
	7	0.93 (0.90)	0.03	0.04	0.93 (0.90)	0.03	0.03	0.98 (0.97)	0.02	0.01
	8	0.92 (0.88)	0.05	0.04	0.94 (0.91)	0.04	0.03	0.99 (0.98)	0.01	0.00

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 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 Chappelle (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

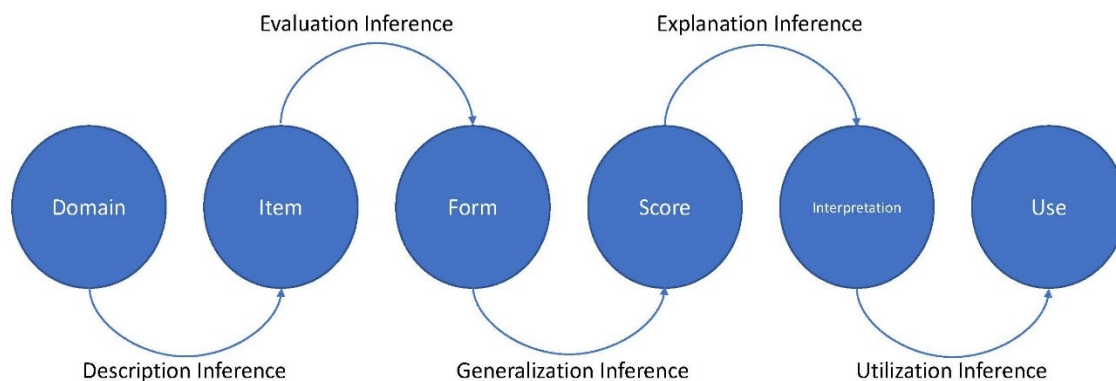
For the RICAS, the overarching validity argument is that the existing design, procedural, and psychometric evidence supports all intended score interpretations and uses of resulting test scores. Each of the interpretation and use inferences is comprised of one or more claims requiring supporting evidence. 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 Chappelle 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 Chappelle's framework.

Figure 9-2 Chappelle (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 2.4.1 for ELA and 2.5.1 for mathematics detail the specific standards addressed by items available for RICAS assessments. Subsections 2.4.3, 2.4.4, and 2.4.5 describe item types, passage types, and cognitive levels for items on the ELA assessments. Subsections 2.5.3 and 2.5.4 describe the item types and cognitive levels for items on the mathematics assessments. Subsection 2.6.1 describes item development and review procedures, and Subsection 2.6.2 describes item field testing and subsequent review, acceptance, and revision 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: Subsections 2.4.2 and 2.4.6 describe the blueprints and test design specifications for ELA, while Subsections 2.5.2 and 2.5.5 cover 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 SY 22-23 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 the SY 22-23 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) find 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 Chappelle et al. (2018):

- Following the Chappelle (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
- 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	1,083	8,564
	4	1,226	8,502
	5	1,187	8,520
	6	1,234	8,494
	7	1,291	8,577
	8	1,126	8,832
Mathematics	3	2,249	7,543
	4	2,248	7,597
	5	2,065	7,752
	6	1,636	8,170
	7	1,691	8,257
	8	1,497	8,574

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	5	32	2	4	4	1
Black on Cream	4	10	0	1	1	0
Black on Light Blue	0	21	0	1	0	0
Black on Light Magenta	0	0	0	0	0	0
White on Black	1	0	1	2	3	1
Yellow on Blue	0	1	1	0	0	0
Dark Gray on Pale Green	0	0	0	0	0	0
Answer Masking	26	88	80	21	10	11
Large Print Test Edition	2	1	0	1	2	1
Screen Reader Edition	0	1	0	1	0	0
Assistive Technology	1	7	13	2	0	1
Braille Test Edition	0	0	0	1	0	1
Human Read Aloud as a Standard Accommodation	0	0	0	0	0	0
Human Read Aloud as a Non-Standard Accommodation	30	33	29	21	24	8
Human Signer as a Standard Accommodation	3	2	1	4	2	4
Human Signer as a Non-Standard Accommodation	1	0	0	0	1	1
Text-to-Speech	160	146	127	73	117	103
Human Scribe as a Standard Accommodation	0	0	0	0	0	0
Human Scribe as a Non-Standard Accommodation	39	44	29	15	12	8
Speech-to-Text as a Standard Accommodation	0	0	0	0	0	0
Speech-to-Text as a Non-Standard Accommodation	97	103	89	76	71	41
Typed Responses	1	0	0	1	0	1
Calculation Device	0	0	0	0	0	0
Spell-checker	38	42	62	34	33	56
Word Prediction	44	74	62	37	42	44
Graphic Organizer/Reference Sheet	777	870	941	902	937	828
Any Other Accommodation	72	79	66	55	45	30
Spanish	0	0	0	0	0	0
Bilingual Dictionary and Glossary	190	213	130	287	303	271

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	8	31	1	3	3	1
Black on Cream	7	10	0	1	1	0
Black on Light Blue	0	20	0	1	0	0
Black on Light Magenta	0	0	0	0	0	0
White on Black	1	1	1	1	2	1
Yellow on Blue	0	0	0	0	0	0
Dark Gray on Pale Green	0	0	0	0	0	0
Answer Masking	30	83	78	23	6	9
Large Print Test Edition	4	6	8	9	4	4
Screen Reader Edition	0	0	0	1	0	0
Assistive Technology	3	6	4	1	0	1
Braille Test Edition	3	6	8	8	4	4
Human Read Aloud as a Standard Accommodation	88	73	59	28	30	9
Human Read Aloud as a Non-Standard Accommodation	0	0	0	0	0	0
Human Signer as a Standard Accommodation	4	2	1	3	3	4
Human Signer as a Non-Standard Accommodation	0	0	0	0	0	0
Text-to-Speech	1,876	1,817	1,610	1,050	1,019	771
Human Scribe as a Standard Accommodation	53	43	25	8	8	3
Human Scribe as a Non-Standard Accommodation	0	0	0	0	0	0
Speech-to-Text as a Standard Accommodation	64	65	54	34	41	29
Speech-to-Text as a Non-Standard Accommodation	0	0	0	0	0	0
Typed Responses	3	6	8	8	4	4
Calculation Device	75	118	102	142	257	255
Spell-checker	0	0	0	0	0	0
Word Prediction	0	0	0	0	0	0
Graphic Organizer/Reference Sheet	776	865	932	832	806	745
Any Other Accommodation	0	0	0	0	0	0
Spanish	128	125	111	115	97	131
Bilingual Dictionary and Glossary	179	238	137	284	318	263

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	58,636	100.00
ELL	9,713	16.56
Economically Disadvantaged	27,599	47.07
African American	5,179	8.83
Asian	1,938	3.31
Hispanic	17,138	29.23
Native American/Alaska Native	451	0.77
White	30,732	52.41
Pacific Islander/Hawaiian	93	0.16
Multiracial	3,105	5.30
Male	30,047	51.24
Female	28,564	48.71
Special Education	9,669	16.49

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

Description	Number Tested	Percent Tested
All Students	59,279	100.00
ELL	10,423	17.58
Economically Disadvantaged	27,980	47.20
African American	5,244	8.85
Asian	2,000	3.37
Hispanic	17,531	29.57
Native American/Alaska Native	467	0.79
White	30,843	52.03
Pacific Islander/Hawaiian	93	0.16
Multiracial	3,101	5.23
Male	30,342	51.19
Female	28,909	48.77
Special Education	9,652	16.28

APPENDIX C

INTERRATER CONSISTENCY

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL028832702	4	932	70.17	28.22	0.78	1.50	0.67
EL909882556#SCORE_TRAIT_Conv	4	927	76.16	22.65	0.81	1.83	0.72
EL909882556#SCORE_TRAIT_Ideadev	5	927	72.82	26.32	0.80	1.83	0.69

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL007459900#SCORE_TRAIT_Conv	4	939	79.98	19.70	0.76	0.64	0.68
EL007459900#SCORE_TRAIT_Ideadev	5	939	77.32	22.26	0.79	0.64	0.70
EL024539092	4	951	81.07	16.93	0.80	2.00	0.75

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL030400392#SCORE_TRAIT_Conv	4	942	69.00	29.30	0.76	2.76	0.64
EL030400392#SCORE_TRAIT_Ideadev	5	942	66.99	31.10	0.74	2.76	0.61
EL624182427#SCORE_TRAIT_Conv	4	952	68.28	31.62	0.79	0.63	0.65
EL624182427#SCORE_TRAIT_Ideadev	5	952	67.02	32.35	0.81	0.63	0.66

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL007051004#SCORE_TRAIT_Conv	4	950	73.47	26.21	0.87	2.21	0.76
EL007051004#SCORE_TRAIT_Ideadev	6	950	68.32	29.79	0.84	2.21	0.71
EL807016586#SCORE_TRAIT_Conv	4	951	66.98	31.76	0.82	2.84	0.68
EL807016586#SCORE_TRAIT_Ideadev	6	951	63.09	35.02	0.84	2.84	0.69

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL006653237#SCORE_TRAIT_Conv	4	956	69.35	29.60	0.86	2.09	0.73
EL006653237#SCORE_TRAIT_Ideadev	6	956	69.46	28.87	0.88	2.09	0.75
EL713375305#SCORE_TRAIT_Conv	4	951	77.92	21.66	0.90	2.31	0.81
EL713375305#SCORE_TRAIT_Ideadev	6	951	83.81	14.09	0.95	2.31	0.87

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL007062902#SCORE_TRAIT_Conv	4	965	75.03	24.56	0.88	1.76	0.78
EL007062902#SCORE_TRAIT_Ideadev	6	965	66.22	32.12	0.89	1.76	0.76
EL007253494#SCORE_TRAIT_Conv	4	942	78.34	21.13	0.90	1.59	0.81
EL007253494#SCORE_TRAIT_Ideadev	6	942	72.40	26.54	0.90	1.59	0.79

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA253711A	4	960	89.69	10.21	0.94	0.10	0.90
MA253711A_ES	4	13	100.00	0.00	1.00	0.00	--
MA286750A	4	962	88.36	11.43	0.95	0.21	0.90
MA286750A_ES	4	13	100.00	0.00	1.00	0.00	--
MA286750A_PA	4	3	100.00	0.00	1.00	0.00	--
MA293460A	4	964	86.93	12.66	0.95	0.41	0.90
MA293460A_ES	4	13	100.00	0.00	1.00	0.00	--
MA310899A	4	962	92.83	6.86	0.95	0.31	0.93
MA310899A_ES	4	13	100.00	0.00	1.00	0.00	--

_ES denotes Spanish items / _PA denotes Paper items.

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA293812	5	956	80.75	17.99	0.91	1.26	0.83
MA293812_ES	5	13	100.00	0.00	1.00	0.00	--
MA303335	5	968	72.11	25.62	0.88	2.27	0.78
MA303335_ES	5	13	100.00	0.00	1.00	0.00	--
MA307317	5	959	76.96	20.33	0.89	2.71	0.81
MA307317_ES	5	13	100.00	0.00	1.00	0.00	--
MA801035466	5	971	70.55	27.19	0.87	2.27	0.76
MA801035466_ES	5	13	100.00	0.00	1.00	0.00	--

_ES denotes Spanish items / _PA denotes Paper items.

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA002343629	5	963	83.70	15.68	0.93	0.62	0.87
MA002343629_ES	5	12	100.00	0.00	1.00	0.00	--
MA005852277	5	974	89.43	10.06	0.95	0.51	0.91
MA005852277_ES	5	12	100.00	0.00	1.00	0.00	--
MA802371654	5	966	81.26	17.91	0.94	0.83	0.87
MA802371654_ES	5	11	100.00	0.00	1.00	0.00	--
MA903746975	5	972	69.65	27.37	0.87	2.98	0.76
MA903746975_ES	5	11	100.00	0.00	1.00	0.00	--
MA903746975_PA	5	1	100.00	0.00	--	0.00	--

_ES denotes Spanish items / _PA denotes Paper items.

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA290253	5	958	82.57	16.18	0.93	1.25	0.86
MA290253_ES	5	12	100.00	0.00	1.00	0.00	--
MA298252	5	952	92.65	6.51	0.97	0.84	0.93
MA298252_ES	5	11	100.00	0.00	1.00	0.00	--
MA800301627	5	958	86.22	13.15	0.96	0.63	0.91
MA800301627_ES	5	10	100.00	0.00	--	0.00	--
MA900378821	5	968	89.57	9.30	0.94	1.14	0.89
MA900378821_ES	5	12	100.00	0.00	1.00	0.00	--

_ES denotes Spanish items / _PA denotes Paper items.

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA002119133	5	975	80.51	18.56	0.94	0.92	0.87
MA002119133_ES	5	9	100.00	0.00	1.00	0.00	--
MA261648	5	973	85.41	13.46	0.96	1.13	0.89
MA261648_ES	5	9	100.00	0.00	1.00	0.00	--
MA717248260	5	966	89.13	10.35	0.95	0.52	0.90
MA717248260_ES	5	8	100.00	0.00	--	0.00	--
MA802907874	5	950	90.53	8.95	0.96	0.53	0.92
MA802907874_ES	5	6	100.00	0.00	1.00	0.00	--

_ES denotes Spanish items / _PA denotes Paper items.

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

Item Number	Number of		Percent		Correlation	Percent of Third Scores	LW Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA010701848	5	978	70.35	25.05	0.87	4.60	0.76
MA010701848_ES	5	14	100.00	0.00	1.00	0.00	--
MA301714	5	950	80.11	18.00	0.91	1.89	0.82
MA301714_ES	5	10	100.00	0.00	--	0.00	--
MA311433	5	974	82.14	16.43	0.91	1.44	0.84
MA311433_ES	5	10	100.00	0.00	1.00	0.00	--
MA800738445	5	978	82.11	16.67	0.94	1.23	0.87
MA800738445_ES	5	10	100.00	0.00	1.00	0.00	--

_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		
		2023	2022	2021
3	Not Meeting Expectations	21.13	19.41	13.76
	Partially Meeting Expectations	42.03	44.03	45.88
	Meeting Expectations	31.54	31.57	35.72
	Exceeding Expectations	5.30	4.98	4.64
4	Not Meeting Expectations	19.91	21.67	16.24
	Partially Meeting Expectations	46.79	49.30	48.30
	Meeting Expectations	29.71	26.49	32.83
	Exceeding Expectations	3.59	2.55	2.63
5	Not Meeting Expectations	20.35	17.74	18.36
	Partially Meeting Expectations	44.55	50.51	48.26
	Meeting Expectations	32.17	28.25	30.00
	Exceeding Expectations	2.94	3.50	3.38
6	Not Meeting Expectations	31.12	32.10	28.21
	Partially Meeting Expectations	36.94	36.72	39.23
	Meeting Expectations	27.16	25.72	26.04
	Exceeding Expectations	4.78	5.47	6.53
7	Not Meeting Expectations	27.50	28.80	26.48
	Partially Meeting Expectations	43.45	42.01	44.83
	Meeting Expectations	24.41	25.74	25.10
	Exceeding Expectations	4.63	3.46	3.59
8	Not Meeting Expectations	31.50	27.54	26.89
	Partially Meeting Expectations	36.28	43.45	44.33
	Meeting Expectations	25.60	24.87	25.60
	Exceeding Expectations	6.62	4.14	3.18

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

Grade	Achievement Level	Percent in Level		
		2023	2022	2021
3	Not Meeting Expectations	25.20	24.84	35.38
	Partially Meeting Expectations	40.33	40.14	39.55
	Meeting Expectations	28.99	30.99	23.04
	Exceeding Expectations	5.47	4.04	2.03
4	Not Meeting Expectations	22.32	23.68	33.89
	Partially Meeting Expectations	41.67	46.13	45.28
	Meeting Expectations	30.66	27.02	19.07
	Exceeding Expectations	5.36	3.17	1.75
5	Not Meeting Expectations	17.90	23.90	28.70
	Partially Meeting Expectations	52.13	50.16	51.03
	Meeting Expectations	27.49	24.31	19.06
	Exceeding Expectations	2.48	1.64	1.21
6	Not Meeting Expectations	24.09	23.77	32.07
	Partially Meeting Expectations	47.18	48.87	50.15
	Meeting Expectations	25.99	25.65	16.46
	Exceeding Expectations	2.74	1.72	1.32
7	Not Meeting Expectations	31.83	29.61	30.80
	Partially Meeting Expectations	42.48	47.67	48.85
	Meeting Expectations	22.34	20.14	18.51
	Exceeding Expectations	3.36	2.57	1.84
8	Not Meeting Expectations	30.86	28.65	36.57
	Partially Meeting Expectations	46.16	50.59	47.38
	Meeting Expectations	19.92	18.40	14.91
	Exceeding Expectations	3.06	2.36	1.14

APPENDIX E

SAMPLE REPORTS

Rhode Island Education

Spring 2023 RICAS Individual Student Report



Name: District:
SASID: School:
Date of Birth: Grade: 3

A State on the Move

To prepare students for lifelong success, Rhode Island is committed to providing the academic, social-emotional, and wrap-around services that our communities need year-round. Additionally, RIDE recognizes that to move Rhode Island forward, we must invest in *all* students. In accordance with the Learning, Equity, & Accelerated Pathways Task Force Report, RIDE and schools statewide are elevating and centering the needs of historically underserved students.

demonstrating that we are moving in the right direction. While a lot of work remains ahead to get our students back up to speed, Rhode Island is on its way to recovery. Together, we can rise to the challenge before us and create a new and better system—one that gives every student the support and opportunities they need in order to succeed, no matter their circumstance.

Join us to
improve education!



Scan the QR code to access important information and resources for your family.

From implementing high-quality curriculum to the approval of a historic \$1.5 billion in local and state school construction referendums in 2022, Rhode Island is reimagining its education system and empowering students to create their own futures. The 2022 RICAS results showed a significant increase in math proficiency,

This report provides your child's results from the 2023 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.

These tests 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 an advocate for your child. For more information on how to better understand the results, visit <https://ride.ri.gov/instruction-assessment/assessment/assessment-results>.



Scan for a personalized video about your child's 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

Rhode Island's Strategic Plan for Public Education: 2022-2027

This comprehensive plan lays out a vision for an education system that offers every Rhode Island student the opportunity to be a **lifelong learner**. Five key priorities of the Strategic Plan are supported by a framework of commitments and measurable goals.

Equity

Excellence in Learning

Engaged Communities

World Class Talent

Governance Structures

Your Child's Overall Results

English Language Arts

Achievement Level

Partially Meeting Expectations

Score
498

(Score range: 440-560)

Details on page 2

Mathematics

Achievement Level

Meeting Expectations

Score
507

(Score range: 440-560)

Details on page 3

Rhode Island Education

Spring 2023 RICAS Individual Student Report



Name: District:
SASID: School:
Date of Birth: Grade: 4

A State on the Move

To prepare students for lifelong success, Rhode Island is committed to providing the academic, social-emotional, and wrap-around services that our communities need year-round. Additionally, RIDE recognizes that to move Rhode Island forward, we must invest in *all* students. In accordance with the Learning, Equity, & Accelerated Pathways Task Force Report, RIDE and schools statewide are elevating and centering the needs of historically underserved students.

demonstrating that we are moving in the right direction. While a lot of work remains ahead to get our students back up to speed, Rhode Island is on its way to recovery. Together, we can rise to the challenge before us and create a new and better system—one that gives every student the support and opportunities they need in order to succeed, no matter their circumstance.

Join us to
improve education!



Scan the QR code to access important information and resources for your family.

From implementing high-quality curriculum to the approval of a historic \$1.5 billion in local and state school construction referendums in 2022, Rhode Island is reimagining its education system and empowering students to create their own futures. The 2022 RICAS results showed a significant increase in math proficiency,

This report provides your child's results from the 2023 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.

These tests 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 an advocate for your child. For more information on how to better understand the results, visit <https://ride.ri.gov/instruction-assessment/assessment/assessment-results>.



Scan for a personalized video about your child's 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

Rhode Island's Strategic Plan for Public Education: 2022-2027

This comprehensive plan lays out a vision for an education system that offers every Rhode Island student the opportunity to be a **lifelong learner**. Five key priorities of the Strategic Plan are supported by a framework of commitments and measurable goals.

Equity

Excellence in Learning

Engaged Communities

World Class Talent

Governance Structures

Your Child's Overall Results

English Language Arts
Achievement Level
Partially Meeting Expectations

Score
484
(Score range: 440-560)
Growth Percentile
12
Details on page 2

Mathematics
Achievement Level
Meeting Expectations

Score
504
(Score range: 440-560)
Growth Percentile
40
Details on page 3

APPENDIX F

2023 REPORTING BUSINESS REQUIREMENTS

Reporting Business Requirements (2023)

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

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

DATA FILE DELIVERABLES.....	13
I. MEGAFILE	13
REPORTING PRODUCTS: INTERNAL TO COGNIA	13
APPENDIX.....	14
ADDENDA	14

Year to Year Change Highlights:

1. West Bay Christian Academy will not be participating in the RICAS in 2023.
2. Student results labels are no longer a reporting deliverable for RICAS.

I. Contract Overview

Contract Code: 104650

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. Accommodations datafile
6. Response Change Analysis from Pearson
7. Printed Individual Student Results Reports
8. 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

- 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.
- Data also includes flags indicating an item will be released.
- 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
 - 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.

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	

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 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:**

- If the student *meets attemptedness* then: **Tested** (PartStatus = 'Z').
- 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 or results label.
- 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.

B. Delivery

- i.* 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.
- ii.* 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.

C. Cover Page

- i.* Title – “Spring 20YY RICAS Parent Guardian Report” where 20YY = test year, e.g. 2018.

- ii. 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”
- iii. SASID – Student ID from eRIDE, no special formatting applied.
- iv. School Name and District Name – School and District names from iCore based on testing school. No special formatting applied.
- v. Grade – Student’s tested grade will be used for all 03-08 reports.
- vi. DOB – DOB from eRIDE, no special formatting. Must be equal to 10 characters in length (MM/DD/YYYY).

D. Reporting Category Display

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

E. Released Item Display

- i. *Subject Title* and *Subject Ordering* follow the same rules as above.
- ii. 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.
- iii. 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.

F. Student Achievement Level and Scaled Score Statements

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

- G. Historical scores- Historical scores are reported in grades 4-8. Up to three years of scores are reported. The years reported in 2023 are 2023, 2022,2021 if available.
- H. Spotlight Videos are being produced for each student receiving a student report.
 - QR codes will be used to provide the link from the student report to the video for the specific student.
 - The SASID is used to link the student’s data to their video and is included in the file from Pearson, with the links, to Cognia.
 - The QR codes are on the front page of the student report.
 - All videos are produced in English with some students also receiving an additional video in a language other than English. The source of language for the second video is a field in eRIDE.
 - The additional languages to be used are: Crioulo, Spanish, Chinese, French, Portuguese, Khmer, Vietnamese, Hindi, Swahili and Arabic. The coding in the links file from Pearson are as follows:
 - English - en
 - Spanish - es
 - Crioulo - kea
 - Chinese - zh
 - Portuguese - pt
 - French - fr
 - Khmer - km
 - Vietnamese - vi
 - Hindi - hi
 - Swahili - sw
 - Arabic - ar
 - The State level growth mean is fixed at 50.

Data File Deliverables

I. Megafile

- i. The megafile contains all the student level results for RICAS. The file follows the ResultsLayout tab in the RICAS_2023_ReportingLayout.xlsx
- ii. The file is formatted as comma separated file, csv.
- iii. 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

Contract Code: [104650] Description RICAS 2023 Admin ID 1	Report Type	Report For	Grade(s)	Report Subtype	Content Code	Qty
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.

Subject	Grade	RepCat Sort Order	Student Report Text: (RepCatText)
ELA	03-08	1	Reading
		2	Language
		3	Writing
Mathematics	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
Mathematics	06-07	1	Ratios and Proportional Relationships
		2	The Number System
		3	Expressions and Equations
		4	Geometry
		5	Statistics and Probability
	08	1	The Number System and Expressions and Equations
		2	Functions
		3	Geometry
		4	Statistics and Probability

Addenda

5/17/23: A Grade 7 Math item had a typo that may have affected how students responded to the item. The UIN is MA002119133_ES. The item is a 4-point 4-part item. RIDE decided to give credit to the affected students for the item. Pearson will provide a list of the affected students.

APPENDIX G
ITEM-LEVEL CLASSICAL STATISTICS

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

Item		Type	Difficulty	Discrimination	Percent Omitted (%)
Number					
EL909469479		MC	0.69	0.57	--
EL909470939		MC	0.48	0.38	0
EL909472828		MC	0.63	0.47	0
EL909473433		MC	0.56	0.48	0
EL909478450		MC	0.79	0.52	0
EL909479176		MC	0.78	0.42	0
EL909480023		MC	0.56	0.43	0
EL909865165		MC	0.70	0.52	0
EL909869416		MC	0.68	0.57	0
EL909882556#SCORE_TRAIT_Conv		WP	0.30	0.64	1
EL909882556#SCORE_TRAIT_Ideadev		WP	0.23	0.63	1
EL911945550		OR	0.64	0.64	0
EL912636232		MC	0.68	0.55	0
EL919652746		MC	0.55	0.60	0
EL015503476		MC	0.55	0.47	0
EL015607246		MC	0.50	0.40	0
EL015611981		MC	0.42	0.27	0
EL015628878		MC	0.53	0.44	0
EL015629502		MC	0.70	0.45	--
EL015630515		MC	0.41	0.20	0
EL019650296		OR	0.72	0.47	6
EL028131763		MC	0.64	0.50	0
EL028160220		MC	0.51	0.38	0
EL028215856		OR	0.47	0.56	0
EL028217140		MC	0.59	0.43	0
EL028217627		MC	0.63	0.45	0
EL028709466		MC	0.49	0.31	0
EL028717847		MC	0.51	0.41	0
EL028753268		OR	0.59	0.62	1
EL028823014		MC	0.57	0.44	0
EL028832702		OR	0.33	0.62	1
EL028907379		MC	0.61	0.47	0
EL028914863		MC	0.47	0.32	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
EL006458075		MC	0.87	0.41	0
EL006549511		MC	0.60	0.49	0
EL007440160		MC	0.77	0.38	0
EL007444742		MC	0.73	0.50	0
EL007446608		MC	0.71	0.41	0
EL007452066		MC	0.83	0.46	0
EL007459900#SCORE_TRAIT_Conv		WP	0.40	0.64	1
EL007459900#SCORE_TRAIT_Ideadev		WP	0.27	0.61	1
EL007464016		OR	0.68	0.33	0
EL009343264		OR	0.56	0.57	0
EL009344832		MC	0.57	0.31	0
EL013314332		MC	0.50	0.32	0
EL014208236		MC	0.68	0.48	0
EL024031609		MC	0.50	0.25	0
EL024132276		MC	0.70	0.51	--
EL024134327		MC	0.61	0.53	0
EL024148759		MC	0.80	0.41	0
EL024437543		MC	0.46	0.32	0
EL024440140		MC	0.66	0.34	--
EL024442344		MC	0.40	0.26	0
EL024455037		MC	0.63	0.28	0
EL024480931		OR	0.50	0.46	--
EL024527106		MC	0.56	0.26	0
EL024532504		MC	0.36	0.19	0
EL024539092		OR	0.47	0.56	0
EL029280667		MC	0.51	0.41	0
EL029323184		OR	0.41	0.44	0
EL029415206		MC	0.60	0.39	0
EL029417376		MC	0.60	0.44	0
EL029429384		MC	0.46	0.26	--
EL030430678		MC	0.54	0.41	0
EL030741768		MC	0.57	0.43	0
EL033943069		MC	0.69	0.44	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
EL029961019		MC	0.69	0.44	0
EL029964018		OR	0.55	0.47	0
EL029974201		MC	0.64	0.55	0
EL029980757		MC	0.53	0.38	0
EL030062229		MC	0.72	0.49	0
EL030080040		MC	0.71	0.46	0
EL030171711		MC	0.64	0.43	0
EL030337822		MC	0.36	0.21	0
EL030400392#SCORE_TRAIT_Conv		WP	0.31	0.69	1
EL030400392#SCORE_TRAIT_Ideadev		WP	0.24	0.67	1
EL030463527		OR	0.58	0.60	0
EL030483081		MC	0.77	0.50	0
EL030659380		MC	0.37	0.27	--
EL624175088		MC	0.75	0.51	0
EL624176168		OR	0.82	0.58	0
EL624176741		MC	0.46	0.32	0
EL624177026		MC	0.77	0.25	0
EL624177447		OR	0.66	0.51	0
EL624178677		MC	0.64	0.31	0
EL624179162		MC	0.77	0.48	0
EL624179855		MC	0.69	0.41	--
EL624180347		MC	0.71	0.35	--
EL624180539		MC	0.73	0.41	0
EL624182427#SCORE_TRAIT_Conv		WP	0.46	0.66	0
EL624182427#SCORE_TRAIT_Ideadev		WP	0.37	0.65	0
EL627148548		MC	0.62	0.46	0
EL033541180		MC	0.83	0.44	0
EL033604260		MC	0.76	0.52	--
EL033646585		OR	0.59	0.64	--
EL033665287		MC	0.54	0.28	0
EL033679189		MC	0.56	0.51	0
EL033800505		MC	0.48	0.28	0
EL033843854		MC	0.43	0.28	--

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
EL806979864		MC	0.67	0.47	0
EL807001596		MC	0.60	0.32	0
EL807002174		MC	0.69	0.34	0
EL807009150		MC	0.67	0.32	0
EL807010236		MC	0.41	0.30	0
EL807011414		MC	0.55	0.34	0
EL807011890		MC	0.63	0.37	0
EL807016586#SCORE_TRAIT_Conv		WP	0.43	0.73	1
EL807016586#SCORE_TRAIT_Ideadev		WP	0.28	0.72	1
EL807061702		MC	0.63	0.22	0
EL807062301		OR	0.42	0.46	0
EL808245411		MC	0.68	0.36	0
EL808246461		OR	0.57	0.27	0
EL006639933		MC	0.47	0.38	0
EL006738734		MC	0.41	0.36	0
EL006742548		MC	0.30	0.30	0
EL007051004#SCORE_TRAIT_Conv		WP	0.45	0.73	1
EL007051004#SCORE_TRAIT_Ideadev		WP	0.24	0.73	1
EL007074213		MC	0.77	0.44	0
EL007074445		MC	0.62	0.43	0
EL007075911		MC	0.66	0.55	0
EL007076177		MC	0.68	0.36	0
EL007077078		MC	0.71	0.47	0
EL007077860		MC	0.55	0.36	0
EL007078526		MC	0.65	0.45	0
EL008181021		OR	0.60	0.56	0
EL008281454		MC	0.50	0.24	0
EL008355		OR	0.62	0.63	0
EL008445593		MC	0.61	0.49	0
EL009438210		OR	0.42	0.38	0
EL009514238		MC	0.55	0.45	0
EL009564267		MC	0.44	0.36	0
EL009978066		MC	0.59	0.47	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
EL713370326		OR	0.60	0.44	0
EL713375305#SCORE_TRAIT_Conv		WP	0.43	0.76	1
EL713375305#SCORE_TRAIT_Ideadev		WP	0.19	0.64	1
EL713378067		MC	0.69	0.37	0
EL713475622		OR	0.44	0.53	0
EL713476504		MC	0.45	0.28	0
EL713479099		MC	0.57	0.39	0
EL713479631		MC	0.71	0.44	0
EL713480064		MC	0.55	0.38	0
EL713480958		MC	0.67	0.41	0
EL713481518		MC	0.58	0.46	0
EL723632935		MC	0.54	0.36	0
EL730170770		MC	0.73	0.48	0
EL006357067		MC	0.67	0.45	0
EL006439224		MC	0.47	0.30	0
EL006446884		MC	0.43	0.36	0
EL006454205		MC	0.72	0.48	0
EL006537445		MC	0.49	0.34	0
EL006544709		MC	0.33	0.33	0
EL006545529		MC	0.48	0.31	0
EL006546235		MC	0.58	0.34	0
EL006560393		MC	0.51	0.33	0
EL006640130		MC	0.50	0.43	0
EL006653237#SCORE_TRAIT_Conv		WP	0.44	0.76	1
EL006653237#SCORE_TRAIT_Ideadev		WP	0.29	0.75	1
EL006977006		MC	0.58	0.38	0
EL006978834		OR	0.44	0.41	0
EL011353608		OR	0.37	0.41	0
EL011362473		MC	0.78	0.41	--
EL011363661		MC	0.66	0.42	0
EL016833358		MC	0.53	0.42	0
EL017655451		OR	0.65	0.58	0
EL113429887		MC	0.67	0.46	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
EL007061131		MC	0.71	0.35	0
EL007061194		MC	0.65	0.44	0
EL007061650		MC	0.71	0.20	0
EL007062053		MC	0.57	0.38	0
EL007062608		OR	0.71	0.38	0
EL007062902#SCORE_TRAIT_Conv		WP	0.56	0.77	2
EL007062902#SCORE_TRAIT_Ideadev		WP	0.37	0.74	2
EL009149967		MC	0.73	0.41	0
EL009246409		MC	0.71	0.46	0
EL009257746		OR	0.56	0.44	0
EL009308236		MC	0.76	0.58	--
EL009308819		MC	0.53	0.31	0
EL009343097		MC	0.81	0.33	0
EL006653570		OR	0.60	0.34	1
EL006655733		MC	0.45	0.45	0
EL007243506		OR	0.51	0.56	0
EL007253045		MC	0.67	0.54	0
EL007253494#SCORE_TRAIT_Conv		WP	0.47	0.78	2
EL007253494#SCORE_TRAIT_Ideadev		WP	0.30	0.77	2
EL007256618		MC	0.64	0.41	0
EL007257202		MC	0.67	0.22	0
EL007257390		MC	0.59	0.34	0
EL007335795		MC	0.62	0.42	0
EL007335808		MC	0.50	0.41	0
EL007350397		MC	0.48	0.34	0
EL007353056		MC	0.49	0.38	0
EL008544460		MC	0.55	0.42	0
EL008553781		OR	0.69	0.45	0
EL009737508		MC	0.53	0.35	0
EL016259168		MC	0.69	0.53	0
EL016259978		MC	0.58	0.30	0
EL016352526		MC	0.57	0.23	--
EL022460231		MC	0.63	0.51	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
MA212474		MC	0.74	0.54	0
MA227485		MC	0.38	0.41	0
MA293460A		OR	0.49	0.71	0
MA310884		OR	0.61	0.60	0
MA310899A		OR	0.44	0.70	1
MA735663821		OR	0.43	0.55	0
MA306285		MC	0.55	0.44	0
MA900425126		OR	0.38	0.47	1
MA000749172		OR	0.60	0.61	0
MA001038775		OR	0.78	0.42	0
MA001047582		OR	0.74	0.50	0
MA001049099		MC	0.43	0.25	0
MA001056175		MC	0.44	0.51	0
MA001344527		OR	0.72	0.23	0
MA001439533		OR	0.72	0.45	0
MA297500		MC	0.46	0.38	0
MA703078093		OR	0.83	0.50	0
MA735736004A		OR	0.57	0.52	0
MA935136577		OR	0.64	0.50	0
MA202994		MC	0.76	0.47	0
MA253641		MC	0.73	0.52	--
MA260962		MC	0.55	0.42	0
MA227232		MC	0.64	0.55	--
MA310880		MC	0.43	0.27	0
MA310889		OR	0.70	0.52	0
MA703056978		OR	0.56	0.56	0
MA207001		MC	0.72	0.51	0
MA286750A		OR	0.51	0.70	0
MA310870		MC	0.62	0.37	0
MA713536927		OR	0.66	0.50	0
MA735756531		OR	0.36	0.54	0
MA735954511		OR	0.47	0.60	0
MA253711A		OR	0.42	0.72	0
MA309747		OR	0.50	0.64	0
MA900372676		MC	0.51	0.45	0
MA001137862		MC	0.68	0.50	0
MA001335228		OR	0.54	0.60	1
MA001338241		OR	0.74	0.49	0
MA001633319		MC	0.57	0.43	0
MA734752477		OR	0.73	0.48	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
MA247691		MC	0.85	0.39	0
MA247705		MC	0.51	0.42	0
MA307060		MC	0.57	0.36	0
MA311568		OR	0.62	0.48	0
MA297614		MC	0.67	0.33	0
MA303335		OR	0.48	0.64	0
MA713677363		OR	0.73	0.48	0
MA800727128		OR	0.47	0.50	0
MA801035466		OR	0.45	0.66	0
MA803956738		OR	0.63	0.55	0
MA311543		MC	0.67	0.44	0
MA903776098		MC	0.38	0.38	0
MA000732007		OR	0.48	0.49	0
MA002034926		MC	0.72	0.43	0
MA002128911		OR	0.65	0.50	0
MA002139080		MC	0.50	0.25	0
MA002334462		OR	0.57	0.57	0
MA003744055		OR	0.65	0.67	1
MA003747173		MC	0.48	0.29	0
MA301798		MC	0.67	0.40	0
MA279759		MC	0.72	0.47	0
MA307075		MC	0.89	0.40	0
MA311567		OR	0.54	0.51	0
MA307317		OR	0.46	0.62	0
MA704653374		OR	0.66	0.49	0
MA293812		OR	0.40	0.64	0
MA307067		MC	0.72	0.55	0
MA736377105		OR	0.49	0.58	0
MA803747806		MC	0.58	0.51	0
MA900756471		OR	0.57	0.59	0
MA900846441		OR	0.32	0.50	0
MA903571693		MC	0.47	0.59	0
MA001750121		MC	0.58	0.47	0
MA001851276		OR	0.63	0.51	0
MA002135528		MC	0.50	0.40	0
MA002140372		MC	0.71	0.38	0
MA002145158		OR	0.71	0.40	0
MA003540652		OR	0.76	0.53	0
MA303324		MC	0.51	0.58	0
MA306990		MC	0.41	0.28	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
MA297992		MC	0.51	0.34	0
MA306458		MC	0.38	0.49	0
MA800650803		MC	0.35	0.36	0
MA801235389		OR	0.61	0.46	0
MA801646735		OR	0.70	0.53	0
MA301169		MC	0.29	0.38	0
MA900982012		MC	0.57	0.24	0
MA904453014		OR	0.56	0.46	0
MA908434516		OR	0.42	0.47	0
MA000846693		OR	0.38	0.54	0
MA000859040		MC	0.70	0.33	--
MA000938134		OR	0.58	0.42	0
MA000953421		MC	0.65	0.50	0
MA000957282		MC	0.57	0.51	--
MA001066377		OR	0.62	0.58	0
MA002343629		OR	0.42	0.67	0
MA002837526		OR	0.65	0.49	0
MA005852277		OR	0.45	0.53	0
MA935150419		OR	0.46	0.54	0
MA301830		MC	0.53	0.31	0
MA298032		OR	0.34	0.54	0
MA301602		MC	0.72	0.46	--
MA306408		MC	0.31	0.27	0
MA704359624		OR	0.69	0.28	0
MA301157		MC	0.65	0.31	0
MA800651876		MC	0.43	0.55	0
MA800975677		OR	0.47	0.57	0
MA804073329		OR	0.34	0.51	0
MA804073907		MC	0.34	0.33	0
MA802371654		OR	0.51	0.72	0
MA900727061		MC	0.61	0.32	0
MA903746975		OR	0.57	0.70	0
MA000846578		MC	0.56	0.44	0
MA000927731		MC	0.39	0.54	0
MA000937699		MC	0.26	0.19	--
MA000955730		OR	0.50	0.51	0
MA000965213		OR	0.42	0.49	0
MA001042212		OR	0.47	0.40	0
MA001142456		OR	0.56	0.40	0
MA006336846		OR	0.62	0.37	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
MA223217		OR	0.44	0.50	0
MA307362		OR	0.64	0.56	0
MA624248796		OR	0.49	0.47	0
MA282127		MC	0.41	0.44	0
MA299673		OR	0.56	0.55	0
MA307338		MC	0.60	0.47	0
MA307363		MC	0.77	0.31	0
MA703149512		OR	0.60	0.44	0
MA736368137		OR	0.56	0.54	0
MA800203270		OR	0.27	0.70	0
MA805109765		OR	0.39	0.38	1
MA805280170		OR	0.59	0.45	0
MA307340		MC	0.37	0.35	0
MA805104566		MC	0.39	0.38	0
MA900378821		OR	0.42	0.67	0
MA001529070		MC	0.53	0.35	0
MA001549477		OR	0.61	0.45	0
MA001604473		OR	0.27	0.55	0
MA003477341		OR	0.37	0.54	0
MA800301627		OR	0.42	0.77	1
MA290253		OR	0.33	0.76	1
MA298252		OR	0.24	0.71	1
MA736063629		OR	0.32	0.49	1
MA805101277		MC	0.37	0.15	0
MA736452404		OR	0.32	0.41	0
MA736509125		MC	0.60	0.44	0
MA805111429		MC	0.51	0.25	0
MA900283487		MC	0.57	0.57	--
MA902758854		OR	0.61	0.33	0
MA001264865		MC	0.47	0.47	0
MA001554177		OR	0.51	0.42	0
MA001577731		OR	0.51	0.51	1
MA001585164		MC	0.60	0.39	--
MA002536621		MC	0.46	0.14	0
MA002538062		MC	0.76	0.42	0
MA003861140		OR	0.50	0.44	0
MA005664640		MC	0.34	0.37	0
MA800385560		OR	0.56	0.57	0
MA900578720		OR	0.39	0.58	0
MA296382		MC	0.44	0.23	0

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
MA304467		MC	0.54	0.54	0
MA306626		MC	0.48	0.16	0
MA624562376		OR	0.41	0.44	0
MA219417		MC	0.33	0.52	0
MA311074		MC	0.31	0.21	0
MA713848056		OR	0.31	0.56	0
MA713849125		OR	0.45	0.57	0
MA259175		MC	0.53	0.21	0
MA298072		MC	0.34	0.47	0
MA804676692		OR	0.22	0.65	0
MA717248260		OR	0.30	0.72	1
MA900557823		MC	0.51	0.39	0
MA900567252		OR	0.30	0.54	1
MA900743031		OR	0.32	0.51	2
MA904000450		MC	0.37	0.22	0
MA000971342		MC	0.48	0.52	0
MA001702061		OR	0.69	0.51	0
MA002046543		OR	0.32	0.55	0
MA002119133		OR	0.42	0.78	1
MA302328		OR	0.39	0.56	0
MA306627		MC	0.30	0.46	--
MA713849162		OR	0.28	0.51	0
MA272764		MC	0.55	0.23	0
MA282221		MC	0.31	0.28	0
MA311092		OR	0.26	0.70	0
MA713848251		OR	0.30	0.41	0
MA713848348		OR	0.34	0.57	1
MA802907874		OR	0.26	0.77	2
MA900336138		MC	0.29	0.48	0
MA900739359		OR	0.41	0.60	0
MA900749529		OR	0.44	0.53	0
MA903983773		MC	0.25	0.29	0
MA904169987		MC	0.30	0.32	0
MA001678587		OR	0.36	0.50	0
MA001759197		OR	0.58	0.35	0
MA005077116		OR	0.35	0.44	0
MA005170212		MC	0.58	0.32	--
MA005207399		MC	0.49	0.44	0
MA900765087		OR	0.56	0.64	0
MA261648		OR	0.31	0.76	2

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

Number	Item	Type	Difficulty	Discrimination	Percent Omitted (%)
MA252991		MC	0.70	0.17	0
MA297517		MC	0.47	0.20	0
MA298198		MC	0.35	0.34	0
MA301714		OR	0.24	0.77	3
MA715919758		OR	0.48	0.32	0
MA800472975		OR	0.28	0.56	0
MA800659905		OR	0.26	0.36	0
MA800738445		OR	0.35	0.75	2
MA901143033		MC	0.47	0.35	0
MA001736920		MC	0.51	0.37	0
MA001737758		OR	0.40	0.52	1
MA001737991		OR	0.44	0.26	0
MA001865159		MC	0.35	0.48	0
MA002177981		OR	0.48	0.52	0
MA002180558		MC	0.52	0.48	0
MA002243883		OR	0.59	0.53	0
MA003936639		OR	0.50	0.58	0
MA908450808		MC	0.32	0.46	0
MA303244		MC	0.34	0.21	0
MA307586		MC	0.46	0.43	0
MA307585		MC	0.70	0.41	0
MA715919661		OR	0.65	0.25	0
MA715919716		OR	0.58	0.54	0
MA715919788		OR	0.27	0.58	0
MA715919853		OR	0.32	0.60	0
MA287538		MC	0.65	0.37	--
MA311433		OR	0.28	0.75	3
MA804535837		MC	0.49	0.22	0
MA297651		OR	0.40	0.52	0
MA804043870		OR	0.21	0.60	1
MA901253257		OR	0.39	0.35	0
MA902308680		OR	0.30	0.62	0
MA905179612		OR	0.17	0.46	1
MA002181298		MC	0.52	0.48	0
MA003128642		OR	0.35	0.48	0
MA003932801		MC	0.54	0.49	0
MA010701848		OR	0.64	0.62	1
MA902278325		OR	0.40	0.52	0
MA908446890		MC	0.64	0.42	0
MA307492		OR	0.57	0.55	0

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	EL909882556#SCORE_TRAIT_Conv	3	36.20	39.91	17.80	4.66		
	EL909882556#SCORE_TRAIT_Ideadev	4	33.38	42.60	17.78	4.20	0.61	
	EL911945550	2	24.13	22.74	52.89			
	EL019650296	2	8.23	29.16	56.97			
	EL028215856	2	41.44	23.47	35.01			
	EL028753268	2	19.78	38.75	40.09			
	EL028832702	3	33.70	35.32	25.77	4.51		
4	EL007459900#SCORE_TRAIT_Conv	3	13.71	56.24	24.80	4.52		
	EL007459900#SCORE_TRAIT_Ideadev	4	21.68	50.92	22.79	3.57	0.31	
	EL007464016	2	29.55	5.36	65.05			
	EL009343264	2	32.60	22.96	44.35			
	EL024480931	2	40.92	18.04	41.04			
	EL024539092	3	12.40	42.03	36.73	8.40		
	EL029323184	2	37.45	43.10	19.35			
EL029964018	2	36.71	15.56	47.70				
5	EL030400392#SCORE_TRAIT_Conv	3	31.97	45.52	17.85	3.94		
	EL030400392#SCORE_TRAIT_Ideadev	4	27.33	50.18	17.99	3.63	0.15	
	EL030463527	2	32.62	19.24	48.12			
	EL624176168	2	13.03	9.73	77.23			
	EL624177447	2	18.74	31.20	49.99			
	EL624182427#SCORE_TRAIT_Conv	3	15.71	40.17	34.71	9.02		
	EL624182427#SCORE_TRAIT_Ideadev	4	13.18	39.25	34.69	11.10	1.39	
6	EL033646585	2	26.19	29.05	44.76			
	EL807016586#SCORE_TRAIT_Conv	3	21.63	38.45	25.02	13.86		
	EL807016586#SCORE_TRAIT_Ideadev	5	19.81	37.89	25.74	10.73	3.25	1.53
	EL807062301	2	54.32	8.02	37.59			
	EL808246461	2	13.05	58.38	28.16			
	EL007051004#SCORE_TRAIT_Conv	3	23.75	30.05	29.55	15.74		
	EL007051004#SCORE_TRAIT_Ideadev	5	33.11	27.04	25.95	10.83	2.13	0.04
7	EL008181021	2	16.36	45.75	37.59			
	EL008355	2	21.20	33.70	45.07			
	EL009438210	2	38.12	39.08	22.75			
	EL713370326	2	20.68	37.65	41.51			
	EL713375305#SCORE_TRAIT_Conv	3	24.86	37.40	18.39	17.95		
	EL713375305#SCORE_TRAIT_Ideadev	5	62.51	8.41	9.77	9.03	6.28	2.62
	EL713475622	2	45.10	22.73	32.14			
8	EL006653237#SCORE_TRAIT_Conv	3	26.19	29.75	25.90	17.13		
	EL006653237#SCORE_TRAIT_Ideadev	5	23.36	31.08	24.47	14.24	4.66	1.15
	EL006978834	2	46.32	18.52	35.12			
	EL011353608	2	44.91	35.74	19.34			
	EL017655451	2	28.32	13.39	58.27			
	EL007062608	2	25.87	6.91	67.21			
	EL007062902#SCORE_TRAIT_Conv	3	11.80	32.02	27.19	27.41		
8	EL007062902#SCORE_TRAIT_Ideadev	5	12.60	31.40	26.06	16.90	8.31	3.15
	EL009257746	2	20.39	46.60	32.96			
	EL006653570	2	14.77	49.37	35.33			
	EL007243506	2	41.28	14.55	44.13			
	EL007253494#SCORE_TRAIT_Conv	3	24.67	26.49	26.68	20.01		
	EL007253494#SCORE_TRAIT_Ideadev	5	24.82	26.73	25.26	14.24	6.11	0.69
	EL008553781	2	11.03	40.58	48.22			

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	MA293460A	3	31.19	18.33	22.93	27.17		
	MA310884	1	38.78	61.14				
	MA310899A	3	22.98	30.02	36.88	9.47		
	MA735663821	1	56.86	43.06				
	MA900425126	1	60.84	38.44				
	MA000749172	1	39.76	60.07				
	MA001038775	1	21.46	78.36				
	MA001047582	1	26.09	73.68				
	MA001344527	1	28.23	71.72				
	MA001439533	1	28.32	71.62				
	MA703078093	1	17.04	82.86				
	MA735736004A	1	42.42	57.39				
	MA935136577	1	35.63	64.06				
	MA310889	1	29.84	70.05				
	MA703056978	1	43.59	56.09				
	MA286750A	3	21.19	24.85	32.49	21.32		
	MA713536927	1	33.48	66.25				
	MA735756531	1	64.08	35.83				
	MA735954511	1	52.85	47.05				
	MA253711A	3	27.34	29.94	31.90	10.50		
MA309747	1	50.30	49.56					
MA001335228	1	44.92	54.38					
MA001338241	1	25.75	74.20					
MA734752477	1	26.88	73.01					
4	MA311568	1	37.75	62.10				
	MA303335	4	12.41	27.12	28.10	20.85	11.27	
	MA713677363	2	11.25	30.90	57.70			
	MA800727128	1	52.57	47.28				
	MA801035466	4	14.96	28.63	26.40	20.02	9.78	
	MA803956738	1	36.49	63.42				
	MA000732007	1	51.78	48.13				
	MA002128911	1	34.79	65.12				
	MA002334462	1	42.68	56.88				
	MA003744055	1	34.82	64.62				
	MA311567	1	45.83	54.07				
	MA307317	4	19.97	12.79	36.54	22.35	8.07	
	MA704653374	2	15.11	37.32	47.43			
	MA293812	4	21.35	21.09	41.38	8.68	7.27	
	MA736377105	1	50.88	48.79				
	MA900756471	1	43.11	56.81				
	MA900846441	1	67.58	32.34				
	MA001851276	1	37.09	62.78				
	MA002145158	1	29.10	70.86				
	MA003540652	1	23.61	76.15				
5	MA801235389	1	39.29	60.63				
	MA801646735	1	29.22	70.38				
	MA904453014	1	43.48	56.41				
	MA908434516	2	34.44	46.79	18.72			
	MA000846693	1	61.85	38.06				
	MA000938134	1	41.62	58.20				
	MA001066377	1	37.66	62.29				
	MA002343629	4	18.74	27.20	27.25	20.15	6.33	
	MA002837526	1	35.12	64.61				
MA005852277	4	17.85	17.35	36.61	22.42	5.64		

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
5	MA935150419	1	54.00	45.83				
	MA298032	1	65.96	33.99				
	MA704359624	1	31.21	68.76				
	MA800975677	1	52.50	47.30				
	MA804073329	1	66.00	33.94				
	MA802371654	4	13.06	21.01	28.48	22.25	14.92	
	MA903746975	4	5.66	27.11	23.26	22.62	21.26	
	MA000955730	1	49.89	49.85				
	MA000965213	2	34.75	46.81	18.37			
	MA001042212	1	52.73	47.14				
MA001142456	1	44.09	55.85					
MA006336846	1	38.09	61.84					
6	MA223217	1	55.44	44.40				
	MA307362	1	36.11	63.71				
	MA624248796	1	50.13	49.43				
	MA299673	1	43.77	56.10				
	MA703149512	1	40.38	59.51				
	MA736368137	1	43.91	55.88				
	MA800203270	2	61.07	22.94	15.76			
	MA805109765	1	60.32	39.07				
	MA805280170	1	41.26	58.66				
	MA900378821	4	7.32	51.02	17.87	14.01	9.38	
	MA001549477	1	38.64	61.32				
	MA001604473	1	72.66	27.19				
	MA003477341	1	62.62	37.16				
	MA800301627	4	25.89	22.01	21.50	13.24	16.17	
	MA290253	4	33.82	26.70	21.47	8.01	9.17	
	MA298252	4	55.25	19.37	7.78	5.80	10.80	
	MA736063629	1	66.94	32.41				
	MA736452404	1	67.31	32.35				
	MA902758854	1	38.67	61.02				
	MA001554177	1	49.32	50.66				
MA001577731	1	48.04	51.39					
MA003861140	1	50.09	49.51					
MA800385560	1	43.40	56.37					
MA900578720	2	40.00	41.11	18.81				
7	MA624562376	2	35.68	47.57	16.72			
	MA713848056	1	69.09	30.86				
	MA713849125	1	54.85	45.00				
	MA804676692	1	78.03	21.61				
	MA717248260	4	24.61	44.15	15.95	10.32	3.57	
	MA900567252	1	69.16	29.59				
	MA900743031	1	66.19	31.55				
	MA001702061	1	30.28	69.34				
	MA002046543	1	68.07	31.51				
	MA002119133	4	27.80	20.43	21.56	14.34	15.02	
	MA302328	1	60.73	38.88				
	MA713849162	1	71.81	28.16				
	MA311092	1	73.47	26.12				
	MA713848251	1	69.28	30.36				
	MA713848348	1	65.60	33.85				
	MA802907874	4	44.20	22.94	18.33	5.77	6.50	
	MA900739359	1	58.38	41.47				
	MA900749529	1	56.34	43.63				
	MA001678587	1	63.11	36.50				
	MA001759197	1	42.10	57.52				
MA005077116	1	65.19	34.78					

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
7	MA900765087	2	17.30	52.24	30.33			
	MA261648	4	48.22	15.41	7.76	14.67	12.26	
8	MA301714	4	51.75	18.33	9.88	8.42	8.13	
	MA715919758	1	51.47	48.37				
	MA800472975	1	71.84	28.02				
	MA800659905	1	74.18	25.57				
	MA800738445	4	34.56	20.75	18.18	15.55	8.99	
	MA001737758	1	59.00	40.42				
	MA001737991	1	55.86	43.95				
	MA002177981	1	51.63	48.32				
	MA002243883	2	21.74	38.97	39.05			
	MA003936639	1	49.94	49.77				
	MA715919661	1	35.44	64.54				
	MA715919716	1	42.05	57.88				
	MA715919788	1	72.49	27.27				
	MA715919853	1	68.05	31.63				
	MA311433	4	31.32	35.62	16.98	9.38	3.71	
	MA297651	1	59.50	40.21				
	MA804043870	1	77.88	21.48				
	MA901253257	1	60.97	38.81				
	MA902308680	1	69.93	29.84				
	MA905179612	1	82.01	16.96				
MA003128642	1	64.80	35.15					
MA010701848	4	7.22	15.57	19.45	24.13	32.34		
MA902278325	2	40.77	37.74	21.39				
MA307492	1	42.99	56.65					

