2025 RICAS Technical Report

Prepared by Cognia and the

Rhode Island Department of Education

1/30/2026

cognia

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

Table of Contents

List of Tables

Chapter 1. Introduction to the Rhode Island Comprehensive Assessment System

Chapter 1 provides an overview of the purpose and organization of this report, including the comparison between the Massachusetts Comprehensive Assessment System (MCAS) and the Rhode Island Comprehensive Assessment System (RICAS). Also provided are updates for the 2025 administrations, the intended interpretations and uses of the RICAS test scores, and the framework for constructing the validity arguments in this report.

1.1 PURPOSES OF THE RICAS AND THIS REPORT

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

- to provide information to parents/guardians and students on Rhode Island student achievement on the state's ELA and mathematics content standards,

- to provide information to support program evaluation and improvement at the school and district level, and

- to provide academic achievement and growth information used as part of the state's school accountability program to inform parents/guardians and the public about the performance of Rhode Island schools.

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 In high school, the PARCC tests have been replaced by the SAT.

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

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

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

The main purpose of this 2025 RICAS Technical Report is to document the technical quality and essential design characteristics of the 2025 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).

1.2 ORGANIZATION OF THIS REPORT

This report provides information regarding the spring 2025 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 2025 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 2025 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	Identical
Administration platform	Identical
<u>Scoring</u> Machine-scored items Hand-scored items	Identical Identical
Psychometric procedures	Identical
<u>Reporting</u> Scaled scores Achievement levels	Identical 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 Arguments Supporting Intended Interpretations and Uses of Test Scores – 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 – Achievement Level Distributions

- Appendix D – Sample Reports

- Appendix E – Item-Level Classical Statistics

- Appendix F – Score Distributions

- Appendix G – Differential Item Functioning Results

- Appendix H – 2024–25 MCAS Equating Report

- Appendix I – Reliability

1.3 UPDATES FOR THE 2025 ADMINISTRATION

In 2025, RICAS was transitioned to new online testing, scoring, and reporting platforms. Administration and reporting were handled through eMetric's Lighthouse suite, with educators using the RICAS Portal and student testing in the RICAS kiosk. Scoring was completed by Measurement Inc. (MI) As part of the transition into a new contract, the Braille and screen reader forms were a reuse of the 2024 forms. In 2025, RICAS began administering bilingual Spanish/English forms in place of the previous Spanish only forms. To ensure a smooth transition, Cognia and its subcontractors completed quality assurance checks throughout the process. Quality assurance checks are detailed throughout each chapter of this report.

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.

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

1.4.1 Intended Interpretations of the RICAS Test Scores

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

1.4.2 Intended Uses of the RICAS Test Scores

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

- Parents/guardians and students can use test scores and their underlying interpretations to monitor academic achievement and participate in decisions regarding student learning to support student growth.

- Educators can use test scores and their underlying interpretations to support curricular planning and identify instructional needs at both the classroom and individual student level.

- School- and district-level administrators can use test scores and their underlying interpretations to support program evaluation and improvements at the school and district levels.

- State-level administrators can use test scores and their underlying interpretations to monitor academic achievement and growth as required by state accountability programs and inform parents/guardians and the public of schools' performances on these metrics.

- Federal administrators can use test scores and their underlying interpretations to verify that ESSA federal accountability requirements are met.

1.4.3 Validation Arguments for RICAS

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

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

Figure 1-1 Logic of Validity Arguments for Tests
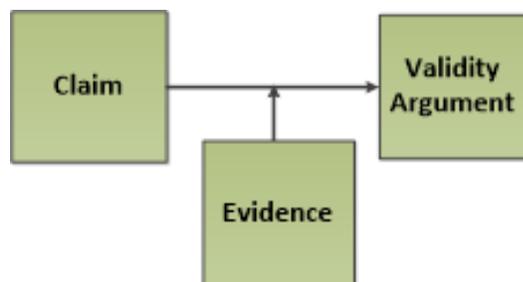
The phrase "intended score interpretations for uses" appears several times in the Standards for Educational and Psychological Testing ("Standards" for short in the following chapters/sections; AERA et al., 2014) and is the core of the field's views on validity and validation.