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	Group		Item Type	Number of Items	Number “Low” Favoring			Number “High” Favoring		
	Reference	Focal			Total	Reference	Focal	Total	Reference	Focal
3	Male	Female	MC	26	1	0	1	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
	Not ELL	ELL	MC	26	6	5	1	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	26	2	2	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
	White	African American	MC	26	2	2	0	1	1	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
		Hispanic / Latino	MC	26	2	1	1	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	26	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
	Online	Paper	MC	26	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
4	Male	Female	MC	26	2	2	0	0	0	0
			OR	5	3	2	1	0	0	0
			WP	2	0	0	0	0	0	0
	Not ELL	ELL	MC	26	4	4	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	2	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	26	2	2	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
	White	African American	MC	26	4	3	1	0	0	0
			OR	5	1	1	0	0	0	0
			WP	2	0	0	0	0	0	0
		Hispanic / Latino	MC	26	2	2	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	2	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	26	2	2	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	2	2	2	0	0	0	0
	Online	Paper	MC	26	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	2	0	0	0	0	0	0
5	Male	Female	MC	24	4	3	1	0	0	0
			OR	5	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Total	Number "Low"		Total	Number "High"	
	Reference	Focal				Favoring Reference	Favoring Focal		Favoring Reference	Favoring Focal
5	Not ELL	ELL	MC	24	3	3	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	4	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	24	2	2	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	4	0	0	0	0	0	0
	White	African American	MC	24	2	2	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	4	0	0	0	0	0	0
		Hispanic / Latino	MC	24	2	2	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	4	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	24	3	3	0	0	0	0
			OR	5	1	1	0	0	0	0
			WP	4	1	1	0	0	0	0
	Online	Paper	MC	24	0	0	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0
Male	Female	MC	24	2	1	1	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	
Not ELL	ELL	MC	24	4	4	0	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	
Not Economically Disadvantaged	Economically Disadvantaged	MC	24	0	0	0	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	
6	African American	MC	24	2	1	1	0	0	0	
		OR	5	1	1	0	0	0	0	
		WP	4	0	0	0	0	0	0	
	White	Hispanic / Latino	MC	24	2	2	0	0	0	0
			OR	5	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0
Students Without Disabilities	Students with Disabilities	MC	24	3	3	0	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	2	2	0	0	0	0	
Online	Paper	MC	24	0	0	0	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	
Male	Female	MC	24	5	3	2	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	1	0	1	0	0	0	
7	Not ELL	ELL	MC	24	3	2	1	0	0	0
			OR	5	0	0	0	0	0	0
			WP	4	0	0	0	0	0	0
Not Economically Disadvantaged	Economically Disadvantaged	MC	24	0	0	0	0	0	0	
		OR	5	0	0	0	0	0	0	
		WP	4	0	0	0	0	0	0	

continued

Grade	Group		Item Type	Number of Items	Total	Number "Low"		Total	Number "High"		
	Reference	Focal				Favoring Reference	Favoring Focal		Favoring Reference	Favoring Focal	
7	White	African American	MC	24	2	1	1	0	0	0	
			OR	5	0	0	0	0	0	0	
			WP	4	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
Male	Female	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 ELL	ELL	MC	24	4	4	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	
8	White	African American	MC	24	0	0	0	0	0	0	
			OR	5	0	0	0	0	0	0	
			WP	4	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	1	1	0	0	0	0	0	
		WP	4	0	0	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	Group		Item Type	Number of Items	Total	Number “Low” Favoring		Total	Number “High” Favoring	
	Reference	Focal				Reference	Focal		Reference	Focal
3	Male	Female	MC	16	2	2	0	0	0	0
			OR	24	4	2	2	0	0	0
	Not ELL	ELL	MC	16	2	1	1	0	0	0
			OR	24	1	1	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	16	0	0	0	0	0	0
			OR	24	0	0	0	0	0	0
	White	African American	MC	16	3	2	1	0	0	0
			OR	24	7	4	3	0	0	0
		Hispanic / Latino	MC	16	2	2	0	0	0	0
			OR	24	2	2	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	16	0	0	0	0	0	0
			OR	24	0	0	0	0	0	0
Online	Paper	MC	16	0	0	0	0	0	0	
		OR	24	0	0	0	0	0	0	
4	Male	Female	MC	20	3	3	0	0	0	0
			OR	20	3	3	0	0	0	0
	Not ELL	ELL	MC	20	2	1	1	0	0	0
			OR	20	4	4	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	20	1	1	0	0	0	0
			OR	20	1	1	0	0	0	0
	White	African American	MC	20	3	2	1	1	0	1
			OR	20	3	2	1	0	0	0
		Hispanic / Latino	MC	20	0	0	0	0	0	0
			OR	20	2	2	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	20	2	0	2	0	0	0
			OR	20	3	2	1	0	0	0
Online	Paper	MC	20	0	0	0	0	0	0	
		OR	20	0	0	0	0	0	0	
5	Male	Female	MC	18	2	2	0	0	0	0
			OR	22	5	4	1	0	0	0
	Not ELL	ELL	MC	18	2	2	0	0	0	0
			OR	22	2	2	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	MC	18	0	0	0	0	0	0
			OR	22	1	1	0	0	0	0
	White	African American	MC	18	0	0	0	0	0	0
			OR	22	2	1	1	1	1	0
		Hispanic / Latino	MC	18	1	1	0	0	0	0
			OR	22	2	2	0	0	0	0
	Students Without Disabilities	Students with Disabilities	MC	18	1	1	0	0	0	0
			OR	22	2	2	0	0	0	0
Online	Paper	MC	18	0	0	0	0	0	0	
		OR	22	0	0	0	0	0	0	

continued

Grade	Group		Item Type	Number of Items	Total	Number "Low"		Number "High"			
	Reference	Focal				Total	Favoring		Total	Favoring	
							Reference	Focal		Reference	Focal
6	Male	Female	MC	16	1	1	0	0	0	0	
			OR	24	4	1	3	0	0	0	
	Not ELL	ELL	MC	16	3	2	1	0	0	0	
			OR	24	2	1	1	0	0	0	
	Not Economically Disadvantaged	Economically Disadvantaged	MC	16	0	0	0	0	0	0	
			OR	24	0	0	0	0	0	0	
	White	African American	MC	16	4	3	1	0	0	0	
			OR	24	2	2	0	0	0	0	
		Hispanic / Latino	MC	16	1	1	0	0	0	0	
			OR	24	1	1	0	0	0	0	
	Students Without Disabilities	Students with Disabilities	MC	16	1	1	0	0	0	0	
			OR	24	5	4	1	0	0	0	
	Online	Paper	MC	16	0	0	0	0	0	0	
			OR	24	0	0	0	0	0	0	
7	Male	Female	MC	17	2	0	2	0	0	0	
			OR	23	6	6	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	5	2	3	0	0	0	
			OR	23	3	3	0	1	1	0	
		Hispanic / Latino	MC	17	0	0	0	0	0	0	
			OR	23	0	0	0	0	0	0	
	Students Without Disabilities	Students with Disabilities	MC	17	6	5	1	0	0	0	
			OR	23	7	6	1	0	0	0	
	Online	Paper	MC	17	0	0	0	0	0	0	
			OR	23	0	0	0	0	0	0	
8	Male	Female	MC	16	1	1	0	0	0	0	
			OR	24	2	1	1	1	1	0	
	Not ELL	ELL	MC	16	1	1	0	0	0	0	
			OR	24	2	2	0	0	0	0	
	Not Economically Disadvantaged	Economically Disadvantaged	MC	16	0	0	0	0	0	0	
			OR	24	0	0	0	0	0	0	
	White	African American	MC	16	3	2	1	0	0	0	
			OR	24	3	3	0	0	0	0	
		Hispanic / Latino	MC	16	1	1	0	0	0	0	
			OR	24	0	0	0	0	0	0	
	Students Without Disabilities	Students with Disabilities	MC	16	1	1	0	0	0	0	
			OR	24	6	4	2	0	0	0	
	Online	Paper	MC	16	0	0	0	0	0	0	
			OR	24	0	0	0	0	0	0	

APPENDIX J
2022–2023 MCAS EQUATING REPORT



Massachusetts Comprehensive Assessment System

2022–2023: EQUATING REPORT

2022–2023 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 2022 - 2023 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
3	2023	60542	14	40	39	7	46	0.3	496.0
	2022	61648	13	41	39	7	46	-6.9	497.1
	2021	50011	8	39	43	10	53	-5.7	500.1
	2019	63602	6	36	48	11	58	5.8	504.8
	2018	43046	6	41	43	10	53	1.0	501.8
	2017	26459	7	41	43	8	52		500.1
4	2023	61836	14	44	36	6	42	1.7	494.7
	2022	62100	14	46	36	4	40	-11.2	493.8
	2021	50867	11	38	45	6	51	-2.7	498.9
	2019	65450	7	39	44	10	54	-0.5	502.6
	2018	69078	7	38	44	10	55	3.5	502.2
	2017	63918	8	41	43	8	51		500.1
5	2023	62316	13	41	41	5	46	3.8	496.0
	2022	63620	11	47	37	5	42	-7.1	495.6
	2021	51362	10	41	41	8	49	-5.0	497.9
	2019	67933	6	39	47	8	54	-1.5	501.9
	2018	69390	6	38	49	7	56	4.4	502.3
	2017	28547	7	42	46	5	51		499.9
6	2023	63574	21	35	36	8	44	1.3	494.0
	2022	63887	20	37	35	8	43	-7.2	494.0
	2021	51319	19	31	37	13	50	-5.7	498.4
	2019	67612	11	33	42	13	56	3.4	502.5
	2018	53988	10	38	42	11	52	-0.7	501.3
	2017	29369	8	39	47	6	53		500.3
7	2023	63711	17	41	34	8	42	-0.3	494.2
	2022	65584	17	40	37	6	42	-3.2	493.7
	2021	51120	17	37	39	7	46	-4.8	495.6
	2019	67462	11	39	42	9	50	3.6	499.8
	2018	66410	13	40	39	8	47	-6.5	497.4
	2017	30209	8	38	48	6	53		500.2
8	2023	65553	20	35	35	10	45	2.2	495.4
	2022	67919	16	40	36	7	43	-0.3	494.8
	2021	50822	15	41	37	7	44	-10.0	496.2
	2019	67350	11	35	42	12	54	1.1	500.6
	2018	69486	13	34	42	10	52	1.1	499.6
	2017	65314	9	40	43	8	51		499.5

Table 1.1.2
Percentage of Students by Achievement Levels Categories
Mathematics

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS
3	2023	51707	12	39	39	10	49	1.2	499.1
	2022	53433	13	39	40	8	48	9.8	497.5
	2021	45242	20	42	32	6	38	-18.0	491.2
	2019	56176	7	37	45	11	56	7.2	503.0
	2018	43501	11	40	40	9	49	-3.2	499.1
	2017	26659	11	37	44	8	52		499.2
4	2023	52554	10	37	43	10	54	4.0	501.1
	2022	53577	10	40	43	7	50	10.9	498.8
	2021	45553	17	44	34	4	39	-17.8	491.7
	2019	57629	6	37	47	10	57	7.7	503.0
	2018	69779	11	40	42	7	49	-1.9	498.0
	2017	64473	10	39	44	6	51		498.7
5	2023	54159	7	45	42	6	48	5.7	499.3
	2022	55635	10	48	38	5	42	3.1	496.5
	2021	46011	13	47	35	5	39	-15.5	493.5
	2019	60444	5	40	48	6	55	8.0	501.7
	2018	70083	9	45	42	5	47	-2.7	497.7
	2017	29285	8	42	42	8	49		499.4
6	2023	56389	11	42	40	8	47	-0.3	498.1
	2022	56939	9	43	42	6	48	8.9	498.2
	2021	46699	16	45	34	5	39	-18.9	493.4
	2019	61719	6	37	46	12	58	9.6	504.0
	2018	54582	9	43	42	6	48	-4.2	498.4
	2017	29704	9	39	46	6	52		499.7
7	2023	57234	16	41	35	9	44	1.7	496.1
	2022	59311	13	45	34	8	42	2.7	495.5
	2021	46839	13	48	32	7	39	-13.4	494.9
	2019	62495	9	39	41	12	53	5.3	501.0
	2018	66925	12	40	40	8	47	-0.9	497.7
	2017	30144	9	43	40	8	48		498.9
8	2023	59572	14	43	34	9	43	2.3	496.4
	2022	62311	12	48	32	8	40	4.4	495.8
	2021	47150	16	48	32	4	36	-15.1	492.0
	2019	62817	8	41	40	11	51	0.4	501.5
	2018	70044	11	39	42	8	51	1.1	498.9
	2017	66077	9	42	40	9	49		500.3

Section 1.2

Raw Scores Associated with Cutpoints

Table 1.2.1
Raw Scores Associated with Cutpoints

Subject	Grade	Cut Point	2022 Actual	2023 Actual	2023 Pred
English Language Arts	3	NM-PM	14	13	11
		PM-ME	28	28	26
		ME-EE	37	38	37
English Language Arts	4	NM-PM	15	16	14
		PM-ME	29	29	29
		ME-EE	39	38	39
English Language Arts	5	NM-PM	16	17	14
		PM-ME	33	32	31
		ME-EE	43	42	42
English Language Arts	6	NM-PM	19	19	17
		PM-ME	32	31	30
		ME-EE	42	41	41
English Language Arts	7	NM-PM	18	16	15
		PM-ME	32	31	31
		ME-EE	44	43	44
English Language Arts	8	NM-PM	22	22	19
		PM-ME	36	34	33
		ME-EE	45	43	44
Mathematics	3	NM-PM	14	14	13
		PM-ME	30	32	31
		ME-EE	43	44	43
Mathematics	4	NM-PM	14	16	16
		PM-ME	33	34	35
		ME-EE	49	48	49
Mathematics	5	NM-PM	13	13	13
		PM-ME	33	32	32
		ME-EE	50	49	49
Mathematics	6	NM-PM	11	12	11
		PM-ME	28	30	29
		ME-EE	48	49	48
Mathematics	7	NM-PM	10	11	10
		PM-ME	25	27	26
		ME-EE	46	47	47
Mathematics	8	NM-PM	12	14	13
		PM-ME	32	32	32
		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	30
English Language Arts	Grade 4	27
English Language Arts	Grade 5	43
English Language Arts	Grade 6	26
English Language Arts	Grade 7	21
English Language Arts	Grade 8	119

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	1.14	-0.21	18	0
English Language Arts	4	1.07	-0.26	18	0
English Language Arts	5	1.14	-0.21	16	0
English Language Arts	6	1.44	-0.36	16	0
English Language Arts	7	1.22	-0.30	16	0
English Language Arts	8	1.41	-0.21	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	28	8
English Language Arts	Grade 4	22	7
English Language Arts	Grade 5	44	8
English Language Arts	Grade 6	17	12
English Language Arts	Grade 7	20	9
English Language Arts	Grade 8	42	12

The Math 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	66
Mathematics	Grade 4	71
Mathematics	Grade 5	50
Mathematics	Grade 6	43
Mathematics	Grade 7	80
Mathematics	Grade 8	33

Table 1.3.1.e
Stocking and Lord Constants for Math

Subject	Grade	Slope	Intercept	Num Eq Items	Num Eq Items Rem
Mathematics	3	1.08	-0.03	20	0
Mathematics	4	1.03	0.10	20	0
Mathematics	5	1.01	-0.02	20	0
Mathematics	6	1.03	-0.11	20	0
Mathematics	7	1.11	-0.14	20	0
Mathematics	8	1.10	-0.19	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	4	IA00289 (EL309792)	beta analysis	retained for equating
English Language Arts	8	IA00063 (EL290814)	beta analysis	retained for equating
Mathematics	4	IA00961 (MA307081)	beta analysis	retained for equating
Mathematics	4	IA01093 (MA623879088)	beta analysis	retained for equating
Mathematics	5	IA00936 (MA306420)	beta analysis	retained for equating
Mathematics	8	IA00865 (MA297656)	beta analysis	retained for equating
Mathematics	8	IA02495 (MA309741)	beta analysis	retained for equating
Mathematics	8	IA05070 (MA804042487)	beta analysis	retained for equating

Section 1.4

Equating Item Summary Statistics

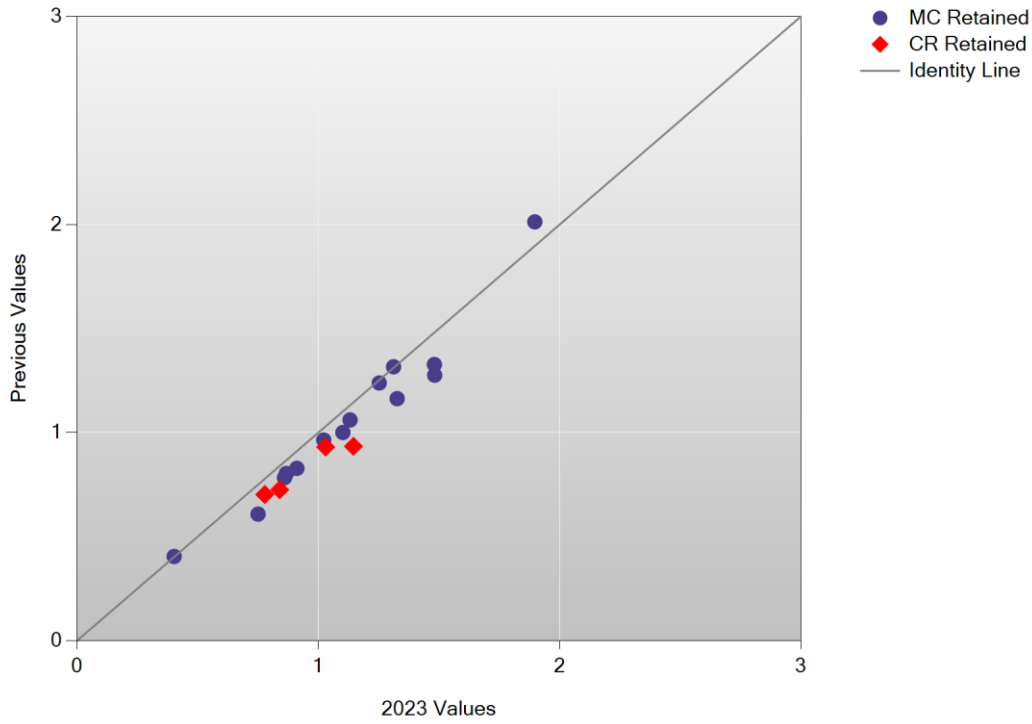
Table 1.4.1
Equating Item Summary Statistics
(2023 a and b values are unscaled parameters before equating for all tests)

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	2023	0.58	0.13	0.50	0.09	1.09	0.34	-0.02	0.52
		Previous	0.58	0.13	0.49	0.09	1.00	0.36	-0.19	0.56
English Language Arts	04	2023	0.64	0.16	0.43	0.09	0.86	0.19	-0.29	0.81
		Previous	0.63	0.16	0.41	0.09	0.80	0.20	-0.63	0.85
English Language Arts	05	2023	0.69	0.14	0.40	0.07	0.86	0.27	-0.50	0.66
		Previous	0.68	0.14	0.40	0.08	0.76	0.26	-0.79	0.75
English Language Arts	06	2023	0.68	0.11	0.41	0.09	0.86	0.22	-0.42	0.67
		Previous	0.68	0.11	0.41	0.08	0.60	0.15	-0.97	0.99
English Language Arts	07	2023	0.71	0.11	0.42	0.08	0.86	0.21	-0.68	0.65
		Previous	0.70	0.11	0.43	0.08	0.70	0.19	-1.17	0.79
English Language Arts	08	2023	0.64	0.13	0.41	0.08	0.89	0.28	-0.29	0.62
		Previous	0.64	0.12	0.40	0.08	0.62	0.20	-0.58	0.79
Mathematics	03	2023	0.65	0.16	0.48	0.11	1.02	0.21	-0.44	0.84
		Previous	0.63	0.16	0.49	0.10	0.91	0.20	-0.50	0.87
Mathematics	04	2023	0.62	0.16	0.51	0.11	1.02	0.27	-0.32	0.61
		Previous	0.59	0.17	0.50	0.11	0.99	0.29	-0.22	0.71
Mathematics	05	2023	0.61	0.19	0.48	0.12	1.01	0.27	-0.32	0.84
		Previous	0.58	0.18	0.48	0.11	0.98	0.29	-0.33	0.81
Mathematics	06	2023	0.56	0.16	0.47	0.14	1.04	0.34	-0.04	0.87
		Previous	0.56	0.16	0.46	0.14	1.03	0.33	-0.11	0.90
Mathematics	07	2023	0.56	0.19	0.53	0.13	1.20	0.32	-0.13	0.72
		Previous	0.56	0.19	0.52	0.13	1.10	0.29	-0.27	0.79
Mathematics	08	2023	0.56	0.14	0.51	0.12	1.21	0.37	0.00	0.59
		Previous	0.56	0.14	0.49	0.12	1.16	0.39	-0.16	0.64

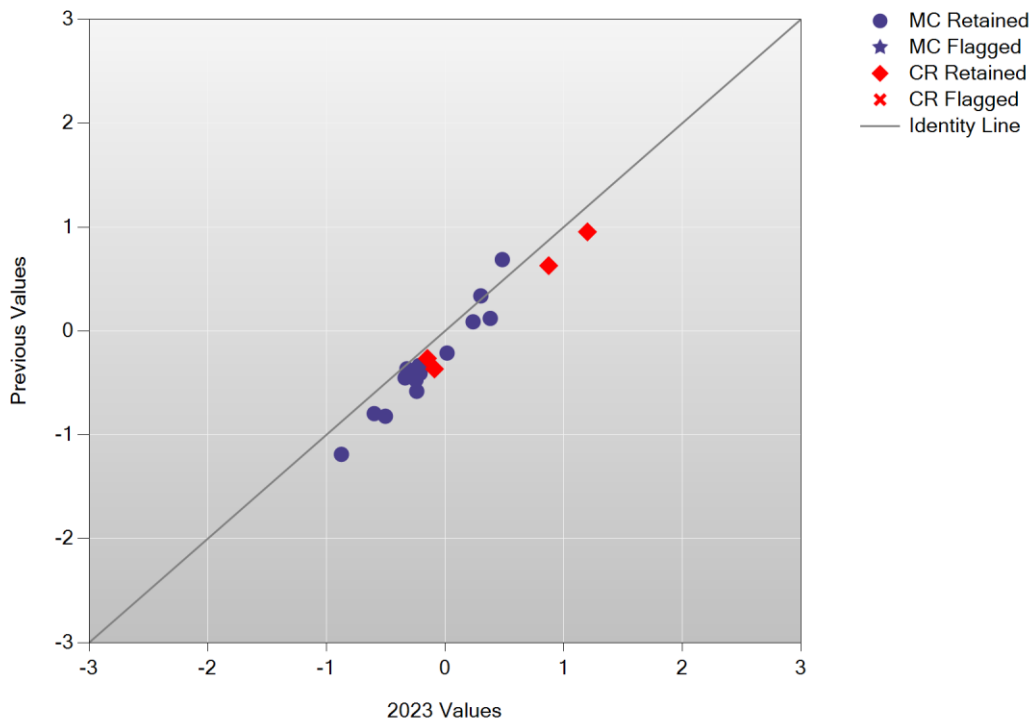
Section 2.1

A/A, B/B, Delta, Test Characteristic Curve, Test Information Function, and Cumulative Scale Score Distribution Plots

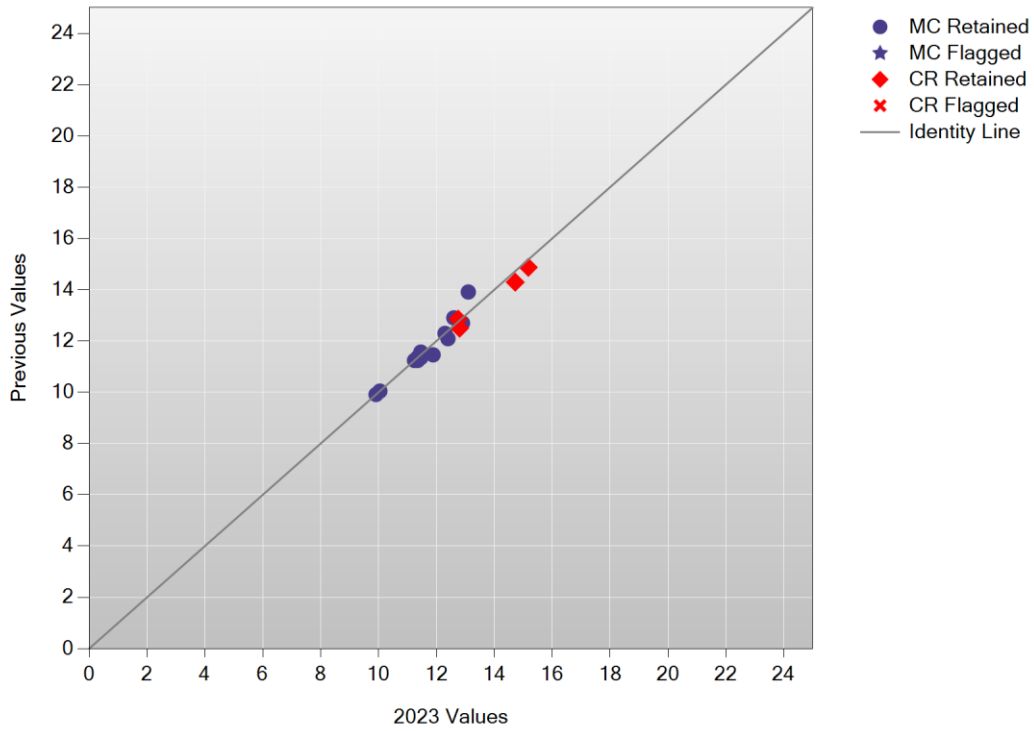
A/A Plot: English Language Arts Grade 3



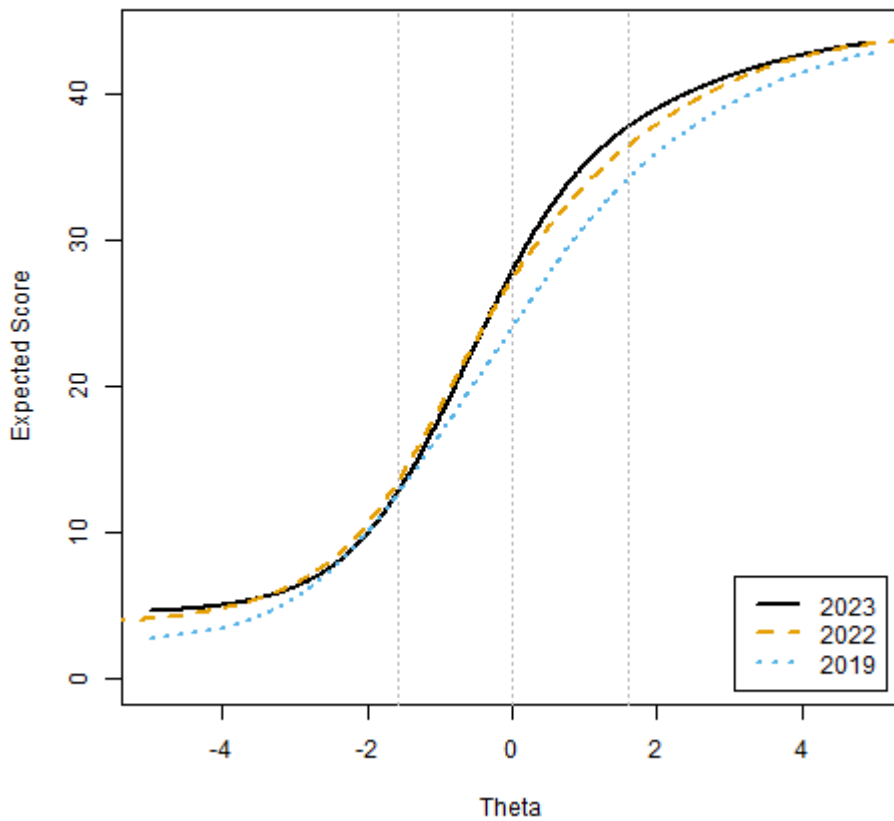
B/B Plot: English Language Arts Grade 3



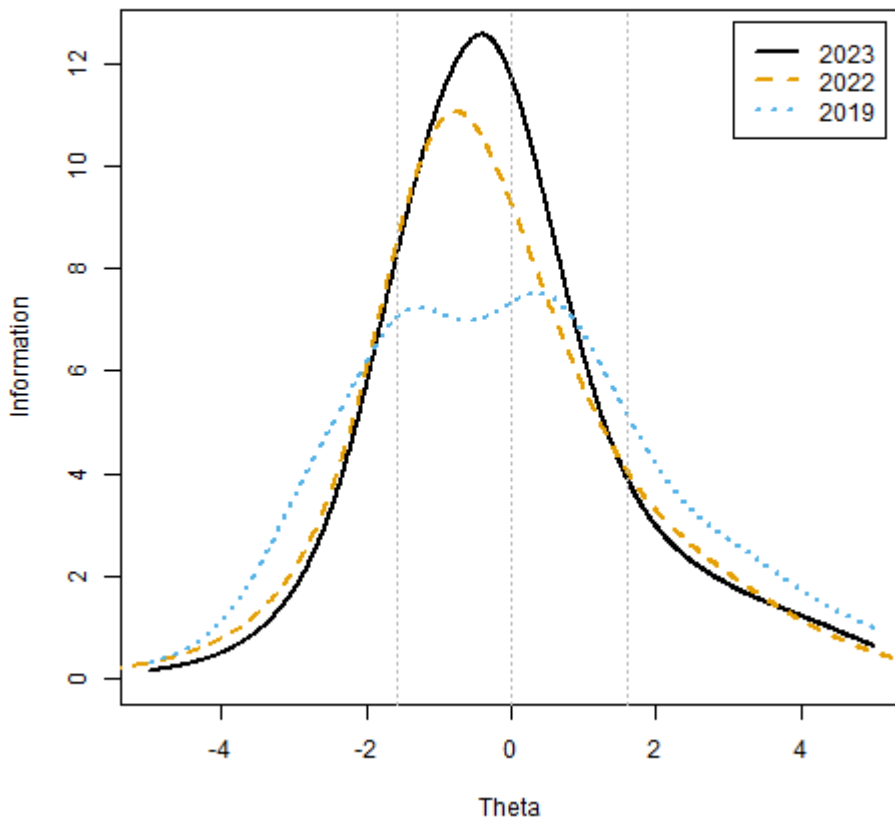
Delta Plot: English Language Arts Grade 3



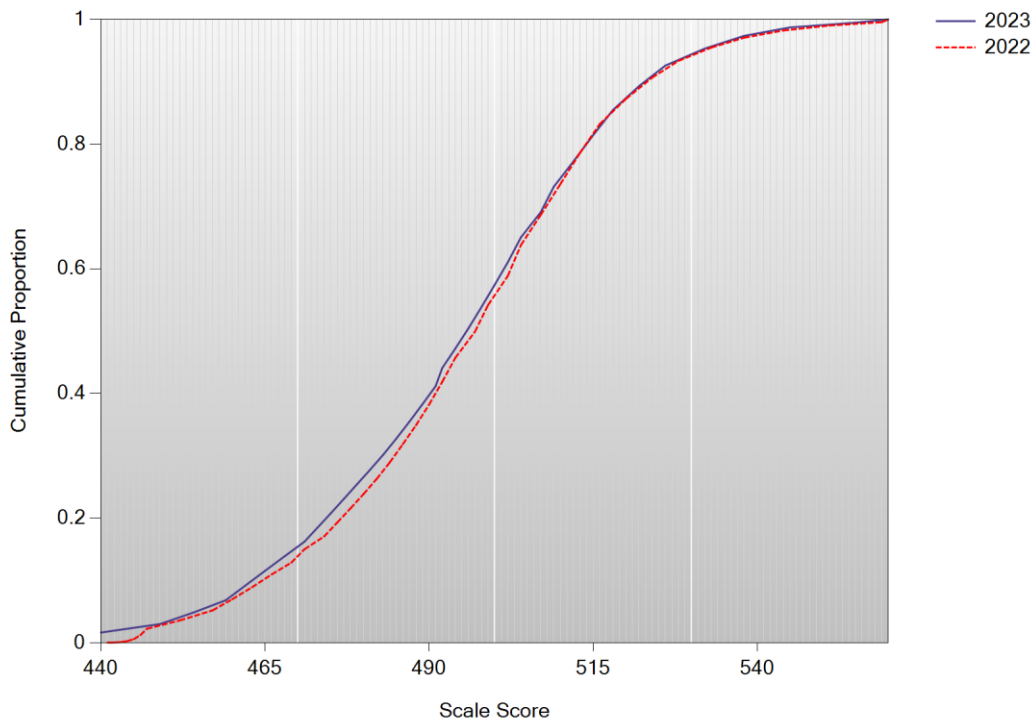
Test Characteristic Curve: ela03



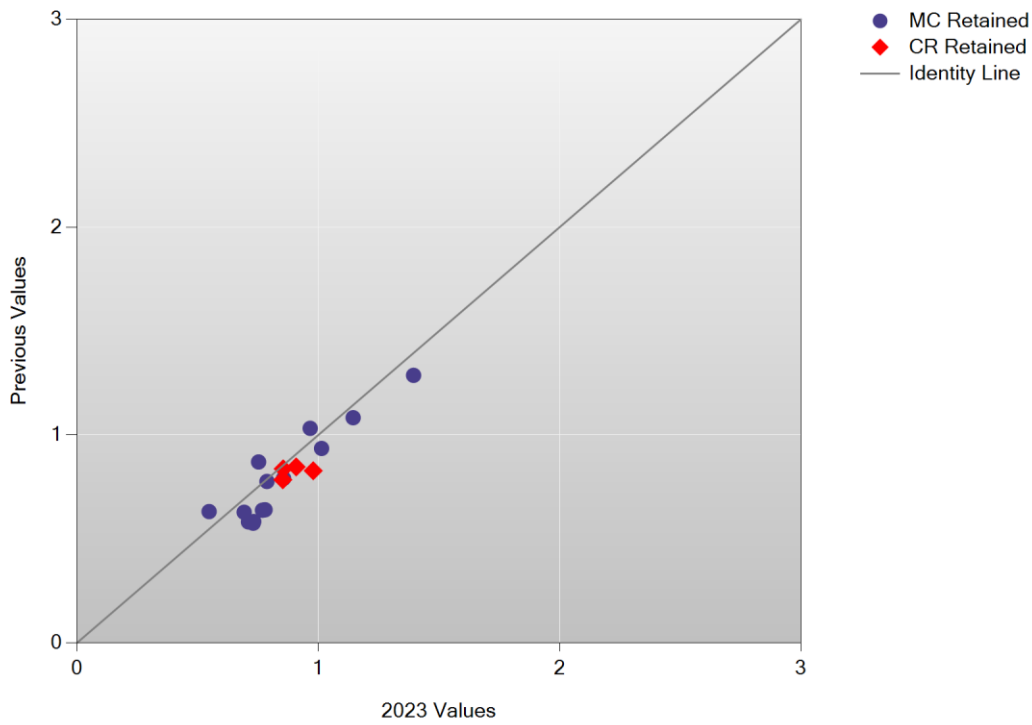
Test Information Function: ela03



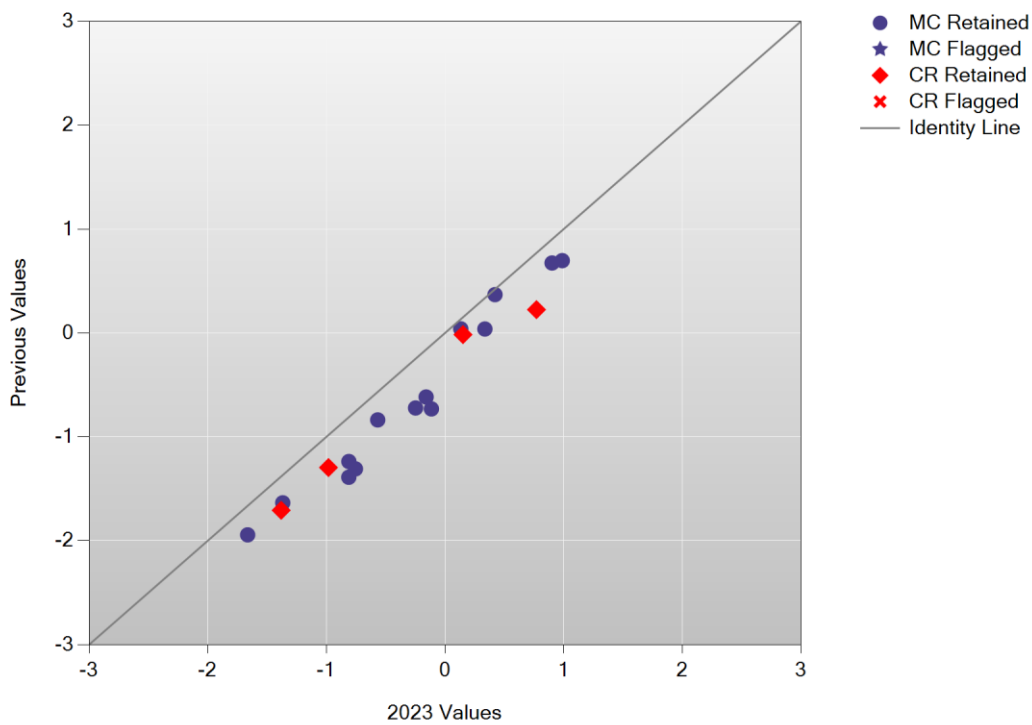
Cumulative Scale Score Distributions: English Language Arts Grade 3



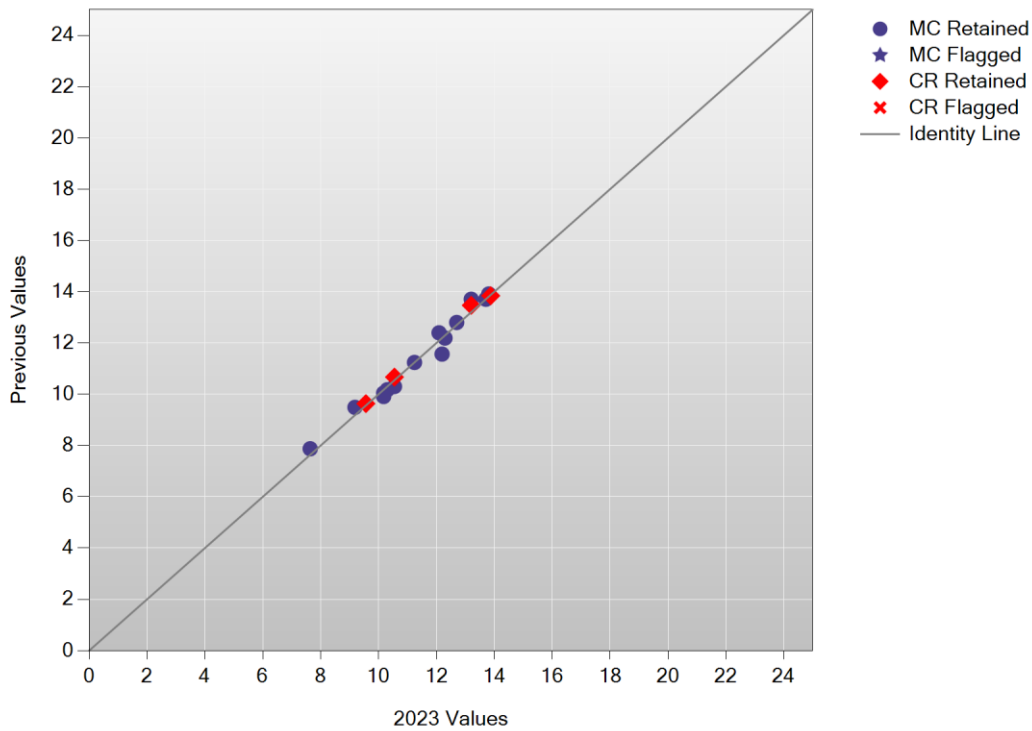
A/A Plot: English Language Arts Grade 4



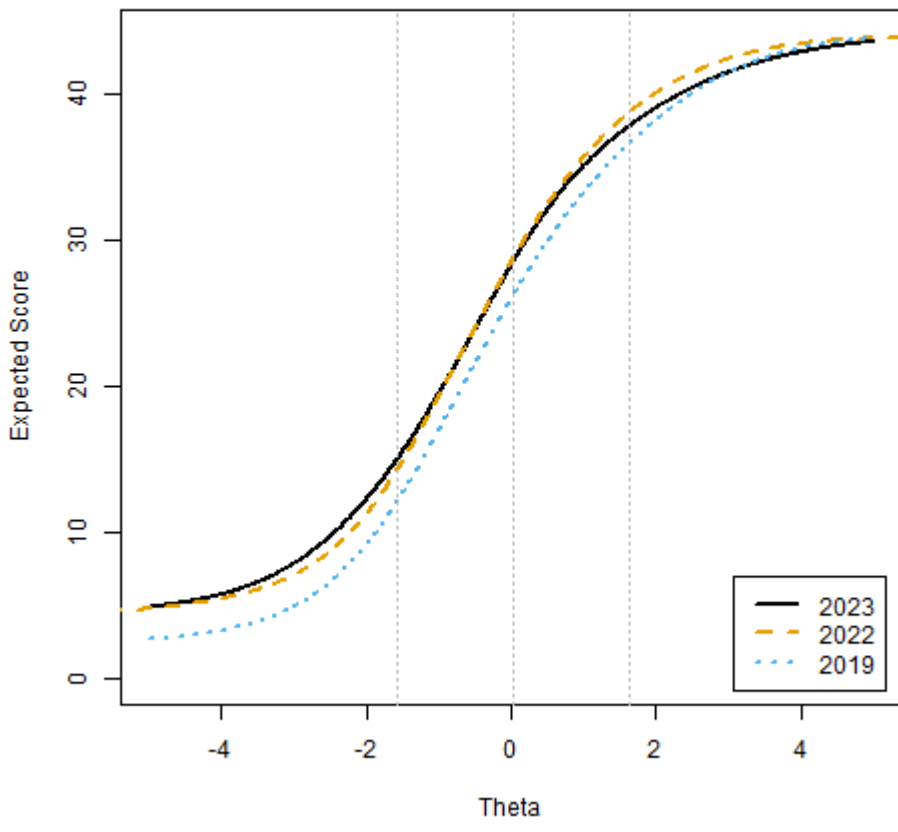
B/B Plot: English Language Arts Grade 4



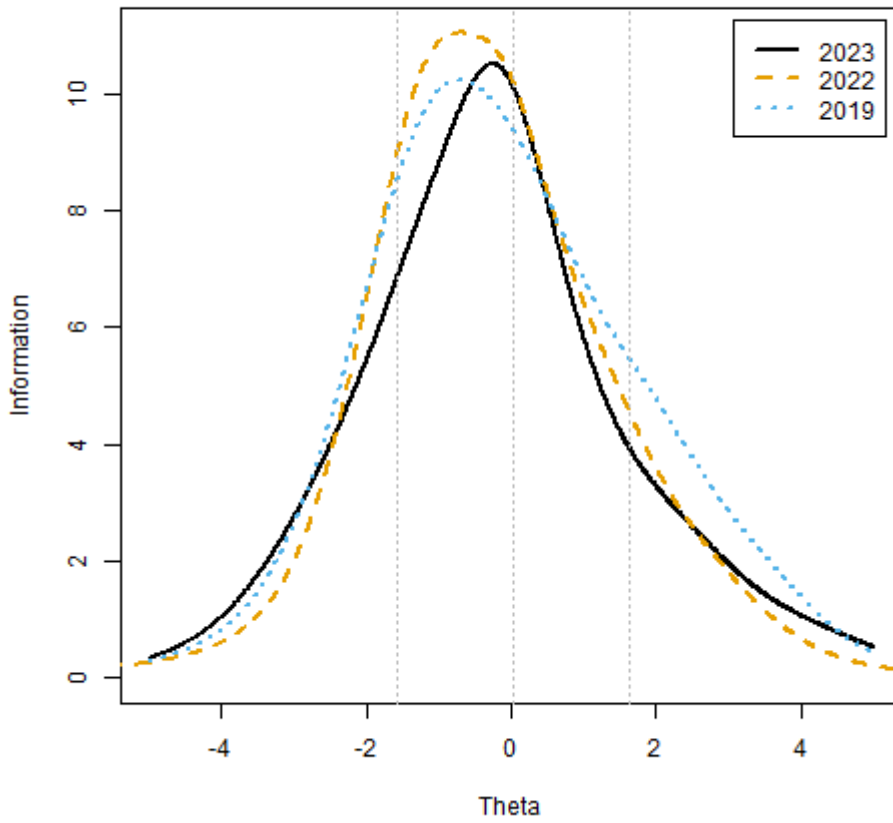
Delta Plot: English Language Arts Grade 4



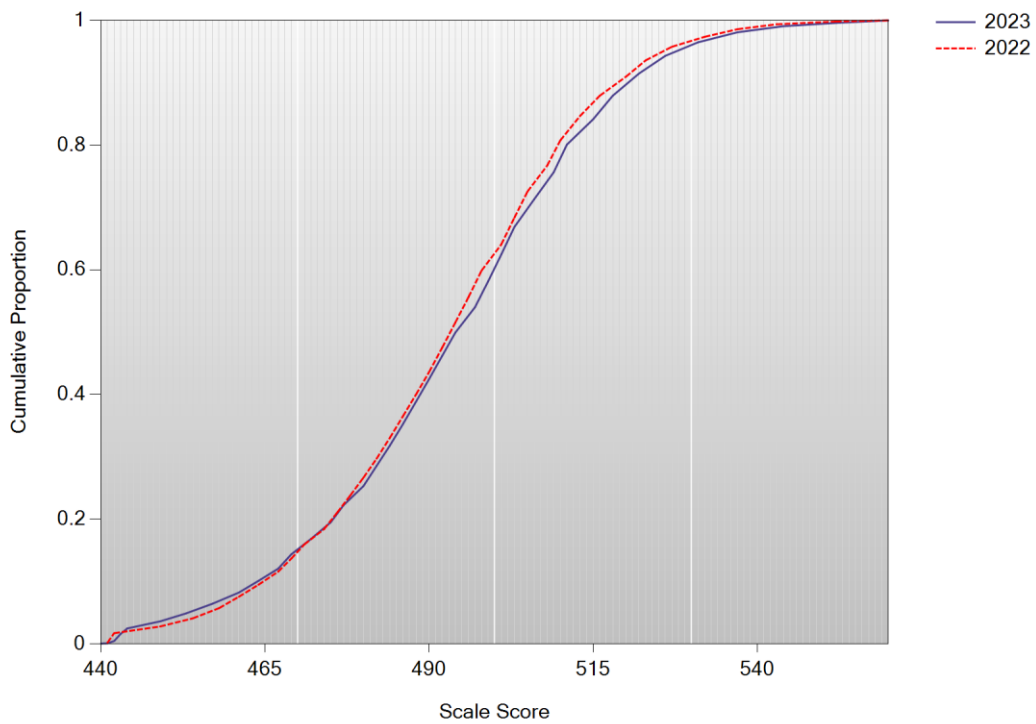
Test Characteristic Curve: ela04



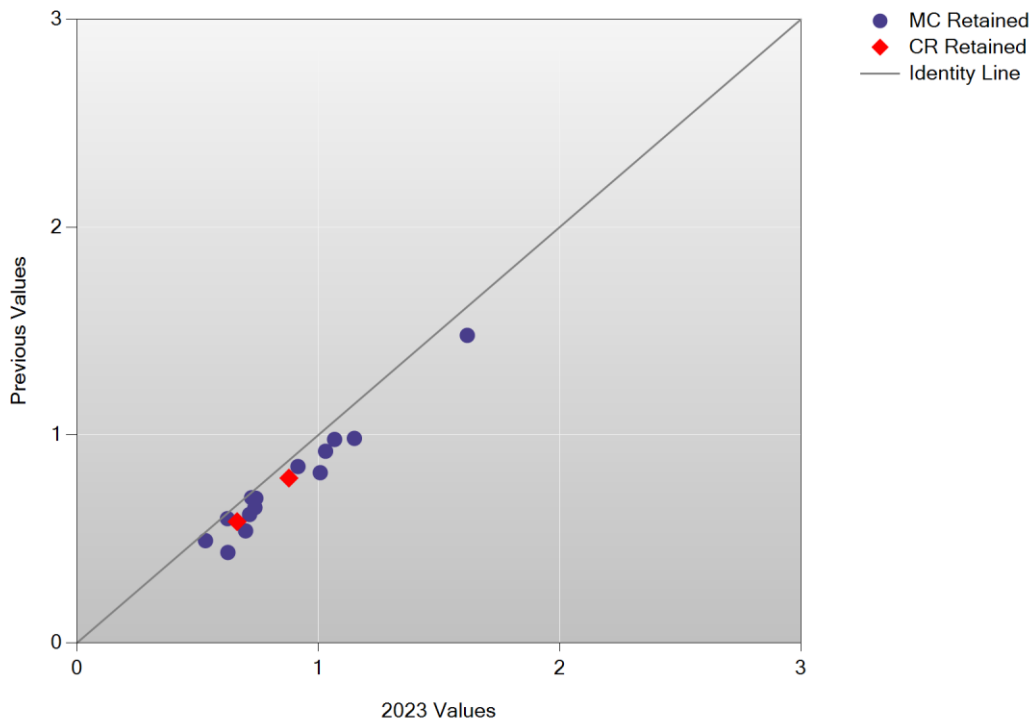
Test Information Function: ela04



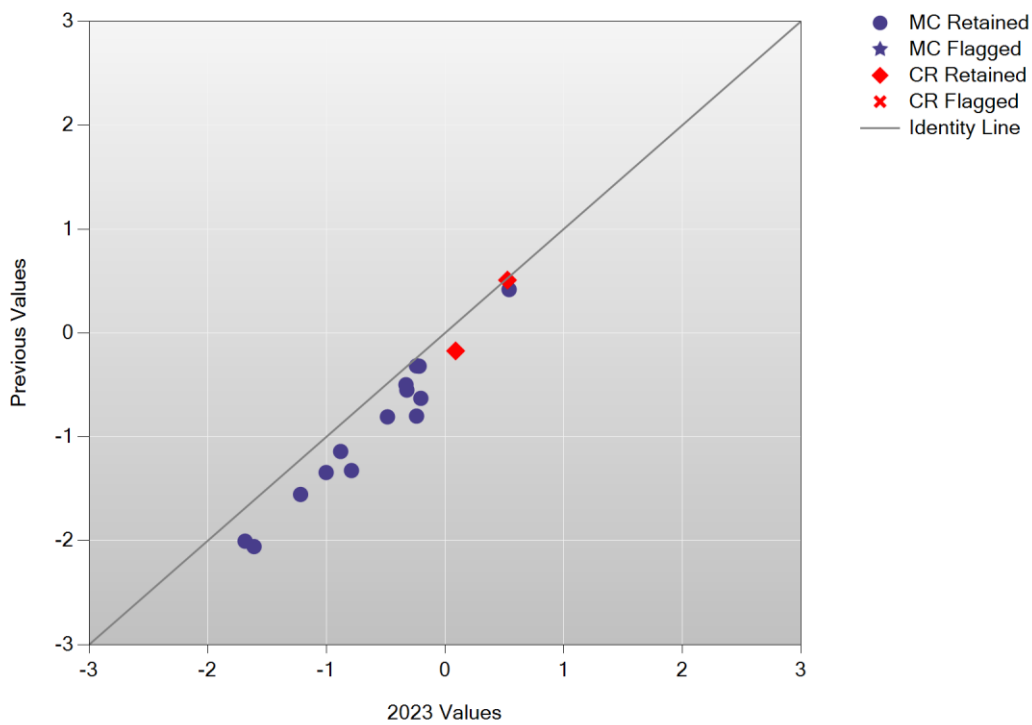
Cumulative Scale Score Distributions: English Language Arts Grade 4



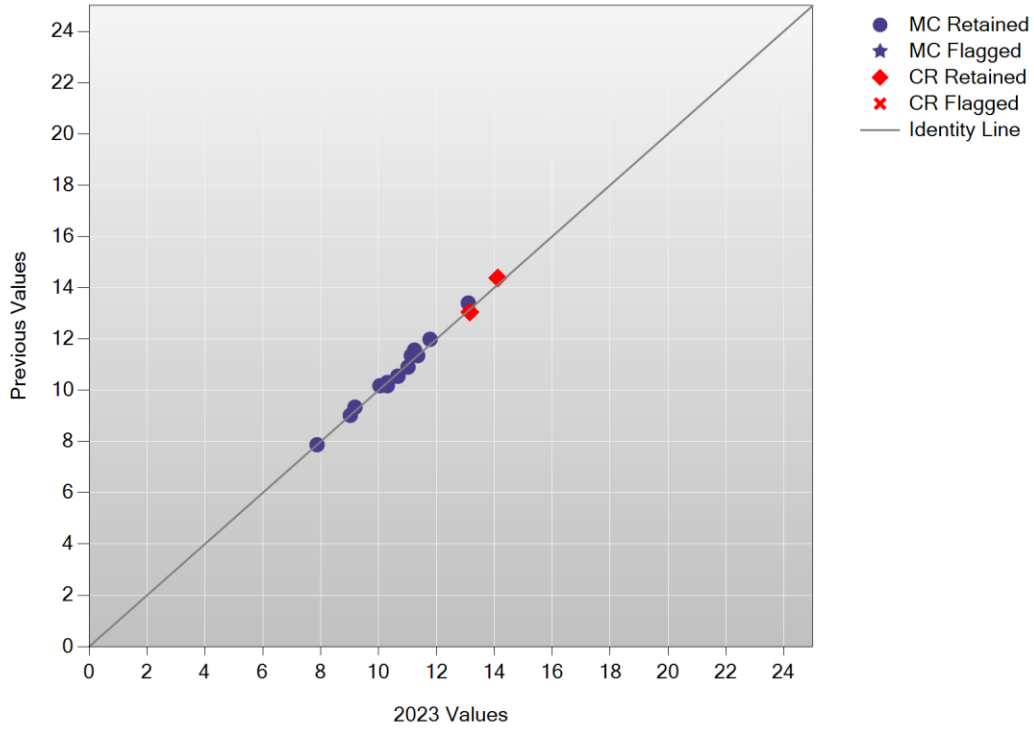
A/A Plot: English Language Arts Grade 5



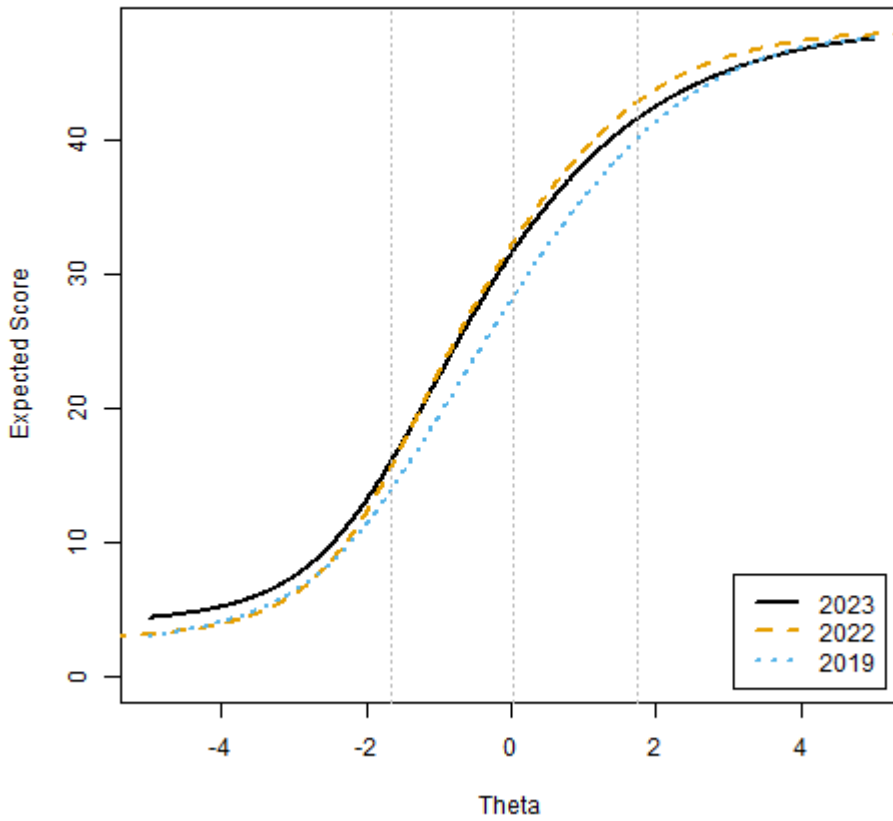
B/B Plot: English Language Arts Grade 5



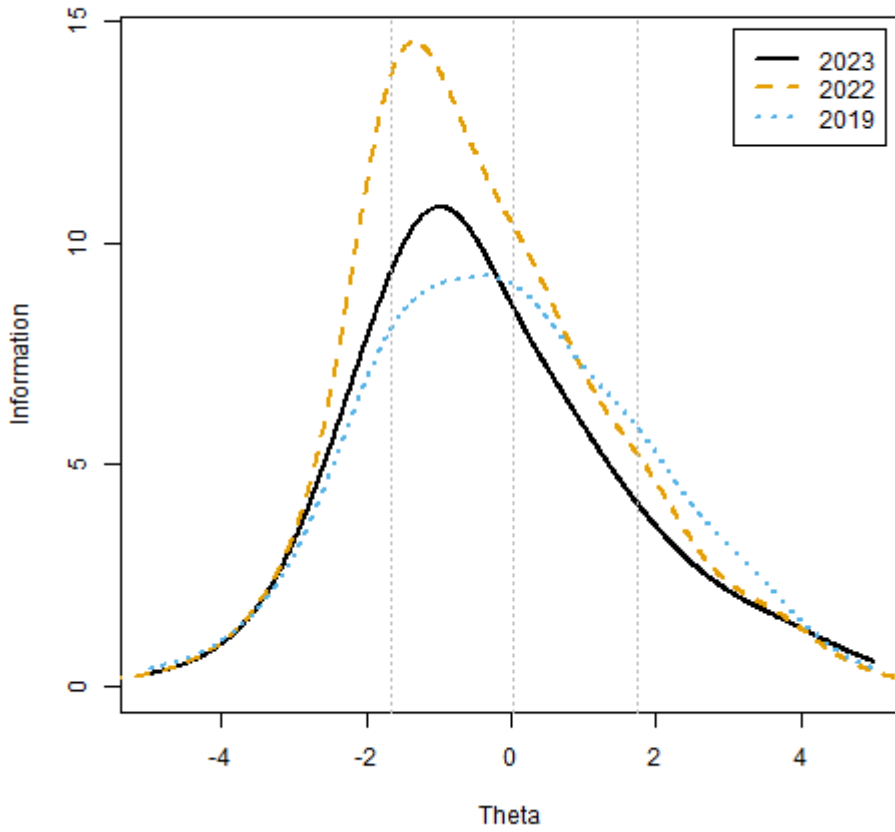
Delta Plot: English Language Arts Grade 5