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

The Standards also provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. These sources address different aspects of supporting evidence for validity arguments but are not considered distinct types of validity. Instead, each contributes to a body of evidence about the individual validity arguments and overall arguments for the validity of intended score interpretations and uses. Moreover, these sources represent only a partial list of potential sources of evidence that informed RICAS design, development, test administration, analysis, and reporting processes that are relevant to the overall validity arguments for intended interpretations and uses of RICAS test scores and related information. Hence, this document will use Chapelle's (2020) framework based on Kane's (2013) work.

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

1) Description Inference: Items sample from the target domain appropriately such that high quality forms can be produced. (Domain to Item)

2) Evaluation Inference: Forms sample from items appropriately such that observed scores reflective of the domain can be produced. (Item to Form)

3) Generalization Inference: Observed scores on individual forms are reliable such that they are reflective of expected scores across forms. (Form to Score)

4) Explanation Inference: Expected scores are associated with classification cuts such that classification decisions are interpretable. (Score to Interpretation)

5) Utilization Inferences: Interpretations of scores and classifications are used as intended and only in ways considered appropriate and fair. (Interpretation to Use)

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

Figure 1-2 Chapelle's (2020) Framework: The Arguments and the Inferential Steps
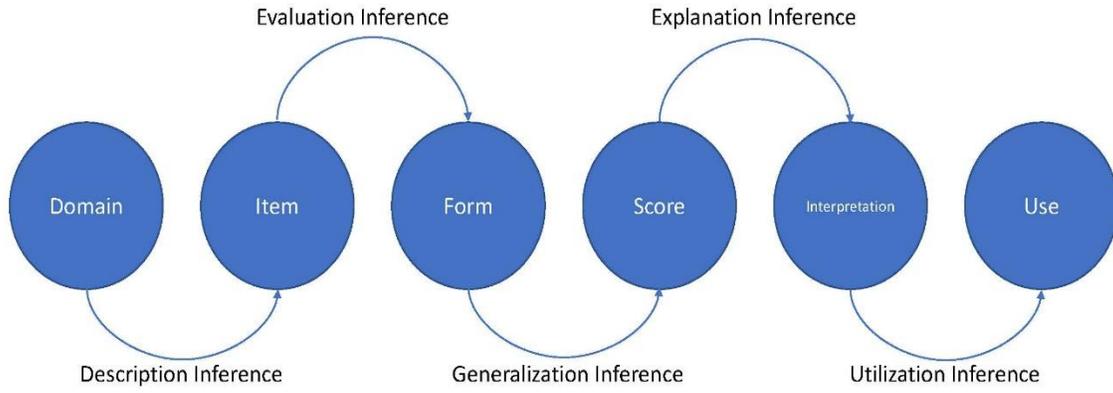
Chapter 2. Test Design and Development

There were no changes in test design or development for the 2025 administration of the RICAS program. The adherence to previous years' blueprints allows for defensible comparisons of where students are relative to grade-level expectations as outlined in the grades 3–8 ELA and mathematics standards despite COVID-related learning disruptions. Chapter 2 is primarily drawn from MCAS technical reporting and relates to the RICAS use of those assessments.

2.1 Appropriateness of Using Massachusetts Standards

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

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

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

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

2.2 Content Standards

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

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

2.3 Performance Standards

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

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

2.4 ELA

2.4.1 ELA Passage Selection

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

- be of interest to and appropriate for students in the grade being addressed;

- have a clear beginning, middle, and end;

- contain appropriate content;

- support the development of a sufficient number of unique assessment items; and

- be free of bias and sensitivity issues.

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

- **Literary passages**—Literary passages represent a variety of genres: poetry, drama, fiction, biographies, memoirs, folktales, fairy tales, myths, legends, narratives, diaries, journal entries, speeches, and essays. Literary passages are not necessarily fictional passages.

- **Informational passages**—Informational passages are reference materials, editorials, encyclopedia articles, and general nonfiction. Informational passages are drawn from a variety of sources, including magazines, newspapers, and books.