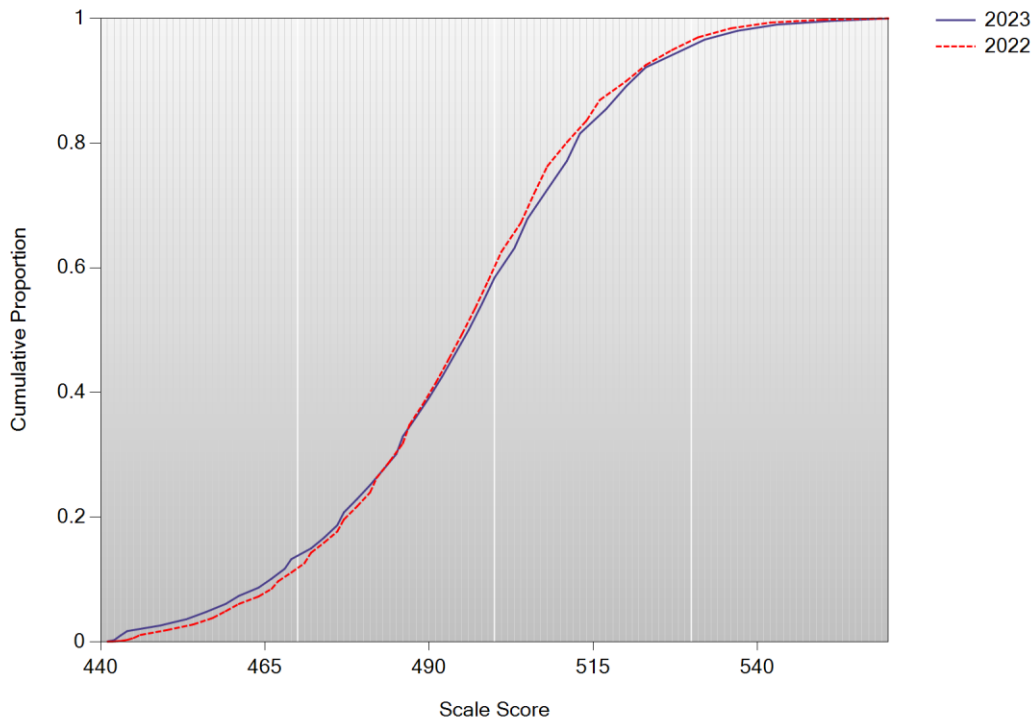
Test Characteristic Curve: ela05



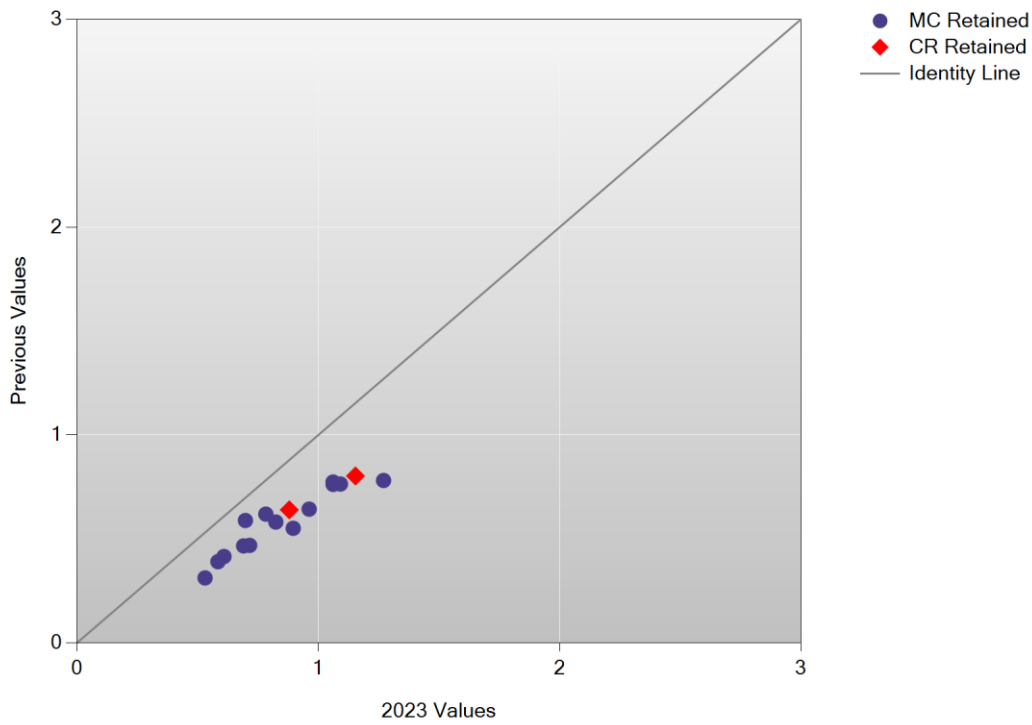
Test Information Function: ela05



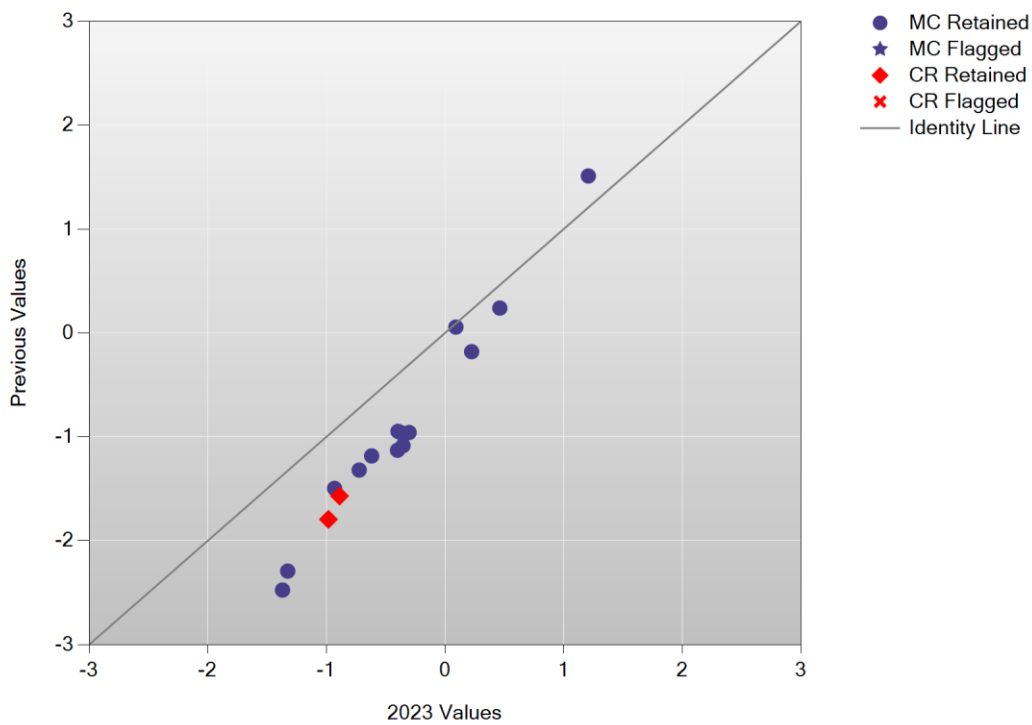
Cumulative Scale Score Distributions: English Language Arts Grade 5



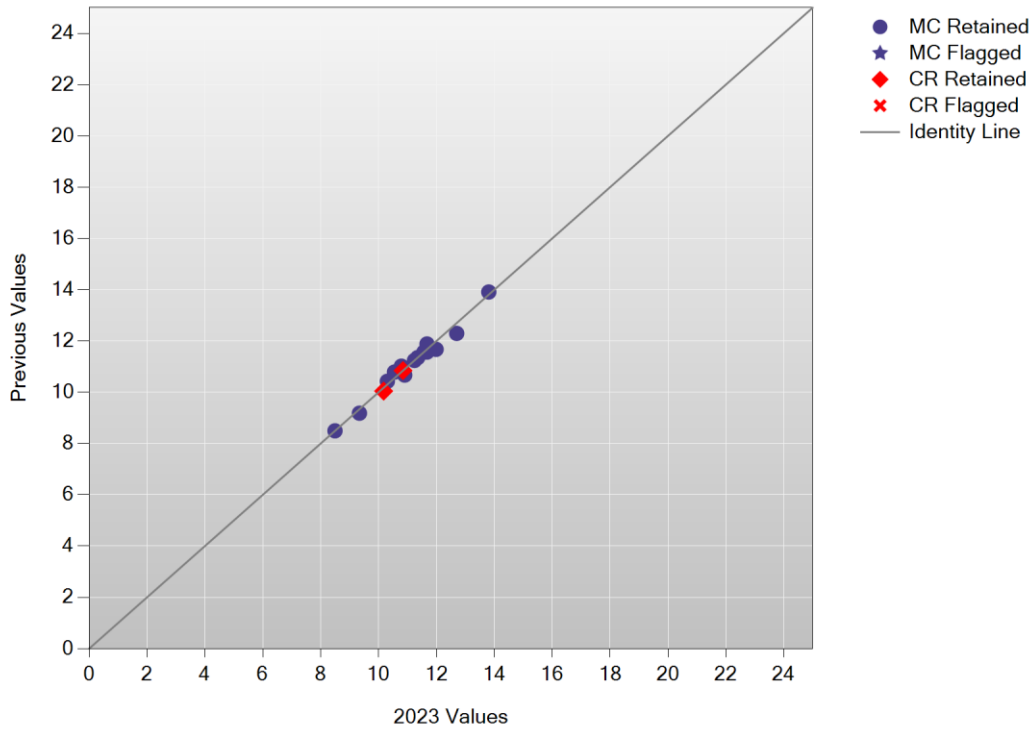
A/A Plot: English Language Arts Grade 6



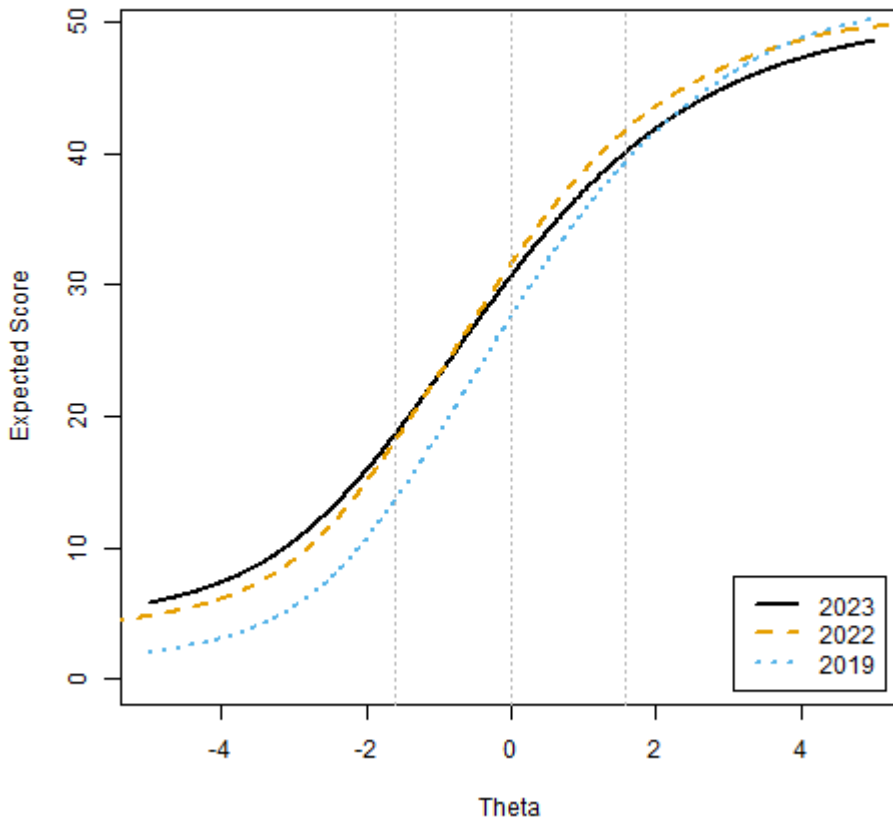
B/B Plot: English Language Arts Grade 6



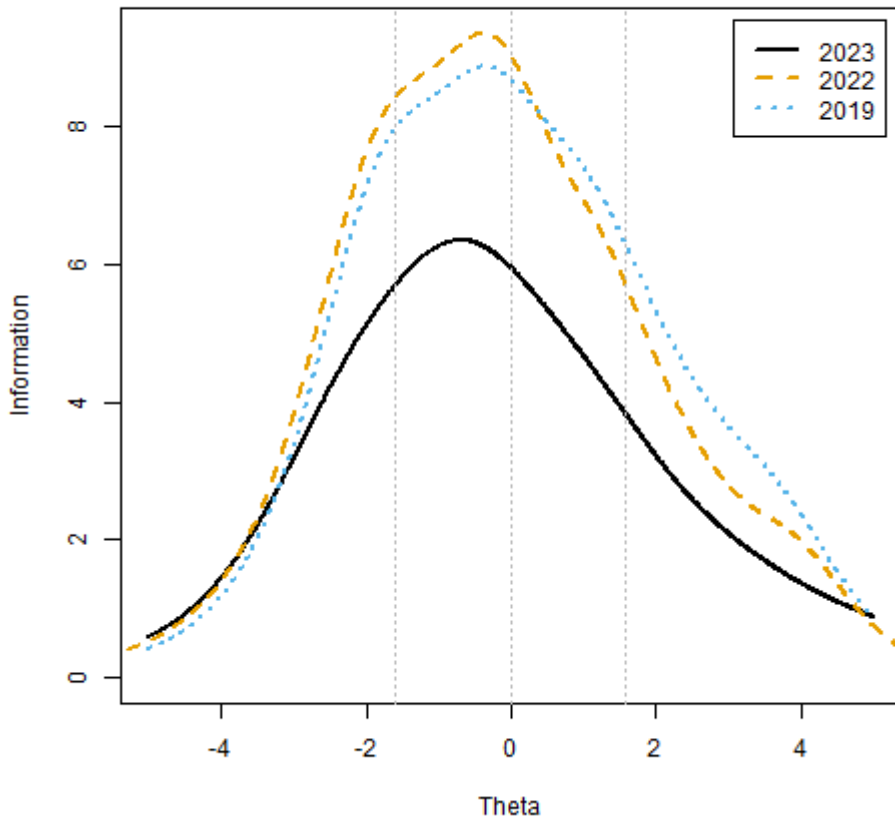
Delta Plot: English Language Arts Grade 6



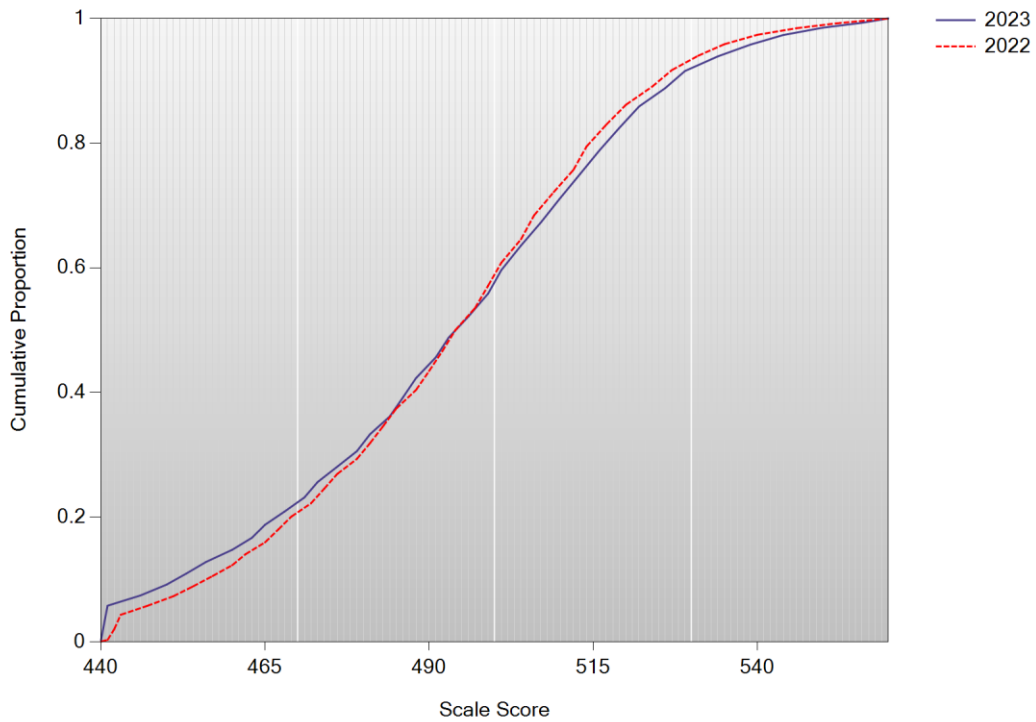
Test Characteristic Curve: ela06



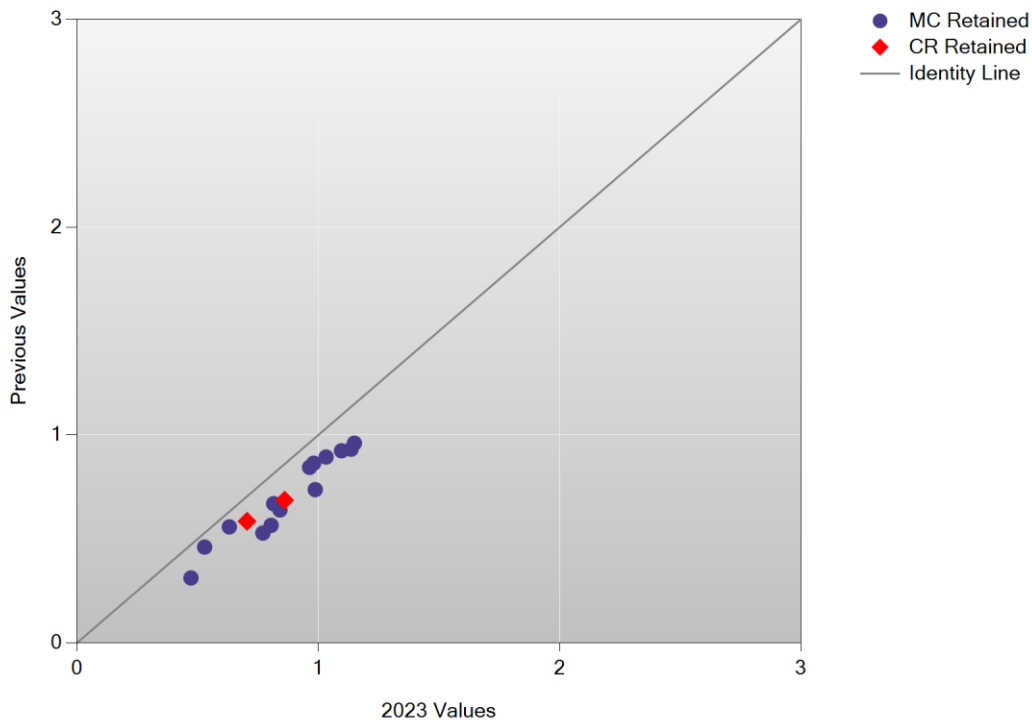
Test Information Function: ela06



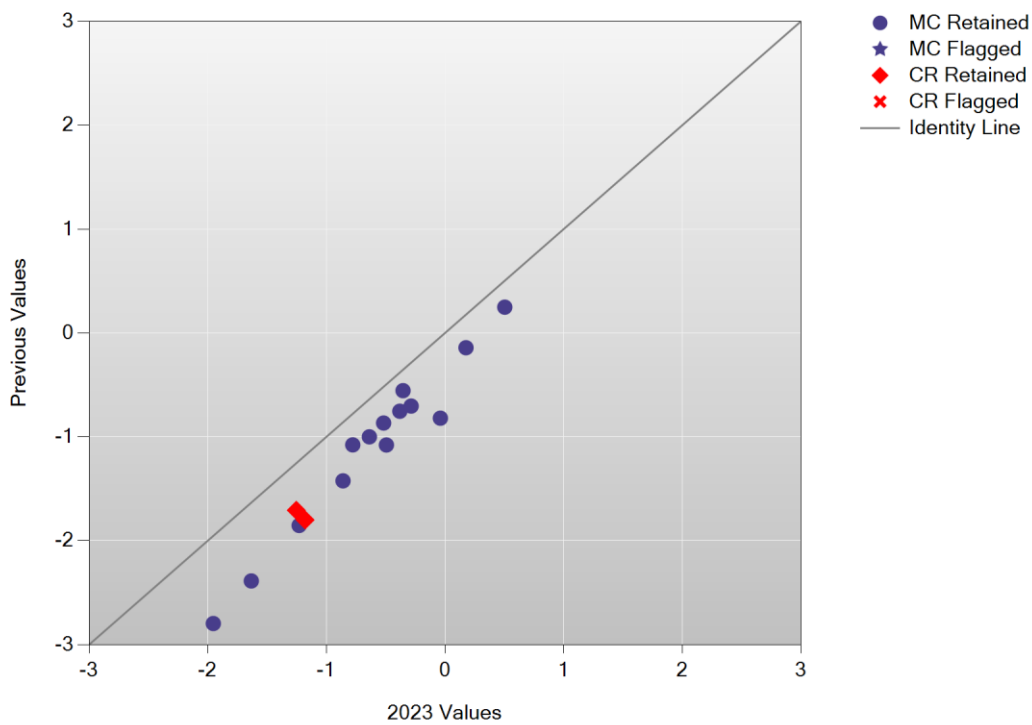
Cumulative Scale Score Distributions: English Language Arts Grade 6



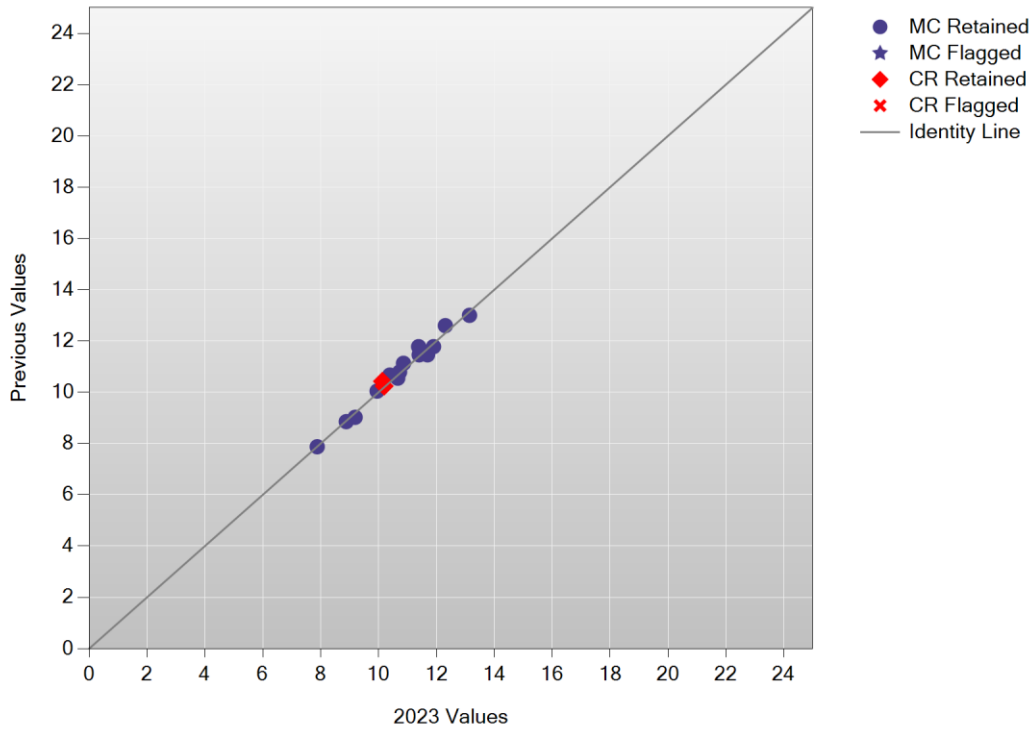
A/A Plot: English Language Arts Grade 7



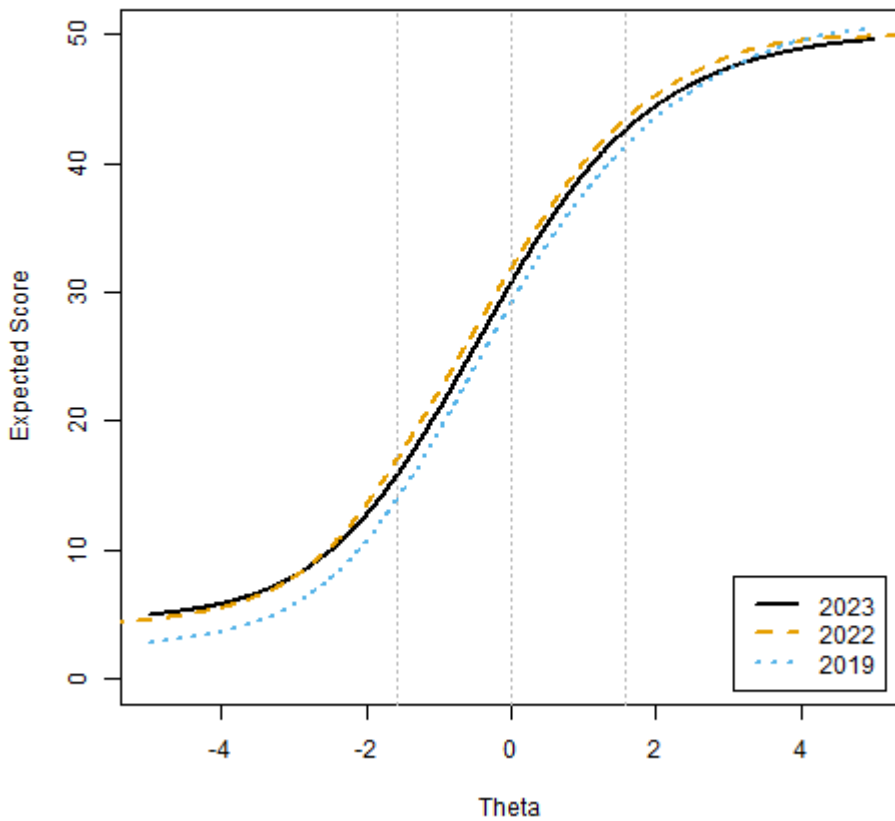
B/B Plot: English Language Arts Grade 7



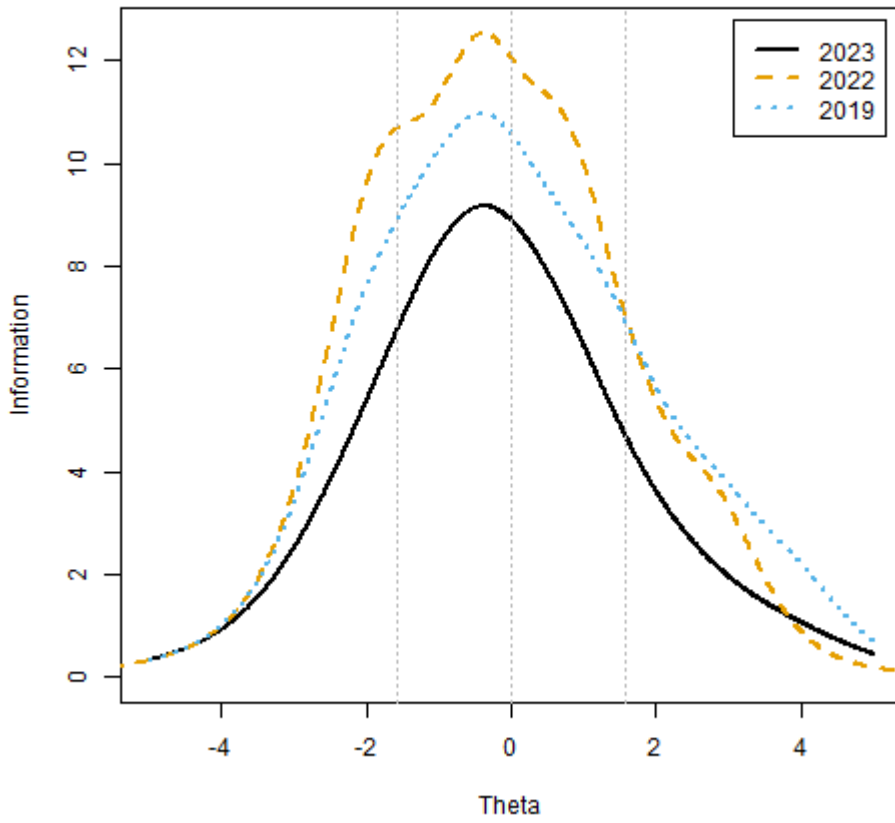
Delta Plot: English Language Arts Grade 7



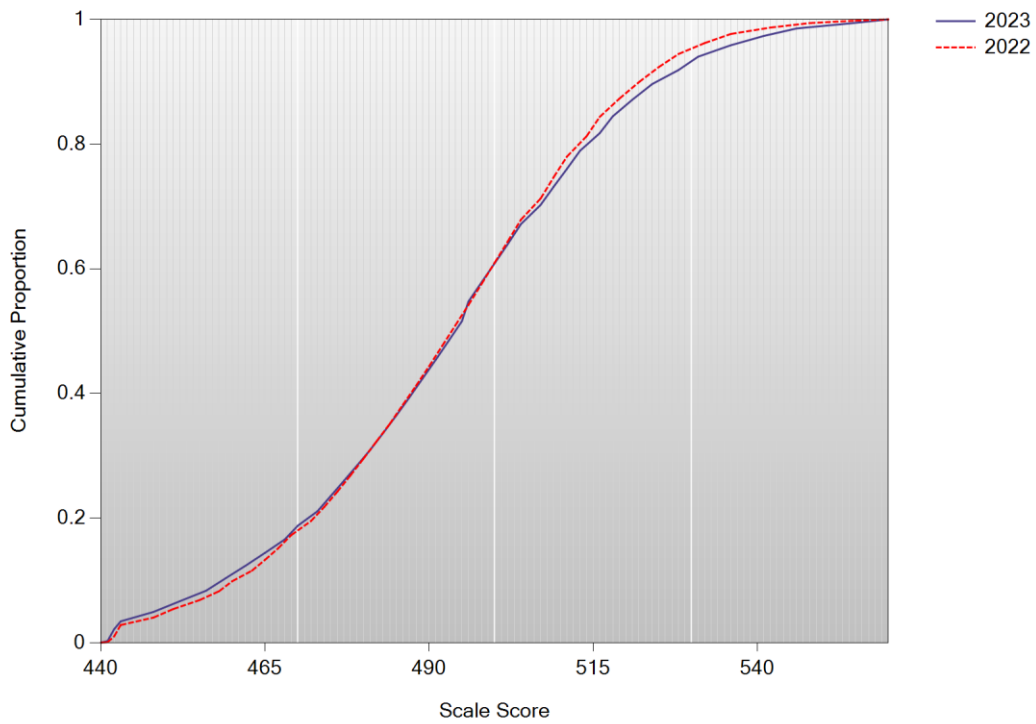
Test Characteristic Curve: ela07



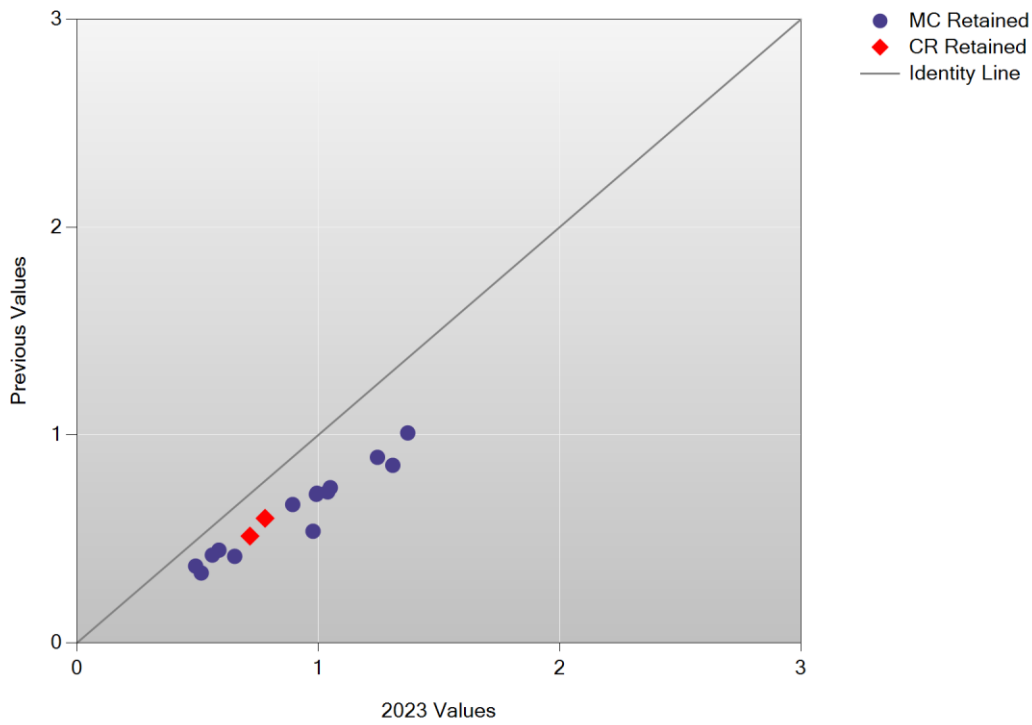
Test Information Function: ela07



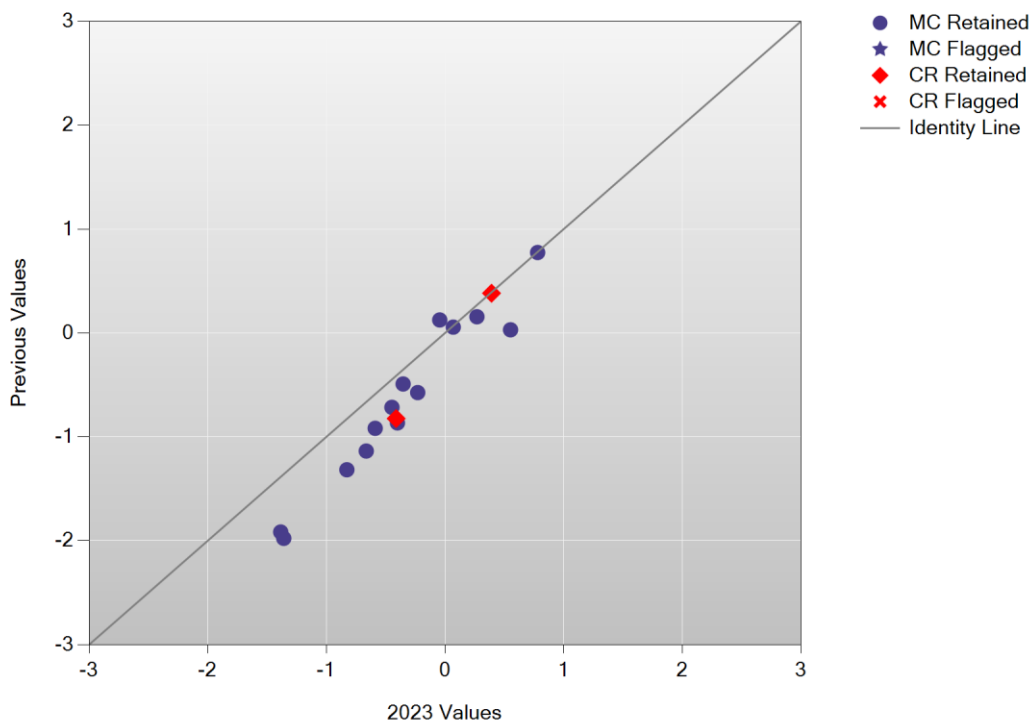
Cumulative Scale Score Distributions: English Language Arts Grade 7



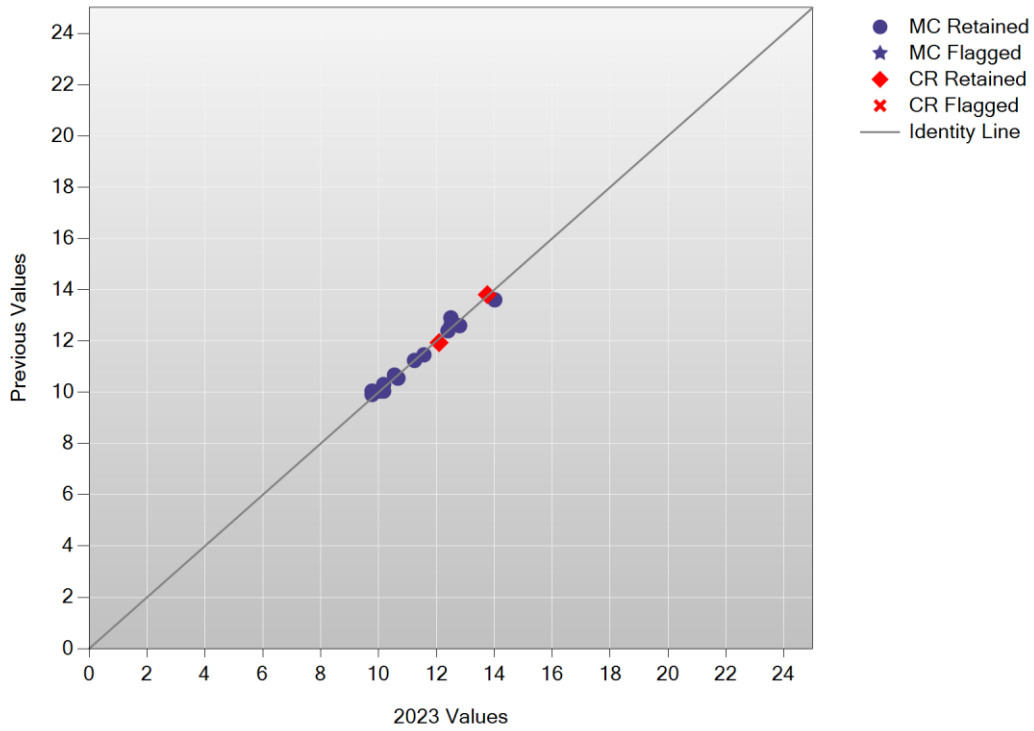
A/A Plot: English Language Arts Grade 8



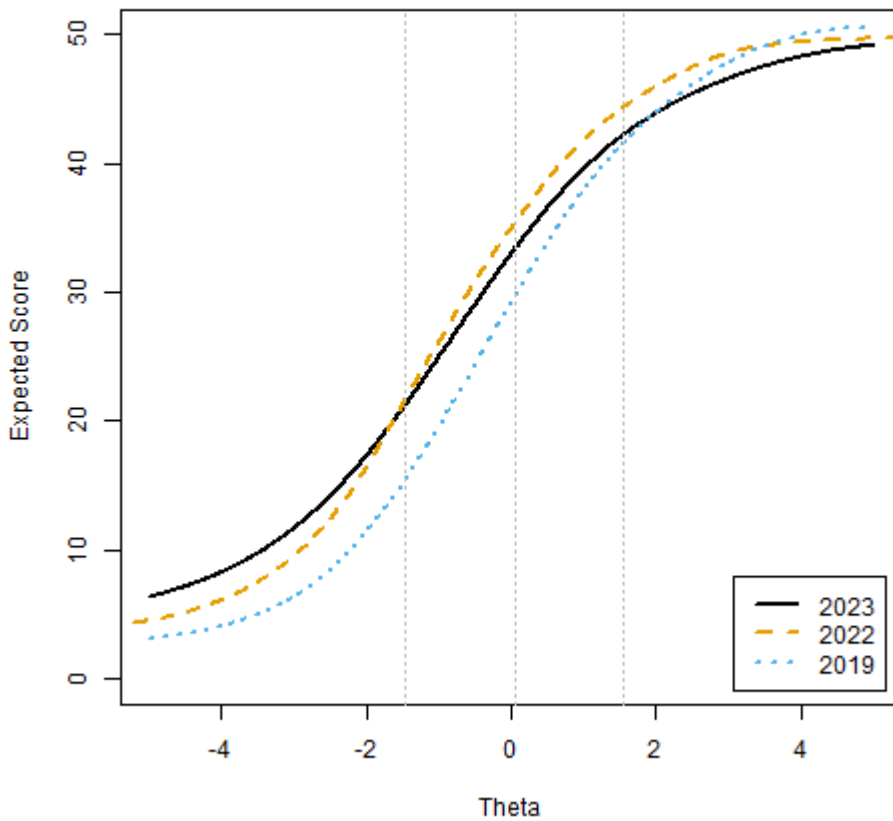
B/B Plot: English Language Arts Grade 8



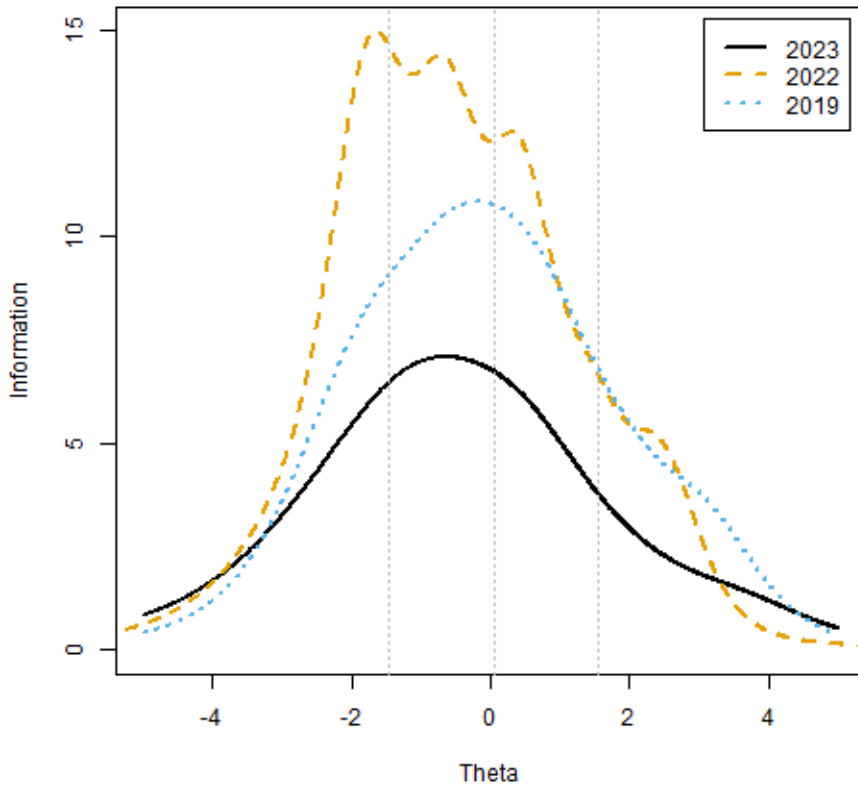
Delta Plot: English Language Arts Grade 8



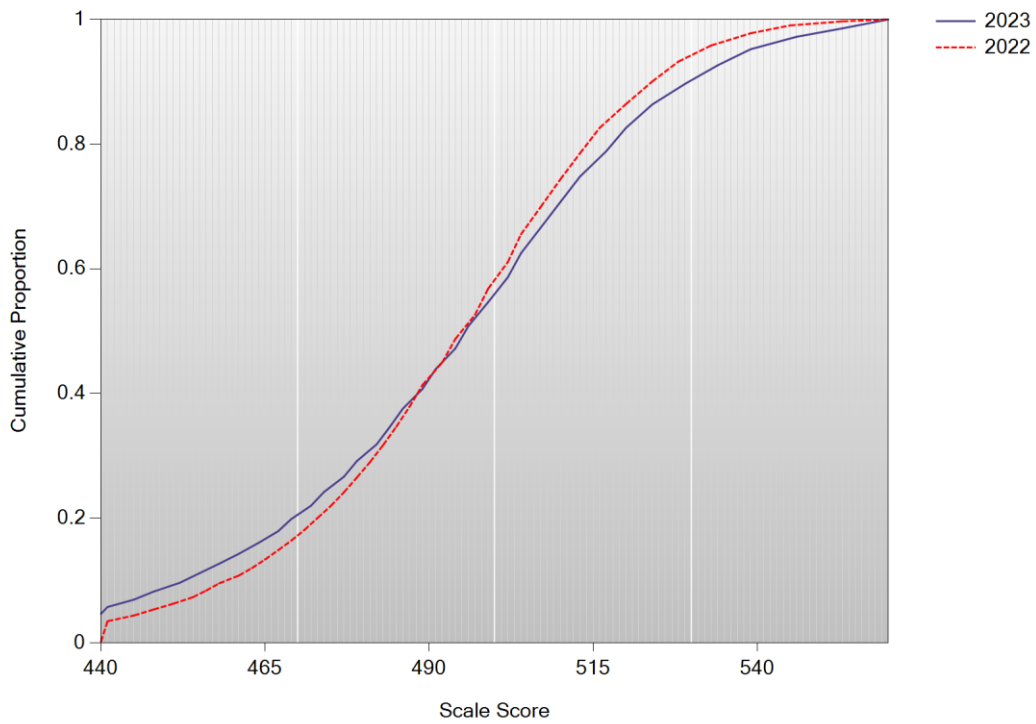
Test Characteristic Curve: ela08



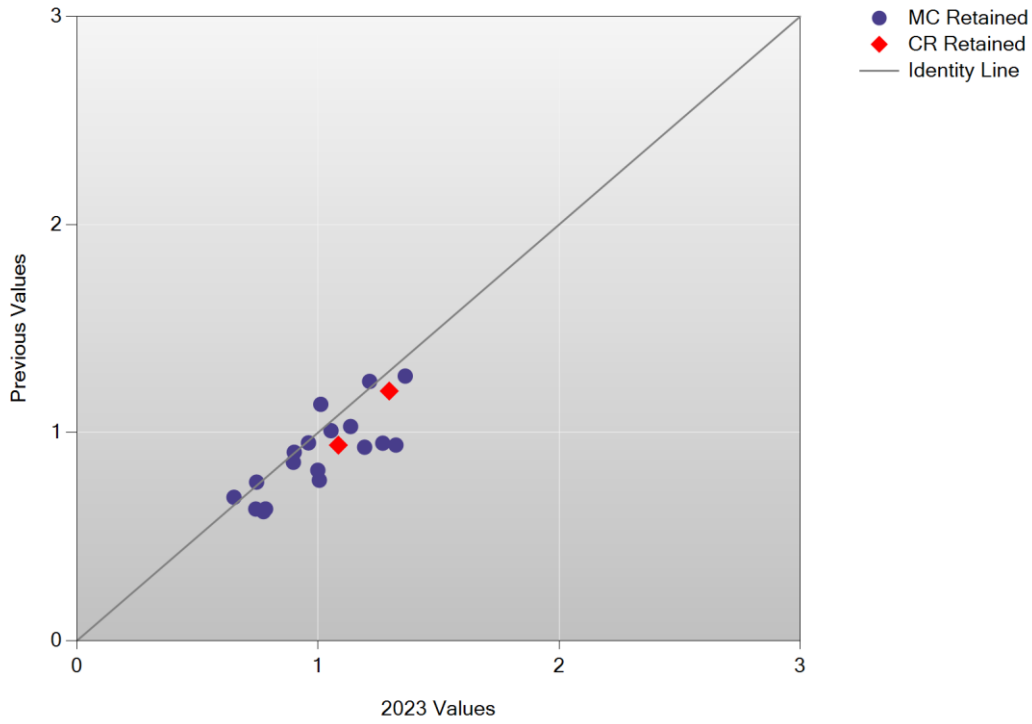
Test Information Function: ela08



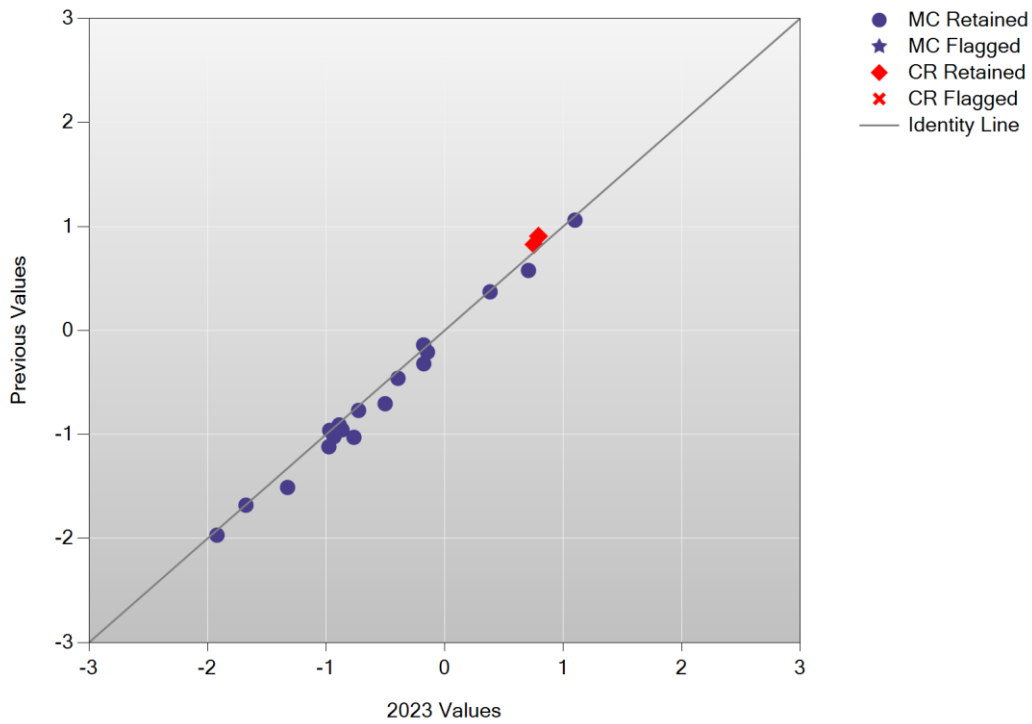
Cumulative Scale Score Distributions: English Language Arts Grade 8



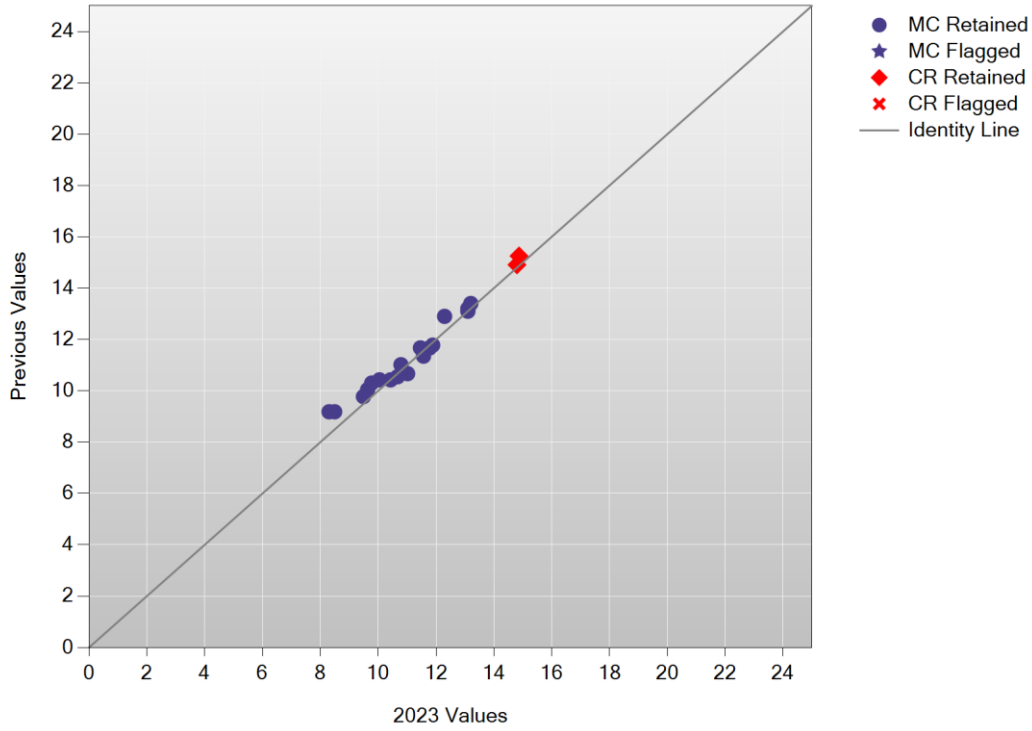
A/A Plot: Mathematics Grade 3



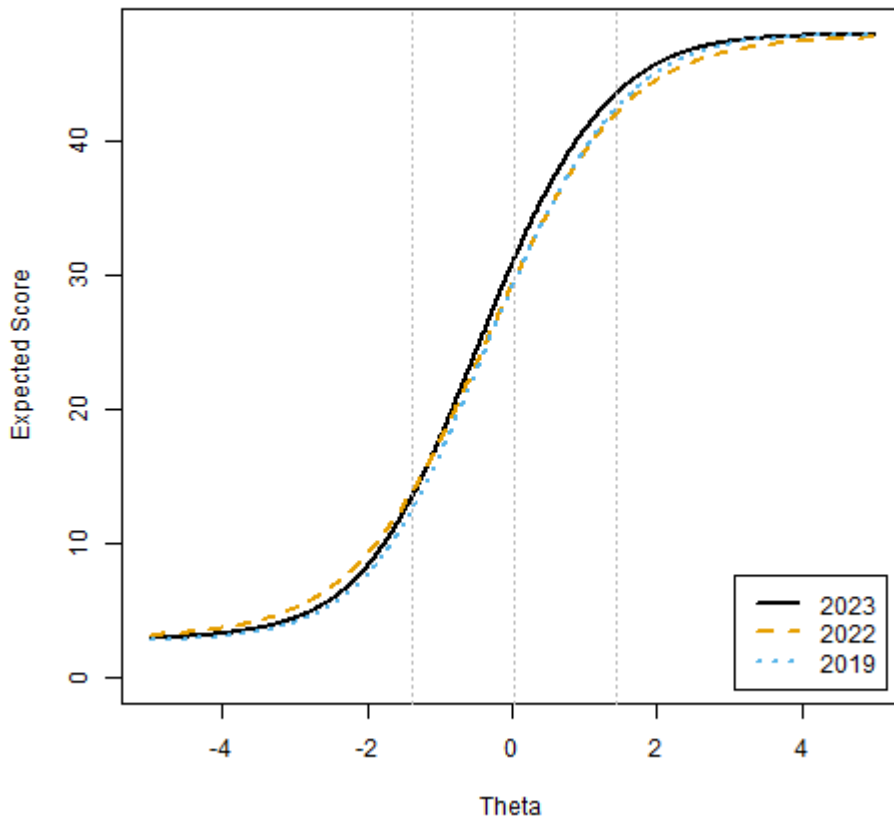
B/B Plot: Mathematics Grade 3



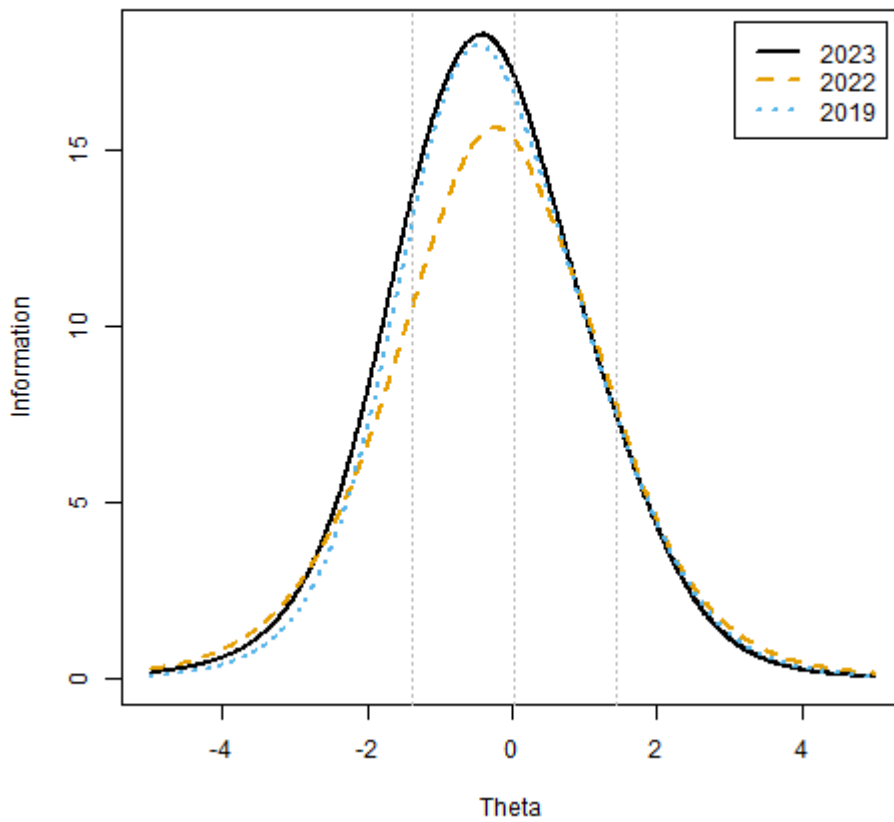
Delta Plot: Mathematics Grade 3



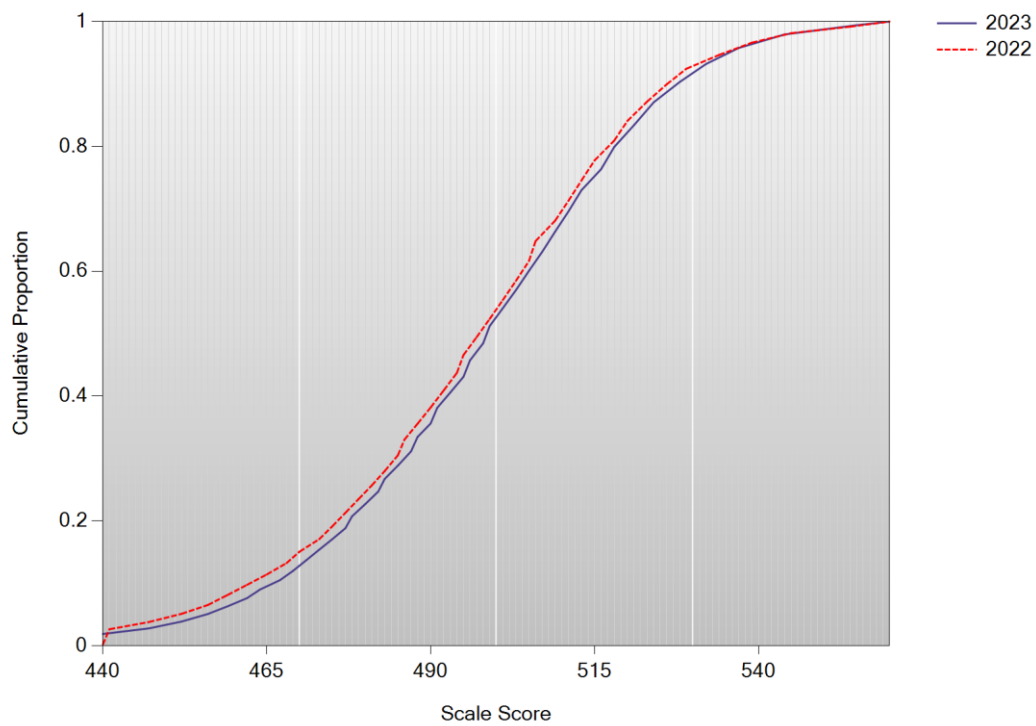
Test Characteristic Curve: mat03



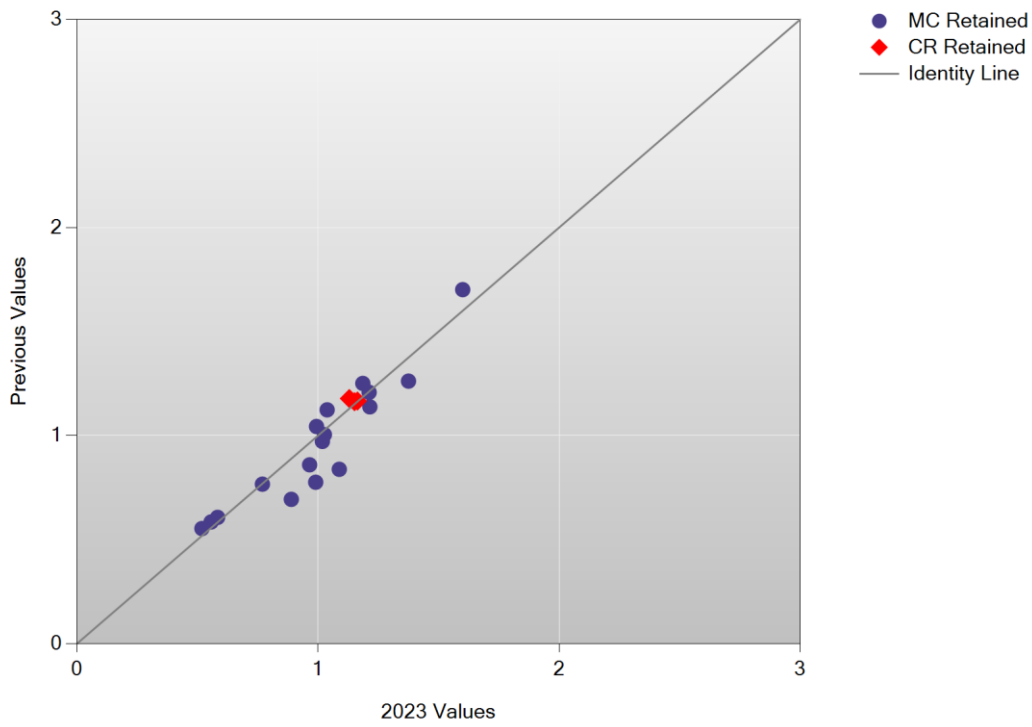
Test Information Function: mat03



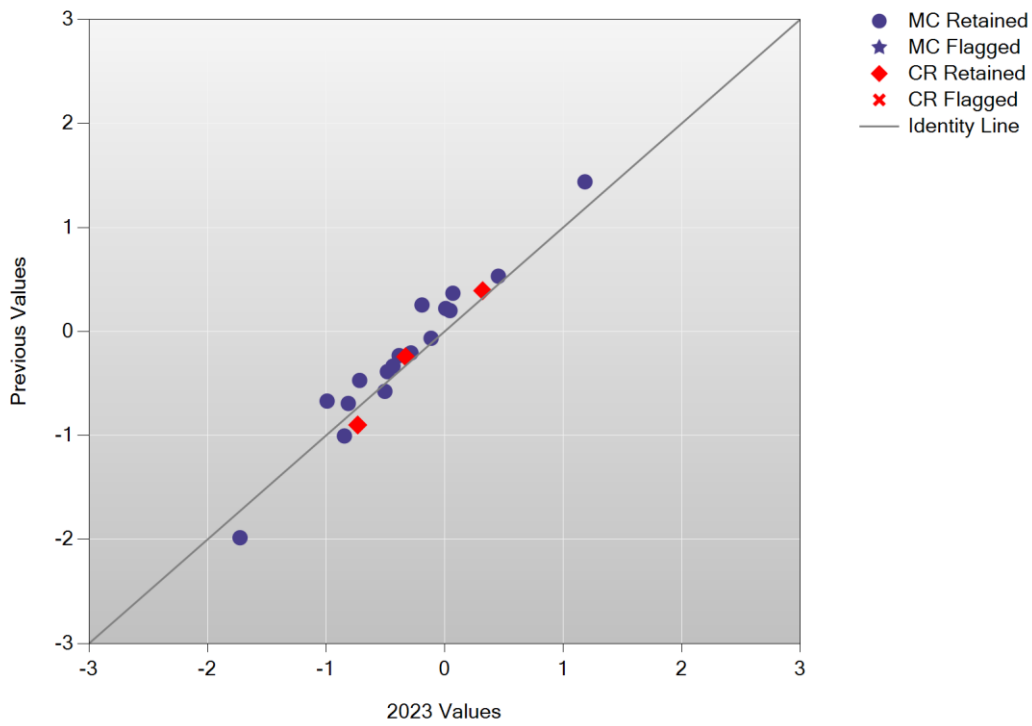
Cumulative Scale Score Distributions: Mathematics Grade 3