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

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

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

2.4.2 ELA Item Design

The 2025 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

- o Key Ideas and Details (Standards 1–3)

- o Craft and Structure (Standards 4–6)

- o Integration of Knowledge and Ideas (Standards 7–9)

- Anchor Standards for Language

- o Conventions of Standard English (Standards 1 and 2)

- o Knowledge of Language (Standard 3)

- o Vocabulary Acquisition and Use (Standards 4–6)

- Anchor Standards for Writing

- o Text Types and Purposes (Standards 1–3)

- o Production and Distribution of Writing (Standards 4–6)

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

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

Reporting Category	% of Points at Each Grade (+/-5%) 3	% of Points at Each Grade (+/-5%) 4	% of Points at Each Grade (+/-5%) 5	% of Points at Each Grade (+/-5%) 6	% of Points at Each Grade (+/-5%) 7	% of Points at Each Grade (+/-5%) 8
Language	25 (27)	25 (20)	25 (25)	25 (24)	25 (24)	25 (22)
Reading	65 (64)	65 (70)	55 (58)	55 (56)	55 (56)	55 (58)
Writing	10 (09)	10 (09)	20 (17)	20 (20)	20 (20)	20 (20)
Total	100	100	100	100	100	100

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

Table 2-2 ELA Item Types and Score Points

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

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

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

- Level I (Identify/Recall)—Level I items require that the student recognize basic information presented in the text. Examples of skills at this level include identifying main ideas/facts/details; recalling and locating details; identifying genre or setting; and identifying definitions, parts of speech, or functions of punctuation. Key words include identify, list, match, recognize, describe, and distinguish.

- Level II (Infer/Analyze)—Level II items require that the student understand a given text by making inferences and drawing conclusions related to the text. Examples of skills at this level include understanding the whole text (Big Picture)/generalizing; interpreting, making connections, visualizing, and forming questions; explaining a character's role/motives; determining whether an

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

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

Each cognitive level is represented in the ELA tests.

2.4.3 ELA Form Design

All items are coded to ELA framework standards. There are no stand-alone items on the tests; all vocabulary, grammar, and mechanics questions are associated with a passage set; more details on the number of items at different grades are 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 2025 ELA test, by grade.

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

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

Table 2-4 Distribution of ELA Common and Matrix Items by Grade and Item Type—Paper-based Test (PBT)[1]

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

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

2.4.4 ELA Reference Materials

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

2.5 MATHEMATICS

2.5.1 Mathematics Item Design

The 2025 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

- o Ratios and Proportional Relationships
- o The Number System
- o Expressions and Equations
- o Geometry
- o Statistics and Probability
- Domains for grade 8
  - o The Number System
  - o Expressions and Equations
  - o Functions
  - o Geometry
  - o Statistics and Probability

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

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

Domain	% of Points at Each Grade (+/-5%) 3	% of Points at Each Grade (+/-5%) 4	% of Points at Each Grade (+/-5%) 5
Operations and Algebraic Thinking	30 (31)	20 (20)	15 (15)
Number and Operations in Base Ten	15 (17)	20 (19)	30 (30)
Number and Operations – Fractions	20 (19)	30 (30)	25 (26)
Geometry	10 (8)	10 (11)	10 (11)
Measurement and Data	25 (25)	20 (20)	20 (18)
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	% of Points at Each Grade (+/-5%) 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 (39)
Functions	20 (20)
Geometry	30 (30)
Statistics and Probability	10 (11)
Total	**100**

The 2025 mathematics tests included several item types, as shown in Table 2-8.

Table 2-8 Mathematics Item Types and Score Points

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

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

- **Level I (Recall and Recognition)**—Level I items require that the student recall mathematical definitions, notations, simple concepts, procedures, and apply common, routine procedures or algorithms (that may involve multiple steps) to solve a well-defined problem.

- **Level II (Analysis and Interpretation)**—Level II items require that the student engage in mathematical reasoning beyond simple recall in a more flexible thought process, and in enhanced organization of thinking skills. These items require a student to make a decision about the approach needed, to represent or model a situation, or to use one or more non-routine procedures to solve a well-defined problem.

- **Level III (Judgment and Synthesis)**—Level III items require that the student perform more abstract reasoning, planning, and evidence-gathering. To answer questions of this cognitive level, a student must engage in reasoning about an open-ended situation with multiple decision points, represent or model unfamiliar mathematical situations, and solve more complex, non-routine, or less well-defined problems.