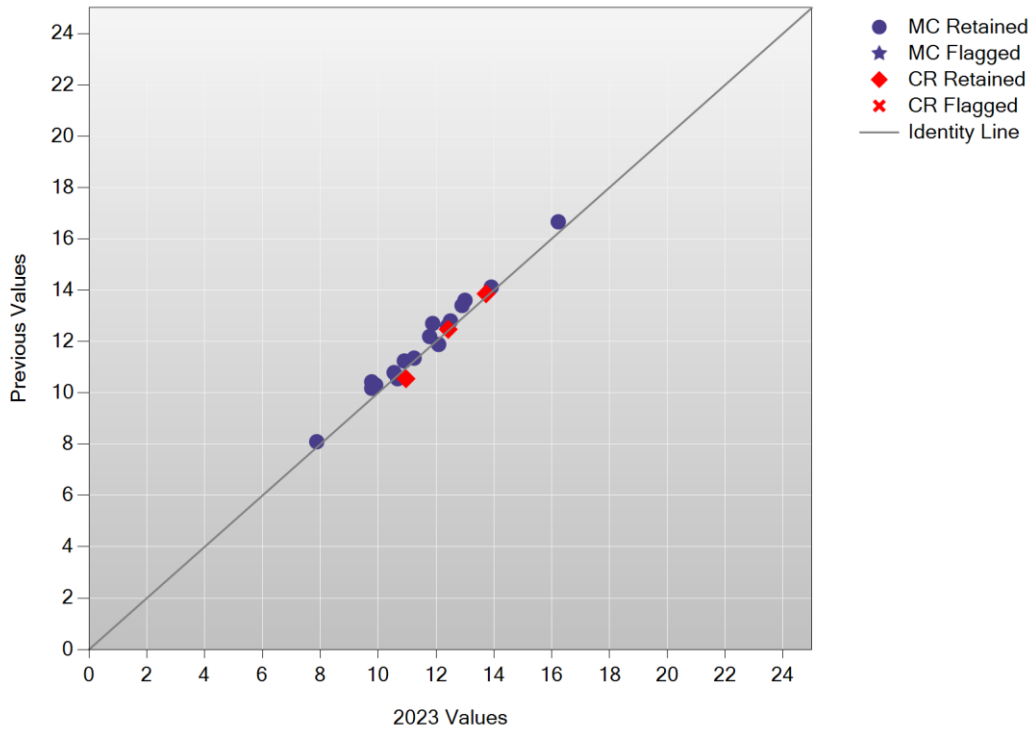
A/A Plot: Mathematics Grade 4



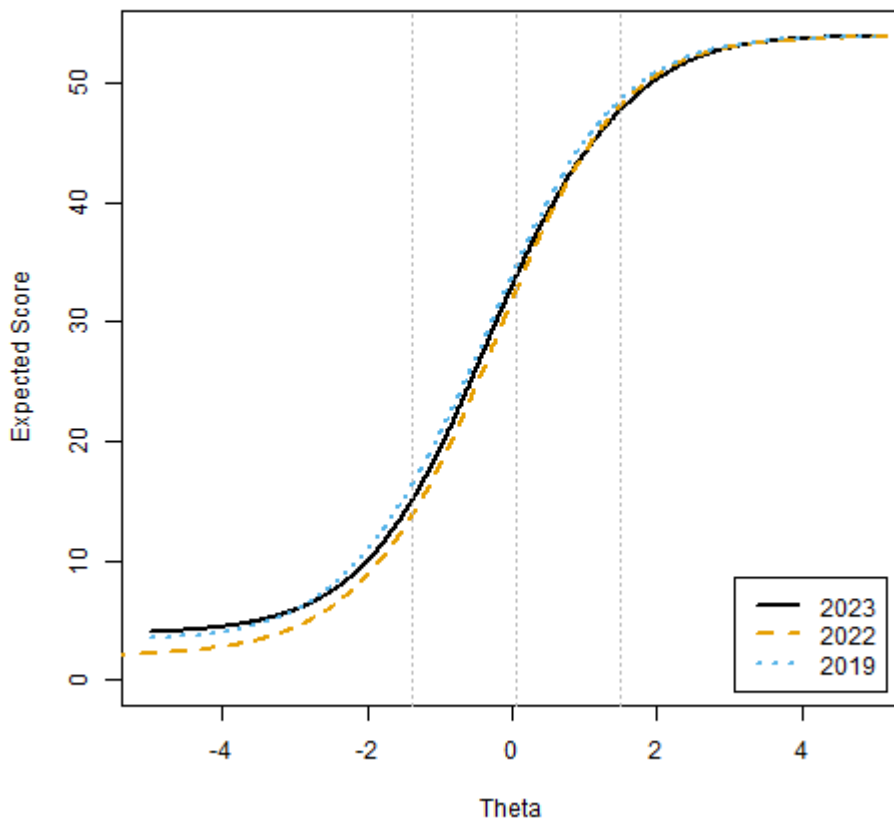
B/B Plot: Mathematics Grade 4



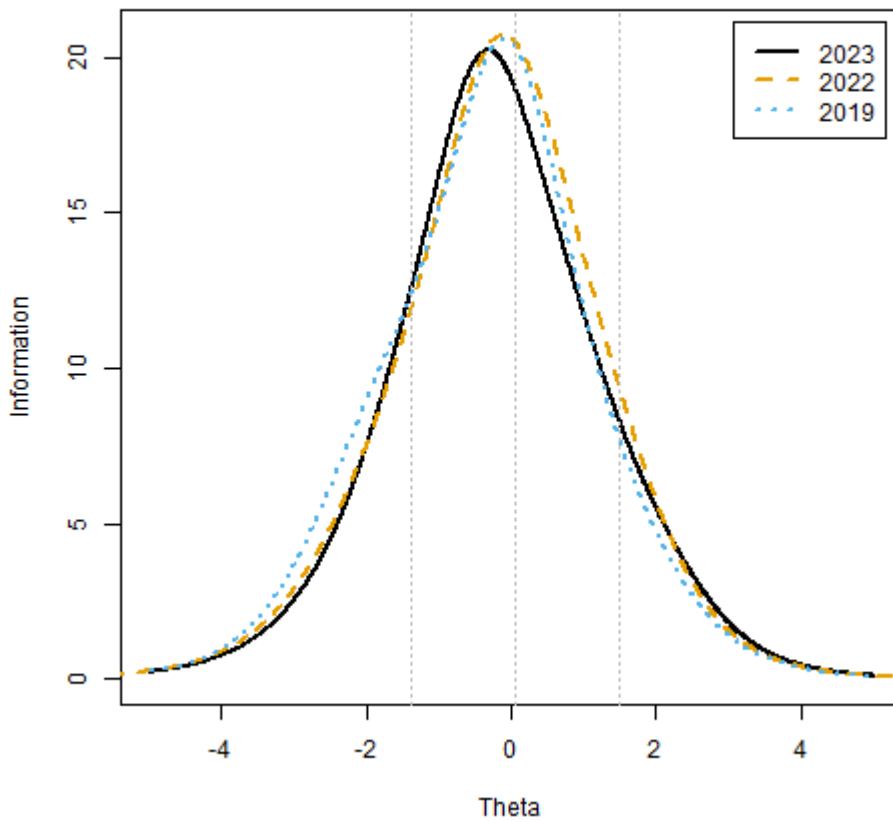
Delta Plot: Mathematics Grade 4



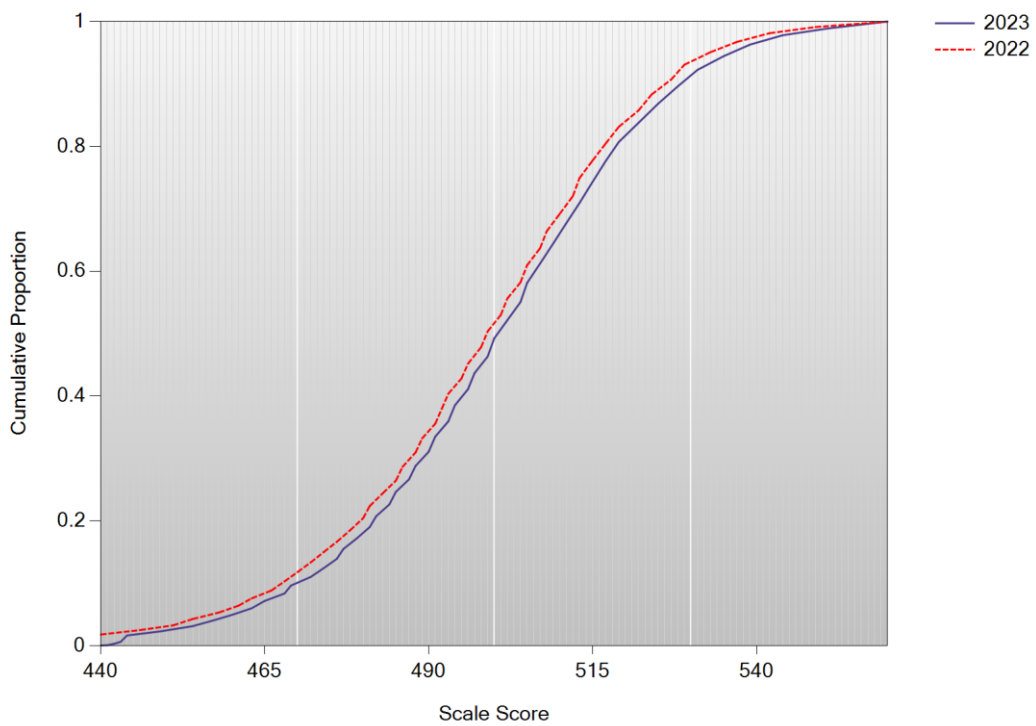
Test Characteristic Curve: mat04



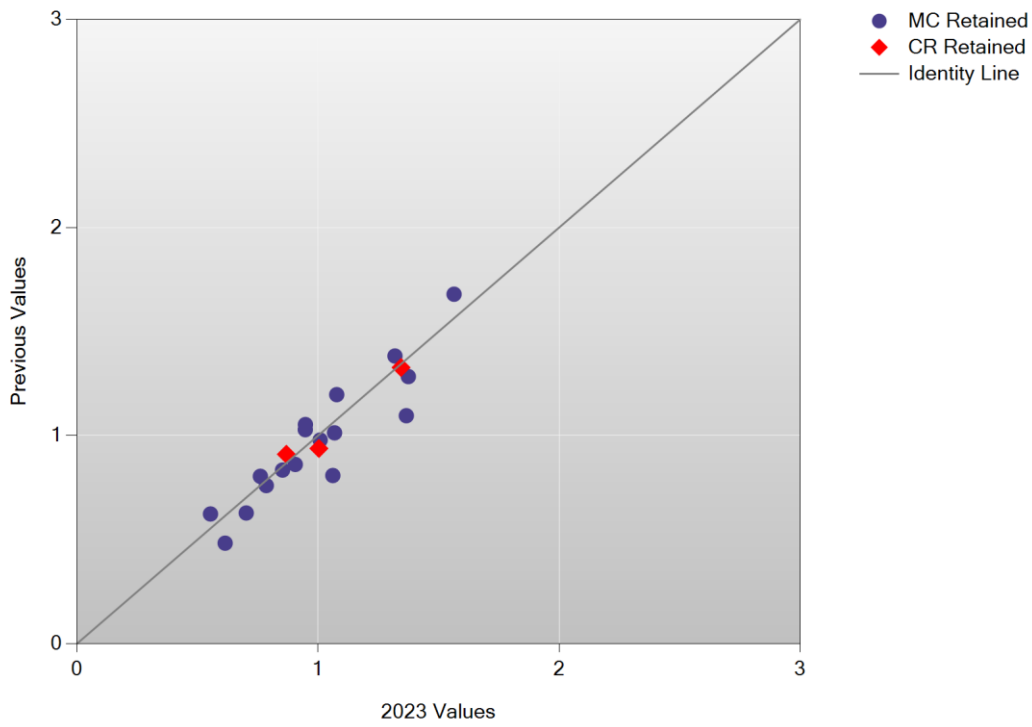
Test Information Function: mat04



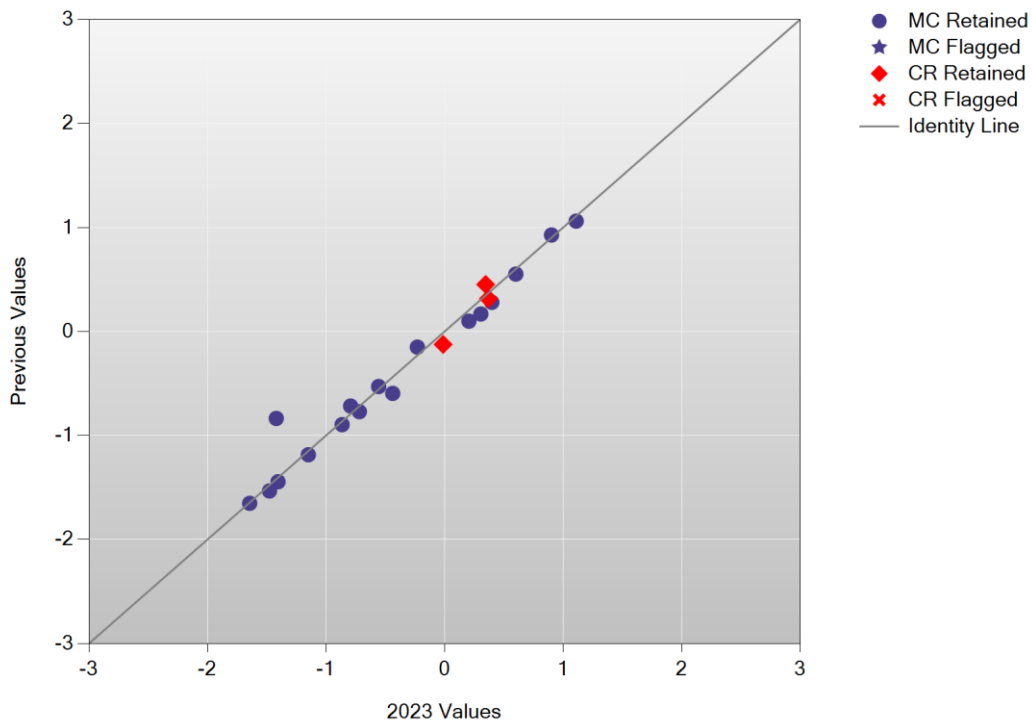
Cumulative Scale Score Distributions: Mathematics Grade 4



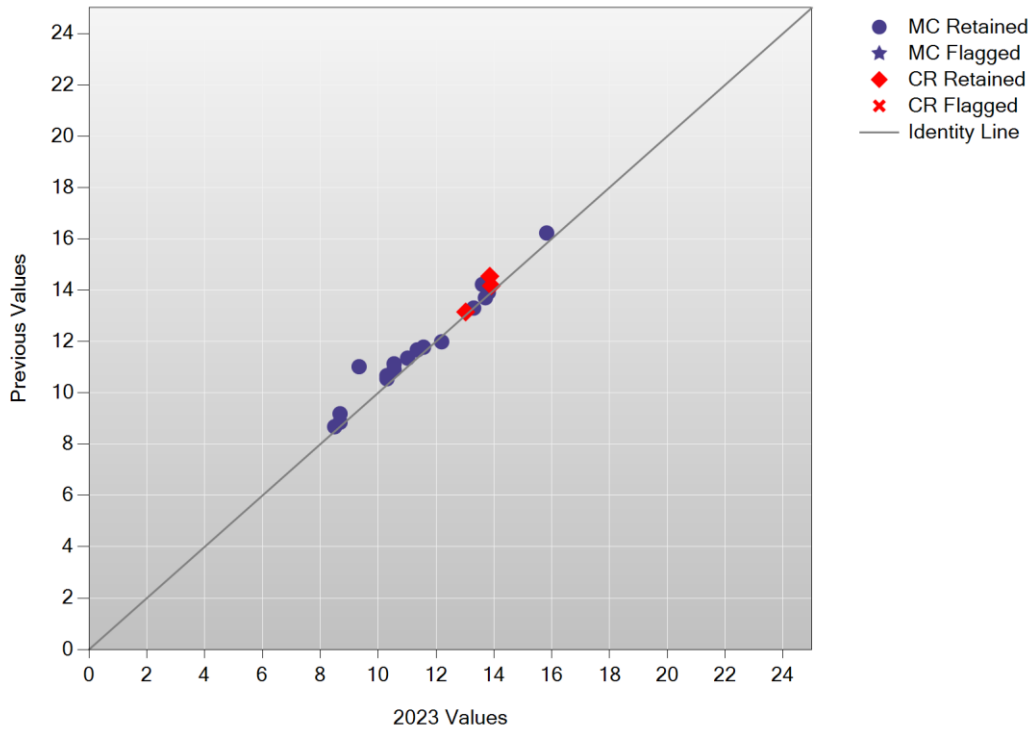
A/A Plot: Mathematics Grade 5



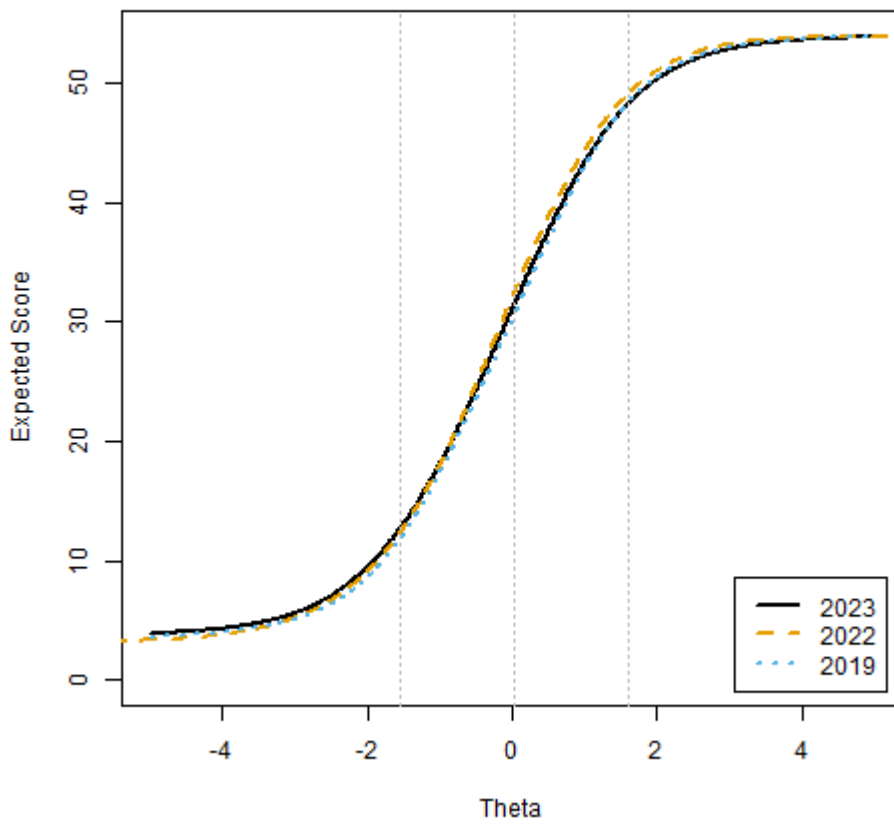
B/B Plot: Mathematics Grade 5



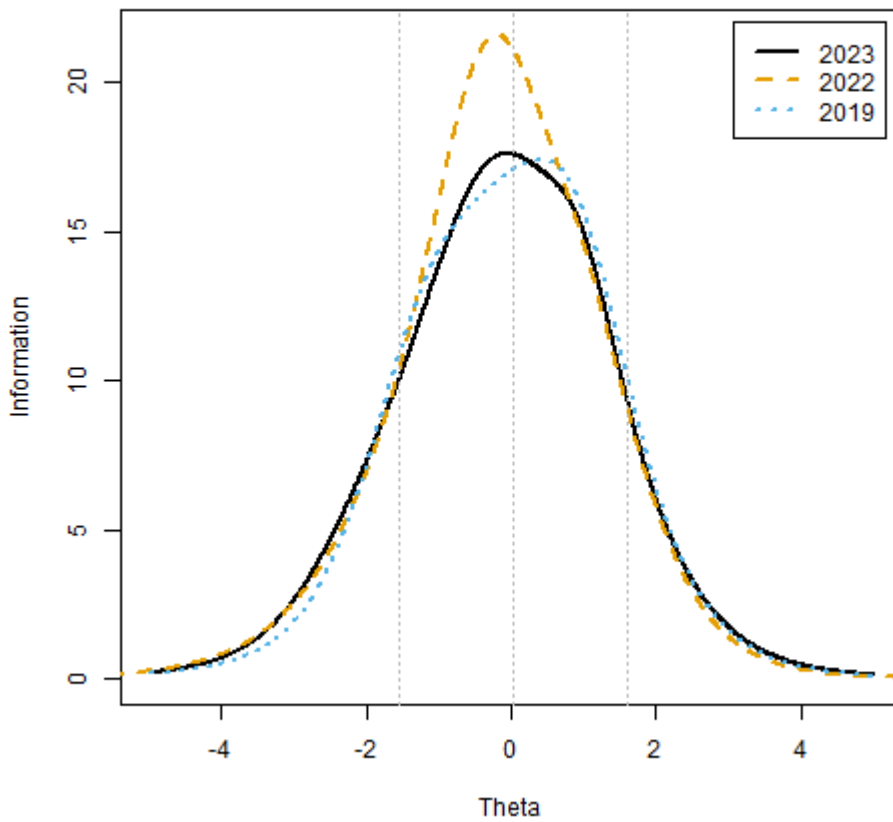
Delta Plot: Mathematics Grade 5



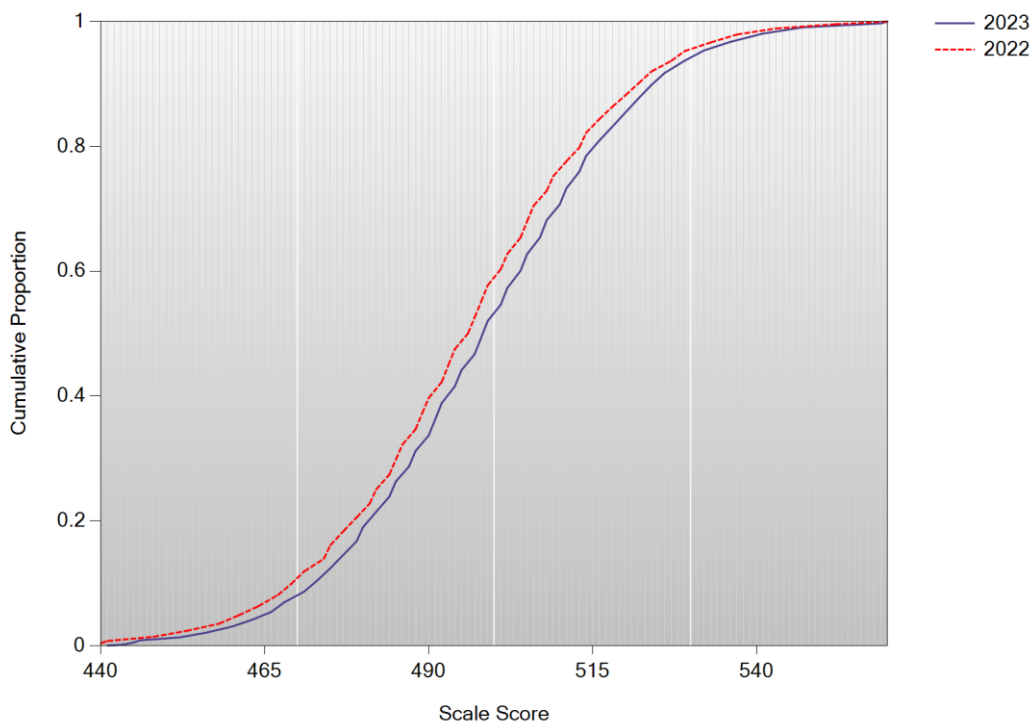
Test Characteristic Curve: mat05



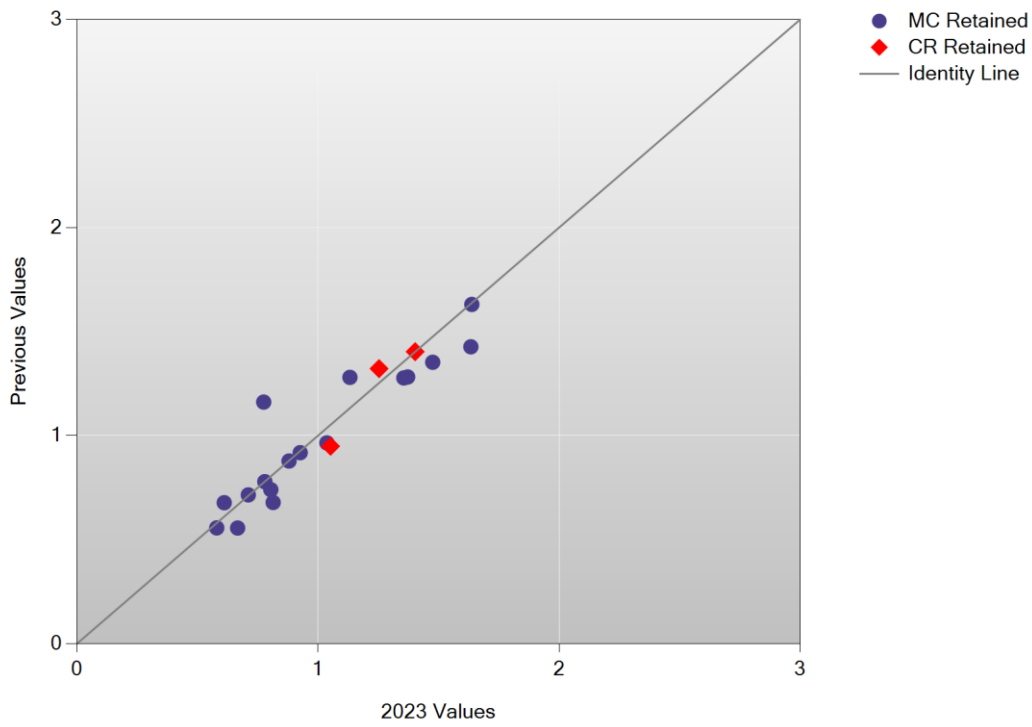
Test Information Function: mat05



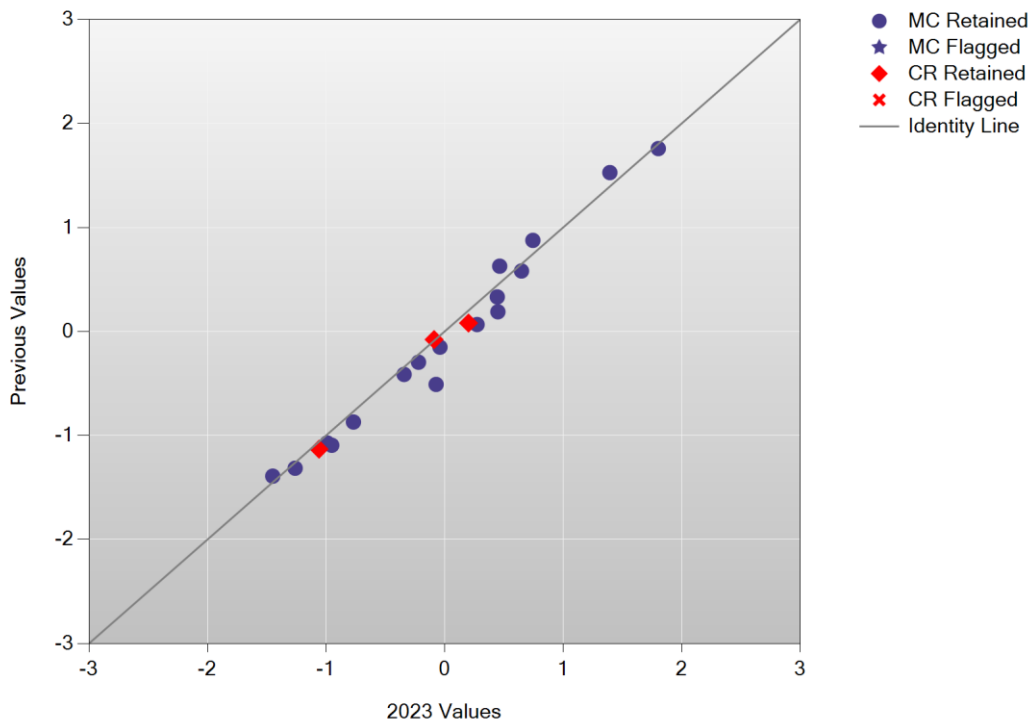
Cumulative Scale Score Distributions: Mathematics Grade 5



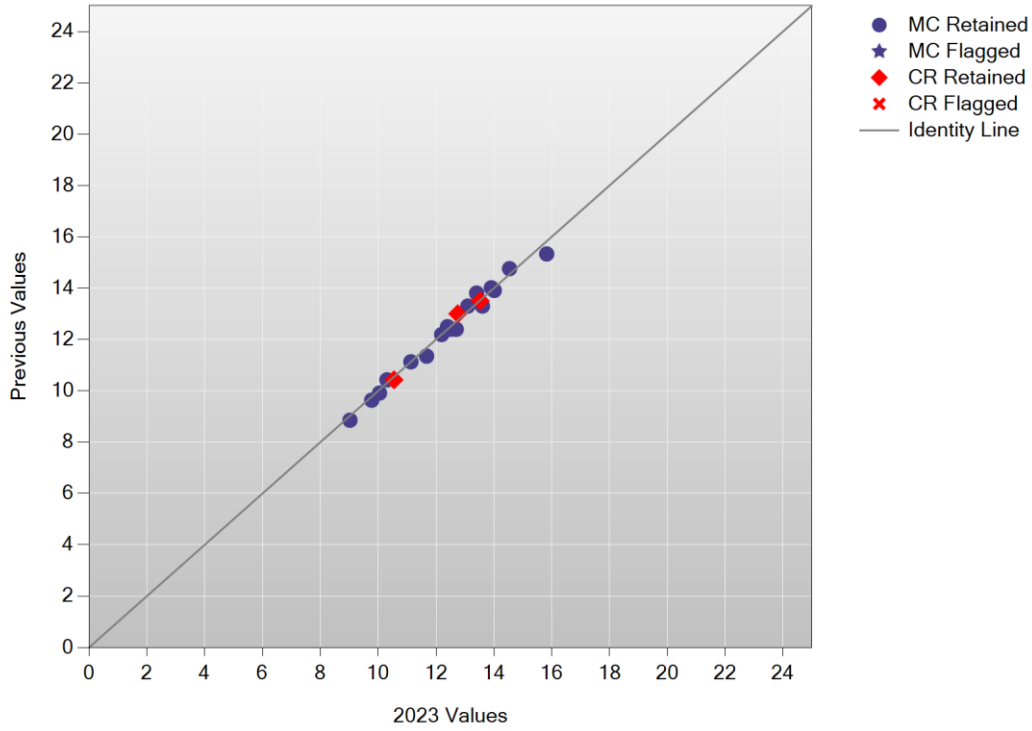
A/A Plot: Mathematics Grade 6



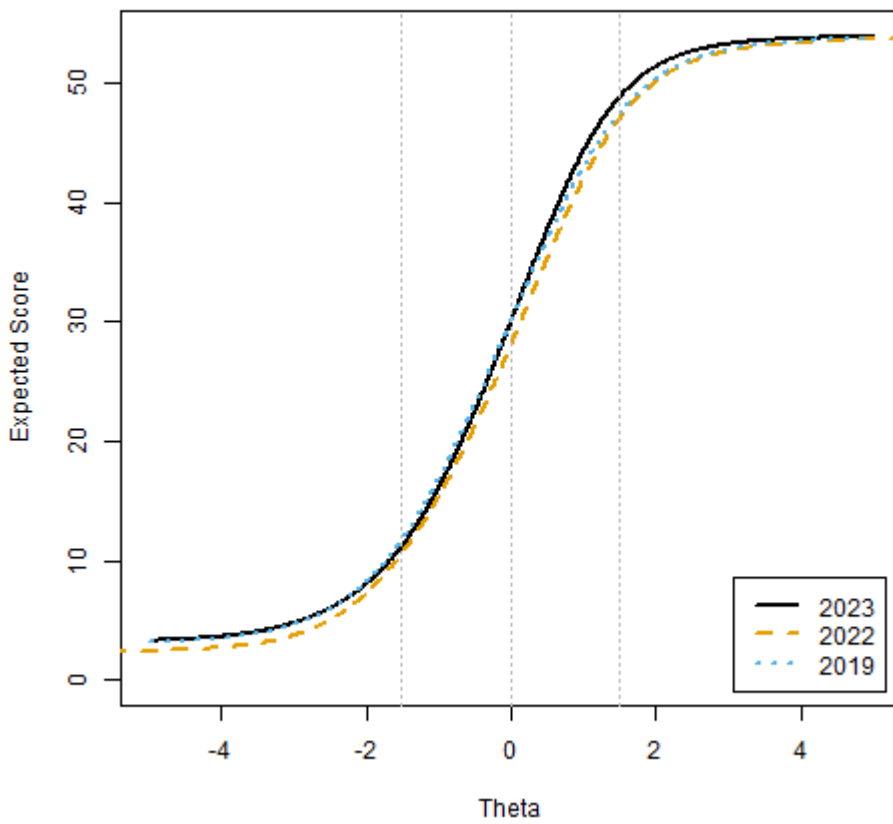
B/B Plot: Mathematics Grade 6



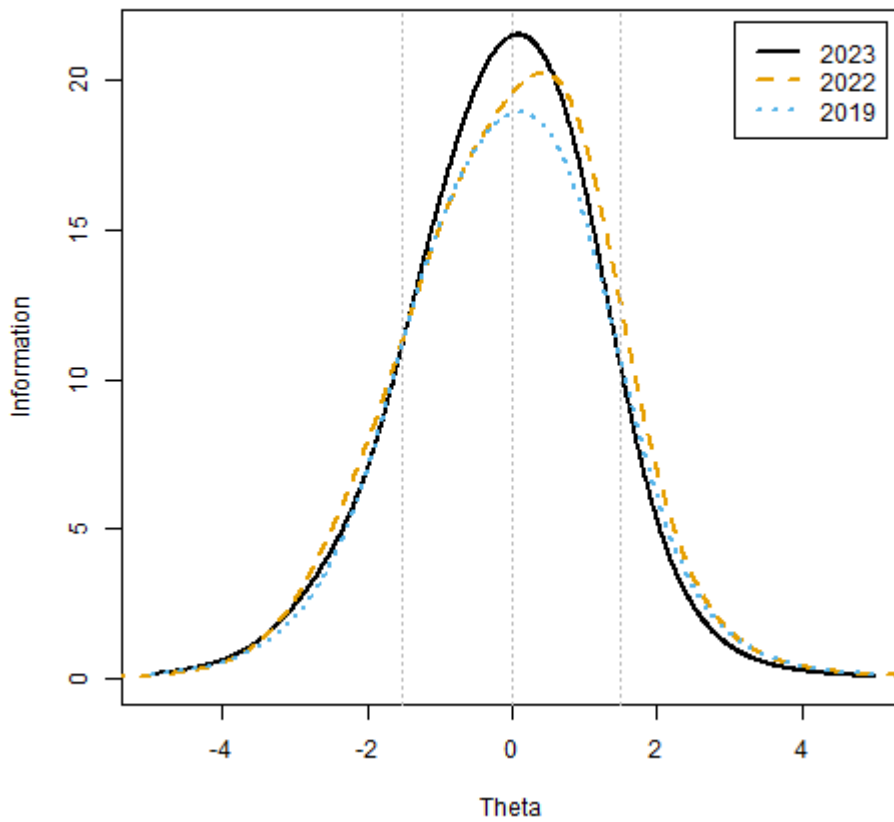
Delta Plot: Mathematics Grade 6



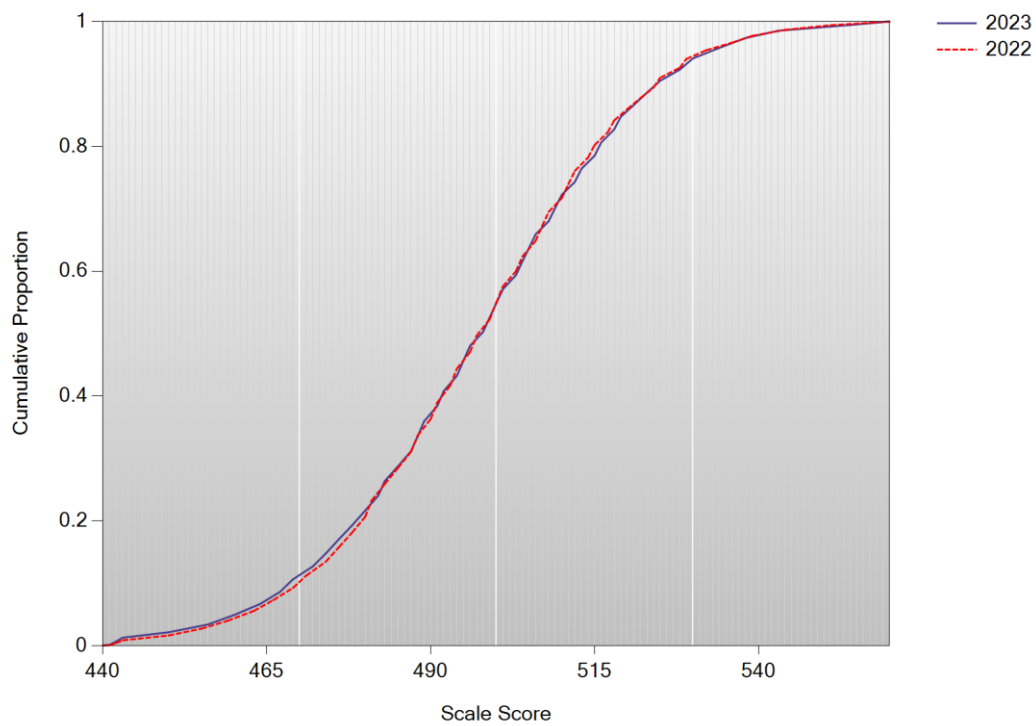
Test Characteristic Curve: mat06



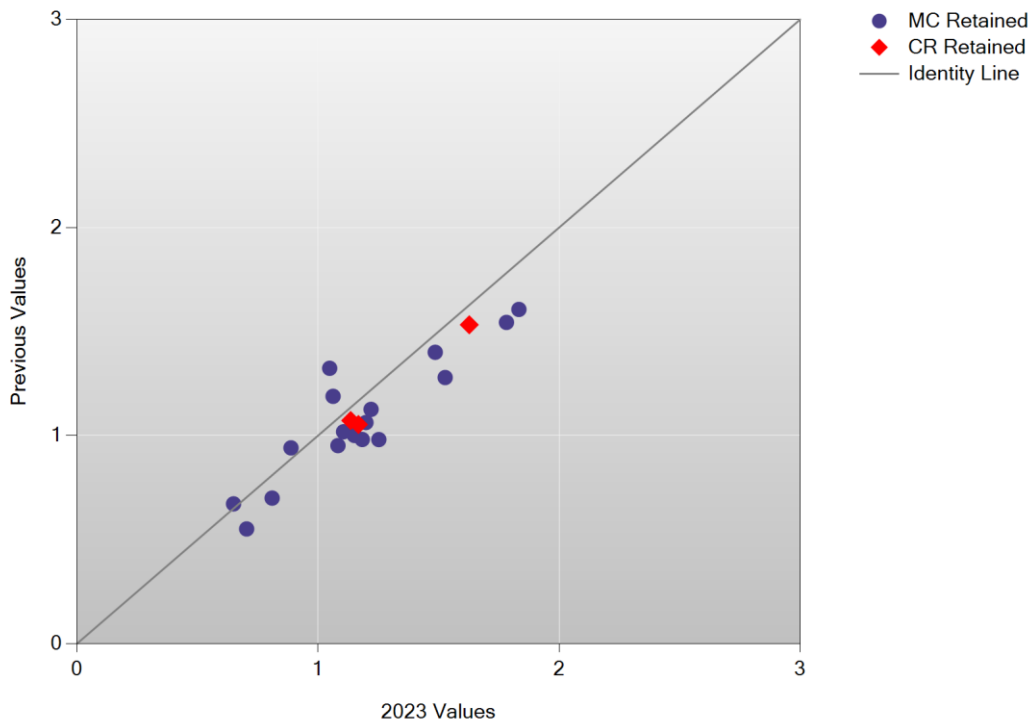
Test Information Function: mat06



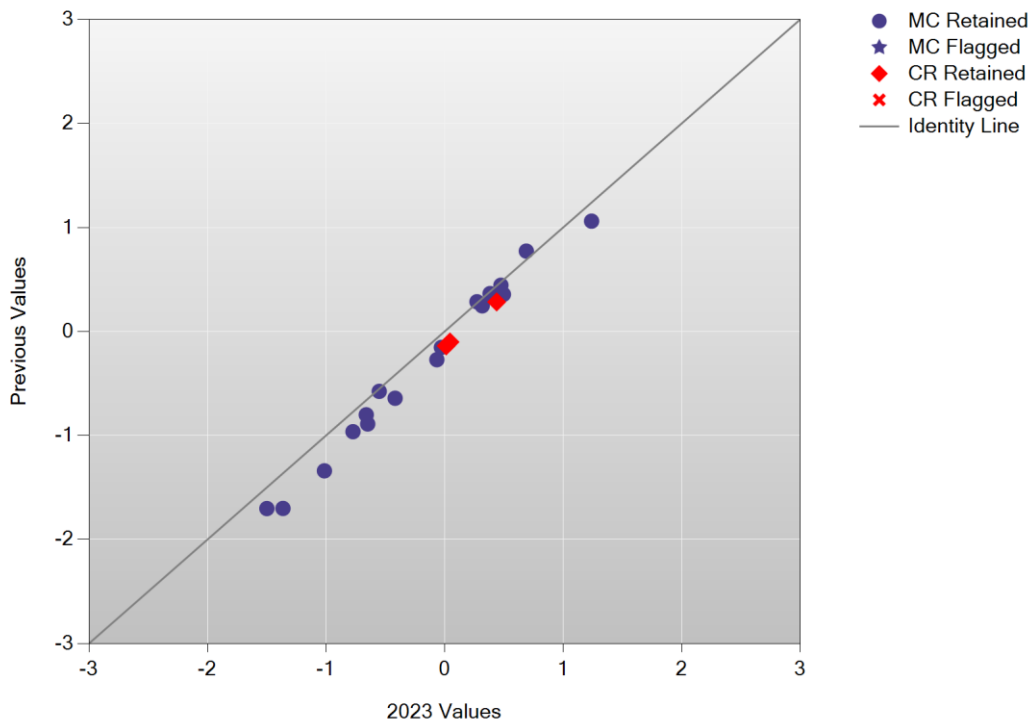
Cumulative Scale Score Distributions: Mathematics Grade 6



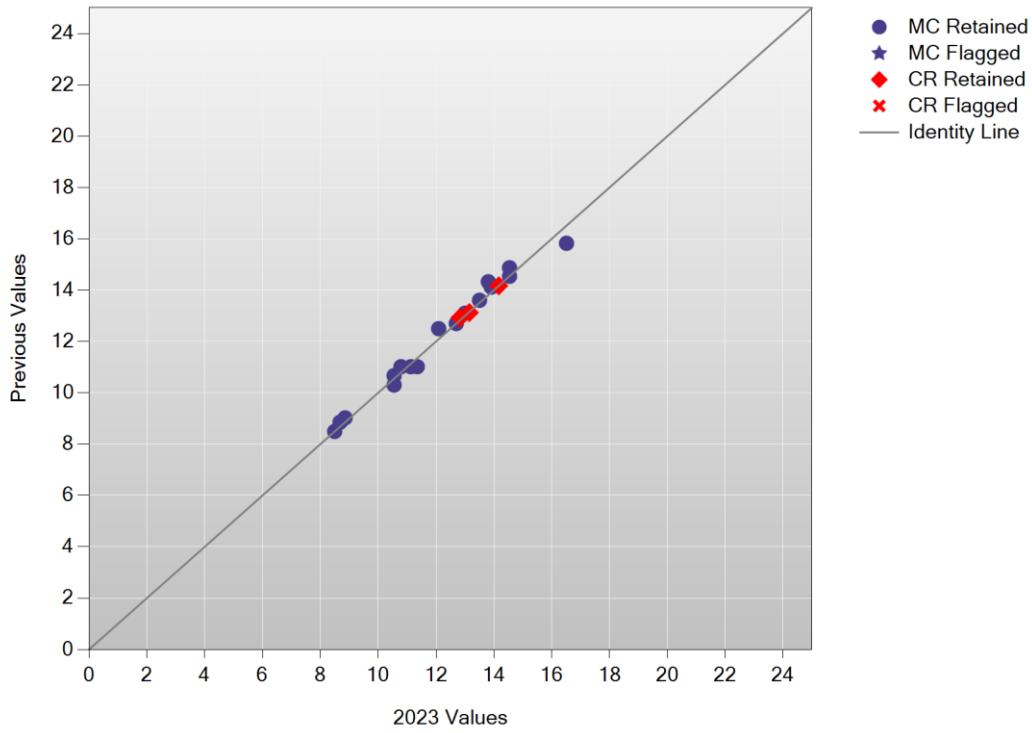
A/A Plot: Mathematics Grade 7



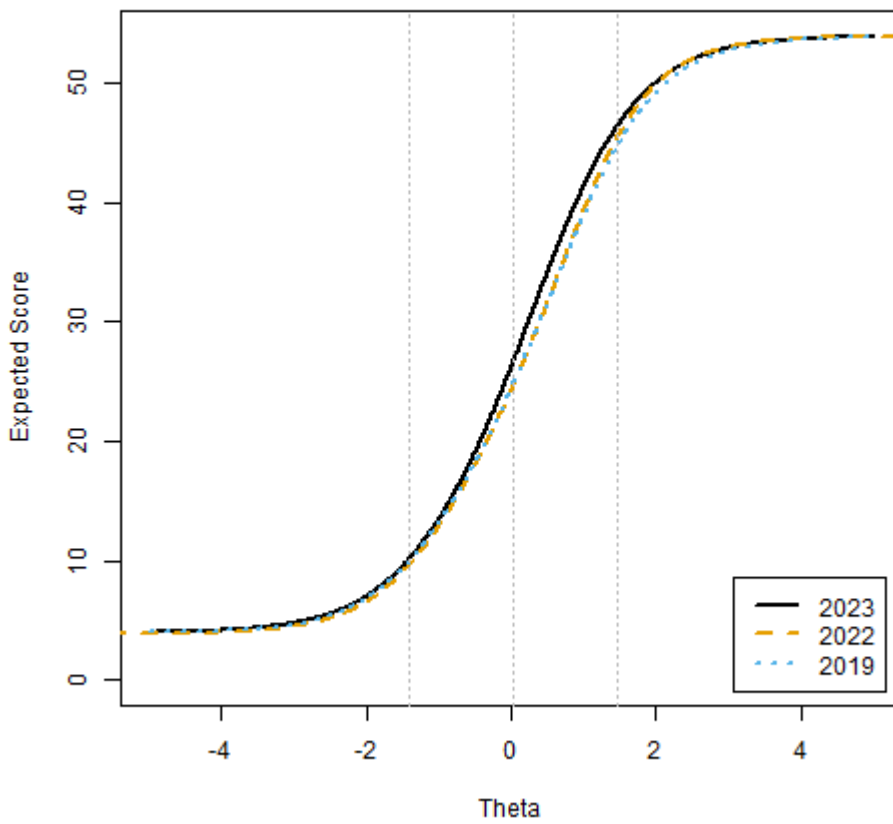
B/B Plot: Mathematics Grade 7



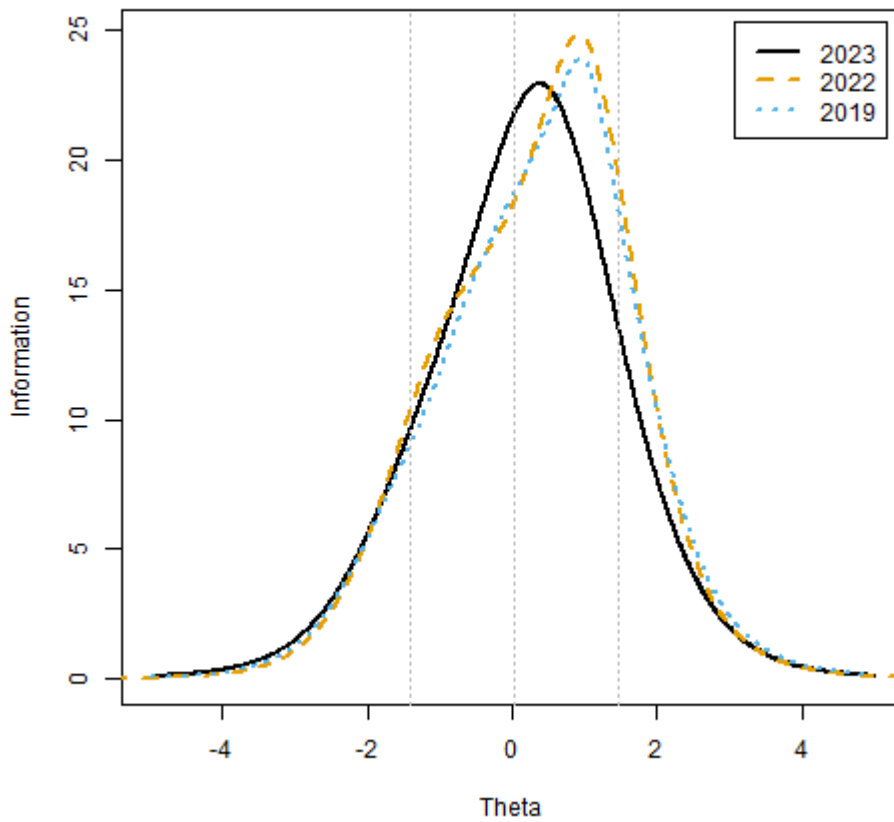
Delta Plot: Mathematics Grade 7



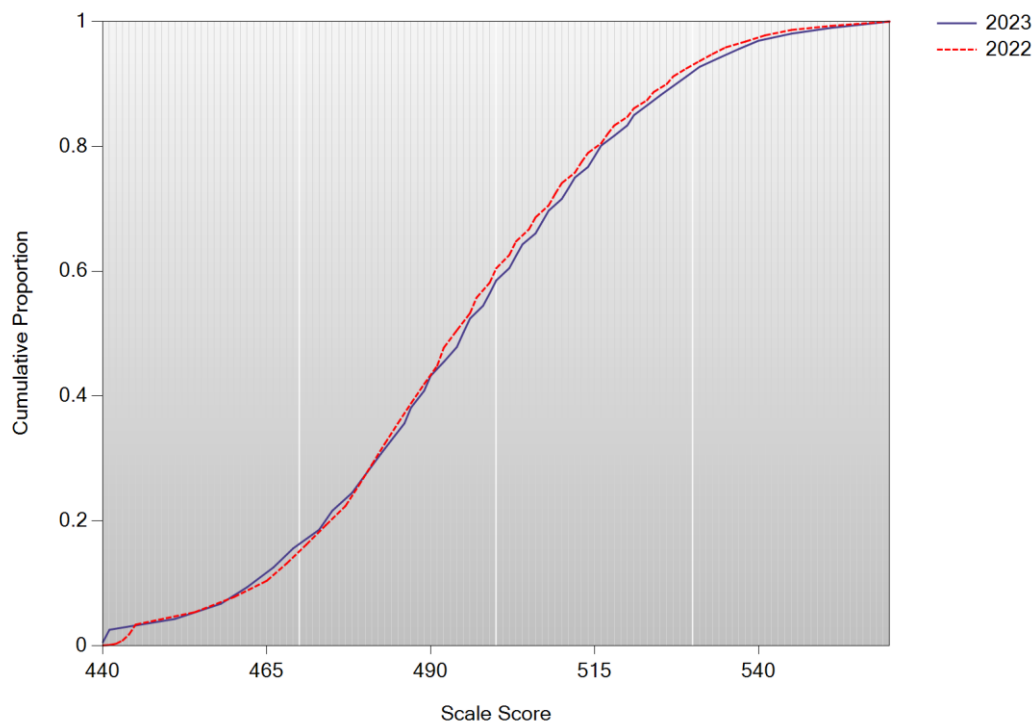
Test Characteristic Curve: mat07



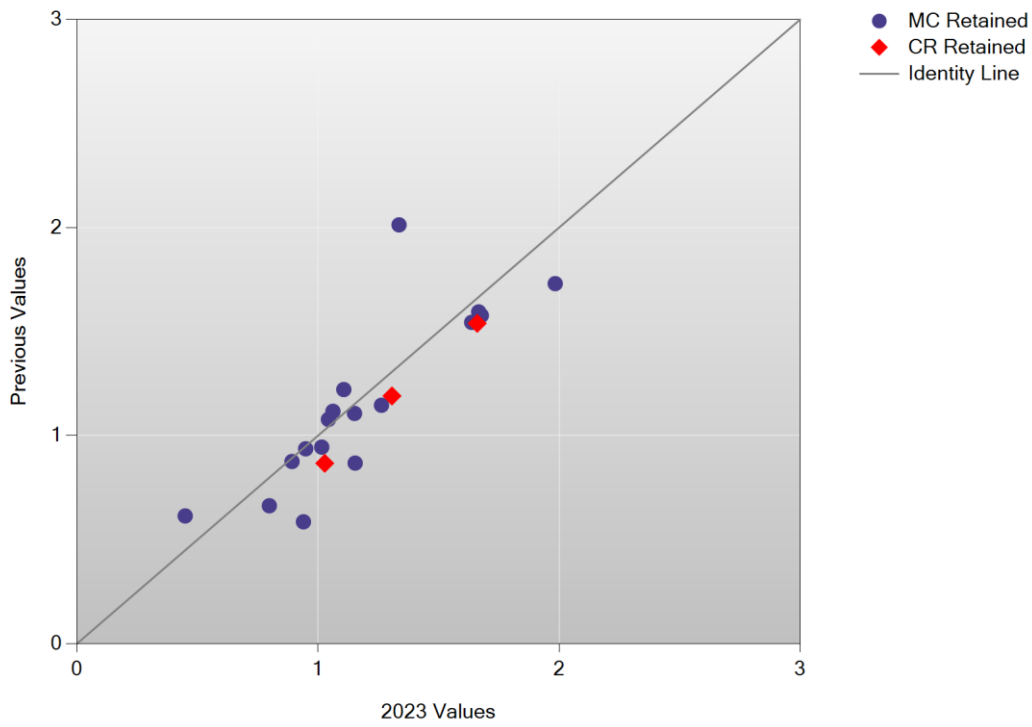
Test Information Function: mat07



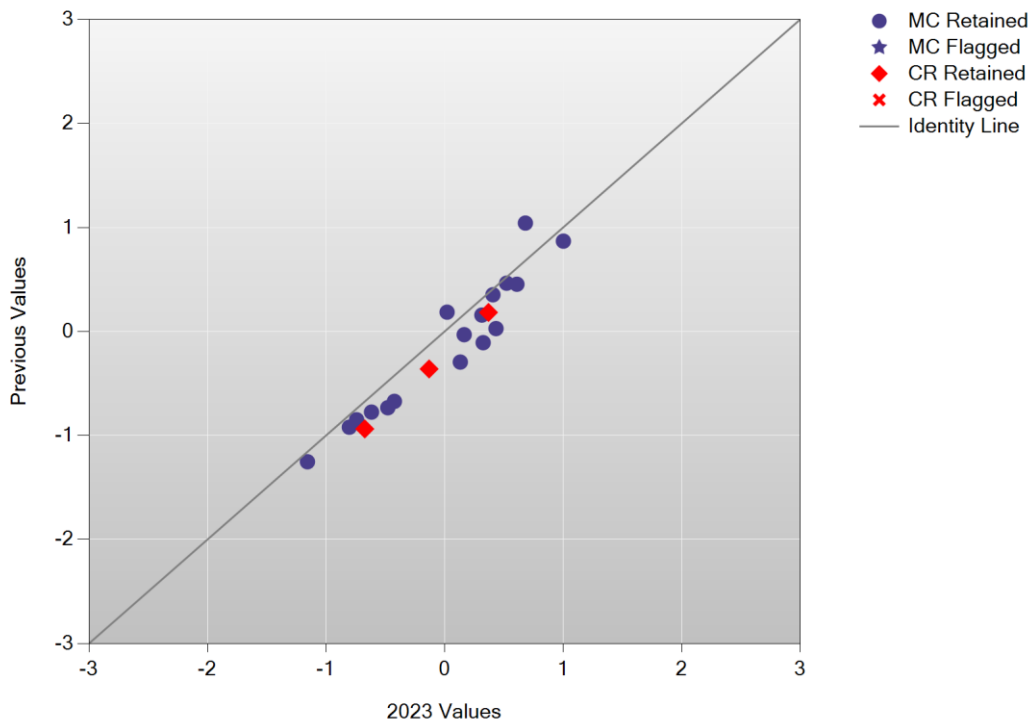
Cumulative Scale Score Distributions: Mathematics Grade 7



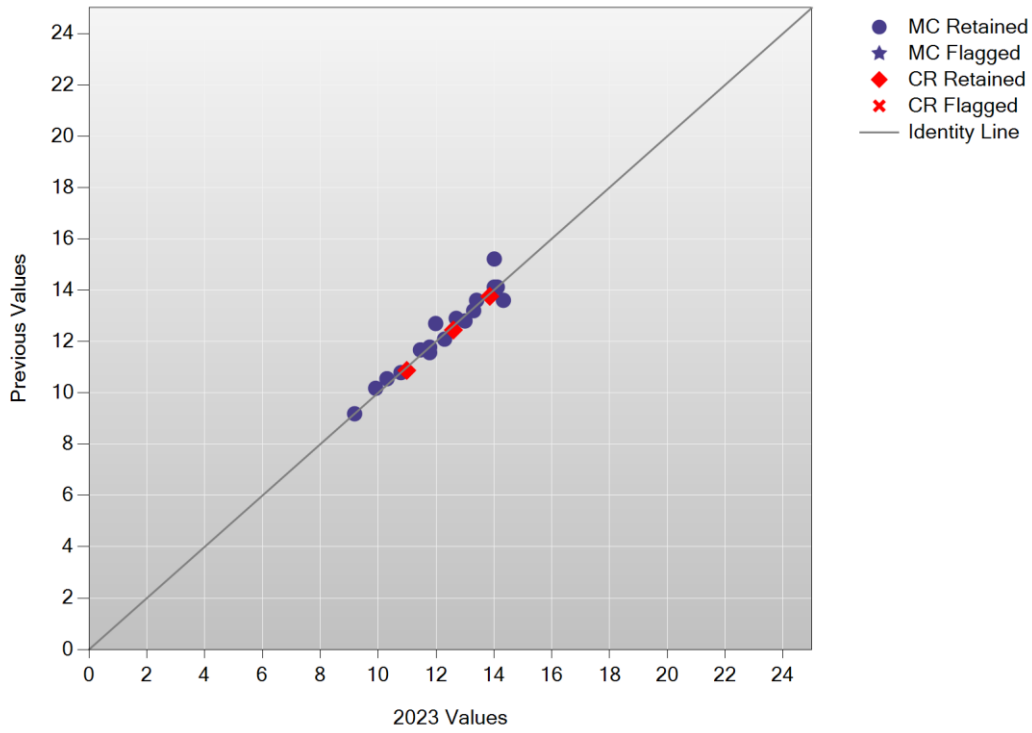
A/A Plot: Mathematics Grade 8



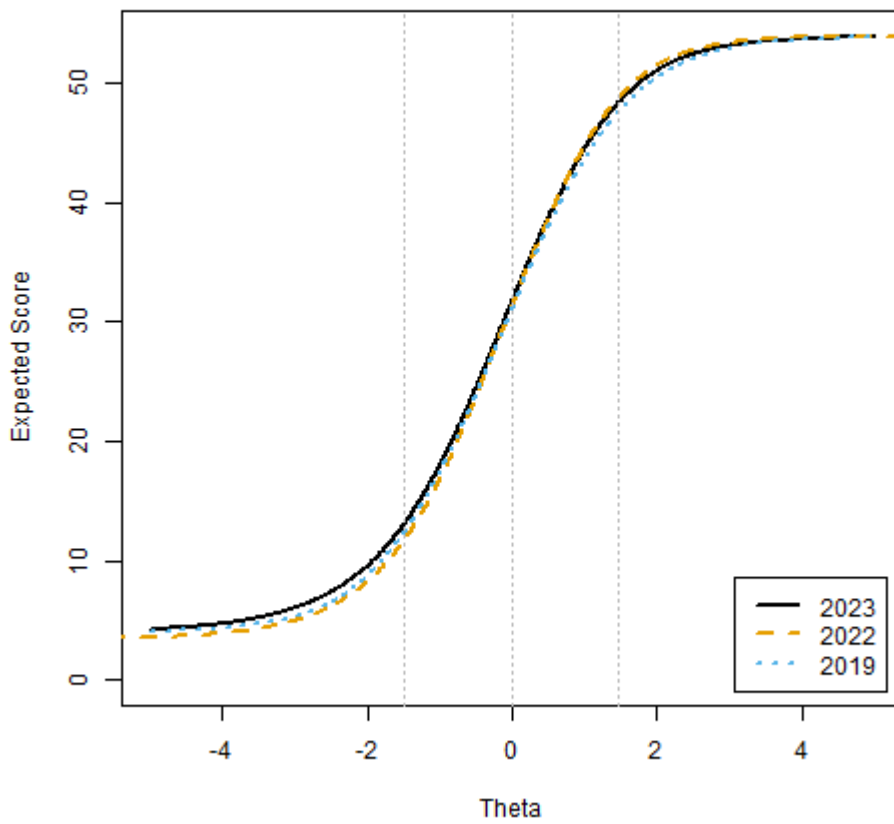
B/B Plot: Mathematics Grade 8



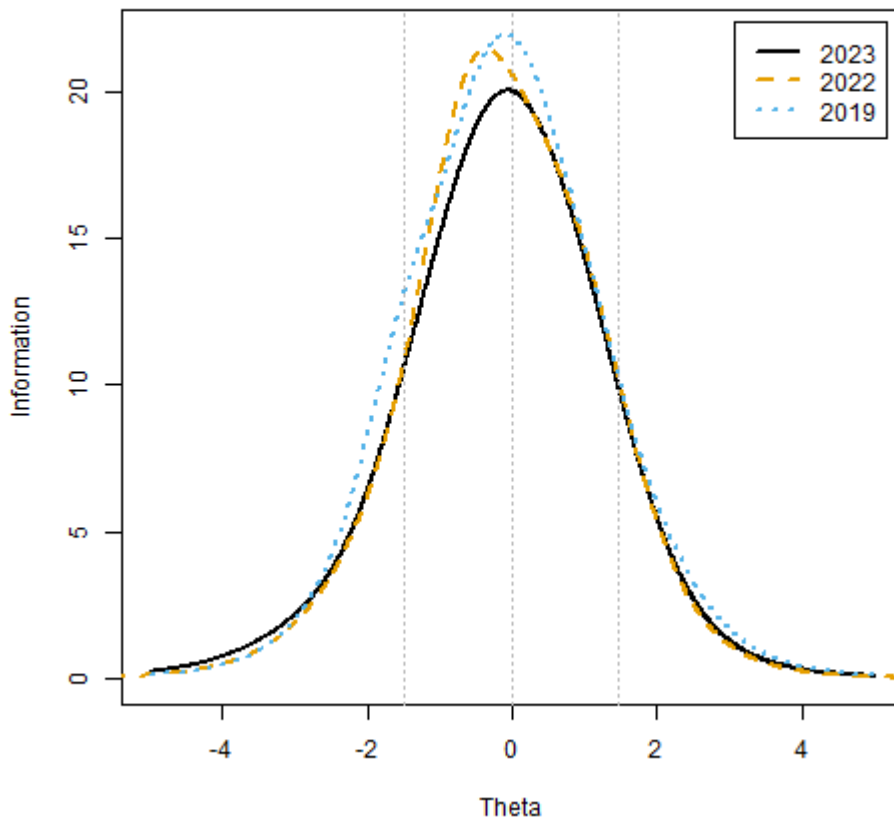
Delta Plot: Mathematics Grade 8



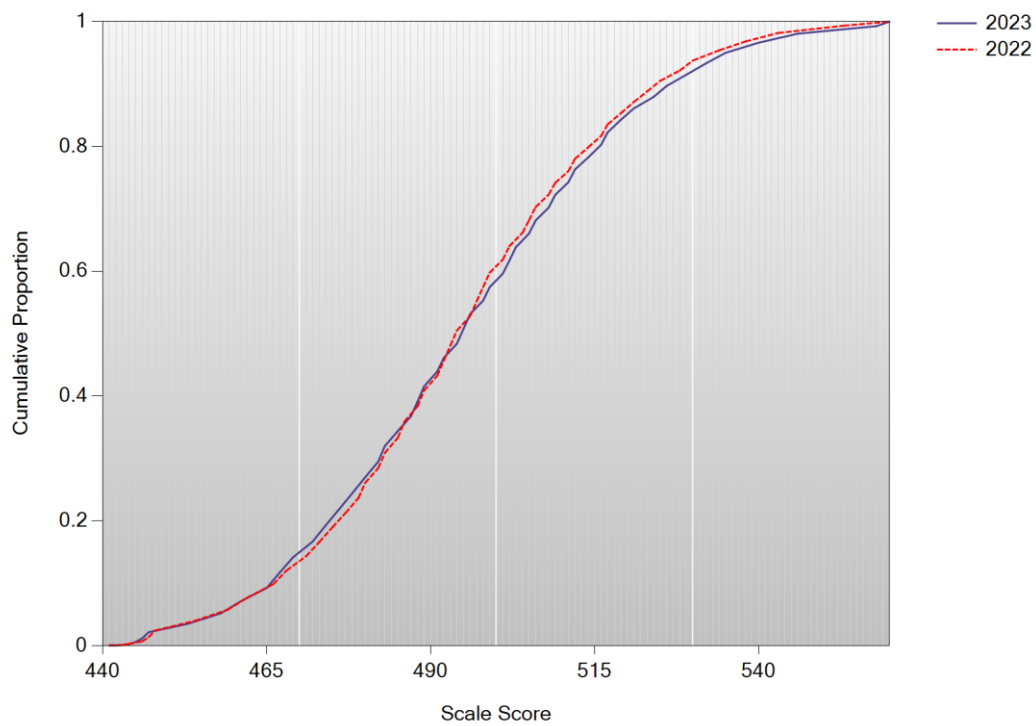
Test Characteristic Curve: mat08



Test Information Function: mat08



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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.173	1.40	10.0	440	1	440	1
1	-3.170	1.41	10.0	440	1	441	1
2	-3.166	1.42	10.0	440	1	442	1
3	-3.162	1.42	10.0	440	1	443	1
4	-3.158	1.43	10.0	440	1	444	1
5	-3.154	1.44	10.0	440	1	445	1
6	-3.150	1.45	10.0	440	1	446	1
7	-2.707	2.51	10.0	449	1	447	1
8	-2.408	3.62	9.9	454	1	452	1
9	-2.178	4.73	8.7	459	1	457	1
10	-1.987	5.78	7.8	462	1	460	1
11	-1.823	6.75	7.2	465	1	463	1
12	-1.677	7.64	6.8	468	1	466	1
13	-1.543	8.44	6.5	471	2	469	1
14	-1.419	9.15	6.2	473	2	471	2
15	-1.302	9.78	6.0	475	2	474	2
16	-1.191	10.35	5.9	477	2	476	2
17	-1.084	10.84	5.7	479	2	478	2
18	-0.981	11.28	5.6	481	2	480	2
19	-0.880	11.65	5.5	483	2	482	2
20	-0.780	11.97	5.4	485	2	484	2
21	-0.682	12.23	5.4	487	2	486	2
22	-0.585	12.41	5.3	489	2	488	2
23	-0.487	12.53	5.3	491	2	490	2
24	-0.390	12.56	5.3	492	2	492	2
25	-0.291	12.49	5.3	494	2	494	2
26	-0.190	12.33	5.4	496	2	497	2
27	-0.087	12.07	5.4	498	2	499	2
28	0.020	11.69	5.5	500	3	502	3
29	0.131	11.21	5.6	502	3	504	3
30	0.247	10.63	5.8	504	3	507	3
31	0.371	9.94	6.0	507	3	510	3
32	0.503	9.17	6.2	509	3	513	3
33	0.646	8.31	6.5	512	3	516	3
34	0.803	7.40	6.9	515	3	520	3
35	0.979	6.45	7.4	518	3	524	3
36	1.179	5.48	8.0	522	3	528	3
37	1.411	4.55	8.8	526	3	533	4

Table 2.2.1 (continued)
 Raw Score to Scale Score Lookup Table
 English Language Arts Grade 3

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	1.685	3.69	9.8	532	4	538	4
39	2.015	2.96	10.0	538	4	544	4
40	2.415	2.38	10.0	545	4	551	4
41	2.905	1.91	10.0	555	4	559	4
42	3.196	1.70	10.0	560	4	560	4
43	3.196	1.70	10.0	560	4	560	4
44	3.196	1.70	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.153	2.42	10.0	440	1	440	1
1	-3.130	2.47	10.0	440	1	440	1
2	-3.106	2.52	10.0	441	1	441	1
3	-3.083	2.57	10.0	441	1	441	1
4	-3.059	2.63	10.0	442	1	441	1
5	-3.036	2.68	10.0	442	1	442	1
6	-3.012	2.73	10.0	443	1	442	1
7	-2.989	2.78	10.0	443	1	442	1
8	-2.965	2.84	10.0	444	1	449	1
9	-2.681	3.51	10.0	449	1	454	1
10	-2.443	4.13	9.3	453	1	458	1
11	-2.236	4.72	8.7	457	1	461	1
12	-2.050	5.28	8.2	461	1	464	1
13	-1.880	5.82	7.8	464	1	467	1
14	-1.723	6.34	7.5	467	1	469	1
15	-1.576	6.84	7.2	469	1	471	2
16	-1.437	7.32	7.0	472	2	474	2
17	-1.305	7.77	6.8	475	2	476	2
18	-1.178	8.21	6.6	477	2	478	2
19	-1.056	8.62	6.4	480	2	480	2
20	-0.937	9.02	6.3	482	2	482	2
21	-0.822	9.40	6.1	484	2	484	2
22	-0.708	9.74	6.0	486	2	486	2
23	-0.596	10.05	5.9	488	2	488	2
24	-0.485	10.29	5.9	490	2	490	2
25	-0.374	10.46	5.8	492	2	492	2
26	-0.262	10.52	5.8	494	2	494	2
27	-0.148	10.47	5.8	497	2	496	2
28	-0.032	10.28	5.9	499	2	498	2
29	0.088	9.96	6.0	501	3	501	3
30	0.213	9.51	6.1	503	3	503	3
31	0.346	8.93	6.3	506	3	505	3
32	0.487	8.25	6.6	509	3	508	3
33	0.639	7.50	6.9	511	3	510	3
34	0.805	6.70	7.3	515	3	513	3
35	0.989	5.90	7.8	518	3	516	3
36	1.194	5.14	8.3	522	3	520	3
37	1.426	4.44	8.9	526	3	523	3

Table 2.2.2 (continued)
 Raw Score to Scale Score Lookup Table
 English Language Arts Grade 4

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	1.689	3.83	9.6	531	4	527	3
39	1.992	3.30	10.0	537	4	532	4
40	2.346	2.80	10.0	544	4	537	4
41	2.773	2.24	10.0	552	4	543	4
42	3.215	1.71	10.0	560	4	552	4
43	3.215	1.71	10.0	560	4	560	4
44	3.215	1.71	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.360	2.14	10.0	440	1	440	1
1	-3.330	2.22	10.0	441	1	441	1
2	-3.300	2.30	10.0	441	1	442	1
3	-3.270	2.39	10.0	442	1	443	1
4	-3.239	2.48	10.0	442	1	444	1
5	-3.209	2.57	10.0	443	1	445	1
6	-3.179	2.66	10.0	443	1	446	1
7	-3.149	2.75	10.0	444	1	450	1
8	-2.870	3.76	9.1	449	1	454	1
9	-2.648	4.70	8.2	453	1	457	1
10	-2.462	5.57	7.5	456	1	459	1
11	-2.299	6.36	7.0	459	1	461	1
12	-2.152	7.08	6.6	461	1	464	1
13	-2.018	7.73	6.3	464	1	466	1
14	-1.892	8.33	6.0	466	1	467	1
15	-1.773	8.86	5.9	468	1	469	1
16	-1.660	9.33	5.7	469	1	471	2
17	-1.551	9.74	5.7	472	2	472	2
18	-1.445	10.09	5.6	474	2	474	2
19	-1.341	10.37	5.5	476	2	476	2
20	-1.239	10.58	5.4	477	2	477	2
21	-1.137	10.73	5.4	479	2	479	2
22	-1.036	10.80	5.4	481	2	481	2
23	-0.935	10.80	5.4	483	2	482	2
24	-0.833	10.74	5.4	485	2	484	2
25	-0.730	10.61	5.4	486	2	486	2
26	-0.626	10.43	5.5	488	2	487	2
27	-0.519	10.19	5.5	490	2	489	2
28	-0.410	9.90	5.6	492	2	491	2
29	-0.297	9.58	5.7	494	2	493	2
30	-0.181	9.22	5.8	496	2	495	2
31	-0.060	8.84	5.9	498	2	497	2
32	0.066	8.45	6.1	500	3	499	2
33	0.197	8.05	6.2	503	3	501	3
34	0.335	7.65	6.4	505	3	504	3
35	0.480	7.25	6.6	508	3	506	3
36	0.633	6.84	6.8	511	3	508	3
37	0.796	6.41	7.0	513	3	511	3

Table 2.2.3 (continued)
 Raw Score to Scale Score Lookup Table
 English Language Arts Grade 5

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.971	5.96	7.2	517	3	514	3
39	1.160	5.49	7.6	520	3	516	3
40	1.366	4.98	7.9	523	3	520	3
41	1.594	4.46	8.4	528	3	523	3
42	1.851	3.92	8.9	532	4	527	3
43	2.145	3.36	9.6	537	4	531	4
44	2.493	2.79	10.0	543	4	536	4
45	2.921	2.25	10.0	551	4	542	4
46	3.430	1.77	10.0	560	4	550	4
47	3.430	1.77	10.0	560	4	560	4
48	3.430	1.77	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.171	2.84	10.0	440	1	440	1
1	-3.166	2.85	10.0	440	1	440	1
2	-3.160	2.86	10.0	440	1	441	1
3	-3.155	2.87	10.0	440	1	441	1
4	-3.150	2.88	10.0	440	1	441	1
5	-3.144	2.89	10.0	441	1	442	1
6	-3.139	2.91	10.0	441	1	442	1
7	-3.133	2.92	10.0	441	1	442	1
8	-3.128	2.93	10.0	441	1	443	1
9	-3.122	2.94	10.0	441	1	443	1
10	-3.117	2.95	10.0	441	1	447	1
11	-2.880	3.44	10.0	446	1	451	1
12	-2.669	3.87	9.6	450	1	454	1
13	-2.479	4.26	9.2	453	1	457	1
14	-2.302	4.59	8.9	456	1	460	1
15	-2.137	4.89	8.6	460	1	462	1
16	-1.980	5.15	8.4	463	1	465	1
17	-1.829	5.38	8.2	465	1	467	1
18	-1.685	5.59	8.0	468	1	469	1
19	-1.544	5.77	7.9	471	2	472	2
20	-1.407	5.93	7.8	473	2	474	2
21	-1.273	6.07	7.7	476	2	476	2
22	-1.140	6.18	7.6	479	2	479	2
23	-1.009	6.27	7.6	481	2	481	2
24	-0.879	6.32	7.5	484	2	483	2
25	-0.750	6.35	7.5	486	2	485	2
26	-0.620	6.35	7.5	488	2	488	2
27	-0.489	6.32	7.6	491	2	490	2
28	-0.358	6.25	7.6	493	2	492	2
29	-0.224	6.16	7.6	496	2	494	2
30	-0.088	6.04	7.7	499	2	497	2
31	0.051	5.90	7.8	501	3	499	2
32	0.193	5.74	7.9	504	3	501	3
33	0.340	5.56	8.1	507	3	504	3
34	0.492	5.37	8.2	510	3	506	3
35	0.649	5.16	8.4	513	3	509	3
36	0.814	4.94	8.5	516	3	512	3
37	0.986	4.70	8.8	519	3	514	3

Table 2.2.4 (continued)
 Raw Score to Scale Score Lookup Table
 English Language Arts Grade 6

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	1.168	4.44	9.0	522	3	517	3
39	1.360	4.16	9.3	526	3	520	3
40	1.567	3.86	9.7	529	3	524	3
41	1.790	3.53	10.0	534	4	527	3
42	2.035	3.18	10.0	539	4	531	4
43	2.306	2.83	10.0	544	4	535	4
44	2.611	2.47	10.0	550	4	540	4
45	2.961	2.13	10.0	556	4	546	4
46	3.150	1.97	10.0	560	4	552	4
47	3.150	1.97	10.0	560	4	560	4
48	3.150	1.97	10.0	560	4	560	4
49	3.150	1.97	10.0	560	4	560	4
50	3.150	1.97	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.131	2.23	10.0	440	1	440	1
1	-3.113	2.27	10.0	440	1	440	1
2	-3.095	2.31	10.0	441	1	441	1
3	-3.078	2.34	10.0	441	1	441	1
4	-3.060	2.38	10.0	441	1	442	1
5	-3.043	2.42	10.0	442	1	442	1
6	-3.025	2.46	10.0	442	1	442	1
7	-3.007	2.50	10.0	442	1	443	1
8	-2.990	2.54	10.0	443	1	443	1
9	-2.713	3.23	10.0	448	1	448	1
10	-2.485	3.87	9.7	452	1	451	1
11	-2.289	4.46	9.0	456	1	455	1
12	-2.115	5.01	8.5	459	1	458	1
13	-1.958	5.52	8.1	462	1	460	1
14	-1.813	5.99	7.8	465	1	463	1
15	-1.677	6.43	7.5	468	1	465	1
16	-1.549	6.84	7.3	470	2	467	1
17	-1.428	7.22	7.1	473	2	469	1
18	-1.311	7.57	6.9	475	2	472	2
19	-1.199	7.89	6.8	477	2	474	2
20	-1.090	8.18	6.7	479	2	476	2
21	-0.984	8.43	6.6	481	2	478	2
22	-0.880	8.65	6.5	483	2	480	2
23	-0.777	8.84	6.4	485	2	482	2
24	-0.676	8.98	6.4	487	2	484	2
25	-0.576	9.08	6.3	489	2	486	2
26	-0.476	9.15	6.3	491	2	488	2
27	-0.377	9.17	6.3	493	2	490	2
28	-0.277	9.15	6.3	495	2	492	2
29	-0.176	9.09	6.3	496	2	494	2
30	-0.074	8.99	6.4	498	2	496	2
31	0.029	8.86	6.4	500	3	498	2
32	0.134	8.69	6.5	502	3	500	3
33	0.242	8.49	6.6	504	3	502	3
34	0.352	8.26	6.6	507	3	504	3
35	0.467	7.99	6.8	509	3	507	3
36	0.585	7.69	6.9	511	3	509	3
37	0.709	7.35	7.0	513	3	511	3

Table 2.2.5 (continued)
 Raw Score to Scale Score Lookup Table
 English Language Arts Grade 7

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.840	6.97	7.2	516	3	514	3
39	0.979	6.55	7.5	518	3	516	3
40	1.127	6.09	7.7	521	3	519	3
41	1.287	5.58	8.1	524	3	522	3
42	1.463	5.05	8.5	528	3	525	3
43	1.659	4.49	9.0	531	4	528	3
44	1.880	3.91	9.7	536	4	532	4
45	2.137	3.33	10.0	541	4	536	4
46	2.442	2.76	10.0	546	4	542	4
47	2.822	2.18	10.0	554	4	548	4
48	3.153	1.79	10.0	560	4	555	4
49	3.153	1.79	10.0	560	4	560	4
50	3.153	1.79	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.964	3.32	10.0	440	1	440	1
1	-2.962	3.32	10.0	440	1	440	1
2	-2.959	3.33	10.0	440	1	440	1
3	-2.957	3.33	10.0	440	1	440	1
4	-2.955	3.33	10.0	440	1	440	1
5	-2.953	3.34	10.0	440	1	441	1
6	-2.950	3.34	10.0	440	1	441	1
7	-2.948	3.35	10.0	440	1	441	1
8	-2.946	3.35	10.0	440	1	441	1
9	-2.943	3.36	10.0	440	1	441	1
10	-2.941	3.36	10.0	440	1	441	1
11	-2.939	3.37	10.0	440	1	445	1
12	-2.937	3.37	10.0	441	1	448	1
13	-2.729	3.80	10.0	445	1	451	1
14	-2.539	4.22	9.7	448	1	454	1
15	-2.364	4.62	9.3	452	1	456	1
16	-2.201	4.99	8.9	455	1	458	1
17	-2.048	5.33	8.6	458	1	461	1
18	-1.901	5.65	8.4	461	1	463	1
19	-1.761	5.94	8.2	464	1	465	1
20	-1.626	6.20	8.0	467	1	467	1
21	-1.495	6.42	7.9	469	1	469	1
22	-1.367	6.61	7.7	472	2	471	2
23	-1.242	6.77	7.6	474	2	473	2
24	-1.118	6.89	7.6	477	2	475	2
25	-0.995	6.98	7.5	479	2	477	2
26	-0.873	7.05	7.5	482	2	479	2
27	-0.752	7.08	7.5	484	2	481	2
28	-0.630	7.09	7.5	486	2	483	2
29	-0.507	7.07	7.5	489	2	485	2
30	-0.383	7.03	7.5	491	2	487	2
31	-0.258	6.97	7.5	494	2	489	2
32	-0.130	6.89	7.6	496	2	492	2
33	0.001	6.78	7.6	499	2	494	2
34	0.135	6.64	7.7	502	3	497	2
35	0.273	6.47	7.8	504	3	499	2
36	0.417	6.25	8.0	507	3	502	3
37	0.567	5.99	8.1	510	3	504	3

Table 2.2.6 (continued)
 Raw Score to Scale Score Lookup Table
 English Language Arts Grade 8

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.726	5.67	8.4	513	3	507	3
39	0.894	5.29	8.6	517	3	510	3
40	1.076	4.87	9.0	520	3	513	3
41	1.275	4.40	9.5	524	3	516	3
42	1.495	3.90	10.0	529	3	520	3
43	1.741	3.40	10.0	534	4	524	3
44	2.022	2.91	10.0	539	4	528	3
45	2.347	2.45	10.0	546	4	533	4
46	2.730	2.05	10.0	553	4	539	4
47	3.066	1.79	10.0	560	4	545	4
48	3.066	1.79	10.0	560	4	553	4
49	3.066	1.79	10.0	560	4	560	4
50	3.066	1.79	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.782	3.14	10.0	440	1	440	1
1	-2.781	3.14	10.0	440	1	440	1
2	-2.780	3.15	10.0	440	1	440	1
3	-2.780	3.15	10.0	440	1	441	1
4	-2.779	3.15	10.0	440	1	441	1
5	-2.778	3.16	10.0	440	1	441	1
6	-2.467	4.72	9.8	447	1	441	1
7	-2.238	6.23	8.6	452	1	447	1
8	-2.055	7.66	7.7	456	1	452	1
9	-1.900	8.99	7.1	459	1	456	1
10	-1.764	10.22	6.7	462	1	459	1
11	-1.643	11.34	6.3	464	1	462	1
12	-1.532	12.36	6.1	467	1	465	1
13	-1.428	13.27	5.9	469	1	468	1
14	-1.331	14.10	5.7	471	2	470	2
15	-1.239	14.83	5.5	473	2	473	2
16	-1.151	15.48	5.4	475	2	475	2
17	-1.066	16.06	5.3	477	2	477	2
18	-0.984	16.56	5.2	478	2	479	2
19	-0.904	17.00	5.2	480	2	481	2
20	-0.825	17.37	5.1	482	2	483	2
21	-0.748	17.67	5.1	483	2	485	2
22	-0.672	17.92	5.0	485	2	486	2
23	-0.597	18.10	5.0	487	2	488	2
24	-0.522	18.22	5.0	488	2	490	2
25	-0.447	18.28	5.0	490	2	492	2
26	-0.372	18.27	5.0	491	2	494	2
27	-0.297	18.20	5.0	493	2	495	2
28	-0.222	18.06	5.0	495	2	497	2
29	-0.145	17.85	5.1	496	2	499	2
30	-0.067	17.57	5.1	498	2	501	3
31	0.012	17.23	5.1	499	2	503	3
32	0.094	16.82	5.2	501	3	505	3
33	0.177	16.34	5.3	503	3	506	3
34	0.264	15.79	5.4	505	3	509	3
35	0.354	15.18	5.5	507	3	511	3
36	0.449	14.51	5.6	509	3	513	3
37	0.548	13.78	5.8	511	3	515	3

Table 2.2.7 (continued)
 Raw Score to Scale Score Lookup Table
 Mathematics Grade 3

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.653	12.99	5.9	513	3	518	3
39	0.766	12.15	6.1	516	3	520	3
40	0.888	11.25	6.4	518	3	523	3
41	1.021	10.29	6.7	521	3	526	3
42	1.169	9.28	7.0	524	3	529	3
43	1.337	8.18	7.5	528	3	534	4
44	1.531	6.97	8.1	532	4	539	4
45	1.769	5.59	9.0	537	4	545	4
46	2.086	3.97	10.0	544	4	554	4
47	2.600	2.05	10.0	555	4	560	4
48	2.837	1.46	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.812	3.15	10.0	440	1	440	1
1	-2.783	3.25	10.0	441	1	440	1
2	-2.755	3.36	10.0	441	1	440	1
3	-2.726	3.47	10.0	442	1	440	1
4	-2.698	3.58	10.0	442	1	440	1
5	-2.670	3.70	10.0	443	1	440	1
6	-2.641	3.82	10.0	444	1	446	1
7	-2.613	3.94	10.0	444	1	451	1
8	-2.364	5.16	9.2	449	1	454	1
9	-2.165	6.35	8.3	454	1	458	1
10	-1.996	7.49	7.6	457	1	461	1
11	-1.849	8.59	7.1	460	1	463	1
12	-1.718	9.63	6.7	463	1	466	1
13	-1.598	10.63	6.4	465	1	468	1
14	-1.488	11.58	6.2	468	1	470	2
15	-1.385	12.50	5.9	469	1	472	2
16	-1.288	13.40	5.7	472	2	474	2
17	-1.196	14.27	5.5	474	2	476	2
18	-1.108	15.11	5.4	476	2	478	2
19	-1.024	15.92	5.2	477	2	480	2
20	-0.943	16.70	5.1	479	2	481	2
21	-0.865	17.42	5.0	481	2	483	2
22	-0.789	18.08	4.9	482	2	485	2
23	-0.715	18.67	4.8	484	2	486	2
24	-0.642	19.17	4.8	485	2	488	2
25	-0.570	19.59	4.7	487	2	489	2
26	-0.499	19.90	4.7	488	2	491	2
27	-0.429	20.11	4.7	490	2	492	2
28	-0.359	20.22	4.7	491	2	493	2
29	-0.289	20.23	4.7	493	2	495	2
30	-0.218	20.15	4.7	494	2	496	2
31	-0.147	19.97	4.7	496	2	498	2
32	-0.075	19.70	4.7	497	2	499	2
33	-0.002	19.35	4.8	499	2	501	3
34	0.073	18.92	4.8	500	3	502	3
35	0.149	18.42	4.9	502	3	504	3
36	0.228	17.85	5.0	504	3	505	3
37	0.309	17.24	5.0	505	3	507	3

Table 2.2.8 (continued)
 Raw Score to Scale Score Lookup Table
 Mathematics Grade 4

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.393	16.58	5.1	507	3	508	3
39	0.480	15.88	5.3	509	3	510	3
40	0.570	15.15	5.4	511	3	512	3
41	0.665	14.39	5.5	513	3	513	3
42	0.765	13.61	5.7	515	3	515	3
43	0.871	12.79	5.9	517	3	517	3
44	0.983	11.93	6.1	519	3	519	3
45	1.104	11.03	6.3	522	3	522	3
46	1.234	10.08	6.6	525	3	524	3
47	1.377	9.10	6.9	528	3	527	3
48	1.534	8.08	7.4	531	4	529	3
49	1.712	7.04	7.9	535	4	533	4
50	1.918	5.95	8.6	539	4	537	4
51	2.166	4.77	9.6	544	4	542	4
52	2.493	3.41	10.0	551	4	549	4
53	2.920	2.04	10.0	560	4	560	4
54	2.920	2.04	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.126	2.25	10.0	440	1	440	1
1	-3.078	2.40	10.0	441	1	440	1
2	-3.029	2.55	10.0	442	1	440	1
3	-2.981	2.71	10.0	443	1	440	1
4	-2.932	2.87	10.0	444	1	440	1
5	-2.884	3.05	10.0	445	1	441	1
6	-2.835	3.23	10.0	446	1	448	1
7	-2.518	4.58	8.9	452	1	453	1
8	-2.278	5.77	7.9	456	1	458	1
9	-2.082	6.83	7.3	460	1	461	1
10	-1.914	7.79	6.8	463	1	464	1
11	-1.765	8.67	6.5	466	1	467	1
12	-1.631	9.51	6.2	468	1	469	1
13	-1.509	10.31	5.9	471	2	471	2
14	-1.396	11.07	5.7	473	2	474	2
15	-1.290	11.81	5.5	475	2	475	2
16	-1.190	12.52	5.4	477	2	477	2
17	-1.095	13.19	5.2	479	2	479	2
18	-1.005	13.83	5.1	480	2	481	2
19	-0.918	14.43	5.0	482	2	482	2
20	-0.834	14.98	4.9	484	2	484	2
21	-0.753	15.48	4.8	485	2	485	2
22	-0.673	15.94	4.8	487	2	486	2
23	-0.596	16.34	4.7	488	2	488	2
24	-0.519	16.69	4.7	490	2	489	2
25	-0.444	16.98	4.6	491	2	490	2
26	-0.370	17.22	4.6	492	2	492	2
27	-0.296	17.40	4.6	494	2	493	2
28	-0.223	17.53	4.5	495	2	494	2
29	-0.150	17.60	4.5	497	2	496	2
30	-0.077	17.63	4.5	498	2	497	2
31	-0.003	17.62	4.5	499	2	498	2
32	0.070	17.58	4.5	501	3	499	2
33	0.145	17.50	4.6	502	3	501	3
34	0.220	17.39	4.6	504	3	502	3
35	0.295	17.27	4.6	505	3	504	3
36	0.373	17.14	4.6	507	3	505	3
37	0.451	17.00	4.6	508	3	506	3

Table 2.2.9 (continued)
 Raw Score to Scale Score Lookup Table
 Mathematics Grade 5

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.531	16.84	4.6	510	3	508	3
39	0.613	16.66	4.7	511	3	509	3
40	0.697	16.44	4.7	513	3	511	3
41	0.783	16.15	4.7	514	3	513	3
42	0.873	15.77	4.8	516	3	514	3
43	0.967	15.24	4.9	518	3	516	3
44	1.067	14.55	5.0	520	3	518	3
45	1.173	13.66	5.2	522	3	520	3
46	1.288	12.57	5.4	524	3	522	3
47	1.415	11.28	5.7	526	3	524	3
48	1.558	9.84	6.1	529	3	527	3
49	1.723	8.29	6.6	532	4	529	3
50	1.920	6.66	7.4	536	4	533	4
51	2.170	4.98	8.5	541	4	537	4
52	2.516	3.27	10.0	547	4	543	4
53	3.101	1.54	10.0	559	4	552	4
54	3.176	1.39	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-3.028	2.38	10.0	440	1	440	1
1	-3.001	2.46	10.0	441	1	441	1
2	-2.974	2.54	10.0	441	1	442	1
3	-2.947	2.62	10.0	442	1	442	1
4	-2.920	2.70	10.0	442	1	443	1
5	-2.893	2.79	10.0	443	1	450	1
6	-2.512	4.20	9.7	450	1	455	1
7	-2.234	5.54	8.4	456	1	459	1
8	-2.014	6.91	7.6	460	1	463	1
9	-1.831	8.28	6.9	464	1	466	1
10	-1.674	9.64	6.4	467	1	469	1
11	-1.537	10.92	6.0	469	1	471	2
12	-1.413	12.11	5.7	472	2	474	2
13	-1.301	13.20	5.5	474	2	476	2
14	-1.198	14.18	5.3	476	2	478	2
15	-1.101	15.07	5.1	478	2	480	2
16	-1.009	15.88	5.0	480	2	481	2
17	-0.923	16.63	4.9	482	2	483	2
18	-0.840	17.31	4.8	483	2	485	2
19	-0.760	17.94	4.7	485	2	487	2
20	-0.683	18.52	4.6	487	2	488	2
21	-0.609	19.05	4.6	488	2	490	2
22	-0.537	19.53	4.5	489	2	491	2
23	-0.466	19.95	4.4	491	2	493	2
24	-0.397	20.33	4.4	492	2	494	2
25	-0.329	20.65	4.4	494	2	496	2
26	-0.263	20.92	4.3	495	2	497	2
27	-0.197	21.14	4.3	496	2	499	2
28	-0.132	21.31	4.3	498	2	500	3
29	-0.067	21.43	4.3	499	2	501	3
30	-0.003	21.51	4.3	500	3	503	3
31	0.061	21.54	4.3	501	3	504	3
32	0.125	21.53	4.3	503	3	506	3
33	0.189	21.48	4.3	504	3	507	3
34	0.254	21.38	4.3	505	3	508	3
35	0.319	21.23	4.3	506	3	510	3
36	0.384	21.04	4.3	508	3	511	3
37	0.451	20.79	4.4	509	3	512	3

Table 2.2.10 (continued)
 Raw Score to Scale Score Lookup Table
 Mathematics Grade 6

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.519	20.48	4.4	510	3	514	3
39	0.588	20.11	4.4	512	3	515	3
40	0.659	19.66	4.5	513	3	517	3
41	0.732	19.12	4.5	515	3	518	3
42	0.808	18.48	4.6	516	3	520	3
43	0.888	17.74	4.7	518	3	522	3
44	0.971	16.88	4.8	519	3	524	3
45	1.061	15.89	5.0	521	3	525	3
46	1.157	14.75	5.2	523	3	528	3
47	1.263	13.46	5.4	525	3	529	3
48	1.381	11.99	5.7	528	3	532	4
49	1.516	10.32	6.2	530	4	536	4
50	1.679	8.46	6.8	534	4	539	4
51	1.886	6.37	7.9	538	4	544	4
52	2.180	4.10	9.8	543	4	551	4
53	2.708	1.77	10.0	554	4	560	4
54	3.011	1.10	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.859	1.84	10.0	440	1	440	1
1	-2.852	1.86	10.0	440	1	441	1
2	-2.844	1.88	10.0	440	1	442	1
3	-2.837	1.90	10.0	440	1	443	1
4	-2.829	1.92	10.0	441	1	444	1
5	-2.822	1.94	10.0	441	1	445	1
6	-2.319	3.80	10.0	451	1	454	1
7	-2.008	5.49	8.9	458	1	460	1
8	-1.777	6.97	7.9	462	1	465	1
9	-1.590	8.27	7.2	466	1	468	1
10	-1.430	9.42	6.8	469	1	471	2
11	-1.290	10.47	6.4	473	2	474	2
12	-1.164	11.45	6.1	475	2	477	2
13	-1.050	12.38	5.9	478	2	479	2
14	-0.944	13.28	5.7	480	2	481	2
15	-0.846	14.14	5.5	482	2	483	2
16	-0.753	14.97	5.4	484	2	485	2
17	-0.666	15.76	5.2	486	2	487	2
18	-0.583	16.53	5.1	487	2	489	2
19	-0.503	17.27	5.0	489	2	491	2
20	-0.427	17.98	4.9	490	2	492	2
21	-0.353	18.66	4.8	492	2	494	2
22	-0.282	19.30	4.7	494	2	496	2
23	-0.213	19.91	4.7	495	2	497	2
24	-0.145	20.48	4.6	496	2	499	2
25	-0.079	21.00	4.5	498	2	500	3
26	-0.014	21.47	4.5	499	2	502	3
27	0.050	21.88	4.4	500	3	503	3
28	0.114	22.23	4.4	502	3	505	3
29	0.177	22.51	4.4	503	3	506	3
30	0.239	22.73	4.4	504	3	508	3
31	0.302	22.86	4.3	506	3	509	3
32	0.364	22.93	4.3	507	3	510	3
33	0.427	22.91	4.3	508	3	512	3
34	0.490	22.81	4.3	510	3	513	3
35	0.554	22.62	4.4	511	3	514	3
36	0.619	22.35	4.4	512	3	516	3
37	0.686	21.99	4.4	514	3	517	3

Table 2.2.11 (continued)
 Raw Score to Scale Score Lookup Table
 Mathematics Grade 7

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.754	21.53	4.5	515	3	518	3
39	0.824	20.99	4.5	516	3	520	3
40	0.896	20.34	4.6	518	3	521	3
41	0.971	19.60	4.7	520	3	523	3
42	1.050	18.75	4.8	521	3	524	3
43	1.132	17.80	4.9	523	3	526	3
44	1.220	16.73	5.1	525	3	527	3
45	1.315	15.56	5.3	527	3	529	3
46	1.417	14.27	5.5	529	3	531	4
47	1.530	12.87	5.8	531	4	533	4
48	1.656	11.36	6.2	534	4	535	4
49	1.802	9.74	6.7	537	4	538	4
50	1.976	8.00	7.3	540	4	541	4
51	2.194	6.13	8.4	545	4	545	4
52	2.495	4.12	10.0	551	4	551	4
53	2.922	2.24	10.0	560	4	560	4
54	2.922	2.24	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)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
0	-2.983	2.23	10.0	440	1	440	1
1	-2.932	2.35	10.0	441	1	441	1
2	-2.881	2.47	10.0	442	1	443	1
3	-2.830	2.61	10.0	443	1	444	1
4	-2.778	2.75	10.0	444	1	446	1
5	-2.727	2.90	10.0	445	1	447	1
6	-2.676	3.06	10.0	446	1	448	1
7	-2.625	3.23	10.0	447	1	454	1
8	-2.336	4.42	9.6	453	1	459	1
9	-2.110	5.68	8.5	458	1	462	1
10	-1.923	6.95	7.6	461	1	466	1
11	-1.764	8.20	7.0	465	1	468	1
12	-1.624	9.38	6.6	467	1	471	2
13	-1.499	10.50	6.2	469	1	473	2
14	-1.385	11.54	5.9	472	2	475	2
15	-1.279	12.52	5.7	474	2	477	2
16	-1.180	13.43	5.5	476	2	479	2
17	-1.087	14.28	5.3	478	2	480	2
18	-0.998	15.07	5.2	480	2	482	2
19	-0.914	15.81	5.1	482	2	483	2
20	-0.832	16.49	5.0	483	2	485	2
21	-0.754	17.11	4.9	485	2	486	2
22	-0.677	17.68	4.8	487	2	488	2
23	-0.603	18.18	4.7	488	2	489	2
24	-0.531	18.63	4.7	489	2	491	2
25	-0.459	19.01	4.6	491	2	492	2
26	-0.389	19.32	4.6	492	2	493	2
27	-0.320	19.58	4.6	494	2	494	2
28	-0.251	19.78	4.5	495	2	496	2
29	-0.183	19.92	4.5	496	2	497	2
30	-0.115	20.00	4.5	498	2	498	2
31	-0.048	20.02	4.5	499	2	499	2
32	0.021	19.98	4.5	501	3	501	3
33	0.089	19.89	4.5	502	3	502	3
34	0.158	19.73	4.5	503	3	504	3
35	0.228	19.52	4.6	505	3	505	3
36	0.299	19.25	4.6	506	3	506	3
37	0.371	18.92	4.6	508	3	508	3

Table 2.2.12 (continued)
 Raw Score to Scale Score Lookup Table
 Mathematics Grade 8

Raw Score	Theta	Information	SE (Scale Score)	2023		2022	
				Scale Score	Achievement Levels	Scale Score	Achievement Levels
38	0.445	18.54	4.7	509	3	509	3
39	0.521	18.09	4.7	511	3	511	3
40	0.599	17.58	4.8	512	3	512	3
41	0.681	17.01	4.9	514	3	514	3
42	0.766	16.36	5.0	516	3	516	3
43	0.855	15.64	5.1	517	3	517	3
44	0.949	14.83	5.2	519	3	519	3
45	1.050	13.93	5.4	521	3	521	3
46	1.159	12.93	5.6	524	3	523	3
47	1.278	11.80	5.9	526	3	525	3
48	1.411	10.54	6.2	529	3	528	3
49	1.563	9.12	6.7	532	4	530	4
50	1.744	7.52	7.4	535	4	534	4
51	1.972	5.73	8.4	540	4	538	4
52	2.291	3.74	10.0	546	4	543	4
53	2.854	1.62	10.0	558	4	553	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	980	0.01619	0.01619
449	NM	823	0.01359	0.02978
454	NM	1098	0.01814	0.04792
459	NM	1210	0.01999	0.06790
462	NM	1431	0.02364	0.09154
465	NM	1433	0.02367	0.11521
468	NM	1420	0.02345	0.13866
471	PM	1403	0.02317	0.16184
473	PM	1377	0.02274	0.18458
475	PM	1371	0.02265	0.20723
477	PM	1405	0.02321	0.23044
479	PM	1421	0.02347	0.25391
481	PM	1413	0.02334	0.27725
483	PM	1460	0.02412	0.30136
485	PM	1570	0.02593	0.32729
487	PM	1646	0.02719	0.35448
489	PM	1710	0.02824	0.38273
491	PM	1743	0.02879	0.41152
492	PM	1758	0.02904	0.44055
494	PM	1896	0.03132	0.47187
496	PM	1965	0.03246	0.50433
498	PM	2085	0.03444	0.53877
500	ME	2145	0.03543	0.57420
502	ME	2185	0.03609	0.61029
504	ME	2424	0.04004	0.65033
507	ME	2429	0.04012	0.69045
509	ME	2480	0.04096	0.73141
512	ME	2465	0.04072	0.77213
515	ME	2546	0.04205	0.81418
518	ME	2436	0.04024	0.85442
522	ME	2337	0.03860	0.89302
526	ME	1989	0.03285	0.92587
532	EE	1651	0.02727	0.95314
538	EE	1237	0.02043	0.97357
545	EE	821	0.01356	0.98713
555	EE	466	0.00770	0.99483
560	EE	313	0.00517	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	4	0.00006	0.00006
441	NM	29	0.00047	0.00053
442	NM	223	0.00361	0.00414
443	NM	711	0.01150	0.01564
444	NM	537	0.00868	0.02432
449	NM	700	0.01132	0.03564
453	NM	802	0.01297	0.04861
457	NM	952	0.01540	0.06401
461	NM	1104	0.01785	0.08186
464	NM	1165	0.01884	0.10070
467	NM	1207	0.01952	0.12022
469	NM	1424	0.02303	0.14325
472	PM	1527	0.02469	0.16794
475	PM	1645	0.02660	0.19455
477	PM	1727	0.02793	0.22248
480	PM	1876	0.03034	0.25281
482	PM	1999	0.03233	0.28514
484	PM	1996	0.03228	0.31742
486	PM	2134	0.03451	0.35193
488	PM	2235	0.03614	0.38807
490	PM	2248	0.03635	0.42443
492	PM	2308	0.03732	0.46175
494	PM	2304	0.03726	0.49901
497	PM	2523	0.04080	0.53981
499	PM	2564	0.04146	0.58128
501	ME	2663	0.04307	0.62435
503	ME	2729	0.04413	0.66848
506	ME	2741	0.04433	0.71280
509	ME	2698	0.04363	0.75644
511	ME	2736	0.04425	0.80068
515	ME	2513	0.04064	0.84132
518	ME	2348	0.03797	0.87929
522	ME	2199	0.03556	0.91486
526	ME	1749	0.02828	0.94314
531	EE	1348	0.02180	0.96494
537	EE	989	0.01599	0.98093
544	EE	605	0.00978	0.99072
552	EE	323	0.00522	0.99594
560	EE	251	0.00406	1.00000

Table 2.3.3
Cumulative Scale Score Distribution
English Language Arts Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
441	NM	16	0.00026	0.00026
442	NM	130	0.00209	0.00234
443	NM	475	0.00762	0.00997
444	NM	429	0.00688	0.01685
449	NM	553	0.00887	0.02572
453	NM	633	0.01016	0.03588
456	NM	727	0.01167	0.04755
459	NM	805	0.01292	0.06047
461	NM	799	0.01282	0.07329
464	NM	826	0.01326	0.08654
466	NM	898	0.01441	0.10095
468	NM	993	0.01593	0.11689
469	NM	961	0.01542	0.13231
472	PM	1069	0.01715	0.14946
474	PM	1079	0.01731	0.16678
476	PM	1235	0.01982	0.18660
477	PM	1280	0.02054	0.20714
479	PM	1330	0.02134	0.22848
481	PM	1406	0.02256	0.25104
483	PM	1534	0.02462	0.27566
485	PM	1570	0.02519	0.30085
486	PM	1770	0.02840	0.32926
488	PM	1910	0.03065	0.35991
490	PM	1957	0.03140	0.39131
492	PM	2126	0.03412	0.42543
494	PM	2297	0.03686	0.46229
496	PM	2309	0.03705	0.49934
498	PM	2574	0.04131	0.54065
500	ME	2730	0.04381	0.58446
503	ME	2916	0.04679	0.63125
505	ME	2942	0.04721	0.67846
508	ME	2904	0.04660	0.72506
511	ME	2876	0.04615	0.77121
513	ME	2749	0.04411	0.81533
517	ME	2453	0.03936	0.85469
520	ME	2242	0.03598	0.89067
523	ME	1909	0.03063	0.92130
528	ME	1542	0.02474	0.94605
532	EE	1242	0.01993	0.96598
537	EE	879	0.01411	0.98009