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

2.5.2 Mathematics Form Design

Form Design by Grade

Grade 3

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

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

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

Grades 4–6

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

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

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

Grades 7–8

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

The matrix portion included five 1-point selected-response or short-answer items, one 2-point selected-response or short-answer 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 SR/MS/SA/TE (1 pt.)	Common SR/MS/SA/TE (2 pt.)	Common CR (3 pt.)	Common CR (4 pt.)	Common Totals # (pt.)	Matrix SR/MS/SA/TE (1 or 2 pt.)	Matrix CR (3 or 4 pt.)	Matrix Totals # (pt.)
3	1	36	0	4	0	40 (48)	3	1	4 (6)
4	1	34	2	0	4	40 (54)	5	1	4 (8–10)
5	1	34	2	0	4	40 (54)	5	1	4 (8–10)
6	1	34	2	0	4	40 (54)	5	1	4 (8–10)
7	1	34	2	0	4	40 (54)	6	2	6 (13–15)
8	1	34	2	0	4	40 (54)	6	2	6 (13–15)

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

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

2.5.3 Mathematics Reference Materials

Rulers were provided to students in grades 3–8. Handheld rulers were provided to students taking the paper version of the mathematics test. Students taking the computer-based mathematics test had access

to two separate computer-based rulers: a centimeter ruler and a 1/8-inch ruler; students were not permitted to use handheld rulers on the computer-based test.

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

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

2.6 ITEM AND TEST DEVELOPMENT PROCESS

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

Table 2-11 Overview of Item and Test Development Process

Phase	Development Step	Detail of the Process	Cognia Test Developers	DESE Test Developers	ADC	BSC	External Experts
Initial Item Design	Selection of reading passages (for ELA only)	Cognia's test developers find potential passages and present them to DESE for initial approval; DESE-approved passages go to Assessment Development Committees (ADCs) composed of experienced educators, and then to a Bias and Sensitivity Committee (BSC) for review and recommendations. ELA items are not developed until passages have been reviewed by an ADC and a BSC. With the ADC and BSC recommendations, DESE makes the final determination as to which passages will be developed and used on a future MCAS/RICAS test.	X	X	X	X	
Initial Item Design	Development of initial item versions	Cognia's test developers generate items and edit items from subcontractors that are aligned to Massachusetts standards and specifications.	X				
Item Review and Refinement	Review of initial item versions by DESE and educators	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.	X	X	X	X	
Item Review and Refinement	Revision of initial items based on review	DESE test developers edit & revise items based on recommendations from ADC & BSC.	X	X			
Item Review and Refinement	Independent expert review of revised items	Experts from higher education and practitioners review all field-tested items for content accuracy. Each item is reviewed by at least two independent expert reviewers. Comments and suggested edits are provided to DESE staff for review.	X	X			X
Operational Field-Testing	Benchmark paper selection for CR and essay scoring	DESE and Cognia test developers meet to determine appropriate benchmark papers for training of scorers of field-tested constructed-response items and essays. Scoring rubrics and notes are reviewed and edited during benchmarking meetings. During the scoring of field-tested items, Cognia contacts DESE test developers with any unforeseen issues.	X	X			
Operational Field-Testing	Item performance review based on data from field tests	ADCs review field-test statistics and recommend items for the common-eligible status, for re-field-testing (with edits, for mathematics, since ELA is passage-based), or for rejection. BSC also reviews items and recommends items to become common-eligible or to be rejected.	X	X	X	X	
Final Form Construction	Test form construction I: Proposition of items for operational use	DESE provides target performance-level cut scores to Cognia's test developers. Cognia proposes sets of common items (items that count toward student scores) and matrix items. Matrix items consist of field-test and equating items, which do not count toward student scores. Each common set of items is delivered with proposed cut scores, including test characteristic curves (TCCs) and test information functions (TIFs).	X	X			
Final Form Construction	Test form construction II: Selection of items for operational use	DESE test developers and editorial staff review and edit proposed sets of items. Cognia and DESE test developers and editorial staff meet to review edits and changes to tests. Psychometricians are available to provide statistical information for changes to the common form. Approved common-eligible items become part of the common item set and are used to determine individual student scores.	X	X			
Public Release	Public release of select common items	Approximately 50% of common items in grades 3–8 are released to the public, and the remaining items are returned to the common-eligible pools to be used on future MCAS/RICAS tests. An item description (a statement specifying the content of the item) is released for each common item (both released and non-released).	X	X			

2.6.1 Item Review and Refinement

DESE Initial Item Review

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

- **Alignment**: Are the items aligned to the standards?