Table 2.3.3 (continued)
 Cumulative Scale Score Distribution
 English Language Arts Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
543	EE	621	0.00997	0.99005
551	EE	358	0.00574	0.99580
560	EE	262	0.00420	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	147	0.00231	0.00231
441	NM	3513	0.05526	0.05757
446	NM	1044	0.01642	0.07399
450	NM	1111	0.01748	0.09147
453	NM	1126	0.01771	0.10918
456	NM	1184	0.01862	0.12780
460	NM	1240	0.01950	0.14731
463	NM	1226	0.01928	0.16659
465	NM	1328	0.02089	0.18748
468	NM	1362	0.02142	0.20891
471	PM	1425	0.02241	0.23132
473	PM	1562	0.02457	0.25589
476	PM	1572	0.02473	0.28062
479	PM	1589	0.02499	0.30561
481	PM	1726	0.02715	0.33276
484	PM	1808	0.02844	0.36120
486	PM	1927	0.03031	0.39151
488	PM	1983	0.03119	0.42270
491	PM	2106	0.03313	0.45583
493	PM	2048	0.03221	0.48805
496	PM	2141	0.03368	0.52172
499	PM	2336	0.03674	0.55847
501	ME	2386	0.03753	0.59600
504	ME	2508	0.03945	0.63545
507	ME	2333	0.03670	0.67215
510	ME	2525	0.03972	0.71186
513	ME	2436	0.03832	0.75018
516	ME	2448	0.03851	0.78869
519	ME	2272	0.03574	0.82443
522	ME	2177	0.03424	0.85867
526	ME	1878	0.02954	0.88821
529	ME	1744	0.02743	0.91564
534	EE	1487	0.02339	0.93903
539	EE	1210	0.01903	0.95806
544	EE	975	0.01534	0.97340
550	EE	742	0.01167	0.98507
556	EE	505	0.00794	0.99302
560	EE	444	0.00698	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	2	0.00003	0.00003
441	NM	161	0.00253	0.00256
442	NM	1208	0.01896	0.02152
443	NM	800	0.01256	0.03408
448	NM	968	0.01519	0.04927
452	NM	1087	0.01706	0.06633
456	NM	1074	0.01686	0.08319
459	NM	1281	0.02011	0.10329
462	NM	1270	0.01993	0.12323
465	NM	1326	0.02081	0.14404
468	NM	1369	0.02149	0.16553
470	PM	1408	0.02210	0.18763
473	PM	1482	0.02326	0.21089
475	PM	1543	0.02422	0.23511
477	PM	1544	0.02423	0.25934
479	PM	1564	0.02455	0.28389
481	PM	1593	0.02500	0.30889
483	PM	1743	0.02736	0.33625
485	PM	1793	0.02814	0.36440
487	PM	1847	0.02899	0.39339
489	PM	1927	0.03025	0.42363
491	PM	1907	0.02993	0.45356
493	PM	1972	0.03095	0.48452
495	PM	1981	0.03109	0.51561
496	PM	2024	0.03177	0.54738
498	PM	1956	0.03070	0.57808
500	ME	1958	0.03073	0.60881
502	ME	1960	0.03076	0.63958
504	ME	2038	0.03199	0.67156
507	ME	1961	0.03078	0.70234
509	ME	1872	0.02938	0.73173
511	ME	1805	0.02833	0.76006
513	ME	1854	0.02910	0.78916
516	ME	1801	0.02827	0.81743
518	ME	1728	0.02712	0.84455
521	ME	1706	0.02678	0.87133
524	ME	1589	0.02494	0.89627
528	ME	1429	0.02243	0.91870
531	EE	1365	0.02142	0.94012
536	EE	1178	0.01849	0.95861

Table 2.3.5 (continued)
 Cumulative Scale Score Distribution
 English Language Arts Grade 7

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
541	EE	945	0.01483	0.97344
546	EE	767	0.01204	0.98548
554	EE	507	0.00796	0.99344
560	EE	418	0.00656	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	3036	0.04631	0.04631
441	NM	711	0.01085	0.05716
445	NM	770	0.01175	0.06891
448	NM	842	0.01284	0.08175
452	NM	930	0.01419	0.09594
455	NM	1005	0.01533	0.11127
458	NM	1002	0.01529	0.12655
461	NM	1041	0.01588	0.14243
464	NM	1150	0.01754	0.15998
467	NM	1230	0.01876	0.17874
469	NM	1280	0.01953	0.19827
472	PM	1400	0.02136	0.21962
474	PM	1458	0.02224	0.24187
477	PM	1582	0.02413	0.26600
479	PM	1661	0.02534	0.29134
482	PM	1751	0.02671	0.31805
484	PM	1830	0.02792	0.34596
486	PM	1927	0.02940	0.37536
489	PM	2081	0.03175	0.40711
491	PM	2079	0.03171	0.43882
494	PM	2172	0.03313	0.47195
496	PM	2361	0.03602	0.50797
499	PM	2511	0.03830	0.54628
502	ME	2614	0.03988	0.58615
504	ME	2528	0.03856	0.62472
507	ME	2684	0.04094	0.66566
510	ME	2697	0.04114	0.70680
513	ME	2691	0.04105	0.74785
517	ME	2666	0.04067	0.78852
520	ME	2446	0.03731	0.82584
524	ME	2463	0.03757	0.86341
529	ME	2169	0.03309	0.89650
534	EE	1963	0.02995	0.92644
539	EE	1685	0.02570	0.95215
546	EE	1309	0.01997	0.97211
553	EE	886	0.01352	0.98563
560	EE	942	0.01437	1.00000

Table 2.3.7
Cumulative Scale Score Distribution
Mathematics Grade 3

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	952	0.01841	0.01841
447	NM	462	0.00893	0.02735
452	NM	567	0.01097	0.03831
456	NM	626	0.01211	0.05042
459	NM	633	0.01224	0.06266
462	NM	685	0.01325	0.07591
464	NM	721	0.01394	0.08985
467	NM	764	0.01478	0.10463
469	NM	771	0.01491	0.11954
471	PM	874	0.01690	0.13644
473	PM	882	0.01706	0.15350
475	PM	876	0.01694	0.17044
477	PM	910	0.01760	0.18804
478	PM	977	0.01889	0.20694
480	PM	1005	0.01944	0.22637
482	PM	1048	0.02027	0.24664
483	PM	1058	0.02046	0.26710
485	PM	1096	0.02120	0.28830
487	PM	1173	0.02269	0.31098
488	PM	1200	0.02321	0.33419
490	PM	1120	0.02166	0.35585
491	PM	1293	0.02501	0.38086
493	PM	1277	0.02470	0.40555
495	PM	1293	0.02501	0.43056
496	PM	1368	0.02646	0.45702
498	PM	1406	0.02719	0.48421
499	PM	1430	0.02766	0.51186
501	ME	1481	0.02864	0.54051
503	ME	1495	0.02891	0.56942
505	ME	1609	0.03112	0.60054
507	ME	1562	0.03021	0.63075
509	ME	1694	0.03276	0.66351
511	ME	1647	0.03185	0.69536
513	ME	1775	0.03433	0.72969
516	ME	1743	0.03371	0.76340
518	ME	1839	0.03557	0.79896
521	ME	1800	0.03481	0.83377
524	ME	1892	0.03659	0.87037
528	ME	1697	0.03282	0.90319
532	EE	1486	0.02874	0.93192

Table 2.3.7 (continued)
 Cumulative Scale Score Distribution
 Mathematics Grade 3

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
537	EE	1334	0.02580	0.95772
544	EE	1117	0.02160	0.97933
555	EE	768	0.01485	0.99418
560	EE	301	0.00582	1.00000

Table 2.3.8
Cumulative Scale Score Distribution
Mathematics Grade 4

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	1	0.00002	0.00002
441	NM	22	0.00042	0.00044
442	NM	117	0.00223	0.00266
443	NM	162	0.00308	0.00575
444	NM	542	0.01031	0.01606
449	NM	337	0.00641	0.02247
454	NM	445	0.00847	0.03094
457	NM	458	0.00871	0.03965
460	NM	493	0.00938	0.04904
463	NM	556	0.01058	0.05961
465	NM	622	0.01184	0.07145
468	NM	618	0.01176	0.08321
469	NM	673	0.01281	0.09602
472	PM	723	0.01376	0.10977
474	PM	747	0.01421	0.12399
476	PM	784	0.01492	0.13890
477	PM	830	0.01579	0.15470
479	PM	876	0.01667	0.17137
481	PM	960	0.01827	0.18963
482	PM	924	0.01758	0.20722
484	PM	986	0.01876	0.22598
485	PM	1070	0.02036	0.24634
487	PM	1043	0.01985	0.26618
488	PM	1134	0.02158	0.28776
490	PM	1194	0.02272	0.31048
491	PM	1272	0.02420	0.33468
493	PM	1292	0.02458	0.35927
494	PM	1346	0.02561	0.38488
496	PM	1359	0.02586	0.41074
497	PM	1357	0.02582	0.43656
499	PM	1396	0.02656	0.46312
500	ME	1530	0.02911	0.49224
502	ME	1530	0.02911	0.52135
504	ME	1520	0.02892	0.55027
505	ME	1597	0.03039	0.58066
507	ME	1660	0.03159	0.61225
509	ME	1683	0.03202	0.64427
511	ME	1719	0.03271	0.67698
513	ME	1684	0.03204	0.70902
515	ME	1787	0.03400	0.74303

Table 2.3.8 (continued)
 Cumulative Scale Score Distribution
 Mathematics Grade 4

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
517	ME	1753	0.03336	0.77638
519	ME	1609	0.03062	0.80700
522	ME	1608	0.03060	0.83760
525	ME	1614	0.03071	0.86831
528	ME	1461	0.02780	0.89611
531	EE	1386	0.02637	0.92248
535	EE	1158	0.02203	0.94451
539	EE	971	0.01848	0.96299
544	EE	790	0.01503	0.97802
551	EE	580	0.01104	0.98906
560	EE	575	0.01094	1.00000

Table 2.3.9
Cumulative Scale Score Distribution
Mathematics Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
441	NM	2	0.00004	0.00004
442	NM	12	0.00022	0.00026
443	NM	44	0.00081	0.00107
444	NM	76	0.00140	0.00247
445	NM	123	0.00227	0.00475
446	NM	191	0.00353	0.00827
452	NM	252	0.00465	0.01292
456	NM	405	0.00748	0.02040
460	NM	533	0.00984	0.03024
463	NM	577	0.01065	0.04090
466	NM	700	0.01292	0.05382
468	NM	841	0.01553	0.06935
471	PM	913	0.01686	0.08621
473	PM	973	0.01797	0.10417
475	PM	1065	0.01966	0.12384
477	PM	1167	0.02155	0.14539
479	PM	1169	0.02158	0.16697
480	PM	1239	0.02288	0.18985
482	PM	1327	0.02450	0.21435
484	PM	1304	0.02408	0.23843
485	PM	1324	0.02445	0.26287
487	PM	1300	0.02400	0.28688
488	PM	1346	0.02485	0.31173
490	PM	1345	0.02483	0.33656
491	PM	1403	0.02591	0.36247
492	PM	1411	0.02605	0.38852
494	PM	1461	0.02698	0.41550
495	PM	1389	0.02565	0.44115
497	PM	1387	0.02561	0.46676
498	PM	1423	0.02627	0.49303
499	PM	1459	0.02694	0.51997
501	ME	1423	0.02627	0.54624
502	ME	1452	0.02681	0.57305
504	ME	1479	0.02731	0.60036
505	ME	1439	0.02657	0.62693
507	ME	1484	0.02740	0.65433
508	ME	1477	0.02727	0.68160
510	ME	1370	0.02530	0.70690
511	ME	1395	0.02576	0.73266
513	ME	1465	0.02705	0.75971

Table 2.3.9 (continued)
 Cumulative Scale Score Distribution
 Mathematics Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
514	ME	1339	0.02472	0.78443
516	ME	1325	0.02447	0.80890
518	ME	1213	0.02240	0.83129
520	ME	1234	0.02278	0.85408
522	ME	1232	0.02275	0.87683
524	ME	1177	0.02173	0.89856
526	ME	1033	0.01907	0.91763
529	ME	1046	0.01931	0.93694
532	EE	906	0.01673	0.95367
536	EE	741	0.01368	0.96736
541	EE	726	0.01340	0.98076
547	EE	533	0.00984	0.99060
559	EE	350	0.00646	0.99706
560	EE	159	0.00294	1.00000

Table 2.3.10
Cumulative Scale Score Distribution
Mathematics Grade 6

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	3	0.00005	0.00005
441	NM	54	0.00096	0.00101
442	NM	295	0.00523	0.00624
443	NM	351	0.00622	0.01247
450	NM	487	0.00864	0.02110
456	NM	702	0.01245	0.03355
460	NM	858	0.01522	0.04877
464	NM	1005	0.01782	0.06659
467	NM	1084	0.01922	0.08581
469	NM	1143	0.02027	0.10608
472	PM	1153	0.02045	0.12653
474	PM	1178	0.02089	0.14742
476	PM	1291	0.02289	0.17032
478	PM	1246	0.02210	0.19241
480	PM	1336	0.02369	0.21611
482	PM	1352	0.02398	0.24008
483	PM	1327	0.02353	0.26362
485	PM	1312	0.02327	0.28688
487	PM	1386	0.02458	0.31146
488	PM	1302	0.02309	0.33455
489	PM	1393	0.02470	0.35925
491	PM	1380	0.02447	0.38373
492	PM	1363	0.02417	0.40790
494	PM	1370	0.02430	0.43219
495	PM	1370	0.02430	0.45649
496	PM	1334	0.02366	0.48015
498	PM	1268	0.02249	0.50263
499	PM	1301	0.02307	0.52571
500	ME	1273	0.02258	0.54828
501	ME	1272	0.02256	0.57084
503	ME	1274	0.02259	0.59343
504	ME	1264	0.02242	0.61585
505	ME	1195	0.02119	0.63704
506	ME	1225	0.02172	0.65876
508	ME	1198	0.02125	0.68001
509	ME	1215	0.02155	0.70156
510	ME	1173	0.02080	0.72236
512	ME	1171	0.02077	0.74312
513	ME	1181	0.02094	0.76407
515	ME	1178	0.02089	0.78496

Table 2.3.10 (continued)
 Cumulative Scale Score Distribution
 Mathematics Grade 6

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
516	ME	1188	0.02107	0.80603
518	ME	1191	0.02112	0.82715
519	ME	1142	0.02025	0.84740
521	ME	1076	0.01908	0.86648
523	ME	1104	0.01958	0.88606
525	ME	1070	0.01898	0.90503
528	ME	1009	0.01789	0.92293
530	EE	1015	0.01800	0.94093
534	EE	942	0.01671	0.95763
538	EE	893	0.01584	0.97347
543	EE	662	0.01174	0.98521
554	EE	524	0.00929	0.99450
560	EE	310	0.00550	1.00000

Table 2.3.11
 Cumulative Scale Score Distribution
 Mathematics Grade 7

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	298	0.00521	0.00521
441	NM	1140	0.01992	0.02512
451	NM	993	0.01735	0.04247
458	NM	1394	0.02436	0.06683
462	NM	1521	0.02658	0.09341
466	NM	1804	0.03152	0.12493
469	NM	1743	0.03045	0.15538
473	PM	1730	0.03023	0.18561
475	PM	1738	0.03037	0.21597
478	PM	1633	0.02853	0.24451
480	PM	1633	0.02853	0.27304
482	PM	1629	0.02846	0.30150
484	PM	1572	0.02747	0.32897
486	PM	1543	0.02696	0.35592
487	PM	1443	0.02521	0.38114
489	PM	1539	0.02689	0.40803
490	PM	1382	0.02415	0.43217
492	PM	1279	0.02235	0.45452
494	PM	1361	0.02378	0.47830
495	PM	1335	0.02333	0.50162
496	PM	1280	0.02236	0.52399
498	PM	1176	0.02055	0.54454
499	PM	1125	0.01966	0.56419
500	ME	1197	0.02091	0.58511
502	ME	1126	0.01967	0.60478
503	ME	1109	0.01938	0.62416
504	ME	1068	0.01866	0.64282
506	ME	1011	0.01766	0.66048
507	ME	1041	0.01819	0.67867
508	ME	1045	0.01826	0.69693
510	ME	1071	0.01871	0.71564
511	ME	963	0.01683	0.73247
512	ME	1005	0.01756	0.75003
514	ME	966	0.01688	0.76690
515	ME	970	0.01695	0.78385
516	ME	989	0.01728	0.80113
518	ME	902	0.01576	0.81689
520	ME	957	0.01672	0.83361
521	ME	931	0.01627	0.84988
523	ME	899	0.01571	0.86559

Table 2.3.11 (continued)
 Cumulative Scale Score Distribution
 Mathematics Grade 7

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
525	ME	911	0.01592	0.88150
527	ME	867	0.01515	0.89665
529	ME	865	0.01511	0.91177
531	EE	887	0.01550	0.92726
534	EE	829	0.01448	0.94175
537	EE	816	0.01426	0.95601
540	EE	757	0.01323	0.96923
545	EE	648	0.01132	0.98055
551	EE	525	0.00917	0.98973
560	EE	588	0.01027	1.00000

Table 2.3.12
Cumulative Scale Score Distribution
Mathematics Grade 8

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
441	NM	2	0.00003	0.00003
442	NM	7	0.00012	0.00015
443	NM	37	0.00062	0.00077
444	NM	86	0.00144	0.00222
445	NM	201	0.00337	0.00559
446	NM	353	0.00593	0.01152
447	NM	583	0.00979	0.02130
453	NM	806	0.01353	0.03483
458	NM	995	0.01670	0.05153
461	NM	1134	0.01904	0.07057
465	NM	1291	0.02167	0.09224
467	NM	1488	0.02498	0.11722
469	NM	1430	0.02400	0.14122
472	PM	1511	0.02536	0.16659
474	PM	1549	0.02600	0.19259
476	PM	1509	0.02533	0.21792
478	PM	1521	0.02553	0.24345
480	PM	1550	0.02602	0.26947
482	PM	1504	0.02525	0.29472
483	PM	1477	0.02479	0.31951
485	PM	1451	0.02436	0.34387
487	PM	1405	0.02358	0.36745
488	PM	1409	0.02365	0.39111
489	PM	1443	0.02422	0.41533
491	PM	1411	0.02369	0.43901
492	PM	1309	0.02197	0.46099
494	PM	1343	0.02254	0.48353
495	PM	1416	0.02377	0.50730
496	PM	1401	0.02352	0.53082
498	PM	1302	0.02186	0.55268
499	PM	1288	0.02162	0.57430
501	ME	1289	0.02164	0.59593
502	ME	1236	0.02075	0.61668
503	ME	1290	0.02165	0.63834
505	ME	1301	0.02184	0.66018
506	ME	1247	0.02093	0.68111
508	ME	1229	0.02063	0.70174
509	ME	1223	0.02053	0.72227
511	ME	1194	0.02004	0.74231
512	ME	1219	0.02046	0.76277

Table 2.3.12 (continued)
 Cumulative Scale Score Distribution
 Mathematics Grade 8

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
514	ME	1143	0.01919	0.78196
516	ME	1213	0.02036	0.80232
517	ME	1200	0.02014	0.82247
519	ME	1190	0.01998	0.84244
521	ME	1094	0.01836	0.86081
524	ME	1092	0.01833	0.87914
526	ME	1053	0.01768	0.89681
529	ME	1070	0.01796	0.91478
532	EE	1084	0.01820	0.93297
535	EE	1009	0.01694	0.94991
540	EE	958	0.01608	0.96599
546	EE	868	0.01457	0.98056
558	EE	717	0.01204	0.99260
560	EE	441	0.00740	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	1.16500	0.96000	0.83140	0.84971	-0.24657	False
IA00288	3	1.02000	0.87000	0.77628	0.73880	-0.19323	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	1.44000	1.44500	0.83057	0.69959	0.00602	False
IA00226	3	1.20000	1.20000	0.78298	0.76349	0.00000	False

Table 2.4.3
Rescore Analysis
Mathematics Grade 3

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01080	3	1.01000	0.97500	0.86233	0.91573	-0.04059	False
IA01081	3	0.78500	0.82000	0.89036	0.89533	0.03931	False

Table 2.4.4
Rescore Analysis
Mathematics Grade 4

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA00789	4	1.41000	1.38500	1.32312	1.41307	-0.01889	False
IA01057	4	2.13000	2.20000	1.20847	1.19041	0.05792	False

Table 2.4.5
Rescore Analysis
Mathematics Grade 5

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01032	4	1.41500	1.45000	1.22486	1.30999	0.02857	False
IA02736	4	1.86000	1.95500	1.44250	1.45726	0.06586	False

Table 2.4.6
Rescore Analysis
Mathematics Grade 6

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA00881	4	1.82000	1.78500	1.64350	1.59105	-0.02130	False
IA00972	4	2.92000	2.90000	1.12245	1.07507	-0.01782	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	2.13500	2.15000	1.15909	1.17661	0.01294	False
IA02722	4	2.22000	2.12500	1.59194	1.56898	-0.05968	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	2.72000	2.70500	1.50764	1.53942	-0.00995	False
IA01066	4	2.12000	2.13000	1.64595	1.67545	0.00608	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	0.78000	9.91123	9.91123	1	False	-0.86431
IA00280 (EL308824)	0.66000	0.66000	11.35015	11.35015	1	False	-0.73854
IA00281 (EL308826)	0.59000	0.56000	12.08982	12.39612	1	False	0.30730
IA00282 (EL308827)	0.65000	0.61000	11.45872	11.88272	1	False	0.91086
IA00283 (EL308835)	0.57000	0.57000	12.29450	12.29450	1	False	-0.65601
IA00284 (EL308837)	0.66000	0.66000	11.35015	11.35015	1	False	-0.73854
IA00285 (EL308838)	0.77000	0.77000	10.04461	10.04461	1	False	-0.85265
IA00286 (EL308842)	0.41000	0.49000	13.91018	13.10028	1	False	3.25869
IA00287 (EL308855)	0.37333	0.33333	14.29215	14.72291	3	False	0.69467
IA00288 (EL308857)	0.32000	0.29333	14.87080	15.17469	3	False	0.05303
IA00443 (EL626042844)	0.66000	0.65000	11.35015	11.45872	1	False	-0.54932
IA00444 (EL626043062)	0.67000	0.67000	11.24035	11.24035	1	False	-0.74814
IA00445 (EL626043435)	0.51500	0.52500	12.84957	12.74917	2	False	-0.13974
IA00446 (EL626049849)	0.53000	0.51000	12.69892	12.89972	1	False	-0.23747
IA00450 (EL626050679)	0.67000	0.66000	11.24035	11.35015	1	False	-0.53399
IA00451 (EL626050927)	0.51000	0.54000	12.89972	12.59827	1	False	0.80144
IA00452 (EL626051097)	0.55000	0.52000	12.49735	12.79939	2	False	0.25178
IA00453 (EL626051328)	0.64000	0.65000	11.56616	11.45872	1	False	-0.21905

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	0.83000	9.48841	9.18334	1	False	0.70702
IA00219 (EL307709)	0.77000	0.76000	10.04461	10.17479	1	False	-0.38548
IA00220 (EL307710)	0.43000	0.48000	13.70550	13.20061	1	False	1.53502
IA00221 (EL307713)	0.52000	0.53000	12.79939	12.69892	1	False	-0.89187
IA00222 (EL307714)	0.76000	0.75000	10.17479	10.30204	1	False	-0.39087
IA00223 (EL307719)	0.56000	0.59000	12.39612	12.08982	1	False	0.42844
IA00224 (EL307724)	0.75000	0.73000	10.30204	10.54875	1	False	0.36485
IA00225 (EL307728)	0.45333	0.48000	13.46898	13.20061	3	False	0.08678
IA00226 (EL307729)	0.41667	0.41333	13.84171	13.87591	3	False	-0.60884
IA00289 (EL309792)	0.64000	0.58000	11.56616	12.19243	1	False	2.85068
IA00407 (EL624647403)	0.41000	0.42000	13.91018	13.80757	1	False	-0.98789
IA00408 (EL624647580)	0.58000	0.57000	12.19243	12.29450	1	False	-0.34889
IA00411 (EL624652450)	0.78000	0.76000	9.91123	10.17479	1	False	0.43125
IA00412 (EL624652621)	0.90000	0.91000	7.87379	7.63698	1	False	0.44126
IA00414 (EL624652989)	0.43000	0.43000	13.70550	13.70550	1	False	-0.83503
IA00415 (EL624653348)	0.67000	0.67000	11.24035	11.24035	1	False	-1.07768
IA00416 (EL624653492)	0.72000	0.73000	10.66863	10.54875	2	False	-0.56131
IA00419 (EL624654711)	0.80000	0.80500	9.63352	9.56153	2	False	-0.75744

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	-1.34700
IA00497 (EL626304969)	0.73000	0.72000	10.54875	10.66863	1	False	0.57721
IA00500 (EL626332335)	0.66000	0.66000	11.35015	11.35015	1	False	-0.75796
IA00501 (EL626332592)	0.84000	0.84000	9.02217	9.02217	1	False	-1.75129
IA00502 (EL626333002)	0.82000	0.83000	9.33854	9.18334	1	False	0.38243
IA00505 (EL626355215)	0.60000	0.62000	11.98661	11.77808	1	False	-0.31145
IA00506 (EL626355557)	0.64000	0.67000	11.56616	11.24035	1	False	1.67163
IA00508 (EL626356291)	0.36500	0.39000	14.38050	14.11728	2	False	-0.84227
IA00638 (EL627351056)	0.66000	0.68000	11.35015	11.12920	1	False	0.23105
IA01669 (EL711809263)	0.76000	0.77000	10.17479	10.04461	1	False	-0.46024
IA01670 (EL711809592)	0.76000	0.75000	10.17479	10.30204	1	False	0.47674
IA01671 (EL711827203)	0.90000	0.90000	7.87379	7.87379	1	False	-1.10590
IA01672 (EL711827807)	0.73000	0.72000	10.54875	10.66863	1	False	0.57721
IA01679 (EL711868011)	0.49500	0.48500	13.05013	13.15043	2	False	1.69129
IA01680 (EL711900602)	0.70000	0.69000	10.90240	11.01660	1	False	0.69129
IA01691 (EL712167015)	0.46000	0.49000	13.40173	13.10028	1	False	0.27724

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	0.75000	10.42662	10.30204	1	False	-0.01385
IA00174 (EL303500)	0.64000	0.63000	11.56616	11.67259	1	False	-0.65780
IA00175 (EL303504)	0.61000	0.63000	11.88272	11.67259	1	False	0.85814
IA00176 (EL303508)	0.71000	0.73000	10.78646	10.54875	1	False	1.02461
IA00177 (EL303510)	0.87000	0.87000	8.49444	8.49444	1	False	-1.26973
IA00178 (EL303513)	0.66000	0.66000	11.35015	11.35015	1	False	-1.06074
IA00179 (EL303514)	0.63000	0.60000	11.67259	11.98661	1	False	1.19165
IA00180 (EL303518)	0.57000	0.53000	12.29450	12.69892	1	False	1.95480
IA00515 (EL626864414)	0.83000	0.82000	9.18334	9.33854	1	False	-0.04704
IA00517 (EL626864724)	0.69000	0.71000	11.01660	10.78646	1	False	0.97368
IA00518 (EL626865003)	0.67000	0.67000	11.24035	11.24035	1	False	-1.06877
IA00520 (EL626865416)	0.41000	0.42000	13.91018	13.80757	1	False	0.04453
IA00522 (EL626865773)	0.72000	0.70000	10.66863	10.90240	1	False	0.54710
IA00523 (EL626865942)	0.64000	0.64000	11.56616	11.56616	1	False	-1.04493
IA00528 (EL626867605)	0.77000	0.76000	10.04461	10.17479	2	False	-0.33393
IA00530 (EL626868748)	0.70500	0.70500	10.84466	10.84466	2	False	-1.09773

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	0.72051	10.54875	10.66252	1	False	0.17679
IA00066 (EL292163)	0.65000	0.62829	11.45872	11.69065	1	False	1.38684
IA00067 (EL292168)	0.50000	0.48588	13.00000	13.14156	1	False	0.76163
IA00068 (EL292170)	0.71000	0.71518	10.78646	10.72565	1	False	-1.39867
IA00069 (EL292172)	0.62000	0.65699	11.77808	11.38290	1	False	1.52278
IA00070 (EL292176)	0.62000	0.60853	11.77808	11.89802	1	False	0.39862
IA00081 (EL293802)	0.68000	0.70373	11.12920	10.85933	1	False	0.45617
IA00082 (EL293804)	0.54000	0.56885	12.59827	12.30619	1	False	0.46340
IA00257 (EL308358)	0.85000	0.84836	8.85427	8.88234	1	False	-0.83962
IA00258 (EL308360)	0.75500	0.75852	10.23876	10.19375	2	False	-1.32675
IA00262 (EL308382)	0.65000	0.65515	11.45872	11.40300	1	False	-1.26156
IA00265 (EL308389)	0.90000	0.89974	7.87379	7.87982	1	False	-1.17423
IA00269 (EL308397)	0.84000	0.82967	9.02217	9.18848	1	False	0.45563
IA00655 (EL628647210)	0.72000	0.74311	10.66863	10.38814	1	False	0.61580
IA00657 (EL628647689)	0.77000	0.77721	10.04461	9.94882	1	False	-1.00103
IA00658 (EL628653398)	0.74000	0.76319	10.42662	10.13353	2	False	0.76420

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	0.79000	10.04461	9.77432	1	False	0.46444
IA00057 (EL290798)	0.78000	0.79000	9.91123	9.77432	1	False	-0.77635
IA00058 (EL290799)	0.75000	0.76000	10.30204	10.17479	1	False	-0.71010
IA00059 (EL290800)	0.67000	0.67000	11.24035	11.24035	1	False	-1.17573
IA00060 (EL290801)	0.77000	0.76000	10.04461	10.17479	1	False	0.45062
IA00061 (EL290805)	0.56000	0.56000	12.39612	12.39612	1	False	-1.02786
IA00062 (EL290808)	0.54000	0.55000	12.59827	12.49735	1	False	-0.04969
IA00063 (EL290814)	0.44000	0.40000	13.60388	14.01339	1	False	1.55263
IA00368 (EL623873883)	0.72000	0.73000	10.66863	10.54875	1	False	-0.63283
IA00371 (EL623951471)	0.60500	0.59000	11.93476	12.08982	2	False	-0.06445
IA00373 (EL623952377)	0.42000	0.42500	13.80757	13.75647	2	False	-0.02219
IA00374 (EL623952612)	0.73000	0.72000	10.54875	10.66863	1	False	0.16237
IA00378 (EL623955555)	0.51000	0.55000	12.89972	12.49735	1	False	2.75458
IA00379 (EL623955757)	0.54000	0.52000	12.59827	12.79939	1	False	0.08738
IA00383 (EL623959265)	0.65000	0.64000	11.45872	11.56616	1	False	-0.30324
IA00699 (EL632808123)	0.77000	0.77000	10.04461	10.04461	1	False	-0.70958

Table 2.5.7
Delta Analysis
Mathematics Grade 3

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00769 (MA203641)	0.83000	0.88000	9.18334	8.30005	1	False	1.57790
IA00799 (MA260559)	0.49000	0.49000	13.10028	13.10028	1	False	-0.68268
IA00834 (MA293457)	0.79000	0.81000	9.77432	9.48841	1	False	-1.18663
IA00838 (MA293524)	0.74000	0.77000	10.42662	10.04461	1	False	-0.71070
IA00850 (MA297405)	0.69000	0.71000	11.01660	10.78646	1	False	-1.27010
IA00852 (MA297438)	0.63000	0.65000	11.67259	11.45872	1	False	-1.22450
IA00924 (MA306310)	0.46000	0.48000	13.40173	13.20061	1	False	-0.76288
IA00925 (MA306315)	0.75000	0.79000	10.30204	9.77432	1	False	0.02711
IA00930 (MA306359)	0.66000	0.64000	11.35015	11.56616	1	False	1.00353
IA00932 (MA306375)	0.48000	0.49000	13.20061	13.10028	1	False	-1.24771
IA00993 (MA310834)	0.63000	0.62000	11.67259	11.77808	1	False	0.31624
IA01019 (MA311277)	0.74000	0.74000	10.42662	10.42662	1	False	0.13604
IA01071 (MA623063509)	0.73000	0.72000	10.54875	10.66863	1	False	0.73703
IA01080 (MA623654449)	0.31667	0.32667	14.90816	14.79655	3	False	-0.77820
IA01081 (MA623656013)	0.28667	0.32000	15.25260	14.87080	3	False	0.76601
IA02323 (MA301611A)	0.83000	0.87000	9.18334	8.49444	1	False	0.54281
IA04760 (MA713752330)	0.77000	0.80000	10.04461	9.63352	1	False	-0.67277
IA04813 (MA735572247)	0.72000	0.69000	10.66863	11.01660	1	False	1.91484
IA04828 (MA735653938)	0.51000	0.57000	12.89972	12.29450	1	False	1.23523
IA04844 (MA735735757)	0.62000	0.61000	11.77808	11.88272	1	False	0.27944

Table 2.5.8
Delta Analysis
Mathematics Grade 4

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00789 (MA250543)	0.41500	0.42750	13.85881	13.73097	4	False	-0.21097
IA00828 (MA287237)	0.76000	0.79000	10.17479	9.77432	1	False	-0.21310
IA00841 (MA293718)	0.71000	0.73000	10.78646	10.54875	1	False	-1.20028
IA00861 (MA297629)	0.89000	0.90000	8.09389	7.87379	1	False	-0.93982
IA00869 (MA297988)	0.18000	0.21000	16.66146	16.22568	1	False	-0.88035
IA00906 (MA301811)	0.73000	0.72000	10.54875	10.66863	1	False	0.72683
IA00958 (MA307055)	0.46000	0.51000	13.40173	12.89972	1	False	-0.07774
IA00961 (MA307081)	0.53000	0.61000	12.69892	11.88272	1	False	1.76423
IA00963 (MA307085)	0.67000	0.70000	11.24035	10.90240	1	False	-0.70292
IA01048 (MA311534)	0.54000	0.56000	12.59827	12.39612	1	False	-0.79230
IA01049 (MA311537)	0.66000	0.67000	11.35015	11.24035	1	False	-0.44462
IA01055 (MA311572)	0.52000	0.55000	12.79939	12.49735	1	False	-1.11038
IA01057 (MA311581)	0.55250	0.55750	12.47208	12.42146	4	False	0.03411
IA01093 (MA623879088)	0.73000	0.69500	10.54875	10.95971	2	False	2.34659
IA02175 (MA286769)	0.75000	0.78000	10.30204	9.91123	1	False	-0.28381
IA02819 (MA713583365)	0.61000	0.59000	11.88272	12.08982	1	False	1.38976
IA02841 (MA713774890)	0.44000	0.50000	13.60388	13.00000	1	False	0.46221
IA02902 (MA714251321)	0.39000	0.41000	14.11728	13.91018	1	False	-0.61762
IA04661 (MA307327)	0.74000	0.79000	10.42662	9.77432	1	False	1.15474
IA04965 (MA800867144)	0.58000	0.62000	12.19243	11.77808	1	False	-0.40454

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)	0.73000	0.75000	10.54875	10.30204	1	False	-0.36791
IA00776 (MA221207)	0.66000	0.69000	11.35015	11.01660	1	False	-0.75606
IA00803 (MA262207)	0.72000	0.75000	10.66863	10.30204	1	False	-0.80042
IA00806 (MA272292)	0.47000	0.47000	13.30108	13.30108	1	False	0.20796
IA00826 (MA287178)	0.85000	0.86000	8.85427	8.67872	1	False	0.05888
IA00872 (MA298003)	0.63000	0.66000	11.67259	11.35015	1	False	-0.75092
IA00880 (MA298106)	0.21000	0.24000	16.22568	15.82521	1	False	-0.25178
IA00885 (MA299556)	0.68000	0.73000	11.12920	10.54875	1	False	-0.15522
IA00936 (MA306420)	0.69000	0.82000	11.01660	9.33854	1	False	3.67778
IA00943 (MA306466)	0.62000	0.64000	11.77808	11.56616	1	False	-0.37481
IA00989 (MA307638)	0.86000	0.87000	8.67872	8.49444	1	False	0.04665
IA01020 (MA311280)	0.43000	0.43000	13.70550	13.70550	1	False	0.16559
IA01029 (MA311337)	0.83000	0.86000	9.18334	8.67872	1	False	-0.62473
IA01032 (MA311366)	0.38500	0.41000	14.16950	13.91018	4	False	-0.79139
IA01149 (MA624347774)	0.41000	0.42000	13.91018	13.80757	1	False	-0.21527
IA01155 (MA624357395)	0.35000	0.41500	14.54128	13.85881	2	False	0.55958
IA02552 (MA311324)	0.38000	0.44000	14.22192	13.60388	1	False	0.30044
IA02736 (MA704359678)	0.48500	0.49750	13.15043	13.02507	4	False	-0.21541
IA04970 (MA800974344)	0.70000	0.73000	10.90240	10.54875	1	False	-0.77957
IA05002 (MA801652356)	0.60000	0.58000	11.98661	12.19243	1	False	1.06661

Table 2.5.10
Delta Analysis
Mathematics Grade 6

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00777 (MA221667)	0.85000	0.84000	8.85427	9.02217	1	False	-0.66392
IA00778 (MA221669)	0.80000	0.79000	9.63352	9.77432	1	False	-0.78936
IA00804 (MA264305)	0.74000	0.75000	10.42662	10.30204	1	False	0.13826
IA00817 (MA280989)	0.47000	0.49000	13.30108	13.10028	1	False	0.42433
IA00818 (MA282268)	0.40000	0.41000	14.01339	13.91018	1	False	-0.37124
IA00819 (MA282277)	0.47000	0.44000	13.30108	13.60388	1	False	0.77408
IA00827 (MA287186)	0.56000	0.55000	12.39612	12.49735	1	False	-0.81368
IA00845 (MA296349)	0.56000	0.53000	12.39612	12.69892	1	False	0.68564
IA00881 (MA298139)	0.45000	0.44750	13.50265	13.52792	4	False	-1.27053
IA00899 (MA301508)	0.33000	0.35000	14.75965	14.54128	1	False	0.41244
IA00972 (MA307339)	0.74000	0.73000	10.42662	10.54875	4	False	-0.85072
IA00992 (MA309941)	0.41000	0.40000	13.91018	14.01339	1	False	-0.65100
IA01058 (MA311658)	0.42000	0.46000	13.80757	13.40173	1	False	1.89995
IA02037 (MA217493)	0.66000	0.63000	11.35015	11.67259	1	False	0.72951
IA02597 (MA311693)	0.78000	0.77000	9.91123	10.04461	1	False	-0.81737
IA02698 (MA703179529)	0.28000	0.24000	15.33137	15.82521	1	False	2.39358
IA04745 (MA703231515)	0.50000	0.52500	13.00000	12.74917	2	False	0.82584
IA04884 (MA736365836)	0.68000	0.68000	11.12920	11.12920	1	False	-0.85706
IA05126 (MA805103779)	0.55000	0.56000	12.49735	12.39612	1	False	-0.23778
IA05135 (MA805171807)	0.58000	0.58000	12.19243	12.19243	1	False	-0.96097

Table 2.5.11
Delta Analysis
Mathematics Grade 7

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00796 (MA259267)	0.44000	0.45000	13.60388	13.50265	1	False	-0.66698
IA00831 (MA288414)	0.72000	0.73000	10.66863	10.54875	1	False	-0.76147
IA00847 (MA296358)	0.53000	0.53000	12.69892	12.69892	1	False	-0.75606
IA00909 (MA301846)	0.87000	0.87000	8.49444	8.49444	1	False	-0.47584
IA00910 (MA301854)	0.35000	0.35000	14.54128	14.54128	1	False	-0.87885
IA00945 (MA306538)	0.69000	0.71000	11.01660	10.78646	1	False	-0.14051
IA00948 (MA306600)	0.84000	0.85000	9.02217	8.85427	1	False	-0.61087
IA00949 (MA306605)	0.49000	0.50000	13.10028	13.00000	1	False	-0.70573
IA01006 (MA311093)	0.85000	0.86000	8.85427	8.67872	1	False	-0.58063
IA01011 (MA311109)	0.39000	0.41000	14.11728	13.91018	1	False	-0.05878
IA01016 (MA311125)	0.55000	0.59000	12.49735	12.08982	1	False	0.92001
IA01017 (MA311135)	0.75000	0.73000	10.30204	10.54875	1	False	0.74130
IA01018 (MA311140)	0.37000	0.42000	14.32741	13.80757	1	False	1.65088
IA01069 (MA316886)	0.50750	0.51750	12.92480	12.82448	4	False	-0.71720
IA01097 (MA623950280)	0.32000	0.35000	14.87080	14.54128	1	False	0.65517
IA01108 (MA624149677)	0.38500	0.38500	14.16950	14.16950	2	False	-0.85408
IA02722 (MA703943185)	0.48750	0.48500	13.12535	13.15043	4	False	-0.64850
IA04486 (MA227988)	0.69000	0.68000	11.01660	11.12920	1	False	-0.03340
IA04538 (MA282218)	0.69000	0.66000	11.01660	11.35015	1	False	1.16452
IA04593 (MA298192)	0.24000	0.19000	15.82521	16.51159	1	False	2.75701

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	0.62000	11.77808	11.77808	1	False	-0.72911
IA00858 (MA297513)	0.76000	0.78000	10.17479	9.91123	1	False	-0.07380
IA00864 (MA297652)	0.70250	0.69250	10.87358	10.98820	4	False	-0.37359
IA00865 (MA297656)	0.53000	0.60000	12.69892	11.98661	1	False	1.38560
IA00903 (MA301674)	0.73000	0.75000	10.54875	10.30204	1	False	-0.15975
IA00905 (MA301702)	0.44000	0.46000	13.60388	13.40173	1	False	-0.52420
IA00979 (MA307472)	0.63000	0.65000	11.67259	11.45872	1	False	-0.35344
IA00985 (MA307570)	0.51000	0.53000	12.89972	12.69892	1	False	-0.48230
IA01033 (MA311384)	0.59000	0.57000	12.08982	12.29450	1	False	0.03372
IA01037 (MA311414)	0.39000	0.39000	14.11728	14.11728	1	False	-0.57377
IA01042 (MA311448)	0.52000	0.50000	12.79939	13.00000	1	False	0.06608
IA01044 (MA311463)	0.64000	0.62000	11.56616	11.77808	1	False	0.02515
IA01066 (MA314812)	0.55500	0.54000	12.44678	12.59827	4	False	-0.13547
IA01125 (MA624247061)	0.42500	0.41500	13.75647	13.85881	2	False	-0.22670
IA02495 (MA309741)	0.44000	0.37000	13.60388	14.32741	1	False	2.01547
IA04665 (MA307399)	0.48000	0.47000	13.20061	13.30108	1	False	-0.27038
IA04678 (MA309738)	0.39000	0.40000	14.11728	14.01339	1	False	-0.91454
IA05057 (MA803856437)	0.83000	0.83000	9.18334	9.18334	1	False	-0.90141
IA05059 (MA803856627)	0.71000	0.71000	10.78646	10.78646	1	False	-0.79496
IA05070 (MA804042487)	0.29000	0.40000	15.21354	14.01339	1	False	2.98739

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)	-1.18760	-0.87340	-0.55775	False
IA00280 (EL308824)	-0.45090	-0.33680	-0.18995	False
IA00281 (EL308826)	0.12140	0.38150	0.20141	False
IA00282 (EL308827)	-0.58060	-0.23860	0.51369	False
IA00283 (EL308835)	-0.21220	0.01710	-0.59857	False
IA00284 (EL308837)	-0.40770	-0.21030	-1.25578	False
IA00285 (EL308838)	-0.79560	-0.59680	-0.94841	False
IA00286 (EL308842)	0.68750	0.48310	2.84228	False
IA00287 (EL308855)	0.62903	0.87387	0.57357	False
IA00288 (EL308857)	0.95470	1.20053	0.95909	False
IA00443 (EL626042844)	-0.82100	-0.50270	-0.08327	False
IA00444 (EL626043062)	-0.47130	-0.24520	-0.93803	False
IA00445 (EL626043435)	-0.26400	-0.14830	-0.42518	False
IA00446 (EL626049849)	0.08860	0.23660	-1.26727	False
IA00450 (EL626050679)	-0.33230	-0.21730	-0.33765	False
IA00451 (EL626050927)	0.33820	0.30180	0.95599	False
IA00452 (EL626051097)	-0.36570	-0.08895	-0.12844	False
IA00453 (EL626051328)	-0.36360	-0.32110	0.68429	False

Table 2.6.2
b/b Analysis
English Language Arts Grade 4

Item Id	Old b	New b	Std Dist	Flag
IA00218 (EL307705)	-1.63480	-1.36810	0.01561	False
IA00219 (EL307709)	-1.30830	-0.75530	0.53035	False
IA00220 (EL307710)	0.36960	0.42110	1.42753	False
IA00221 (EL307713)	0.03850	0.33620	-1.46460	False
IA00222 (EL307714)	-1.38880	-0.81150	0.78469	False
IA00223 (EL307719)	0.04010	0.13280	1.12192	False
IA00224 (EL307724)	-0.83660	-0.56750	-0.53412	False
IA00225 (EL307728)	-0.01553	0.15157	0.21907	False
IA00226 (EL307729)	0.22513	0.77117	1.44036	False
IA00289 (EL309792)	-0.73060	-0.11490	1.69771	False
IA00407 (EL624647403)	0.69700	0.98800	-1.47171	False
IA00408 (EL624647580)	-0.61600	-0.15950	-0.23717	False
IA00411 (EL624652450)	-1.23730	-0.81050	-1.01637	False
IA00412 (EL624652621)	-1.94330	-1.66390	0.05607	False
IA00414 (EL624652989)	0.67400	0.90240	-1.00344	False
IA00415 (EL624653348)	-0.72190	-0.24810	-0.08772	False
IA00416 (EL624653492)	-1.29545	-0.98225	-0.79216	False
IA00419 (EL624654711)	-1.70685	-1.38085	-0.68600	False

Table 2.6.3
b/b Analysis
English Language Arts Grade 5

Item Id	Old b	New b	Std Dist	Flag
IA00495 (EL626304658)	-1.32450	-0.78860	1.24122	False
IA00497 (EL626304969)	-0.62920	-0.20410	0.97243	False
IA00500 (EL626332335)	-0.80010	-0.24030	2.46249	False
IA00501 (EL626332592)	-2.05630	-1.61000	-1.16952	False
IA00502 (EL626333002)	-1.55440	-1.21770	-0.58302	False
IA00505 (EL626355215)	-0.31660	-0.24020	0.73092	False
IA00506 (EL626355557)	-0.49820	-0.32940	-0.18018	False
IA00508 (EL626356291)	0.50945	0.52635	0.09440	False
IA00638 (EL627351056)	-0.31920	-0.21860	0.41474	False
IA01669 (EL711809263)	-1.14070	-0.88000	-0.28961	False
IA01670 (EL711809592)	-0.80730	-0.48550	-0.70357	False
IA01671 (EL711827203)	-2.00420	-1.68540	0.43006	False
IA01672 (EL711827807)	-1.34330	-1.00210	-1.00680	False
IA01679 (EL711868011)	-0.17220	0.08940	-0.40572	False
IA01680 (EL711900602)	-0.54940	-0.32200	-0.86835	False
IA01691 (EL712167015)	0.41820	0.54010	-1.13950	False

Table 2.6.4
b/b Analysis
English Language Arts Grade 6

Item Id	Old b	New b	Std Dist	Flag
IA00173 (EL303496)	-1.31980	-0.72360	-0.11057	False
IA00174 (EL303500)	-1.08450	-0.35410	1.11142	False
IA00175 (EL303504)	0.05690	0.09160	1.99640	False
IA00176 (EL303508)	-1.49590	-0.93120	1.56865	False
IA00177 (EL303510)	-2.47490	-1.37010	-0.39480	False
IA00178 (EL303513)	-0.95990	-0.37020	-0.77446	False
IA00179 (EL303514)	-0.18000	0.22420	0.53263	False
IA00180 (EL303518)	0.24010	0.46220	-0.31321	False
IA00515 (EL626864414)	-2.29240	-1.32600	-1.09959	False
IA00517 (EL626864724)	-1.18330	-0.61830	-0.36323	False
IA00518 (EL626865003)	-1.12900	-0.40060	0.79950	False
IA00520 (EL626865416)	1.51200	1.20810	-0.58371	False
IA00522 (EL626865773)	-0.95790	-0.30290	0.46945	False
IA00523 (EL626865942)	-0.94700	-0.39630	-1.43052	False
IA00528 (EL626867605)	-1.79435	-0.98285	-1.24707	False
IA00530 (EL626868748)	-1.56920	-0.88885	-0.16089	False

Table 2.6.5
b/b Analysis
English Language Arts Grade 7

Item Id	Old b	New b	Std Dist	Flag
IA00065 (EL292160)	-1.07830	-0.49410	0.28880	False
IA00066 (EL292163)	-0.81990	-0.03900	2.98005	False
IA00067 (EL292168)	0.24900	0.50380	-0.72789	False
IA00068 (EL292170)	-0.99950	-0.63780	0.08614	False
IA00069 (EL292172)	-0.55490	-0.35390	0.98455	False
IA00070 (EL292176)	-0.70340	-0.28570	-0.81208	False
IA00081 (EL293802)	-0.86670	-0.51780	-0.03542	False
IA00082 (EL293804)	-0.14160	0.17670	-0.79907	False
IA00257 (EL308358)	-1.85270	-1.22970	-0.81706	False
IA00258 (EL308360)	-1.79950	-1.17955	-0.74535	False
IA00262 (EL308382)	-0.75300	-0.38100	-0.51677	False
IA00265 (EL308389)	-2.79690	-1.95330	-0.24585	False
IA00269 (EL308397)	-2.38720	-1.63330	-0.42701	False
IA00655 (EL628647210)	-1.07690	-0.77780	0.93304	False
IA00657 (EL628647689)	-1.42330	-0.86010	-0.62796	False
IA00658 (EL628653398)	-1.70655	-1.25400	0.48189	False

Table 2.6.6
b/b Analysis
English Language Arts Grade 8

Item Id	Old b	New b	Std Dist	Flag
IA00056 (EL290795)	-1.91590	-1.38520	-0.38054	False
IA00057 (EL290798)	-1.97670	-1.35960	-0.76962	False
IA00058 (EL290799)	-0.71570	-0.44760	-0.45879	False
IA00059 (EL290800)	-0.49090	-0.35340	0.32782	False
IA00060 (EL290801)	-0.86570	-0.40120	0.19333	False
IA00061 (EL290805)	0.05680	0.06920	0.34393	False
IA00062 (EL290808)	0.15650	0.26900	-0.86242	False
IA00063 (EL290814)	0.77490	0.78060	-0.68233	False
IA00368 (EL623873883)	-0.91740	-0.58870	-0.60772	False
IA00371 (EL623951471)	-0.82355	-0.41210	-0.23527	False
IA00373 (EL623952377)	0.38275	0.39235	-0.35165	False
IA00374 (EL623952612)	-1.13630	-0.66400	-0.33024	False
IA00378 (EL623955555)	0.12560	-0.04470	1.98938	False
IA00379 (EL623955757)	0.03080	0.55220	2.74224	False
IA00383 (EL623959265)	-0.57360	-0.23020	-0.35047	False
IA00699 (EL632808123)	-1.31650	-0.82770	-0.56763	False

Table 2.6.7
b/b Analysis
Mathematics Grade 3

Item Id	Old b	New b	Std Dist	Flag
IA00769 (MA203641)	-1.96785	-1.92040	0.13183	False
IA00799 (MA260559)	0.37165	0.38390	-1.03904	False
IA00834 (MA293457)	-1.02068	-0.93390	-1.39267	False
IA00838 (MA293524)	-0.96025	-0.96950	0.48479	False
IA00850 (MA297405)	-0.76843	-0.72590	-0.73934	False
IA00852 (MA297438)	-0.46005	-0.39330	-1.35136	False
IA00924 (MA306310)	1.06135	1.09990	-0.69949	False
IA00925 (MA306315)	-0.95483	-0.86560	-1.28871	False
IA00930 (MA306359)	-0.70412	-0.50150	1.25461	False
IA00932 (MA306375)	0.57731	0.70770	0.80323	False
IA00993 (MA310834)	-0.32025	-0.17550	0.37167	False
IA01019 (MA311277)	-0.90789	-0.88930	-0.13220	False
IA01071 (MA623063509)	-1.11798	-0.97640	-0.34069	False
IA01080 (MA623654449)	0.90717	0.79200	1.15619	False
IA01081 (MA623656013)	0.82747	0.75467	0.34651	False
IA02323 (MA301611A)	-1.67972	-1.67600	0.80060	False
IA04760 (MA713752330)	-1.50904	-1.32490	0.22047	False
IA04813 (MA735572247)	-1.02722	-0.76520	2.21840	False
IA04828 (MA735653938)	-0.13916	-0.17710	0.41097	False
IA04844 (MA735735757)	-0.20812	-0.14470	-1.21577	False

Table 2.6.8
b/b Analysis
Mathematics Grade 4

Item Id	Old b	New b	Std Dist	Flag
IA00789 (MA250543)	0.39264	0.32140	0.19720	False
IA00828 (MA287237)	-1.00453	-0.84530	0.63616	False
IA00841 (MA293718)	-0.20587	-0.28400	-0.93557	False
IA00861 (MA297629)	-1.98288	-1.72590	0.14776	False
IA00869 (MA297988)	1.43941	1.18430	-0.27173	False
IA00906 (MA301811)	-0.47012	-0.71660	0.89852	False
IA00958 (MA307055)	0.36789	0.07010	0.07476	False
IA00961 (MA307081)	0.25488	-0.19080	2.12184	False
IA00963 (MA307085)	-0.38578	-0.48250	-1.12181	False
IA01048 (MA311534)	0.20108	0.04570	-1.19000	False
IA01049 (MA311537)	-0.23132	-0.38380	-0.69460	False
IA01055 (MA311572)	-0.06492	-0.11410	-0.32705	False
IA01057 (MA311581)	-0.24249	-0.32923	-1.10722	False
IA01093 (MA623879088)	-0.89879	-0.73315	0.90123	False
IA02175 (MA286769)	-0.69178	-0.81270	-0.28408	False
IA02819 (MA713583365)	-0.57544	-0.50380	0.29114	False
IA02841 (MA713774890)	0.22164	0.00960	-0.74185	False
IA02902 (MA714251321)	0.53170	0.45370	0.35584	False
IA04661 (MA307327)	-0.66913	-0.99140	2.19433	False
IA04965 (MA800867144)	-0.33270	-0.43500	-1.14485	False

Table 2.6.9
b/b Analysis
Mathematics Grade 5

Item Id	Old b	New b	Std Dist	Flag
IA00771 (MA204911)	-1.18544	-1.14990	-0.26895	False
IA00776 (MA221207)	-0.77091	-0.71900	-0.28627	False
IA00803 (MA262207)	-0.71683	-0.79320	-0.30460	False
IA00806 (MA272292)	0.16842	0.30560	0.08547	False
IA00826 (MA287178)	-1.44487	-1.40520	-0.13547	False
IA00872 (MA298003)	-0.59493	-0.43800	0.54304	False
IA00880 (MA298106)	1.06174	1.11020	-0.69602	False
IA00885 (MA299556)	-0.52914	-0.55660	-0.65079	False
IA00936 (MA306420)	-0.83569	-1.42050	3.98800	False
IA00943 (MA306466)	0.09940	0.20500	-0.15781	False
IA00989 (MA307638)	-1.53346	-1.47740	0.03791	False
IA01020 (MA311280)	0.55217	0.60080	-0.81529	False
IA01029 (MA311337)	-1.65213	-1.64470	-0.33203	False
IA01032 (MA311366)	0.31569	0.37018	-0.67579	False
IA01149 (MA624347774)	0.28042	0.39840	-0.12074	False
IA01155 (MA624357395)	0.45272	0.34610	0.39637	False
IA02552 (MA311324)	0.92804	0.90200	-0.11108	False
IA02736 (MA704359678)	-0.12435	-0.01200	-0.01549	False
IA04970 (MA800974344)	-0.89563	-0.86460	-0.41717	False
IA05002 (MA801652356)	-0.15010	-0.22960	-0.06328	False

Table 2.6.10
b/b Analysis
Mathematics Grade 6

Item Id	Old b	New b	Std Dist	Flag
IA00777 (MA221667)	-1.39100	-1.45010	0.93006	False
IA00778 (MA221669)	-1.07161	-0.99080	-0.72912	False
IA00804 (MA264305)	-1.31464	-1.26100	-0.33555	False
IA00817 (MA280989)	0.33230	0.44460	-0.41010	False
IA00818 (MA282268)	0.87654	0.74490	0.83104	False
IA00819 (MA282277)	0.06705	0.27390	0.52148	False
IA00827 (MA287186)	-0.15051	-0.03900	-0.60911	False
IA00845 (MA296349)	0.18948	0.44960	1.15353	False
IA00881 (MA298139)	0.08142	0.20220	-0.41608	False
IA00899 (MA301508)	0.58200	0.64890	-0.80919	False
IA00972 (MA307339)	-1.13141	-1.05963	-0.60664	False
IA00992 (MA309941)	1.75926	1.80310	-0.59777	False
IA01058 (MA311658)	0.62836	0.46490	1.27760	False
IA02037 (MA217493)	-0.50899	-0.07170	2.81975	False
IA02597 (MA311693)	-1.09414	-0.95220	-0.64765	False
IA02698 (MA703179529)	1.52821	1.39440	0.59791	False
IA04745 (MA703231515)	-0.07723	-0.08855	-0.11151	False
IA04884 (MA736365836)	-0.87050	-0.76920	-1.00485	False
IA05126 (MA805103779)	-0.29546	-0.22020	-0.97429	False
IA05135 (MA805171807)	-0.41273	-0.34190	-0.87951	False

Table 2.6.11
b/b Analysis
Mathematics Grade 7

Item Id	Old b	New b	Std Dist	Flag
IA00796 (MA259267)	0.36479	0.38380	0.00009	False
IA00831 (MA288414)	-0.57561	-0.55170	1.87147	False
IA00847 (MA296358)	-0.27118	-0.06520	0.03125	False
IA00909 (MA301846)	-1.70161	-1.36350	0.11234	False
IA00910 (MA301854)	0.35704	0.49630	-0.20844	False
IA00945 (MA306538)	-0.96347	-0.77320	-1.21566	False
IA00948 (MA306600)	-1.33959	-1.01400	0.58306	False
IA00949 (MA306605)	-0.15288	-0.02730	-1.40887	False
IA01006 (MA311093)	-1.70269	-1.49980	0.04948	False
IA01011 (MA311109)	0.77310	0.68930	1.55176	False
IA01016 (MA311125)	0.24681	0.31730	-0.95954	False
IA01017 (MA311135)	-0.80092	-0.66110	-0.37437	False
IA01018 (MA311140)	0.28610	0.27290	0.92266	False
IA01069 (MA316886)	-0.13840	0.01365	-0.95464	False
IA01097 (MA623950280)	0.44499	0.47540	-0.43702	False
IA01108 (MA624149677)	0.28427	0.43955	0.01403	False
IA02722 (MA703943185)	-0.10107	0.04570	-0.99977	False
IA04486 (MA227988)	-0.64182	-0.41750	-0.32101	False
IA04538 (MA282218)	-0.88898	-0.64870	-0.46833	False
IA04593 (MA298192)	1.06123	1.24020	2.21152	False

Table 2.6.12
b/b Analysis
Mathematics Grade 8

Item Id	Old b	New b	Std Dist	Flag
IA00849 (MA296757)	-0.67172	-0.42370	-0.64723	False
IA00858 (MA297513)	-0.92047	-0.80400	-0.18698	False
IA00864 (MA297652)	-0.93695	-0.67328	-0.69607	False
IA00865 (MA297656)	0.18595	0.02110	1.49715	False
IA00903 (MA301674)	-0.84896	-0.74230	-0.15130	False
IA00905 (MA301702)	0.35314	0.40760	-0.53651	False
IA00979 (MA307472)	-0.29387	0.13240	1.17505	False
IA00985 (MA307570)	0.46466	0.52340	-0.65198	False
IA01033 (MA311384)	-0.03049	0.16520	-0.65574	False
IA01037 (MA311414)	0.15802	0.31440	-0.86734	False
IA01042 (MA311448)	-0.10728	0.32520	1.35992	False
IA01044 (MA311463)	-0.73317	-0.47920	-0.63825	False
IA01066 (MA314812)	-0.36007	-0.13078	-0.59274	False
IA01125 (MA624247061)	0.18351	0.37075	-0.57980	False
IA02495 (MA309741)	0.02810	0.43390	1.22151	False
IA04665 (MA307399)	0.45471	0.61020	-0.66745	False
IA04678 (MA309738)	0.86892	1.00240	-0.56987	False
IA05057 (MA803856437)	-1.25271	-1.15710	0.22788	False
IA05059 (MA803856627)	-0.77502	-0.61670	-0.65453	False
IA05070 (MA804042487)	1.04291	0.68160	2.61427	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	0.78000	ON03	ON03	0.00181	False
IA00280 (EL308824)	2	0.66000	0.66000	ON03	ON03	0.01213	False
IA00281 (EL308826)	2	0.59000	0.56000	ON03	ON03	-0.02565	False
IA00282 (EL308827)	2	0.65000	0.61000	ON03	ON03	-0.03644	False
IA00283 (EL308835)	2	0.57000	0.57000	ON03	ON03	0.00748	False
IA00284 (EL308837)	2	0.66000	0.66000	ON03	ON03	0.00326	False
IA00285 (EL308838)	2	0.77000	0.77000	ON03	ON03	0.00105	False
IA00286 (EL308842)	2	0.41000	0.49000	ON03	ON03	0.09562	True
IA00287 (EL308855)	4	1.12000	1.00000	ON03	ON03	-0.02951	False
IA00288 (EL308857)	4	0.96000	0.88000	ON03	ON03	-0.01706	False
IA00443 (EL626042844)	2	0.66000	0.65000	ON04	ON04	-0.00156	False
IA00444 (EL626043062)	2	0.67000	0.67000	ON04	ON04	0.00933	False
IA00445 (EL626043435)	3	1.03000	1.05000	ON04	ON04	0.02025	False
IA00446 (EL626049849)	2	0.53000	0.51000	ON04	ON04	-0.01027	False
IA00450 (EL626050679)	2	0.67000	0.66000	ON04	ON04	0.00993	False
IA00451 (EL626050927)	2	0.51000	0.54000	ON04	ON04	0.03056	False
IA00452 (EL626051097)	3	1.10000	1.04000	ON04	ON04	-0.01435	False
IA00453 (EL626051328)	2	0.64000	0.65000	ON04	ON04	0.02038	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	0.83000	ON03	ON03	0.01254	False
IA00219 (EL307709)	2	0.77000	0.76000	ON03	ON03	-0.01018	False
IA00220 (EL307710)	2	0.43000	0.48000	ON03	ON03	0.04259	False
IA00221 (EL307713)	2	0.52000	0.53000	ON03	ON03	0.00436	False
IA00222 (EL307714)	2	0.76000	0.75000	ON03	ON03	-0.00578	False
IA00223 (EL307719)	2	0.56000	0.59000	ON03	ON03	0.01970	False
IA00224 (EL307724)	2	0.75000	0.73000	ON03	ON03	-0.01615	False
IA00225 (EL307728)	4	1.36000	1.44000	ON03	ON03	0.01917	False
IA00226 (EL307729)	4	1.25000	1.24000	ON03	ON03	-0.01178	False
IA00289 (EL309792)	2	0.64000	0.58000	ON03	ON03	-0.05748	True
IA00407 (EL624647403)	2	0.41000	0.42000	ON04	ON04	-0.00533	False
IA00408 (EL624647580)	2	0.58000	0.57000	ON04	ON04	-0.01509	False
IA00411 (EL624652450)	2	0.78000	0.76000	ON04	ON04	-0.02399	False
IA00412 (EL624652621)	2	0.90000	0.91000	ON04	ON04	0.01028	False
IA00414 (EL624652989)	2	0.43000	0.43000	ON04	ON04	-0.00387	False
IA00415 (EL624653348)	2	0.67000	0.67000	ON04	ON04	-0.01364	False
IA00416 (EL624653492)	3	1.44000	1.46000	ON04	ON04	0.00631	False
IA00419 (EL624654711)	3	1.60000	1.61000	ON04	ON04	0.00563	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	-0.01253	False
IA00497 (EL626304969)	2	0.73000	0.72000	ON03	ON03	-0.01012	False
IA00500 (EL626332335)	2	0.66000	0.66000	ON03	ON03	-0.00490	False
IA00501 (EL626332592)	2	0.84000	0.84000	ON03	ON03	-0.00450	False
IA00502 (EL626333002)	2	0.82000	0.83000	ON03	ON03	0.00188	False
IA00505 (EL626355215)	2	0.60000	0.62000	ON03	ON03	0.01512	False
IA00506 (EL626355557)	2	0.64000	0.67000	ON03	ON03	0.01867	False
IA00508 (EL626356291)	3	0.73000	0.78000	ON03	ON03	0.01949	False
IA00638 (EL627351056)	2	0.66000	0.68000	ON03	ON03	0.01137	False
IA01669 (EL711809263)	2	0.76000	0.77000	ON04	ON04	0.00565	False
IA01670 (EL711809592)	2	0.76000	0.75000	ON04	ON04	-0.01361	False
IA01671 (EL711827203)	2	0.90000	0.90000	ON04	ON04	0.00270	False
IA01672 (EL711827807)	2	0.73000	0.72000	ON04	ON04	-0.01497	False
IA01679 (EL711868011)	3	0.99000	0.97000	ON04	ON04	-0.01713	False
IA01680 (EL711900602)	2	0.70000	0.69000	ON04	ON04	-0.01796	False
IA01691 (EL712167015)	2	0.46000	0.49000	ON04	ON04	0.02173	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	0.75000	ON04	ON04	0.00197	False
IA00174 (EL303500)	2	0.64000	0.63000	ON04	ON04	-0.00988	False
IA00175 (EL303504)	2	0.61000	0.63000	ON04	ON04	0.03108	False
IA00176 (EL303508)	2	0.71000	0.73000	ON04	ON04	0.02470	False
IA00177 (EL303510)	2	0.87000	0.87000	ON04	ON04	-0.00571	False
IA00178 (EL303513)	2	0.66000	0.66000	ON04	ON04	-0.00245	False
IA00179 (EL303514)	2	0.63000	0.60000	ON04	ON04	-0.03293	False
IA00180 (EL303518)	2	0.57000	0.53000	ON04	ON04	-0.04121	False
IA00515 (EL626864414)	2	0.83000	0.82000	ON03	ON03	0.00010	False
IA00517 (EL626864724)	2	0.69000	0.71000	ON03	ON03	0.03048	False
IA00518 (EL626865003)	2	0.67000	0.67000	ON03	ON03	-0.00034	False
IA00520 (EL626865416)	2	0.41000	0.42000	ON03	ON03	0.00940	False
IA00522 (EL626865773)	2	0.72000	0.70000	ON03	ON03	-0.01567	False
IA00523 (EL626865942)	2	0.64000	0.64000	ON03	ON03	0.00973	False
IA00528 (EL626867605)	3	1.54000	1.52000	ON03	ON03	-0.00381	False
IA00530 (EL626868748)	3	1.41000	1.41000	ON03	ON03	0.00486	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	0.72051	ON04	ON04	-0.02028	False
IA00066 (EL292163)	2	0.65000	0.62829	ON04	ON04	-0.03191	False
IA00067 (EL292168)	2	0.50000	0.48588	ON04	ON04	-0.00818	False
IA00068 (EL292170)	2	0.71000	0.71518	ON04	ON04	0.00436	False
IA00069 (EL292172)	2	0.62000	0.65699	ON04	ON04	0.03377	False
IA00070 (EL292176)	2	0.62000	0.60853	ON04	ON04	-0.01409	False
IA00081 (EL293802)	2	0.68000	0.70373	ON04	ON04	0.02489	False
IA00082 (EL293804)	2	0.54000	0.56885	ON04	ON04	0.02950	False
IA00257 (EL308358)	2	0.85000	0.84836	ON03	ON03	-0.00959	False
IA00258 (EL308360)	3	1.51000	1.51705	ON03	ON03	-0.00391	False
IA00262 (EL308382)	2	0.65000	0.65515	ON03	ON03	0.00053	False
IA00265 (EL308389)	2	0.90000	0.89974	ON03	ON03	-0.00639	False
IA00269 (EL308397)	2	0.84000	0.82967	ON03	ON03	-0.01947	False
IA00655 (EL628647210)	2	0.72000	0.74311	ON03	ON03	0.01708	False
IA00657 (EL628647689)	2	0.77000	0.77721	ON03	ON03	-0.00660	False
IA00658 (EL628653398)	3	1.48000	1.52639	ON03	ON03	0.01185	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	0.79000	ON04	ON04	0.03591	False
IA00057 (EL290798)	2	0.78000	0.79000	ON04	ON04	0.01306	False
IA00058 (EL290799)	2	0.75000	0.76000	ON04	ON04	0.01272	False
IA00059 (EL290800)	2	0.67000	0.67000	ON04	ON04	0.00331	False
IA00060 (EL290801)	2	0.77000	0.76000	ON04	ON04	-0.01902	False
IA00061 (EL290805)	2	0.56000	0.56000	ON04	ON04	0.00570	False
IA00062 (EL290808)	2	0.54000	0.55000	ON04	ON04	0.00728	False
IA00063 (EL290814)	2	0.44000	0.40000	ON04	ON04	-0.05278	True
IA00368 (EL623873883)	2	0.72000	0.73000	ON03	ON03	0.01526	False
IA00371 (EL623951471)	3	1.21000	1.18000	ON03	ON03	-0.01466	False
IA00373 (EL623952377)	3	0.84000	0.85000	ON03	ON03	0.00548	False
IA00374 (EL623952612)	2	0.73000	0.72000	ON03	ON03	-0.01059	False
IA00378 (EL623955555)	2	0.51000	0.55000	ON03	ON03	0.03766	False
IA00379 (EL623955757)	2	0.54000	0.52000	ON03	ON03	-0.02420	False
IA00383 (EL623959265)	2	0.65000	0.64000	ON03	ON03	-0.00713	False
IA00699 (EL632808123)	2	0.77000	0.77000	ON03	ON03	0.00064	False

Table 2.7.7
Beta Analysis
Mathematics Grade 3

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00769 (MA203641)	2	0.83000	0.88000	ON19	ON19	0.03796	False
IA00799 (MA260559)	2	0.49000	0.49000	ON05	ON05	-0.00792	False
IA00834 (MA293457)	2	0.79000	0.81000	ON03	ON03	-0.01471	False
IA00838 (MA293524)	2	0.74000	0.77000	ON07	ON07	0.01914	False
IA00850 (MA297405)	2	0.69000	0.71000	ON09	ON09	0.01376	False
IA00852 (MA297438)	2	0.63000	0.65000	ON06	ON06	0.00331	False
IA00924 (MA306310)	2	0.46000	0.48000	ON14	ON14	0.01234	False
IA00925 (MA306315)	2	0.75000	0.79000	ON02	ON02	0.02920	False
IA00930 (MA306359)	2	0.66000	0.64000	ON04	ON04	-0.04767	False
IA00932 (MA306375)	2	0.48000	0.49000	ON13	ON13	0.00173	False
IA00993 (MA310834)	2	0.63000	0.62000	ON18	ON18	-0.03100	False
IA01019 (MA311277)	2	0.74000	0.74000	ON17	ON17	-0.00705	False
IA01071 (MA623063509)	2	0.73000	0.72000	ON16	ON16	-0.01847	False
IA01080 (MA623654449)	4	0.95000	0.98000	ON21	ON21	0.00039	False
IA01081 (MA623656013)	4	0.86000	0.96000	ON08	ON08	0.00980	False
IA02323 (MA301611A)	2	0.83000	0.87000	ON15	ON15	0.01100	False
IA04760 (MA713752330)	2	0.77000	0.80000	ON11	ON11	0.01359	False
IA04813 (MA735572247)	2	0.72000	0.69000	ON12	ON12	-0.03889	False
IA04828 (MA735653938)	2	0.51000	0.57000	ON10	ON10	0.02588	False
IA04844 (MA735735757)	2	0.62000	0.61000	ON20	ON20	-0.02290	False

Table 2.7.8
Beta Analysis
Mathematics Grade 4

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00789 (MA250543)	5	1.66000	1.71000	ON20	ON20	-0.00600	False
IA00828 (MA287237)	2	0.76000	0.79000	ON06	ON06	0.00209	False
IA00841 (MA293718)	2	0.71000	0.73000	ON08	ON08	0.00094	False
IA00861 (MA297629)	2	0.89000	0.90000	ON21	ON21	-0.00679	False
IA00869 (MA297988)	2	0.18000	0.21000	ON16	ON16	0.00972	False
IA00906 (MA301811)	2	0.73000	0.72000	ON09	ON09	-0.02302	False
IA00958 (MA307055)	2	0.46000	0.51000	ON10	ON10	0.02794	False
IA00961 (MA307081)	2	0.53000	0.61000	ON13	ON13	0.06627	True
IA00963 (MA307085)	2	0.67000	0.70000	ON15	ON15	0.00835	False
IA01048 (MA311534)	2	0.54000	0.56000	ON12	ON12	-0.00296	False
IA01049 (MA311537)	2	0.66000	0.67000	ON18	ON18	0.00484	False
IA01055 (MA311572)	2	0.52000	0.55000	ON19	ON19	-0.01331	False
IA01057 (MA311581)	5	2.21000	2.23000	ON11	ON11	0.00099	False
IA01093 (MA623879088)	3	1.46000	1.39000	ON14	ON14	-0.06007	True
IA02175 (MA286769)	2	0.75000	0.78000	ON03	ON03	0.01259	False
IA02819 (MA713583365)	2	0.61000	0.59000	ON17	ON17	-0.03460	False
IA02841 (MA713774890)	2	0.44000	0.50000	ON05	ON05	0.03494	False
IA02902 (MA714251321)	2	0.39000	0.41000	ON04	ON04	-0.00018	False
IA04661 (MA307327)	2	0.74000	0.79000	ON07	ON07	0.02060	False
IA04965 (MA800867144)	2	0.58000	0.62000	ON02	ON02	0.01522	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	0.73000	0.75000	ON08	ON08	0.00127	False
IA00776 (MA221207)	2	0.66000	0.69000	ON20	ON20	-0.00314	False
IA00803 (MA262207)	2	0.72000	0.75000	ON05	ON05	0.00079	False
IA00806 (MA272292)	2	0.47000	0.47000	ON23	ON23	-0.02252	False
IA00826 (MA287178)	2	0.85000	0.86000	ON06	ON06	-0.01229	False
IA00872 (MA298003)	2	0.63000	0.66000	ON07	ON07	0.00290	False
IA00880 (MA298106)	2	0.21000	0.24000	ON04	ON04	-0.00626	False
IA00885 (MA299556)	2	0.68000	0.73000	ON18	ON18	0.01118	False
IA00936 (MA306420)	2	0.69000	0.82000	ON09	ON09	0.09850	True
IA00943 (MA306466)	2	0.62000	0.64000	ON11	ON11	-0.00584	False
IA00989 (MA307638)	2	0.86000	0.87000	ON12	ON12	-0.00494	False
IA01020 (MA311280)	2	0.43000	0.43000	ON13	ON13	-0.02726	False
IA01029 (MA311337)	2	0.83000	0.86000	ON17	ON17	0.00483	False
IA01032 (MA311366)	5	1.54000	1.64000	ON19	ON19	-0.00854	False
IA01149 (MA624347774)	2	0.41000	0.42000	ON16	ON16	-0.01537	False
IA01155 (MA624357395)	3	0.70000	0.83000	ON03	ON03	0.02620	False
IA02552 (MA311324)	2	0.38000	0.44000	ON14	ON14	0.03076	False
IA02736 (MA704359678)	5	1.94000	1.99000	ON02	ON02	-0.02408	False
IA04970 (MA800974344)	2	0.70000	0.73000	ON10	ON10	0.00232	False
IA05002 (MA801652356)	2	0.60000	0.58000	ON19	ON15	0.03161	False

Table 2.7.10
Beta Analysis
Mathematics Grade 6

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00777 (MA221667)	2	0.85000	0.84000	ON10	ON10	-0.01390	False
IA00778 (MA221669)	2	0.80000	0.79000	ON11	ON11	-0.00538	False
IA00804 (MA264305)	2	0.74000	0.75000	ON17	ON17	-0.00487	False
IA00817 (MA280989)	2	0.47000	0.49000	ON14	ON14	0.00970	False
IA00818 (MA282268)	2	0.40000	0.41000	ON05	ON05	0.00829	False
IA00819 (MA282277)	2	0.47000	0.44000	ON03	ON03	-0.02477	False
IA00827 (MA287186)	2	0.56000	0.55000	ON09	ON09	-0.01653	False
IA00845 (MA296349)	2	0.56000	0.53000	ON16	ON16	-0.01955	False
IA00881 (MA298139)	5	1.80000	1.79000	ON03	ON03	-0.00455	False
IA00899 (MA301508)	2	0.33000	0.35000	ON06	ON06	0.00756	False
IA00972 (MA307339)	5	2.96000	2.92000	ON02	ON02	-0.00705	False
IA00992 (MA309941)	2	0.41000	0.40000	ON02	ON02	-0.00518	False
IA01058 (MA311658)	2	0.42000	0.46000	ON13	ON13	0.04330	False
IA02037 (MA217493)	2	0.66000	0.63000	ON03	ON03	-0.02956	False
IA02597 (MA311693)	2	0.78000	0.77000	ON04	ON04	0.00179	False
IA02698 (MA703179529)	2	0.28000	0.24000	06	ON15	0.02843	False
IA04745 (MA703231515)	3	1.00000	1.05000	ON03	ON03	0.02932	False
IA04884 (MA736365836)	2	0.68000	0.68000	ON12	ON12	0.00331	False
IA05126 (MA805103779)	2	0.55000	0.56000	ON07	ON07	0.01207	False
IA05135 (MA805171807)	2	0.58000	0.58000	ON08	ON08	0.01130	False

Table 2.7.11
Beta Analysis
Mathematics Grade 7

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00796 (MA259267)	2	0.44000	0.45000	ON08	ON08	0.00721	False
IA00831 (MA288414)	2	0.72000	0.73000	ON18	ON18	0.00873	False
IA00847 (MA296358)	2	0.53000	0.53000	ON17	ON17	-0.01687	False
IA00909 (MA301846)	2	0.87000	0.87000	ON13	ON13	0.00272	False
IA00910 (MA301854)	2	0.35000	0.35000	ON07	ON07	-0.01418	False
IA00945 (MA306538)	2	0.69000	0.71000	ON21	ON21	0.02100	False
IA00948 (MA306600)	2	0.84000	0.85000	ON11	ON11	0.00366	False
IA00949 (MA306605)	2	0.49000	0.50000	ON05	ON05	0.00588	False
IA01006 (MA311093)	2	0.85000	0.86000	ON20	ON20	0.00979	False
IA01011 (MA311109)	2	0.39000	0.41000	ON12	ON12	0.01785	False
IA01016 (MA311125)	2	0.55000	0.59000	ON19	ON19	0.02234	False
IA01017 (MA311135)	2	0.75000	0.73000	ON14	ON14	-0.01428	False
IA01018 (MA311140)	2	0.37000	0.42000	ON10	ON10	0.03499	False
IA01069 (MA316886)	5	2.03000	2.07000	ON03	ON03	-0.00073	False
IA01097 (MA623950280)	2	0.32000	0.35000	ON06	ON06	0.01919	False
IA01108 (MA624149677)	3	0.77000	0.77000	ON04	ON04	-0.01757	False
IA02722 (MA703943185)	5	1.95000	1.94000	ON02	ON02	-0.00334	False
IA04486 (MA227988)	2	0.69000	0.68000	ON09	ON09	-0.00233	False
IA04538 (MA282218)	2	0.69000	0.66000	ON16	ON16	-0.02638	False
IA04593 (MA298192)	2	0.24000	0.19000	0102	ON15	-0.03773	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	0.62000	ON17	ON17	-0.00824	False
IA00858 (MA297513)	2	0.76000	0.78000	ON04	ON04	0.02203	False
IA00864 (MA297652)	5	2.81000	2.77000	ON03	ON03	-0.00570	False
IA00865 (MA297656)	2	0.53000	0.60000	ON11	ON11	0.06613	True
IA00903 (MA301674)	2	0.73000	0.75000	ON13	ON13	0.00976	False
IA00905 (MA301702)	2	0.44000	0.46000	ON07	ON07	0.00912	False
IA00979 (MA307472)	2	0.63000	0.65000	ON13	ON13	0.01897	False
IA00985 (MA307570)	2	0.51000	0.53000	ON18	ON18	-0.00342	False
IA01033 (MA311384)	2	0.59000	0.57000	ON08	ON08	-0.01183	False
IA01037 (MA311414)	2	0.39000	0.39000	ON10	ON10	0.00026	False
IA01042 (MA311448)	2	0.52000	0.50000	ON05	ON05	-0.04553	False
IA01044 (MA311463)	2	0.64000	0.62000	ON15	ON15	-0.00589	False
IA01066 (MA314812)	5	2.22000	2.16000	ON02	ON02	-0.00571	False
IA01125 (MA624247061)	3	0.85000	0.83000	ON14	ON14	-0.00690	False
IA02495 (MA309741)	2	0.44000	0.37000	ON15	ON15	-0.06741	True
IA04665 (MA307399)	2	0.48000	0.47000	ON16	ON16	-0.03019	False
IA04678 (MA309738)	2	0.39000	0.40000	ON04	ON04	-0.00306	False
IA05057 (MA803856437)	2	0.83000	0.83000	ON12	ON12	-0.00165	False
IA05059 (MA803856627)	2	0.71000	0.71000	ON09	ON09	0.00568	False
IA05070 (MA804042487)	2	0.29000	0.40000	ON04	ON04	0.10769	True

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)
IA09406 (EL015503476)	1.06961	0.00000	-0.20230	0.00000	0.21540	0.00000
IA09410 (EL015607246)	1.09647	0.00000	0.17350	0.00000	0.25430	0.00000
IA09411 (EL015611981)	0.49324	0.00000	0.67290	0.00000	0.14920	0.00000
IA09413 (EL015628878)	1.19706	0.00000	0.01890	0.00000	0.23060	0.00000
IA09414 (EL015629502)	0.82634	0.00000	-1.05020	0.00000	0.22350	0.00000
IA09415 (EL015630515)	0.34874	0.00000	1.16130	0.00000	0.14010	0.00000
IA09540 (EL028131763)	1.05173	0.00000	-0.63480	0.00000	0.21450	0.00000
IA09543 (EL028160220)	0.78989	0.00000	0.15690	0.00000	0.23380	0.00000
IA09548 (EL028217140)	0.58242	0.00000	-0.62910	0.00000	0.06020	0.00000
IA09549 (EL028217627)	0.88477	0.00000	-0.43430	0.00000	0.26350	0.00000
IA09602 (EL028709466)	0.69859	0.00000	0.45830	0.00000	0.24870	0.00000
IA09603 (EL028717847)	0.66814	0.00000	-0.57490	0.00000	0.08020	0.00000
IA09605 (EL028823014)	0.82704	0.00000	-0.22520	0.00000	0.21730	0.00000
IA09611 (EL028907379)	1.18031	0.00000	-0.29150	0.00000	0.27800	0.00000
IA09614 (EL028914863)	0.63892	0.00000	0.23000	0.00000	0.21440	0.00000
IA09826 (EL909469479)	1.16208	0.00000	-1.06660	0.00000	0.10130	0.00000
IA09827 (EL909470939)	0.85850	0.00000	0.33710	0.00000	0.23580	0.00000
IA09828 (EL909472828)	0.83827	0.00000	-0.74870	0.00000	0.15680	0.00000
IA09829 (EL909473433)	0.82263	0.00000	-0.47020	0.00000	0.10470	0.00000
IA09831 (EL909478450)	1.22122	0.00000	-1.52500	0.00000	0.11040	0.00000
IA09832 (EL909479176)	0.83445	0.00000	-1.52760	0.00000	0.20110	0.00000
IA09833 (EL909480023)	0.56531	0.00000	-0.79240	0.00000	0.00650	0.00000
IA09835 (EL909865165)	1.13686	0.00000	-0.87310	0.00000	0.23530	0.00000
IA09836 (EL909869416)	0.99882	0.00000	-1.23690	0.00000	0.00620	0.00000
IA09846 (EL912636232)	1.03239	0.00000	-1.08140	0.00000	0.08390	0.00000
IA09847 (EL919652746)	1.35926	0.00000	-0.36560	0.00000	0.13980	0.00000

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)
IA09439 (EL019650296)	0.60135	0.00000	-1.66395	0.00000	0.89665	0.00000	-0.89665	0.00000	0.00000	0.00000
IA09547 (EL028215856)	0.76508	0.00000	-0.29635	0.00000	0.54835	0.00000	-0.54835	0.00000	0.00000	0.00000
IA09604 (EL028753268)	0.88777	0.00000	-0.94790	0.00000	0.86670	0.00000	-0.86670	0.00000	0.00000	0.00000
IA09608 (EL028832702)	0.71470	0.00000	1.00717	0.00000	2.11047	0.00000	0.28657	0.00000	-2.39703	0.00000
IA09838A (EL909882556#SCORE_TRAIT_Conv)	0.77984	0.00882	0.65813	0.01091	1.81453	0.01713	0.00783	0.01629	-1.82237	0.02949
IA09838D (EL909882556#SCORE_TRAIT_Ideadev)	0.82334	0.00911	1.46605	0.01717	2.75175	0.02185	0.82665	0.02067	-0.89345	0.03050
IA09840 (EL911945550)	1.00394	0.00000	-1.07355	0.00000	0.49495	0.00000	-0.49495	0.00000	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA09608 (EL028832702)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09838A (EL909882556#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09838D (EL909882556#SCORE_TRAIT_Ideadev)	-2.68495	0.06376	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)
IA08906 (EL006458075)	0.90941	0.00000	-2.21740	0.00000	0.04800	0.00000
IA08926 (EL006549511)	1.04803	0.00000	-0.41950	0.00000	0.17380	0.00000
IA09165 (EL007440160)	0.60517	0.00000	-1.87480	0.00000	0.01720	0.00000
IA09168 (EL007444742)	1.08348	0.00000	-1.17900	0.00000	0.12930	0.00000
IA09170 (EL007446608)	0.62816	0.00000	-1.41940	0.00000	0.02740	0.00000
IA09172 (EL007452066)	1.17972	0.00000	-1.57190	0.00000	0.19250	0.00000
IA09235 (EL009344832)	0.72575	0.00000	-0.17090	0.00000	0.23140	0.00000
IA09359 (EL013314332)	0.66384	0.00000	0.18260	0.00000	0.20750	0.00000
IA09370 (EL014208236)	1.01793	0.00000	-0.94110	0.00000	0.15010	0.00000
IA09446 (EL024031609)	0.39018	0.00000	0.20110	0.00000	0.14970	0.00000
IA09447 (EL024132276)	1.13804	0.00000	-0.97200	0.00000	0.18360	0.00000
IA09449 (EL024134327)	1.14874	0.00000	-0.49270	0.00000	0.13330	0.00000
IA09453 (EL024148759)	0.67784	0.00000	-2.04860	0.00000	0.03350	0.00000
IA09454 (EL024437543)	0.52728	0.00000	0.02380	0.00000	0.10540	0.00000
IA09455 (EL024440140)	0.47713	0.00000	-1.34260	0.00000	0.02990	0.00000
IA09456 (EL024442344)	0.73422	0.00000	0.71270	0.00000	0.19320	0.00000
IA09459 (EL024455037)	0.35720	0.00000	-1.37420	0.00000	0.04180	0.00000
IA09462 (EL024527106)	0.38372	0.00000	-0.52480	0.00000	0.11310	0.00000
IA09463 (EL024532504)	0.67913	0.00000	1.24720	0.00000	0.22750	0.00000
IA09618 (EL029280667)	0.96143	0.00000	-0.07630	0.00000	0.21080	0.00000
IA09623 (EL029415206)	1.20071	0.00000	-0.13550	0.00000	0.31820	0.00000
IA09624 (EL029417376)	1.35438	0.00000	-0.31450	0.00000	0.29870	0.00000
IA09625 (EL029429384)	0.77484	0.00000	0.56180	0.00000	0.25430	0.00000
IA09647 (EL030430678)	1.23627	0.00000	-0.02300	0.00000	0.28440	0.00000
IA09655 (EL030741768)	1.18107	0.00000	-0.15410	0.00000	0.25440	0.00000
IA09673 (EL033943069)	0.96755	0.00000	-0.71200	0.00000	0.24670	0.00000

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)
IA09177A (EL007459900#SCORE_TRAIT_Conv)	0.91182	0.01058	0.08870	0.01131	2.45650	0.02315	-0.28880	0.01503	-2.16770	0.02780
IA09177D (EL007459900#SCORE_TRAIT_Ideadev)	0.82387	0.00964	1.32518	0.01872	3.17928	0.02529	0.84528	0.02187	-1.21303	0.03519
IA09178 (EL007464016)	0.42234	0.00000	-1.74475	0.00000	0.17395	0.00000	-0.17395	0.00000	0.00000	0.00000
IA09234 (EL009343264)	0.87190	0.00000	-0.66810	0.00000	0.47850	0.00000	-0.47850	0.00000	0.00000	0.00000
IA09460 (EL024480931)	0.66655	0.00000	-0.46665	0.00000	0.41175	0.00000	-0.41175	0.00000	0.00000	0.00000
IA09465 (EL024539092)	0.70323	0.00000	-0.31373	0.00000	2.36107	0.00000	0.10907	0.00000	-2.47013	0.00000
IA09619 (EL029323184)	0.64027	0.00000	-0.13125	0.00000	1.07315	0.00000	-1.07315	0.00000	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA09177A (EL007459900#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09177D (EL007459900#SCORE_TRAIT_Ideadev)	-2.81153	0.06822	0.00000	0.00000	n/a	n/a
IA09465 (EL024539092)	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)
IA01339 (EL624175088)	0.87454	0.00000	-1.48780	0.00000	0.09500	0.00000
IA01341 (EL624176741)	0.42287	0.00000	-0.00190	0.00000	0.04620	0.00000
IA01342 (EL624177026)	0.26185	0.00000	-3.09710	0.00000	0.03400	0.00000
IA01344 (EL624178677)	0.47554	0.00000	-0.77190	0.00000	0.19040	0.00000
IA01345 (EL624179162)	0.85603	0.00000	-1.69160	0.00000	0.07410	0.00000
IA01346 (EL624179855)	0.57066	0.00000	-1.29120	0.00000	0.06040	0.00000
IA01348 (EL624180347)	0.65326	0.00000	-0.85290	0.00000	0.31190	0.00000
IA01349 (EL624180539)	0.87084	0.00000	-0.97360	0.00000	0.32450	0.00000
IA01387 (EL627148548)	0.98742	0.00000	-0.64460	0.00000	0.22670	0.00000
IA09629 (EL029961019)	0.79553	0.00000	-0.93320	0.00000	0.21550	0.00000
IA09632 (EL029974201)	1.09359	0.00000	-0.76900	0.00000	0.14060	0.00000
IA09634 (EL029980757)	0.73398	0.00000	-0.14600	0.00000	0.19580	0.00000
IA09637 (EL030062229)	0.92357	0.00000	-1.08780	0.00000	0.20340	0.00000
IA09638 (EL030080040)	0.70364	0.00000	-1.32730	0.00000	0.04630	0.00000
IA09639 (EL030171711)	0.72569	0.00000	-0.92760	0.00000	0.10760	0.00000
IA09643 (EL030337822)	0.64327	0.00000	1.32270	0.00000	0.21250	0.00000
IA09650 (EL030483081)	0.95026	0.00000	-1.38560	0.00000	0.21460	0.00000
IA09652 (EL030659380)	0.89430	0.00000	0.82610	0.00000	0.18520	0.00000
IA09659 (EL033541180)	0.82610	0.00000	-1.98660	0.00000	0.10900	0.00000
IA09660 (EL033604260)	1.17925	0.00000	-1.15570	0.00000	0.22780	0.00000
IA09663 (EL033665287)	0.43316	0.00000	-0.27560	0.00000	0.13480	0.00000
IA09664 (EL033679189)	0.73780	0.00000	-0.63230	0.00000	0.05290	0.00000
IA09668 (EL033800505)	0.28954	0.00000	-0.27400	0.00000	0.01230	0.00000
IA09671 (EL033843854)	1.12651	0.00000	0.66240	0.00000	0.27850	0.00000

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)
IA01340 (EL624176168)	1.13516	0.00000	-1.84640	0.00000	0.28310	0.00000	-0.28310	0.00000	0.00000	0.00000
IA01343 (EL624177447)	0.60905	0.00000	-1.32425	0.00000	0.92935	0.00000	-0.92935	0.00000	0.00000	0.00000
IA01350A (EL624182427#SCORE_TRAIT_Conv)	0.84762	0.00899	-0.30340	0.00893	1.84900	0.02036	-0.04490	0.01319	-1.80410	0.01920
IA01350D (EL624182427#SCORE_TRAIT_Ideadev)	0.82257	0.00876	0.51043	0.01243	2.76723	0.02313	0.89883	0.01596	-0.77548	0.02003
IA09630 (EL029964018)	0.58795	0.00000	-0.79505	0.00000	0.39235	0.00000	-0.39235	0.00000	0.00000	0.00000
IA09645A (EL030400392#SCORE_TRAIT_Conv)	0.85814	0.00935	0.48343	0.00941	1.72723	0.01605	-0.10457	0.01447	-1.62267	0.02439
IA09645D (EL030400392#SCORE_TRAIT_Ideadev)	0.83633	0.00917	1.25553	0.01502	2.70733	0.02081	0.66523	0.01879	-0.80828	0.02705
IA09649 (EL030463527)	0.81346	0.00000	-0.84670	0.00000	0.42830	0.00000	-0.42830	0.00000	0.00000	0.00000
IA09661 (EL033646585)	0.98383	0.00000	-0.92450	0.00000	0.62430	0.00000	-0.62430	0.00000	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA01350A (EL624182427#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA01350D (EL624182427#SCORE_TRAIT_Ideadev)	-2.89058	0.04361	0.00000	0.00000	n/a	n/a
IA09645A (EL030400392#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09645D (EL030400392#SCORE_TRAIT_Ideadev)	-2.56428	0.05470	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)
IA03973 (EL806979864)	0.62034	0.00000	-1.52900	0.00000	0.08280	0.00000
IA03974 (EL807001596)	0.52205	0.00000	-0.38080	0.00000	0.27790	0.00000
IA03975 (EL807002174)	0.34139	0.00000	-2.02870	0.00000	0.02940	0.00000
IA03977 (EL807009150)	0.50370	0.00000	-0.87910	0.00000	0.33410	0.00000
IA03979 (EL807010236)	0.76919	0.00000	0.69010	0.00000	0.22430	0.00000
IA03981 (EL807011414)	0.37025	0.00000	-0.72710	0.00000	0.10170	0.00000
IA03982 (EL807011890)	0.54897	0.00000	-0.67320	0.00000	0.24070	0.00000
IA04006 (EL807061702)	0.19453	0.00000	-2.42500	0.00000	0.04900	0.00000
IA04141 (EL808245411)	0.70994	0.00000	-0.53070	0.00000	0.37100	0.00000
IA08942 (EL006639933)	0.66455	0.00000	0.15750	0.00000	0.21240	0.00000
IA08960 (EL006738734)	0.41922	0.00000	-0.01390	0.00000	0.05610	0.00000
IA08964 (EL006742548)	0.43186	0.00000	1.16040	0.00000	0.07560	0.00000
IA09078 (EL007074213)	0.62381	0.00000	-1.99180	0.00000	0.15380	0.00000
IA09079 (EL007074445)	0.50312	0.00000	-1.35130	0.00000	0.05370	0.00000
IA09081 (EL007075911)	0.75262	0.00000	-1.64430	0.00000	0.02950	0.00000
IA09082 (EL007076177)	0.46655	0.00000	-1.35100	0.00000	0.21310	0.00000
IA09084 (EL007077078)	0.91834	0.00000	-1.26660	0.00000	0.27560	0.00000
IA09086 (EL007077860)	0.58430	0.00000	-0.31260	0.00000	0.20680	0.00000
IA09087 (EL007078526)	0.58801	0.00000	-1.32100	0.00000	0.12630	0.00000
IA09186 (EL008281454)	0.26008	0.00000	0.07310	0.00000	0.15830	0.00000
IA09190 (EL008445593)	0.72169	0.00000	-1.00420	0.00000	0.15020	0.00000
IA09239 (EL009514238)	0.59012	0.00000	-0.67510	0.00000	0.14980	0.00000
IA09242 (EL009564267)	0.35750	0.00000	-0.35830	0.00000	0.02340	0.00000
IA09254 (EL009978066)	0.64539	0.00000	-0.91640	0.00000	0.14300	0.00000

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)
IA03984A (EL807016586#SCORE_TRAIT_Conv)	0.77084	0.00794	-0.43287	0.00887	1.83833	0.02025	-0.11027	0.01413	-1.72807	0.01829
IA03984D (EL807016586#SCORE_TRAIT_Ideadev)	0.79976	0.00794	0.77488	0.01041	3.20388	0.02105	1.26028	0.01507	-0.20432	0.01713
IA04007 (EL807062301)	0.46302	0.00000	-0.18550	0.00000	0.22650	0.00000	-0.22650	0.00000	0.00000	0.00000
IA04142 (EL808246461)	0.27901	0.00000	-1.86625	0.00000	3.25905	0.00000	-3.25905	0.00000	0.00000	0.00000
IA09048A (EL007051004#SCORE_TRAIT_Conv)	0.73486	0.00770	-0.57440	0.00922	1.78280	0.02112	0.02920	0.01482	-1.81200	0.01881
IA09048D (EL007051004#SCORE_TRAIT_Ideadev)	0.76955	0.00776	1.47268	0.01865	3.34588	0.02425	1.75298	0.02170	0.05118	0.02503
IA09185 (EL008181021)	0.57331	0.00000	-1.69175	0.00000	1.53755	0.00000	-1.53755	0.00000	0.00000	0.00000
IA09188 (EL008355)	0.77266	0.00000	-1.42415	0.00000	0.90245	0.00000	-0.90245	0.00000	0.00000	0.00000
IA09237 (EL009438210)	0.31440	0.00000	-0.19545	0.00000	1.88105	0.00000	-1.88105	0.00000	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA03984A (EL807016586#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA03984D (EL807016586#SCORE_TRAIT_Ideadev)	-1.52582	0.02534	-2.73402	0.04056	0.00000	0.00000
IA09048A (EL007051004#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09048D (EL007051004#SCORE_TRAIT_Ideadev)	-1.56992	0.03730	-3.58012	0.08575	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)
IA01781 (EL713378067)	0.41881	0.00000	-1.95130	0.00000	0.01160	0.00000
IA01791 (EL713476504)	0.57872	0.00000	0.49710	0.00000	0.20410	0.00000
IA01793 (EL713479099)	0.87790	0.00000	-0.29150	0.00000	0.23630	0.00000
IA01794 (EL713479631)	0.85397	0.00000	-1.07190	0.00000	0.27300	0.00000
IA01795 (EL713480064)	0.60982	0.00000	-0.40250	0.00000	0.17530	0.00000
IA01797 (EL713480958)	0.58207	0.00000	-1.31520	0.00000	0.07040	0.00000
IA01798 (EL713481518)	0.89847	0.00000	-0.46660	0.00000	0.21040	0.00000
IA01964 (EL723632935)	0.59248	0.00000	-0.18310	0.00000	0.20920	0.00000
IA01973 (EL730170770)	1.01640	0.00000	-1.25480	0.00000	0.21290	0.00000
IA08887 (EL006357067)	0.86114	0.00000	-0.85190	0.00000	0.24240	0.00000
IA08897 (EL006439224)	0.40988	0.00000	0.10130	0.00000	0.15890	0.00000
IA08903 (EL006446884)	0.66226	0.00000	0.17790	0.00000	0.17550	0.00000
IA08905 (EL006454205)	0.81287	0.00000	-1.54300	0.00000	0.08150	0.00000
IA08913 (EL006537445)	0.41952	0.00000	-0.33260	0.00000	0.06070	0.00000
IA08919 (EL006544709)	0.90958	0.00000	0.88220	0.00000	0.21010	0.00000
IA08922 (EL006545529)	0.49124	0.00000	-0.39990	0.00000	0.13030	0.00000
IA08923 (EL006546235)	0.55750	0.00000	-0.66960	0.00000	0.19680	0.00000
IA08928 (EL006560393)	0.79242	0.00000	0.22860	0.00000	0.29540	0.00000
IA08943 (EL006640130)	0.88101	0.00000	-0.07710	0.00000	0.20670	0.00000
IA09023 (EL006977006)	0.84815	0.00000	-0.26420	0.00000	0.29910	0.00000
IA09298 (EL011362473)	0.62845	0.00000	-2.24710	0.00000	0.02570	0.00000
IA09299 (EL011363661)	0.69747	0.00000	-0.99970	0.00000	0.21660	0.00000
IA09434 (EL016833358)	0.83286	0.00000	-0.26690	0.00000	0.20860	0.00000
IA09818 (EL113429887)	0.90805	0.00000	-0.96150	0.00000	0.22570	0.00000

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)
IA01776 (EL713370326)	0.45426	0.00000	-1.41715	0.00000	1.29335	0.00000	-1.29335	0.00000	0.00000	0.00000
IA01779A (EL713375305#SCORE_TRAIT_Conv)	0.98301	0.00999	-0.53300	0.00649	1.30260	0.01525	-0.20460	0.01110	-1.09800	0.01226
IA01779D (EL713375305#SCORE_TRAIT_Ideadev)	0.85726	0.00952	0.81368	0.00715	1.19008	0.01189	0.83168	0.01245	0.23998	0.01362
IA01788 (EL713475622)	0.66725	0.00000	-0.30075	0.00000	0.59135	0.00000	-0.59135	0.00000	0.00000	0.00000
IA08948A (EL006653237#SCORE_TRAIT_Conv)	0.86138	0.00888	-0.55247	0.00742	1.35863	0.01728	0.02813	0.01229	-1.38677	0.01454
IA08948D (EL006653237#SCORE_TRAIT_Ideadev)	0.85461	0.00841	0.72594	0.01147	3.09444	0.02168	1.37284	0.01529	-0.03846	0.01653
IA09028 (EL006978834)	0.49730	0.00000	-0.47815	0.00000	0.50925	0.00000	-0.50925	0.00000	0.00000	0.00000
IA09297 (EL011353608)	0.46814	0.00000	0.28525	0.00000	1.23695	0.00000	-1.23695	0.00000	0.00000	0.00000
IA09437 (EL017655451)	0.84697	0.00000	-1.33035	0.00000	0.33205	0.00000	-0.33205	0.00000	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA01779A (EL713375305#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA01779D (EL713375305#SCORE_TRAIT_Ideadev)	-0.59402	0.01752	-1.66772	0.02725	0.00000	0.00000
IA08948A (EL006653237#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA08948D (EL006653237#SCORE_TRAIT_Ideadev)	-1.39926	0.02437	-3.02956	0.04877	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)
IA08951 (EL006655733)	0.79571	0.00000	-0.04280	0.00000	0.14300	0.00000
IA09049 (EL007061131)	0.36673	0.00000	-2.31380	0.00000	0.01520	0.00000
IA09050 (EL007061194)	0.50870	0.00000	-1.40660	0.00000	0.04270	0.00000
IA09054 (EL007061650)	0.29359	0.00000	-0.61170	0.00000	0.41620	0.00000
IA09058 (EL007062053)	0.69588	0.00000	-0.19040	0.00000	0.24910	0.00000
IA09127 (EL007253045)	0.95773	0.00000	-1.07280	0.00000	0.17980	0.00000
IA09131 (EL007256618)	0.54386	0.00000	-1.20140	0.00000	0.13300	0.00000
IA09133 (EL007257202)	0.18436	0.00000	-2.72570	0.00000	0.04970	0.00000
IA09134 (EL007257390)	0.49859	0.00000	-0.69550	0.00000	0.16760	0.00000
IA09138 (EL007335795)	0.41534	0.00000	-1.45940	0.00000	0.00990	0.00000
IA09139 (EL007335808)	0.89459	0.00000	0.07270	0.00000	0.24000	0.00000
IA09141 (EL007350397)	0.64256	0.00000	0.05730	0.00000	0.20310	0.00000
IA09145 (EL007353056)	0.86978	0.00000	0.37970	0.00000	0.22700	0.00000
IA09194 (EL008544460)	0.51411	0.00000	-0.70520	0.00000	0.09560	0.00000
IA09218 (EL009149967)	0.44844	0.00000	-2.24180	0.00000	0.01140	0.00000
IA09222 (EL009246409)	0.81511	0.00000	-1.17170	0.00000	0.23340	0.00000
IA09226 (EL009308236)	1.03780	0.00000	-1.71090	0.00000	0.07350	0.00000
IA09227 (EL009308819)	0.32934	0.00000	-0.50140	0.00000	0.03150	0.00000
IA09233 (EL009343097)	0.40664	0.00000	-3.25630	0.00000	0.01840	0.00000
IA09245 (EL009737508)	0.88083	0.00000	0.14520	0.00000	0.28790	0.00000
IA09426 (EL016259168)	0.82916	0.00000	-1.46820	0.00000	0.12560	0.00000
IA09427 (EL016259978)	0.42404	0.00000	-0.15520	0.00000	0.23020	0.00000
IA09430 (EL016352526)	0.16514	0.00000	-1.30970	0.00000	0.02290	0.00000
IA09443 (EL022460231)	0.90165	0.00000	-0.90080	0.00000	0.20030	0.00000

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)
IA08949 (EL006653570)	0.31117	0.00000	-1.60220	0.00000	2.46880	0.00000	-2.46880	0.00000	0.00000	0.00000
IA09059 (EL007062608)	0.37901	0.00000	-2.24165	0.00000	0.30075	0.00000	-0.30075	0.00000	0.00000	0.00000
IA09060A (EL007062902#SCORE_TRAIT_Conv)	0.78912	0.00811	-1.19327	0.00950	1.66493	0.02489	-0.06457	0.01562	-1.60037	0.01484
IA09060D (EL007062902#SCORE_TRAIT_Ideadev)	0.76796	0.00747	0.30384	0.00932	3.07344	0.02426	1.39484	0.01567	-0.04396	0.01449
IA09122 (EL007243506)	0.67014	0.00000	-0.64940	0.00000	0.33070	0.00000	-0.33070	0.00000	0.00000	0.00000
IA09130A (EL007253494#SCORE_TRAIT_Conv)	0.88618	0.00894	-0.74167	0.00764	1.43813	0.01837	0.02393	0.01315	-1.46207	0.01381
IA09130D (EL007253494#SCORE_TRAIT_Ideadev)	0.87701	0.00847	0.70690	0.00959	2.76460	0.01865	1.44720	0.01437	0.03640	0.01467
IA09197 (EL008553781)	0.38372	0.00000	-2.34500	0.00000	2.06490	0.00000	-2.06490	0.00000	0.00000	0.00000
IA09224 (EL009257746)	0.42099	0.00000	-1.12025	0.00000	1.77075	0.00000	-1.77075	0.00000	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA09060A (EL007062902#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09060D (EL007062902#SCORE_TRAIT_Ideadev)	-1.51256	0.02045	-2.91176	0.03381	0.00000	0.00000
IA09130A (EL007253494#SCORE_TRAIT_Conv)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA09130D (EL007253494#SCORE_TRAIT_Ideadev)	-1.34850	0.02110	-2.89970	0.03929	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)
IA00793 (MA253641)	1.12128	0.02082	-0.79544	0.02332	0.21380	0.01140
IA00801 (MA260962)	0.60921	0.01475	-0.51911	0.04804	0.10020	0.01780
IA02033 (MA212474)	1.14535	0.02044	-1.01840	0.02353	0.14340	0.01280
IA02052 (MA227232)	1.16909	0.01952	-0.45936	0.01692	0.17760	0.00790
IA02056 (MA227485)	0.69692	0.01583	0.54518	0.02050	0.08730	0.00740
IA02515 (MA310880)	0.89842	0.02689	1.12776	0.01800	0.24630	0.00510
IA04473 (MA207001)	0.89782	0.01578	-0.94705	0.02885	0.09690	0.01430
IA04681 (MA310870)	1.08609	0.02694	0.41678	0.01865	0.34330	0.00610
IA07601 (MA306285)	0.84340	0.01653	-0.17707	0.02353	0.15510	0.00940
IA07801 (MA900372676)	1.09520	0.02163	0.19132	0.01572	0.22380	0.00600
IA09906 (MA001049099)	0.66375	0.02136	0.78854	0.02841	0.24390	0.00880
IA09912 (MA001056175)	0.90953	0.01556	0.12192	0.01518	0.06990	0.00600
IA09920 (MA001137862)	0.76355	0.01306	-0.87916	0.03058	0.01750	0.01440
IA10018 (MA001633319)	0.70440	0.01480	-0.36002	0.03221	0.09270	0.01280
IA10334 (MA297500)	1.11792	0.02504	0.33740	0.01594	0.25850	0.00590
IA12370 (MA202994)	0.84903	0.01735	-1.02296	0.03861	0.17800	0.01780
IA02516 (MA310884)	0.99486	0.01019	-0.54406	0.00878	0.00000	0.00000
IA02521 (MA310889)	0.85662	0.00987	-1.12750	0.01247	0.00000	0.00000
IA02682 (MA703056978)	0.87815	0.00938	-0.36827	0.00933	0.00000	0.00000
IA04754 (MA713536927)	0.72153	0.00851	-1.09584	0.01323	0.00000	0.00000
IA04834 (MA735663821)	0.88433	0.00900	0.08201	0.00889	0.00000	0.00000
IA04851 (MA735756531)	0.93333	0.00992	0.30671	0.00878	0.00000	0.00000
IA04860 (MA735954511)	0.97046	0.01014	-0.06429	0.00846	0.00000	0.00000
IA07645 (MA309747)	1.20021	0.01225	-0.33519	0.00792	0.00000	0.00000
IA07814 (MA900425126)	0.71367	0.00791	-0.00996	0.00976	0.00000	0.00000
IA09854 (MA000749172)	0.97556	0.01025	-0.65055	0.00933	0.00000	0.00000
IA09894 (MA001038775)	0.70522	0.00949	-1.41825	0.01605	0.00000	0.00000
IA09904 (MA001047582)	0.84248	0.00965	-1.35589	0.01366	0.00000	0.00000
IA09961 (MA001335228)	1.02841	0.01041	-0.20234	0.00824	0.00000	0.00000
IA09964 (MA001338241)	0.69540	0.00813	-1.31902	0.01507	0.00000	0.00000
IA09967 (MA001344527)	0.31317	0.00645	-2.20937	0.04414	0.00000	0.00000
IA09977 (MA001439533)	0.72766	0.00857	-1.20168	0.01377	0.00000	0.00000
IA10392 (MA703078093)	1.14421	0.01393	-1.51054	0.01247	0.00000	0.00000
IA10405 (MA734752477)	0.78138	0.00965	-1.15472	0.01301	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	a	SE(a)	b	SE(b)	c	SE(c)
IA10407 (MA735736004A)	0.81748	0.00878	-0.36827	0.00943	0.00000	0.00000
IA10480 (MA935136577)	0.71031	0.00829	-0.93132	0.01247	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)
IA02202 (MA293460A)	1.07563	0.00884	-0.24944	0.00518	0.84075	0.01124	0.01851	0.00955	-0.85926	0.01020
IA04548 (MA286750A)	1.07064	0.00867	-0.21535	0.00567	1.12156	0.01262	0.07483	0.00974	-1.19639	0.01148
IA04686 (MA310899A)	1.07270	0.00884	-0.02196	0.00633	1.41093	0.01303	0.22886	0.01022	-1.63979	0.01437
IA07524 (MA253711A)	1.14822	0.00916	0.10912	0.00567	1.22296	0.01129	0.09836	0.00956	-1.32132	0.01291

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA02202 (MA293460A)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA04548 (MA286750A)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA04686 (MA310899A)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA07524 (MA253711A)	0.00000	0.00000	n/a	n/a	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)
IA00787 (MA247691)	0.99750	0.02482	-1.35704	0.05133	0.28770	0.02610
IA00812 (MA279759)	0.72728	0.01091	-0.98136	0.03012	0.00700	0.01440
IA00960 (MA307075)	1.29635	0.02931	-1.47885	0.03715	0.26380	0.02390
IA02072 (MA247705)	0.70575	0.01539	0.02727	0.02929	0.11320	0.01140
IA02432 (MA307060)	0.81618	0.02153	0.25051	0.02867	0.29240	0.00980
IA04580 (MA297614)	0.47756	0.01165	-1.08599	0.08228	0.03430	0.02990
IA04653 (MA307067)	1.36202	0.02409	-0.67171	0.01677	0.13680	0.00970
IA05045 (MA803747806)	1.24988	0.02278	-0.23425	0.01646	0.20590	0.00800
IA07659 (MA311543)	1.02232	0.02005	-0.37086	0.02401	0.24370	0.01070
IA08145 (MA903571693)	1.19813	0.01625	0.04984	0.01025	0.04450	0.00420
IA08183 (MA903776098)	1.44802	0.02937	0.75855	0.01066	0.18960	0.00380
IA10047 (MA001750121)	1.19399	0.02374	0.04777	0.01646	0.24610	0.00710
IA10077 (MA002034926)	1.09855	0.02414	-0.39787	0.02587	0.34460	0.01070
IA10097 (MA002135528)	0.55930	0.01602	0.20787	0.04523	0.12000	0.01530
IA10101 (MA002139080)	0.97478	0.03255	1.08683	0.01956	0.36630	0.00560
IA10104 (MA002140372)	1.10367	0.02477	-0.28206	0.02494	0.36500	0.00990
IA10224 (MA003747173)	0.46177	0.01511	0.04601	0.06965	0.08040	0.02160
IA12450 (MA301798)	0.61565	0.01261	-0.84517	0.04988	0.04130	0.02110
IA12462 (MA303324)	1.56816	0.02437	0.04404	0.00983	0.11870	0.00450
IA12478 (MA306990)	1.38133	0.03408	0.96119	0.01232	0.25950	0.00400
IA01054 (MA311567)	0.97852	0.01034	-0.40801	0.00869	0.00000	0.00000
IA02582 (MA311568)	0.79720	0.00903	-0.70918	0.01097	0.00000	0.00000
IA04887 (MA736377105)	1.06725	0.01091	-0.26230	0.00807	0.00000	0.00000
IA04942 (MA800727128)	0.74665	0.00841	0.06329	0.00942	0.00000	0.00000
IA05066 (MA803956738)	0.99074	0.01051	-0.64232	0.00942	0.00000	0.00000
IA07924 (MA900756471)	1.10605	0.01113	-0.36030	0.00807	0.00000	0.00000
IA07963 (MA900846441)	0.87866	0.00966	0.43959	0.00880	0.00000	0.00000
IA09849 (MA000732007)	0.77988	0.00852	-0.16056	0.00931	0.00000	0.00000
IA10060 (MA001851276)	0.84833	0.00932	-0.74747	0.01076	0.00000	0.00000
IA10093 (MA002128911)	0.82856	0.00949	-0.71528	0.01087	0.00000	0.00000
IA10111 (MA002145158)	0.61946	0.00818	-1.42100	0.01863	0.00000	0.00000
IA10151 (MA002334462)	1.00647	0.01057	-0.46131	0.00880	0.00000	0.00000
IA10213 (MA003540652)	1.09622	0.01346	-1.21174	0.01201	0.00000	0.00000
IA10222 (MA003744055)	1.71103	0.01818	-0.53893	0.00735	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)
IA02457 (MA307317)	1.00738	0.00790	0.02081	0.00564	1.48299	0.01383	0.87539	0.01115	-0.50427	0.00993
IA02742 (MA704653374)	0.63309	0.00659	-1.00449	0.01176	1.06416	0.02426	-1.06416	0.01553	0.00000	0.00000
IA04568 (MA293812)	0.99914	0.00801	0.40161	0.00578	1.69169	0.01323	0.83891	0.01036	-0.94489	0.01222
IA04621 (MA303335)	0.95785	0.00756	-0.08846	0.00598	1.96074	0.01712	0.40008	0.01046	-0.60804	0.01029
IA04757 (MA713677363)	0.66229	0.00716	-1.38291	0.01295	0.96434	0.02691	-0.96434	0.01677	0.00000	0.00000
IA04975 (MA801035466)	1.09054	0.00858	0.05410	0.00537	1.71169	0.01468	0.46719	0.00978	-0.52841	0.00961