- **Content**: Is the content accurate? Does the item elicit a response that shows a depth of understanding of the content area?

- **Contexts**: Are contexts grade-level appropriate? Are they realistic? Are they interesting to students?

- **Grade-level appropriateness**: Are the content, language, and contexts appropriate for the grade level?

- **Creativity**: Does the item demonstrate creativity regarding approaches to items and contexts?

- **Distractors**: Have the distractors for selected-response items been chosen based on plausible construct-related errors? What are the distractor rationales?

- **Mechanics**: How well are the items written? Are they grammatically correct? Do they follow the conventions of item writing? Is the wording grade-level appropriate and accessible for all students?

- **Technology**: Are the items scored appropriately? Is the item making the best use of the technology? Is there another type of item that is more appropriate?

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

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

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

ADC Passage Review (ELA Only)

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

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

---

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

ADC Item Review

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

BSC Passage and Item Review

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

External Content Expert Item Review

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

Editing of Recommended Items

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

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

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

2.6.2 Operational Field-Testing of Items

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

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

Scoring of Field-Tested Items

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

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

Data Review of Field-Tested Items

Data Review by DESE

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

Data Review by ADCs

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

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

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

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

Data Review by BSCs

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

- Female compared with male [gender/sex]

- African American/Black compared with White [ethnicity I]

- Hispanic or Latino/a compared with White [ethnicity II]

- Current or former ELLs compared with non-ELLs [language status]

2.6.3 Final Form Construction

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

- **Content coverage/match to test design and blueprints**. The test designs and blueprints stipulate a specific number of items per item type and per reporting category for each content area. A broad coverage of standards and cognitive skills is expected. The previous year's common test should also be considered, and items should not be duplicated.

- **Item difficulty and complexity**. Item statistics drawn from the data analysis of items are used to ensure similar levels of difficulty and complexity from year to year as well as high-quality psychometric characteristics. Items can be "reused" if they have not been released and not used the previous year. When an item is reused in the common portion of the test, the latest usage statistics accompany that item.

- **"Clueing" items**. Items are reviewed for any information that might "clue" or help the student answer another item.

- **Item types**. A variety of item types, including approximately 20–30% technology-enhanced items, are selected to populate the common slots.

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

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

- **Key patterns**. The sequence of keys (correct answers) is reviewed to ensure that the key order appears random.

- **Option balance**. Items are balanced across forms so that each form contains a roughly equivalent number of key options (As, Bs, Cs, and Ds).

- **"Clueing" items**. Items are reviewed for any information that might "clue" or help the student answer another item.

- **Item types**. A variety of item types should populate the matrix slots.

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

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

2.6.4 Special Edition Test Forms

Students with Disabilities

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

- **Paper**—Form 1 of the operational CBT was produced to appear on paper. Items that used interactions not assessable on paper (typically technology enhanced items) were replaced with items that asked similar questions in a paper assessable manner. The grades 3–8 tests were administered to most students on the computer and to some students with accommodations on a paper form.

- **Large-print**—Form 1 of the operational test was translated into a large-print edition. The large-print edition contains all common and matrix items found in the paper form.

- **Braille**—This form included only the common items found in the operational test with the following characteristics:

  o 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.

  o 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.

  o 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).

For 2025, the Braille and screen reader forms were a reuse of the 2024 forms.

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 2025 RICAS with each accommodation.