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA02457 (MA307317)	-1.85412	0.01487	0.00000	0.00000	n/a	n/a
IA04568 (MA293812)	-1.58571	0.01541	0.00000	0.00000	n/a	n/a
IA04621 (MA303335)	-1.75277	0.01434	0.00000	0.00000	n/a	n/a
IA04975 (MA801035466)	-1.65047	0.01315	0.00000	0.00000	n/a	n/a

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)
IA00871 (MA297992)	1.06449	0.02427	0.07599	0.02025	0.29230	0.00790
IA02322 (MA301602)	0.96549	0.01908	-1.08193	0.03597	0.11560	0.02120
IA02385 (MA306408)	1.81114	0.04393	0.95962	0.01008	0.21840	0.00300
IA02398 (MA306458)	0.89764	0.01622	0.33607	0.01471	0.06630	0.00600
IA04604 (MA301157)	1.03270	0.03045	0.34867	0.02418	0.46080	0.00700
IA04931 (MA800650803)	1.25497	0.02736	0.86329	0.01189	0.19090	0.00380
IA04932 (MA800651876)	1.22738	0.02024	0.01784	0.01209	0.08830	0.00560
IA05072 (MA804073907)	1.43700	0.03226	0.83789	0.01129	0.21780	0.00360
IA07576 (MA301169)	0.78900	0.01692	0.71616	0.01592	0.06480	0.00600
IA07890 (MA900727061)	0.51673	0.01645	-0.40498	0.07809	0.11770	0.02640
IA07974 (MA900982012)	0.35191	0.01651	-0.54857	0.17896	0.11250	0.04460
IA09862 (MA000846578)	0.93550	0.01867	-0.12414	0.02237	0.18930	0.00960
IA09867 (MA000859040)	1.02103	0.03110	0.08939	0.02942	0.48460	0.00860
IA09870 (MA000927731)	0.89408	0.01231	0.01815	0.01290	0.00310	0.00520
IA09876 (MA000937699)	1.12336	0.02830	0.97907	0.01401	0.22420	0.00430
IA09882 (MA000953421)	1.41162	0.02637	-0.38533	0.01632	0.25880	0.00820
IA09884 (MA000957282)	1.15941	0.01972	-0.34714	0.01703	0.15300	0.00860
IA12452 (MA301830)	0.85534	0.02386	0.43653	0.02519	0.31520	0.00830
IA00876 (MA298032)	0.86006	0.00933	0.41618	0.00887	0.00000	0.00000
IA02734 (MA704359624)	0.39089	0.00677	-1.85805	0.03194	0.00000	0.00000
IA04971 (MA800975677)	0.95586	0.00998	-0.28728	0.00846	0.00000	0.00000
IA04983 (MA801235389)	0.72547	0.00846	-0.76996	0.01139	0.00000	0.00000
IA04999 (MA801646735)	1.17225	0.01383	-0.96081	0.00937	0.00000	0.00000
IA05071 (MA804073329)	0.88859	0.00922	0.19792	0.00836	0.00000	0.00000
IA08213 (MA904453014)	0.75406	0.00863	-0.68400	0.01058	0.00000	0.00000
IA09863 (MA000846693)	0.91706	0.00957	0.08949	0.00816	0.00000	0.00000
IA09877 (MA000938134)	0.66643	0.00823	-0.99114	0.01340	0.00000	0.00000
IA09883 (MA000955730)	0.81030	0.00863	-0.34915	0.00917	0.00000	0.00000
IA09898 (MA001042212)	0.56795	0.00712	-0.16323	0.01098	0.00000	0.00000
IA09918 (MA001066377)	1.15346	0.01237	-0.73469	0.00877	0.00000	0.00000
IA09926 (MA001142456)	0.59450	0.00758	-0.55341	0.01169	0.00000	0.00000
IA10188 (MA002837526)	0.91922	0.01044	-0.82155	0.01008	0.00000	0.00000
IA10284 (MA006336846)	0.52163	0.00758	-0.84382	0.01562	0.00000	0.00000
IA10481 (MA935150419)	0.95825	0.01003	-0.22883	0.00836	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)
IA07728 (MA802371654)	1.24949	0.00939	-0.30580	0.00486	1.52968	0.01397	0.49862	0.00950	-0.56347	0.00865
IA08179 (MA903746975)	1.14605	0.00898	-0.67760	0.00571	1.67380	0.01863	0.27968	0.01072	-0.51326	0.00932
IA08243 (MA908434516)	0.67629	0.00683	0.07825	0.00945	1.12825	0.01599	-1.12825	0.01673	0.00000	0.00000
IA09886 (MA000965213)	0.68650	0.00694	0.15585	0.00935	1.13207	0.01560	-1.13207	0.01676	0.00000	0.00000
IA10152 (MA002343629)	1.14238	0.00881	0.04311	0.00508	1.54011	0.01274	0.54452	0.00927	-0.37014	0.00911
IA10278 (MA005852277)	0.83381	0.00712	-0.02317	0.00664	1.62901	0.01716	0.90056	0.01284	-0.46776	0.01108

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA07728 (MA802371654)	-1.46483	0.01055	0.00000	0.00000	n/a	n/a
IA08179 (MA903746975)	-1.44022	0.01021	0.00000	0.00000	n/a	n/a
IA10152 (MA002343629)	-1.71449	0.01358	0.00000	0.00000	n/a	n/a
IA10278 (MA005852277)	-2.06181	0.01734	0.00000	0.00000	n/a	n/a

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)
IA02145 (MA282127)	1.22191	0.02110	0.41646	0.01144	0.16020	0.00420
IA02464 (MA307338)	0.88150	0.01409	-0.73768	0.02422	0.08650	0.01170
IA02470 (MA307363)	0.55614	0.01728	-1.72886	0.12863	0.18480	0.05050
IA05125 (MA805101277)	1.19579	0.04227	1.40764	0.01752	0.32770	0.00350
IA07619 (MA307340)	0.61084	0.01551	0.40058	0.02824	0.07530	0.01040
IA07694 (MA736509125)	0.71015	0.01363	-0.95484	0.04071	0.03610	0.01940
IA07740 (MA805104566)	1.38608	0.02641	0.41522	0.01144	0.22120	0.00420
IA07742 (MA805111429)	0.93346	0.02424	0.35750	0.02226	0.36290	0.00700
IA07772 (MA900283487)	1.11228	0.01677	-0.48259	0.01494	0.07170	0.00740
IA09952 (MA001264865)	0.84094	0.01517	-0.13278	0.01989	0.06960	0.00840
IA09990 (MA001529070)	1.14485	0.02367	0.09809	0.01618	0.27880	0.00630
IA10012 (MA001585164)	1.07544	0.02464	-0.15175	0.02195	0.34940	0.00820
IA10159 (MA002536621)	0.51091	0.03092	1.69035	0.04277	0.34420	0.01000
IA10161 (MA002538062)	0.76793	0.01278	-1.71278	0.04628	0.03000	0.02730
IA10276 (MA005664640)	1.24838	0.02264	0.40110	0.01154	0.17590	0.00430
IA12436 (MA296382)	1.23178	0.03040	0.72802	0.01412	0.32580	0.00430
IA00782 (MA223217)	0.83792	0.00873	-0.29934	0.00866	0.00000	0.00000
IA00974 (MA307362)	1.07680	0.01118	-1.05821	0.00969	0.00000	0.00000
IA01129 (MA624248796)	0.61421	0.00724	-0.53072	0.01103	0.00000	0.00000
IA02284 (MA299673)	0.93877	0.00953	-0.73118	0.00928	0.00000	0.00000
IA02691 (MA703149512)	0.72110	0.00799	-0.85507	0.01113	0.00000	0.00000
IA04875 (MA736063629)	0.87277	0.00884	0.13436	0.00876	0.00000	0.00000
IA04885 (MA736368137)	1.00254	0.01032	-0.70490	0.00855	0.00000	0.00000
IA05128 (MA805109765)	0.46990	0.00667	-0.13783	0.01319	0.00000	0.00000
IA05141 (MA805280170)	0.75316	0.00833	-0.91577	0.01113	0.00000	0.00000
IA07692 (MA736452404)	0.66018	0.00753	0.25670	0.01062	0.00000	0.00000
IA08104 (MA902758854)	0.47446	0.00719	-1.23177	0.01855	0.00000	0.00000
IA09997 (MA001549477)	0.74654	0.00861	-1.05131	0.01196	0.00000	0.00000
IA10001 (MA001554177)	0.65990	0.00764	-0.41271	0.01010	0.00000	0.00000
IA10011 (MA001577731)	0.90682	0.00953	-0.56669	0.00866	0.00000	0.00000
IA10014 (MA001604473)	1.15472	0.01084	0.25289	0.00763	0.00000	0.00000
IA10210 (MA003477341)	0.94447	0.00941	-0.13361	0.00794	0.00000	0.00000
IA10230 (MA003861140)	0.68688	0.00759	-0.32108	0.00979	0.00000	0.00000
IA10414 (MA800385560)	1.00938	0.01078	-0.61729	0.00876	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)
IA04565 (MA290253)	1.42578	0.01032	0.19043	0.00405	1.25833	0.00928	0.32589	0.00783	-0.53440	0.00873
IA04596 (MA298252)	1.43496	0.01112	0.35992	0.00373	0.73337	0.00776	0.14816	0.00785	-0.25019	0.00830
IA04912 (MA800203270)	1.45281	0.01181	0.32900	0.00516	0.46560	0.00837	-0.46560	0.00947	0.00000	0.00000
IA07810 (MA900378821)	1.17280	0.00907	-0.17772	0.00570	2.31002	0.01784	-0.04566	0.00920	-0.73641	0.00977
IA10413 (MA800301627)	1.45903	0.01055	-0.17138	0.00384	1.08822	0.00966	0.31080	0.00781	-0.42272	0.00781
IA10435 (MA900578720)	0.97339	0.00793	-0.04549	0.00660	0.80804	0.01130	-0.80804	0.01155	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA04565 (MA290253)	-1.04983	0.01022	0.00000	0.00000	n/a	n/a
IA04596 (MA298252)	-0.63133	0.00933	0.00000	0.00000	n/a	n/a
IA07810 (MA900378821)	-1.52795	0.01232	0.00000	0.00000	n/a	n/a
IA10413 (MA800301627)	-0.97629	0.00891	0.00000	0.00000	n/a	n/a

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)
IA00920 (MA304467)	1.18324	0.02198	-0.36628	0.01702	0.17490	0.00730
IA00952 (MA306626)	1.19069	0.03830	1.15262	0.01714	0.41290	0.00370
IA00953 (MA306627)	0.92129	0.01585	0.33553	0.01402	0.08710	0.00500
IA02039 (MA219417)	0.95305	0.01516	0.24050	0.01313	0.07100	0.00480
IA02526 (MA311074)	1.12317	0.02615	1.07628	0.01402	0.22650	0.00340
IA04501 (MA259175)	0.56784	0.02346	0.61316	0.05030	0.40090	0.01140
IA04513 (MA272764)	1.03832	0.03154	0.88200	0.01981	0.44740	0.00450
IA04541 (MA282221)	1.26951	0.02969	1.00651	0.01280	0.22810	0.00330
IA04587 (MA298072)	1.33952	0.02367	0.53215	0.01068	0.15100	0.00340
IA07778 (MA900336138)	1.41908	0.02420	0.64710	0.00990	0.12450	0.00290
IA07843 (MA900557823)	0.67853	0.01585	-0.13238	0.03450	0.17520	0.01210
IA08188 (MA903983773)	0.74441	0.01902	1.00796	0.01869	0.15570	0.00560
IA08189 (MA904000450)	0.74832	0.02251	1.18366	0.02103	0.24160	0.00560
IA08199 (MA904169987)	0.90719	0.01939	0.73656	0.01558	0.18040	0.00480
IA09888 (MA000971342)	1.16226	0.02198	-0.12514	0.01580	0.17180	0.00620
IA10256 (MA005170212)	0.91443	0.02309	0.19555	0.02459	0.38570	0.00720
IA10270 (MA005207399)	1.13157	0.02002	-0.01743	0.01558	0.23900	0.00600
IA02876 (MA713848056)	0.88801	0.00861	0.24006	0.00879	0.00000	0.00000
IA02885 (MA713849125)	0.96425	0.00930	-0.42013	0.00835	0.00000	0.00000
IA02887 (MA713849162)	0.81336	0.00803	0.26265	0.00935	0.00000	0.00000
IA04689 (MA311092)	1.44803	0.01442	0.35489	0.00734	0.00000	0.00000
IA04770 (MA713848251)	0.60705	0.00676	0.20668	0.01090	0.00000	0.00000
IA04773 (MA713848348)	0.90291	0.00866	0.02296	0.00846	0.00000	0.00000
IA05115 (MA804676692)	1.37840	0.01374	0.47106	0.00768	0.00000	0.00000
IA07848 (MA900567252)	0.86017	0.00851	0.28112	0.00901	0.00000	0.00000
IA07897 (MA900739359)	1.03055	0.01014	-0.27347	0.00812	0.00000	0.00000
IA07905 (MA900743031)	0.89202	0.00856	0.20456	0.00901	0.00000	0.00000
IA07908 (MA900749529)	0.81346	0.00814	-0.21517	0.00890	0.00000	0.00000
IA10027 (MA001678587)	0.70505	0.00724	0.01473	0.00957	0.00000	0.00000
IA10032 (MA001702061)	1.17013	0.01342	-1.33714	0.01035	0.00000	0.00000
IA10050 (MA001759197)	0.50640	0.00713	-1.09334	0.01536	0.00000	0.00000
IA10081 (MA002046543)	0.88415	0.00888	0.25719	0.00879	0.00000	0.00000
IA10252 (MA005077116)	0.68577	0.00719	0.05545	0.00990	0.00000	0.00000
IA12459 (MA302328)	0.94385	0.00919	-0.07563	0.00823	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)
IA01160 (MA624562376)	0.57339	0.00544	0.00827	0.01115	1.37090	0.01871	-1.37090	0.01989	0.00000	0.00000
IA07679 (MA717248260)	1.16152	0.00845	0.30521	0.00542	1.90985	0.01218	0.32487	0.00922	-0.49022	0.01042
IA07730 (MA802907874)	1.33048	0.00962	0.45877	0.00445	1.25222	0.00907	0.39107	0.00840	-0.55008	0.00995
IA10090 (MA002119133)	1.40962	0.00999	-0.23823	0.00406	1.11071	0.01009	0.39344	0.00839	-0.39527	0.00820
IA10445 (MA900765087)	1.05168	0.00888	-0.90890	0.00805	1.08019	0.01607	-1.08019	0.01144	0.00000	0.00000
IA12412 (MA261648)	1.31764	0.00993	0.17009	0.00398	0.83703	0.00855	0.23203	0.00816	-0.10757	0.00835

Item ID	Parameters and Measures of Standard Error						
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)	
IA07679 (MA717248260)	-1.74450	0.01566	0.00000	0.00000	n/a	n/a	
IA07730 (MA802907874)	-1.09321	0.01198	0.00000	0.00000	n/a	n/a	
IA10090 (MA002119133)	-1.10888	0.00948	0.00000	0.00000	n/a	n/a	
IA12412 (MA261648)	-0.96149	0.01036	0.00000	0.00000	n/a	n/a	

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)
IA00792 (MA252991)	1.56012	0.05282	0.60178	0.01809	0.63240	0.00340
IA00859 (MA297517)	0.95291	0.02814	0.92010	0.01875	0.35830	0.00470
IA02480 (MA307585)	1.00019	0.02095	-0.87045	0.03088	0.29480	0.01340
IA04553 (MA287538)	0.80660	0.01908	-0.67104	0.03971	0.30600	0.01440
IA04594 (MA298198)	0.64952	0.01636	0.62958	0.02371	0.13920	0.00780
IA05094 (MA804535837)	1.32586	0.03395	0.66278	0.01456	0.39130	0.00390
IA08003 (MA901143033)	1.35086	0.02921	0.25821	0.01346	0.31960	0.00460
IA10039 (MA001736920)	0.83773	0.01855	-0.17294	0.02658	0.25550	0.00950
IA10064 (MA001865159)	0.96762	0.01572	-0.02900	0.01412	0.07910	0.00570
IA10132 (MA002180558)	1.13963	0.02089	-0.29206	0.01677	0.22030	0.00710
IA10134 (MA002181298)	1.00995	0.01929	-0.33949	0.01985	0.20080	0.00840
IA10231 (MA003932801)	0.87568	0.01455	-0.79049	0.02471	0.07170	0.01200
IA10477 (MA908446890)	1.00344	0.02132	-0.63883	0.02757	0.31530	0.01100
IA10478 (MA908450808)	1.03644	0.01594	0.07611	0.01279	0.10480	0.00480
IA12461 (MA303244)	0.88405	0.02431	1.09315	0.01732	0.23050	0.00450
IA12480 (MA307586)	0.58023	0.01151	-0.36287	0.03541	0.01180	0.01350
IA02938 (MA715919661)	0.34726	0.00618	-1.64142	0.02691	0.00000	0.00000
IA02940 (MA715919716)	0.95382	0.00981	-1.00821	0.00993	0.00000	0.00000
IA02943 (MA715919788)	1.00344	0.00970	0.10953	0.00849	0.00000	0.00000
IA02947 (MA715919853)	1.15295	0.01109	-0.17195	0.00761	0.00000	0.00000
IA04779 (MA715919758)	0.47992	0.00624	-0.64258	0.01313	0.00000	0.00000
IA07568 (MA297651)	0.81902	0.00821	-0.29349	0.00893	0.00000	0.00000
IA07701 (MA800472975)	0.91203	0.00895	0.12652	0.00871	0.00000	0.00000
IA07707 (MA800659905)	0.59873	0.00656	0.73226	0.01335	0.00000	0.00000
IA07733 (MA804043870)	1.00792	0.00981	0.21828	0.00816	0.00000	0.00000
IA08010 (MA901253257)	0.53557	0.00618	0.08141	0.01180	0.00000	0.00000
IA08088 (MA902308680)	1.14794	0.01082	-0.12562	0.00761	0.00000	0.00000
IA08226 (MA905179612)	0.98601	0.00917	0.62627	0.00949	0.00000	0.00000
IA10041 (MA001737758)	0.86518	0.00874	-0.40280	0.00893	0.00000	0.00000
IA10042 (MA001737991)	0.38782	0.00565	-0.18430	0.01445	0.00000	0.00000
IA10130 (MA002177981)	0.82776	0.00826	-0.74560	0.00960	0.00000	0.00000
IA10201 (MA003128642)	0.57938	0.00650	-0.27540	0.01070	0.00000	0.00000
IA10234 (MA003936639)	1.04539	0.01045	-0.68560	0.00849	0.00000	0.00000
IA12479 (MA307492)	1.08169	0.01093	-0.94832	0.00904	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)
IA04612 (MA301714)	1.49568	0.01109	0.33851	0.00393	0.88458	0.00818	0.25445	0.00798	-0.25578	0.00857
IA04717 (MA311433)	1.30576	0.00938	0.28498	0.00485	1.69865	0.01105	0.29479	0.00866	-0.55669	0.00988
IA07709 (MA800738445)	1.34030	0.00938	-0.13560	0.00415	1.09100	0.01005	0.41300	0.00848	-0.30712	0.00828
IA10143 (MA002243883)	0.85052	0.00800	-1.11945	0.00897	0.70661	0.01795	-0.70661	0.01266	0.00000	0.00000
IA10290 (MA010701848)	0.92280	0.00725	-1.44190	0.00694	1.52132	0.02197	0.47648	0.01504	-0.45530	0.01166
IA10455 (MA902278325)	0.76050	0.00688	-0.23167	0.00761	0.83279	0.01349	-0.83279	0.01286	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error					
	d3	SE(d3)	d4	SE(d4)	d5	SE(d5)
IA04612 (MA301714)	-0.88325	0.01027	0.00000	0.00000	n/a	n/a
IA04717 (MA311433)	-1.43675	0.01327	0.00000	0.00000	n/a	n/a
IA07709 (MA800738445)	-1.19688	0.01015	0.00000	0.00000	n/a	n/a
IA10290 (MA010701848)	-1.54250	0.01088	0.00000	0.00000	n/a	n/a

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
60542	0.91	0.61	Overall	0.82	0.74	0.09	0.09
			Cut 1	0.95	0.93	0.02	0.03
			Cut 2	0.91	0.87	0.05	0.04
			Cut 3	0.96	0.94	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.84	0.73		
			Perf 2	0.81	0.75		
			Perf 3	0.83	0.77		
			Perf 4	0.77	0.61		

Table 2.9.2
 DAC Results
 English Language Arts Grade 4

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
61836	0.90	0.59	Overall	0.81	0.73	0.10	0.09
			Cut 1	0.95	0.92	0.02	0.03
			Cut 2	0.90	0.86	0.05	0.05
			Cut 3	0.96	0.95	0.03	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.83	0.72		
			Perf 2	0.82	0.76		
			Perf 3	0.80	0.73		
			Perf 4	0.74	0.54		

Table 2.9.3
 DAC Results
 English Language Arts Grade 5

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
62316	0.91	0.61	Overall	0.82	0.75	0.10	0.08
			Cut 1	0.95	0.93	0.02	0.03
			Cut 2	0.91	0.87	0.05	0.04
			Cut 3	0.96	0.95	0.03	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.83	0.74		
			Perf 2	0.81	0.75		
			Perf 3	0.84	0.78		
			Perf 4	0.72	0.52		

Table 2.9.4
 DAC Results
 English Language Arts Grade 6

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
63574	0.91	0.56	Overall	0.78	0.69	0.11	0.11
			Cut 1	0.94	0.91	0.03	0.03
			Cut 2	0.91	0.87	0.05	0.05
			Cut 3	0.94	0.91	0.03	0.03
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.85	0.77		
			Perf 2	0.77	0.69		
			Perf 3	0.78	0.70		
			Perf 4	0.64	0.48		

Table 2.9.5
 DAC Results
 English Language Arts Grade 7

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
63711	0.91	0.60	Overall	0.81	0.73	0.10	0.09
			Cut 1	0.94	0.92	0.03	0.03
			Cut 2	0.91	0.87	0.05	0.05
			Cut 3	0.96	0.94	0.03	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.83	0.73		
			Perf 2	0.81	0.75		
			Perf 3	0.79	0.72		
			Perf 4	0.78	0.62		

Table 2.9.6
 DAC Results
 English Language Arts Grade 8

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
65553	0.91	0.57	Overall	0.78	0.69	0.11	0.11
			Cut 1	0.94	0.91	0.03	0.03
			Cut 2	0.91	0.87	0.04	0.05
			Cut 3	0.94	0.91	0.03	0.03
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.85	0.76		
			Perf 2	0.78	0.69		
			Perf 3	0.77	0.69		
			Perf 4	0.70	0.55		

Table 2.9.7
 DAC Results
 Mathematics Grade 3

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
51707	0.93	0.65	Overall	0.83	0.76	0.08	0.08
			Cut 1	0.96	0.94	0.02	0.02
			Cut 2	0.92	0.88	0.04	0.04
			Cut 3	0.96	0.94	0.02	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.84	0.74		
			Perf 2	0.84	0.78		
			Perf 3	0.83	0.78		
			Perf 4	0.80	0.67		

Table 2.9.8
 DAC Results
 Mathematics Grade 4

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
52554	0.93	0.66	Overall	0.84	0.78	0.08	0.08
			Cut 1	0.96	0.95	0.01	0.02
			Cut 2	0.92	0.89	0.04	0.04
			Cut 3	0.96	0.94	0.02	0.02
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.85	0.74		
			Perf 2	0.83	0.77		
			Perf 3	0.85	0.80		
			Perf 4	0.82	0.70		

Table 2.9.9
 DAC Results
 Mathematics Grade 5

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
54159	0.93	0.67	Overall	0.86	0.80	0.08	0.06
			Cut 1	0.97	0.95	0.02	0.02
			Cut 2	0.92	0.88	0.05	0.04
			Cut 3	0.97	0.96	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.76	0.66		
			Perf 2	0.86	0.82		
			Perf 3	0.87	0.82		
			Perf 4	0.82	0.68		

Table 2.9.10
 DAC Results
 Mathematics Grade 6

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
56389	0.94	0.69	Overall	0.86	0.80	0.07	0.07
			Cut 1	0.96	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	0.85	0.75		
			Perf 2	0.86	0.81		
			Perf 3	0.86	0.81		
			Perf 4	0.83	0.72		

Table 2.9.11
 DAC Results
 Mathematics Grade 7

N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
57234	0.94	0.68	Overall	0.85	0.78	0.08	0.08
			Cut 1	0.95	0.93	0.02	0.03
			Cut 2	0.93	0.90	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	0.86	0.78		
			Perf 2	0.85	0.79		
			Perf 3	0.84	0.78		
			Perf 4	0.84	0.73		

Table 2.9.12
 DAC Results
 Mathematics Grade 8

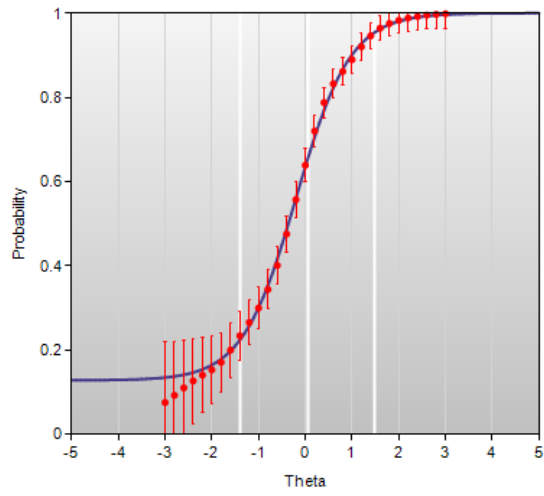
N	Reliability	Kappa		Accuracy	Consistency	F Pos	F Neg
59572	0.94	0.67	Overall	0.84	0.78	0.09	0.07
			Cut 1	0.95	0.92	0.03	0.03
			Cut 2	0.93	0.90	0.04	0.03
			Cut 3	0.97	0.96	0.02	0.01
			Cut 4	1.00	1.00	0.00	0.00
			Perf 1	0.80	0.72		
			Perf 2	0.85	0.79		
			Perf 3	0.85	0.79		
			Perf 4	0.85	0.73		

Section 2.10

Fit Plots of Watchlist Items

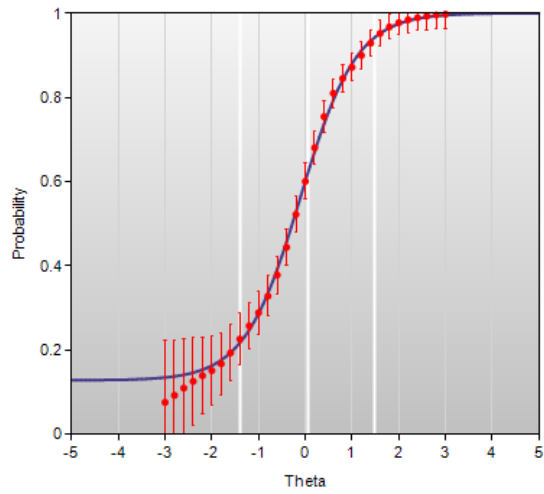
Initial Calibration

Mathematics Grade 4: IA00961



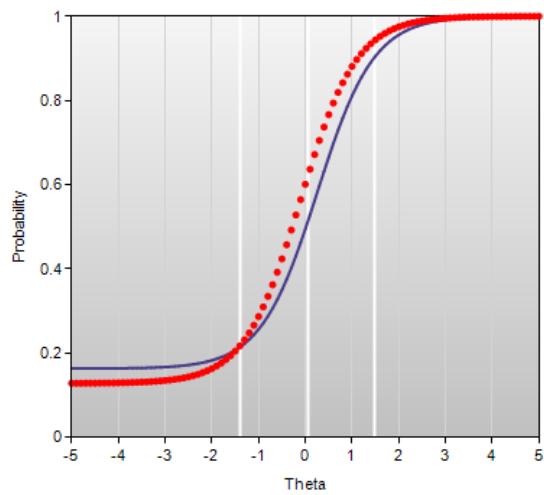
Final Calibration

Mathematics Grade 4: IA00961



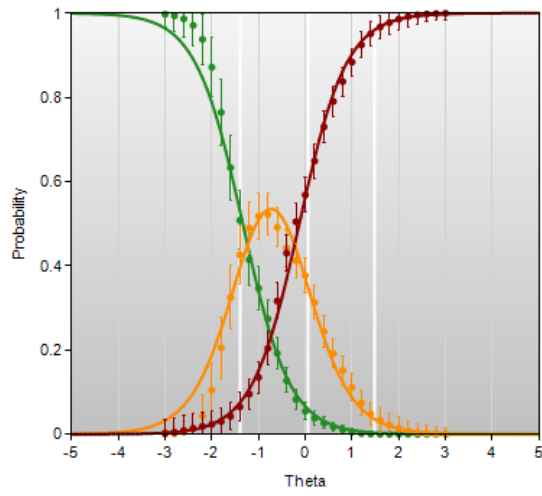
Beta Chart

Mathematics Grade 4: IA00961 (MA307081)



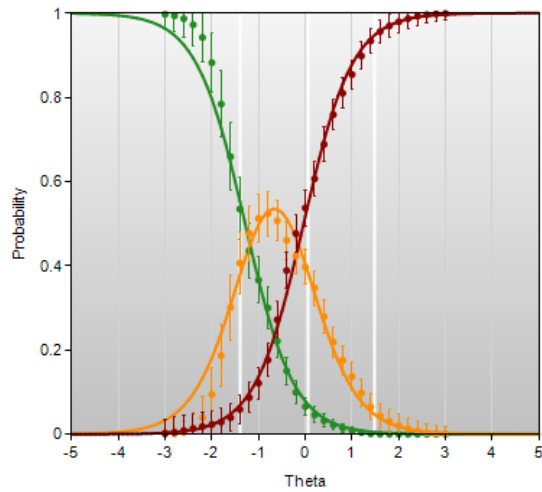
Initial Calibration

Mathematics Grade 4: IA01093



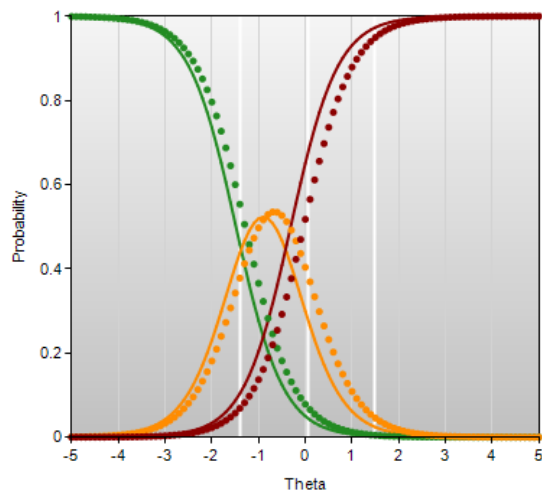
Final Calibration

Mathematics Grade 4: IA01093



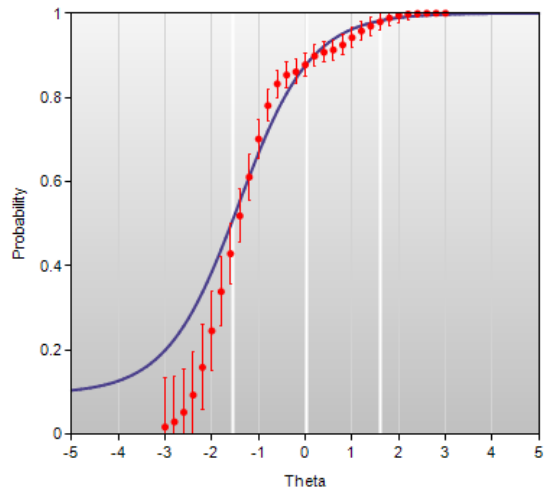
Beta Chart

Mathematics Grade 4: IA01093 (MA623879088)



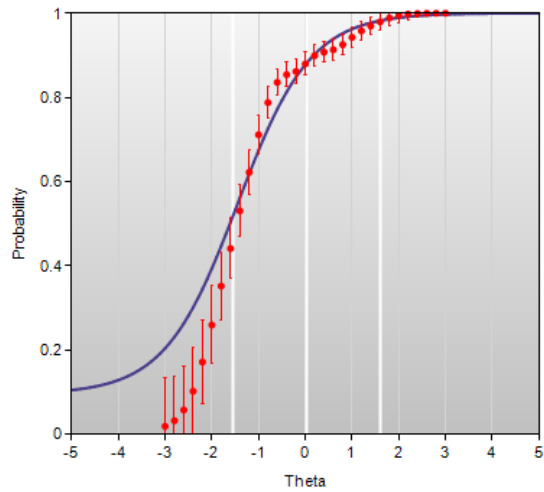
Initial Calibration

Mathematics Grade 5: IA00936



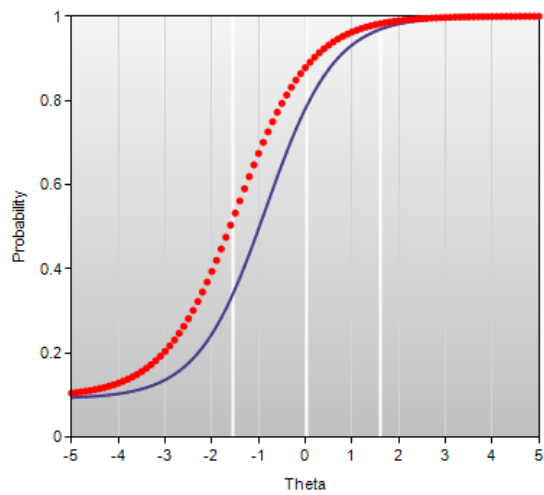
Final Calibration

Mathematics Grade 5: IA00936



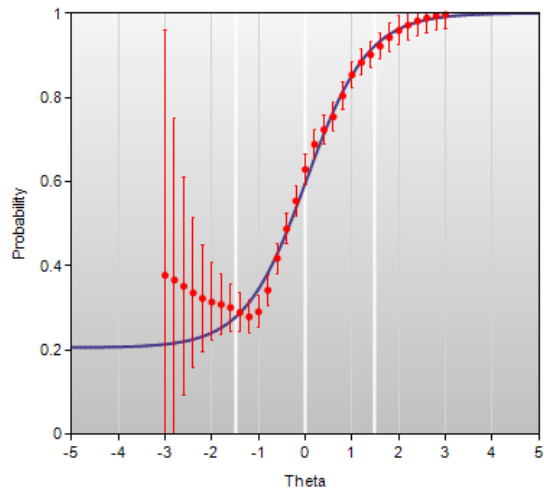
Beta Chart

Mathematics Grade 5: IA00936 (MA306420)



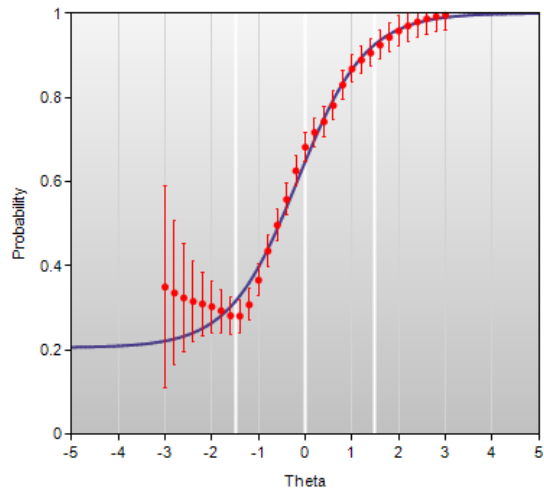
Initial Calibration

Mathematics Grade 8: IA00865



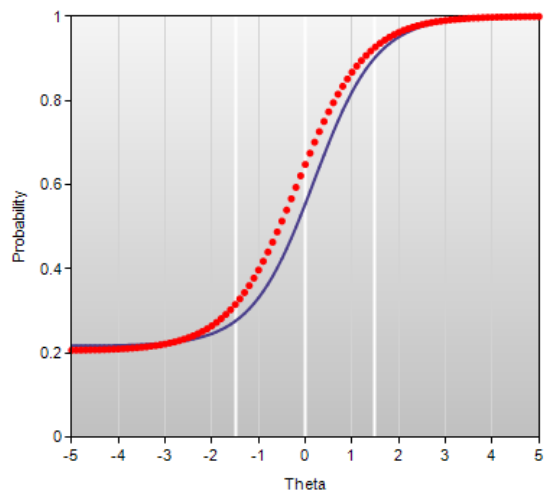
Final Calibration

Mathematics Grade 8: IA00865



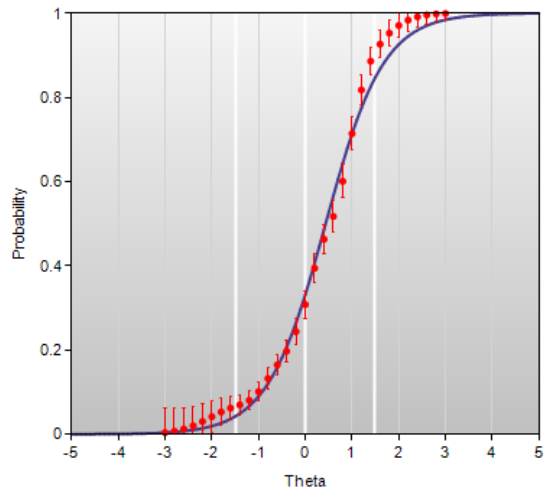
Beta Chart

Mathematics Grade 8: IA00865 (MA297656)



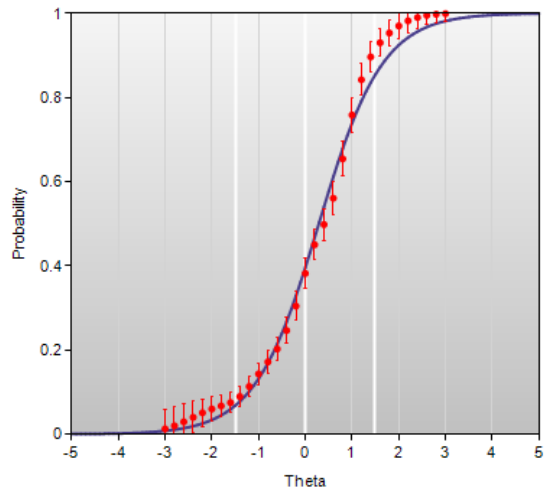
Initial Calibration

Mathematics Grade 8: IA02495



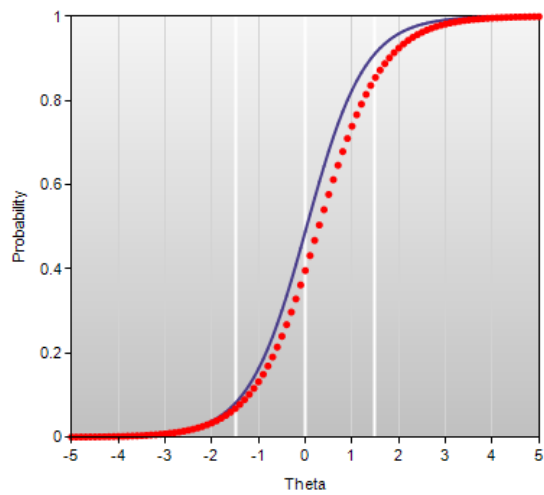
Final Calibration

Mathematics Grade 8: IA02495



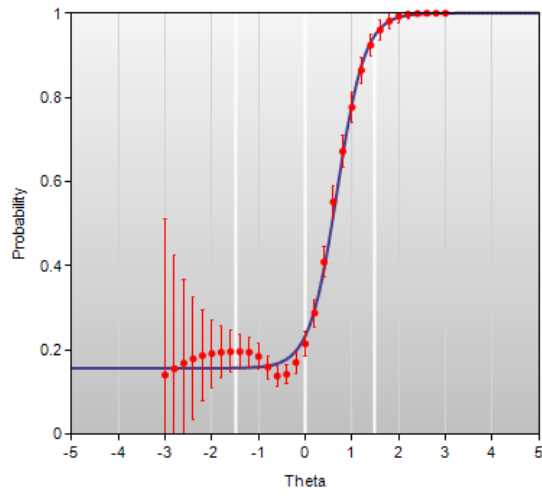
Beta Chart

Mathematics Grade 8: IA02495 (MA309741)



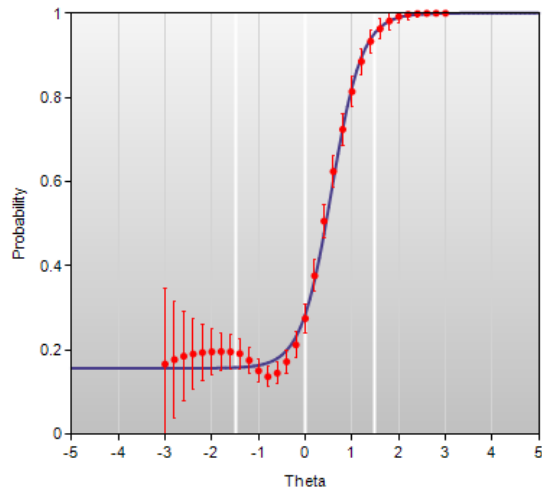
Initial Calibration

Mathematics Grade 8: IA05070



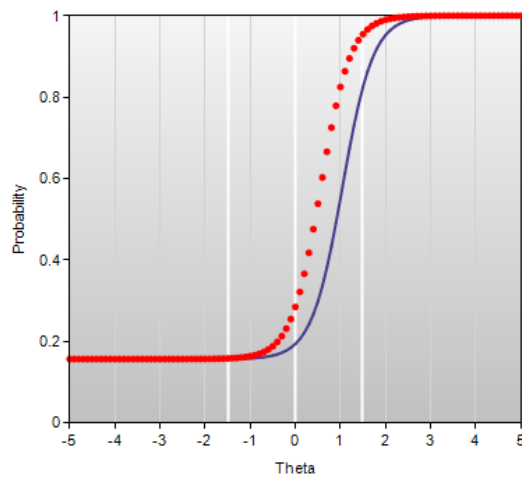
Final Calibration

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	Raw Score			Alpha	SEM
			Maximum	Mean	Standard Deviation		
3	All Students	9,647	44	22.70	9.82	0.91	2.87
	ELL	1,591	44	16.20	8.22	0.88	2.89
	Economically Disadvantaged	4,674	44	18.74	8.92	0.89	2.91
	African American	814	44	19.28	8.87	0.89	2.92
	Asian	303	44	26.93	9.28	0.91	2.78
	Hispanic	2,892	44	18.73	9.15	0.90	2.90
	Native American/Alaska Native	91	44	16.30	7.99	0.87	2.89
	White	4,986	44	25.44	9.40	0.91	2.82
	Pacific Islander/Hawaiian	18	44	18.22	9.66	0.91	2.90
	Multiracial	543	44	22.70	9.52	0.91	2.86
	Male	4,970	44	22.49	9.80	0.92	2.85
	Female	4,676	44	22.93	9.85	0.91	2.88
	Special Education	1,817	44	15.06	8.19	0.88	2.80
4	All Students	9,728	44	23.71	8.59	0.88	2.93
	ELL	1,476	44	18.76	8.20	0.87	3.00
	Economically Disadvantaged	4,667	44	20.35	8.06	0.86	2.99
	African American	815	44	21.08	7.95	0.86	2.98
	Asian	341	44	27.09	8.06	0.88	2.84
	Hispanic	2,896	44	20.59	8.16	0.87	2.98
	Native American/Alaska Native	77	44	18.49	7.55	0.84	3.01
	White	5,018	44	25.77	8.28	0.88	2.88
	Pacific Islander/Hawaiian	12	44	23.33	10.36	0.92	2.94
	Multiracial	569	44	23.97	8.46	0.88	2.90
	Male	4,876	44	23.13	8.57	0.88	2.91
	Female	4,851	44	24.29	8.57	0.88	2.93
	Special Education	1,532	44	15.09	6.92	0.82	2.93
5	All Students	9,707	48	26.14	9.85	0.91	2.96
	ELL	1,532	48	20.50	9.55	0.90	3.05
	Economically Disadvantaged	4,706	48	22.49	9.55	0.90	3.02
	African American	845	48	23.47	9.34	0.90	3.00
	Asian	323	48	29.36	9.95	0.92	2.88
	Hispanic	2,870	48	22.53	9.64	0.90	3.03
	Native American/Alaska Native	77	48	20.06	10.29	0.91	3.05
	White	5,072	48	28.59	9.26	0.90	2.89
	Pacific Islander/Hawaiian	13	48	24.92	12.29	0.95	2.85
	Multiracial	507	48	25.45	9.31	0.90	2.99
	Male	4,965	48	24.82	9.88	0.91	2.97
	Female	4,740	48	27.52	9.62	0.91	2.94
	Special Education	1,559	48	15.81	8.03	0.86	2.96
6	All Students	9,728	50	24.50	10.16	0.90	3.18
	ELL	1,594	50	17.99	8.97	0.88	3.10
	Economically Disadvantaged	4,578	50	20.38	9.32	0.89	3.12
	African American	966	50	20.37	9.44	0.89	3.12
	Asian	320	50	27.23	9.80	0.89	3.21

continued

Grade	Subgroup	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard Deviation		
6	Hispanic	2,735	50	20.44	9.54	0.89	3.13
	Native American/Alaska Native	75	50	19.23	9.23	0.89	3.09
	White	5,129	50	27.40	9.59	0.89	3.17
	Pacific Islander/Hawaiian	21	50	23.10	9.14	0.88	3.18
	Multiracial	482	50	23.97	10.12	0.90	3.18
	Male	4,990	50	23.09	10.00	0.90	3.17
	Female	4,732	50	25.97	10.11	0.90	3.18
	Special Education	1,592	50	14.84	7.29	0.84	2.95
	All Students	9,868	50	23.75	10.89	0.90	3.40
	ELL	1,758	50	17.08	9.13	0.88	3.19
	Economically Disadvantaged	4,545	50	19.32	9.44	0.88	3.23
7	African American	832	50	19.42	9.29	0.88	3.23
	Asian	319	50	27.63	11.41	0.91	3.46
	Hispanic	2,878	50	19.19	9.66	0.89	3.24
	Native American/Alaska Native	65	50	17.83	8.26	0.84	3.26
	White	5,263	50	26.82	10.59	0.90	3.43
	Pacific Islander/Hawaiian	18	50	23.89	11.81	0.92	3.40
	Multiracial	493	50	23.24	10.97	0.90	3.39
	Male	5,108	50	22.31	10.65	0.90	3.32
	Female	4,756	50	25.30	10.93	0.90	3.44
	Special Education	1,620	50	13.79	7.01	0.82	2.97
	All Students	9,958	50	27.22	10.61	0.90	3.29
ELL	1,762	50	20.90	10.01	0.90	3.24	
Economically Disadvantaged	4,429	50	22.99	9.87	0.89	3.24	
8	African American	907	50	23.32	10.16	0.90	3.22
	Asian	332	50	30.01	10.66	0.90	3.30
	Hispanic	2,867	50	23.19	10.20	0.90	3.25
	Native American/Alaska Native	66	50	21.08	9.80	0.90	3.10
	White	5,264	50	30.08	10.00	0.89	3.26
	Pacific Islander/Hawaiian	11	50	25.09	8.71	0.86	3.21
	Multiracial	511	50	26.28	9.93	0.89	3.26
	Male	5,138	50	25.52	10.64	0.90	3.28
	Female	4,809	50	29.01	10.28	0.90	3.28
	Special Education	1,549	50	17.48	8.30	0.86	3.10

Table K-2. Subgroup Reliabilities—Mathematics

Grade	Subgroup	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard		
3	All Students	9,792	48	24.54	12.45	0.94	3.03
	ELL	1,732	48	17.78	11.21	0.93	2.95
	Economically Disadvantaged	4,777	48	19.62	11.33	0.93	3.00
	African American	832	48	19.96	11.34	0.93	3.00
	Asian	318	48	30.66	11.87	0.94	2.94
	Hispanic	2,972	48	19.93	11.52	0.93	3.01
	Native American/Alaska Native	95	48	15.94	10.07	0.92	2.92
	White	5,012	48	27.90	12.03	0.94	3.00
	Pacific Islander/Hawaiian	18	48	20.28	11.73	0.93	3.14
	Multiracial	545	48	23.73	12.00	0.94	3.04
	Male	5,036	48	25.53	12.64	0.94	3.01
	Female	4,755	48	23.48	12.16	0.94	3.05
	Special Education	1,823	48	15.76	10.90	0.93	2.88
4	All Students	9,845	54	27.46	12.86	0.93	3.33
	ELL	1,584	54	20.84	12.01	0.93	3.27
	Economically Disadvantaged	4,738	54	22.21	11.55	0.92	3.30
	African American	827	54	22.69	11.65	0.92	3.31
	Asian	353	54	33.48	13.22	0.94	3.23
	Hispanic	2,964	54	22.41	11.67	0.92	3.31
	Native American/Alaska Native	77	54	19.94	11.41	0.92	3.27
	White	5,043	54	30.92	12.37	0.93	3.29
	Pacific Islander/Hawaiian	12	54	25.50	15.21	0.96	3.13
	Multiracial	569	54	27.47	12.90	0.93	3.37
	Male	4,937	54	28.13	13.12	0.94	3.33
	Female	4,907	54	26.79	12.55	0.93	3.33
	Special Education	1,542	54	15.82	10.31	0.91	3.11
5	All Students	9,817	54	24.68	12.05	0.92	3.39
	ELL	1,649	54	19.03	10.51	0.90	3.26
	Economically Disadvantaged	4,771	54	20.12	10.38	0.90	3.30
	African American	849	54	20.55	10.51	0.90	3.30
	Asian	337	54	31.29	12.78	0.93	3.38
	Hispanic	2,938	54	20.10	10.24	0.90	3.30
	Native American/Alaska Native	80	54	16.94	10.23	0.90	3.18
	White	5,093	54	27.84	12.04	0.92	3.40
	Pacific Islander/Hawaiian	13	54	23.85	15.01	0.95	3.31
	Multiracial	507	54	23.25	12.08	0.92	3.35
	Male	5,025	54	25.16	12.51	0.93	3.38
	Female	4,788	54	24.17	11.54	0.91	3.38
	Special Education	1,554	54	14.54	8.71	0.88	3.07
6	All Students	9,806	54	22.44	12.67	0.93	3.45
	ELL	1,689	54	15.20	10.08	0.91	3.08
	Economically Disadvantaged	4,617	54	17.03	10.33	0.91	3.18
	African American	978	54	16.91	10.39	0.91	3.15
	Asian	328	54	29.30	14.10	0.94	3.57
Hispanic	2,778	54	16.96	10.27	0.90	3.18	

continued

Grade	Subgroup	Number of Students	Raw Score		Alpha	SEM
			Maximum	Mean		
6	Native American/Alaska Native	76	54	16.14	9.73	3.12
	White	5,144	54	26.17	12.62	3.54
	Pacific Islander/Hawaiian	21	54	22.81	15.25	3.64
	Multiracial	481	54	21.78	12.62	3.41
	Male	5,014	54	22.68	12.97	3.45
	Female	4,786	54	22.19	12.34	3.43
	Special Education	1,578	54	11.92	7.56	2.81
	All Students	9,948	54	19.09	12.46	3.38
	ELL	1,870	54	12.43	9.00	2.94
	Economically Disadvantaged	4,584	54	13.92	9.41	3.07
7	African American	844	54	13.70	9.54	3.04
	Asian	328	54	25.12	14.27	3.53
	Hispanic	2,928	54	13.91	9.45	3.08
	Native American/Alaska Native	69	54	12.55	7.48	3.06
	White	5,275	54	22.60	12.87	3.48
	Pacific Islander/Hawaiian	18	54	18.17	13.81	3.36
	Multiracial	486	54	18.43	12.61	3.35
	Male	5,137	54	19.67	12.92	3.39
	Female	4,806	54	18.45	11.92	3.37
	Special Education	1,606	54	9.65	6.35	2.66
8	All Students	10,071	54	22.08	12.24	3.44
	ELL	1,899	54	15.43	9.47	3.16
	Economically Disadvantaged	4,493	54	16.91	9.53	3.26
	African American	914	54	16.55	9.86	3.27
	Asian	336	54	26.97	13.71	3.48
	Hispanic	2,951	54	17.03	9.96	3.26
	Native American/Alaska Native	70	54	16.24	9.66	3.21
	White	5,276	54	25.79	12.31	3.48
	Pacific Islander/Hawaiian	11	54	21.36	14.51	3.45
	Multiracial	513	54	20.42	11.52	3.39
Male	5,193	54	22.19	12.48	3.44	
Female	4,867	54	21.94	11.95	3.43	
Special Education	1,549	54	12.89	7.13	2.96	

Table K-3. Reliabilities by Reporting Categories, Grade, and Content Area—ELA

Grade	Item Reporting Category	Label	Number of Items	Raw Score			Alpha	SEM
				Maximum	Mean	Standard Deviation		
3	1	Reading	24	28	15.64	6.61	0.88	2.31
	2	Language	8	12	6.14	3.04	0.75	1.53
	3	Writing	1	4	0.92	0.86	--	--
4	1	Reading	25	28	16.72	5.73	0.83	2.35
	2	Language	7	12	5.93	2.81	0.68	1.59
	3	Writing	1	4	1.07	0.79	--	--
5	1	Reading	22	26	16.62	5.96	0.85	2.27
	2	Language	9	14	7.10	3.19	0.74	1.62
	3	Writing	2	8	2.43	1.50	0.71	0.81
6	1	Reading	25	30	16.69	6.41	0.85	2.50
	2	Language	6	10	5.20	2.57	0.68	1.46
	3	Writing	2	10	2.61	2.02	0.80	0.90
7	1	Reading	24	28	15.20	6.11	0.84	2.45
	2	Language	7	12	6.18	3.34	0.75	1.66
	3	Writing	2	10	2.38	2.37	0.75	1.18
8	1	Reading	25	29	17.74	6.12	0.85	2.40
	2	Language	6	11	6.19	3.02	0.71	1.62
	3	Writing	2	10	3.28	2.35	0.85	0.90

Table K-4. Reliabilities by Reporting Categories, Grade, and Content Area—Mathematics

Grade	Item Reporting Category	Label	Number of Items	Raw Score			Alpha	SEM
				Maximum	Mean	Standard Deviation		
3	1	Operations and Algebraic Thinking	13	15	8.20	4.29	0.83	1.76
	2	Number and Operations in Base Ten	6	8	3.83	2.40	0.77	1.15
	3	Number and Operations-Fractions	7	9	4.87	2.66	0.75	1.33
	4	Measurement and Data	10	12	5.80	3.27	0.80	1.46
	5	Geometry	4	4	1.83	1.25	0.51	0.87
4	1	Operations and Algebraic Thinking	7	11	5.52	2.94	0.72	1.55
	2	Number and Operations in Base Ten	7	11	6.02	2.98	0.73	1.54
	3	Number and Operations-Fractions	13	16	8.22	4.37	0.84	1.73
	4	Measurement and Data	8	11	4.62	2.66	0.66	1.56
	5	Geometry	5	5	3.08	1.57	0.68	0.88
5	1	Operations and Algebraic Thinking	5	8	3.23	1.81	0.41	1.39
	2	Number and Operations in Base Ten	13	16	7.81	4.44	0.83	1.81
	3	Number and Operations-Fractions	11	14	6.29	3.41	0.73	1.76
	4	Measurement and Data	7	11	4.86	2.76	0.71	1.49
	5	Geometry	4	5	2.48	1.44	0.54	0.97
6	1	Ratios and Proportional Relationships	7	11	4.27	3.08	0.76	1.50
	2	The Number System	8	11	5.40	3.02	0.69	1.69
	3	Expressions and Equations	12	16	6.54	4.16	0.83	1.71
	4	Geometry	5	8	2.28	2.15	0.54	1.46
	5	Statistics and Probability	8	8	3.95	1.98	0.59	1.26
7	1	Ratios and Proportional Relationships	7	11	4.36	2.80	0.64	1.69
	2	The Number System	8	11	3.63	3.10	0.71	1.67
	3	Expressions and Equations	12	13	5.34	3.43	0.82	1.46
	4	Geometry	5	8	2.14	1.97	0.53	1.35
	5	Statistics and Probability	8	11	3.62	2.62	0.70	1.43
8	1	Number System & Expressions/Equations	17	20	8.19	4.61	0.79	2.12
	2	Functions	7	11	4.36	2.85	0.68	1.62
	3	Geometry	12	16	5.68	4.00	0.83	1.66
	4	Statistics and Probability	4	7	3.86	2.11	0.58	1.37