Spanish Edition Test Forms

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

2.7 QUALITY CHECK

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

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

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

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

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

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

Evidence: Subsections 2.4.2 for ELA and 2.5.1 for mathematics explicitly state that items across all grades within those content areas "measured the . . . learning standards as articulated in the Rhode Island Core Standards," detailing the specific standards addressed by items available for RICAS assessments. Subsections 2.4.1 and 2.4.2 describe passage types, item types, and cognitive levels for items on the ELA assessments. Subsection 2.5.1 describes the item types and cognitive levels for items on the mathematics assessments. Subsection 2.6.1 describes item review and refinement procedures, and subsection 2.6.2 describes item field-testing and subsequent scoring and review processes. Together, these subsections describe an overall process of item development that ensures items effectively target the expected knowledge and abilities of the grades being assessed.

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

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

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

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

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

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

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

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

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

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

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

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

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

Chapter 3. Test Administration

Rhode Island test administrations were standardized, and the design was not significantly changed from previous years.

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 2020–2021and subsequent administrations. For more information, consult Appendix B of the 2021, 2022, 2023 and 2024 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 2025, 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	3/12	3/24–4/25	4/28
Mathematics	5/27	4/21	4/28–5/23	5/27

Testing Times

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

Table 3-2 ELA Recommended Testing Times, Grades 3−8

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

Table 3-3 shows the recommended testing times for the 2025 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 Recommended Testing Time (min)	Session 2 Recommended Testing Time (min)	Total 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 2025 Test Coordinator's Manual* (TCM), *Grades 3–8* and the *2025 Test Administrator's Manuals* (TAMs). In spring 2025, 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, 2023, 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, 2025* for details on how ELL students participate in spring 2025 RICAS.

- Students with significant cognitive disabilities who are eligible for the alternate assessment, the DLM assessment. For more information, refer to the DLM page of the RIDE website: https://www.ride.ri.gov/InstructionAssessment/Assessment/DLMAssessments.aspx.

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

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

3.4 SPANISH EDITION TEST FORMS

Spanish-Speaking Students

Spanish editions of the spring grades 3–8 mathematics tests were available to any ELL student with a low level of English proficiency who was receiving or had received mathematics instruction in Spanish. The Spanish edition of the grades 3–8 mathematics tests contained all common and matrix items found in Form 1 of the operational test. Beginning with the Spring 2025 administration, this is a bilingual form that included the content presented in both Spanish and English in a stacked presentation.

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

- A set of translation rules or parameters was generated, taking the following into consideration: vocabulary, usage, and consistency over the years. These rules were provided to both translators.

- The first translator translated from English to Spanish. The second translator proofread the work of the first translator. Discrepancies between the two translations were resolved by the first translator.

- The Publishing Department reviewed the graphics in Spanish.

- The script that the teacher read when administering the test was also translated into Spanish and was included as Appendix A of the *Test Administrator's Manual*.

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

3.5 ADMINISTRATION PROCEDURES

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

- understanding and enforcing the test security requirements and test administration protocols

- ensuring that students participate in testing according to the requirements in section 3.3 of this report

- coordinating the school's test administration schedule and ensuring that tests are scheduled during the prescribed testing window, and in the prescribed order

- ensuring that accommodations are properly administered and that transcriptions, if required for any accommodation, are properly completed

- completing the Principal's Certification of Proper Test Administration (PCPA) and ensuring the accuracy of information provided on the form

- providing RIDE with the school's correct contact information

- ensuring that all students have access to the appropriate grade level reference materials for the mathematics assessment as referenced in Chapter 2, subsection 2.5.3

- ensuring that all students who require supplemental supports (mathematics supplemental reference sheets or ELA graphic organizers) have access to them during testing

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

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

3.6 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING TEST ADMINISTRATION

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

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

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

Chapter 4. Scoring

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

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

4.1 PREPARATION OF STUDENT RESPONSE BOOKLETS

Scoring of the 2025 RICAS tests was conducted by Measurement Inc (MI).

For paper-based tests, MI scanned each RICAS student-answer booklet for uploading into their 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.

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 MI for review in advance of scoring. Content specialists at MI 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, MI, and DESE collaborated on the establishment of final scoring decisions.

- Weekly meetings between the Cognia and MI 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 in the Lighthouse platform. Student responses with multiple marks (possible only on paper-based tests) and blank responses were assigned zero points.

4.5 HAND-SCORED ITEMS

Measurement Incorporated (MI) served as the primary scoring vendor for the Spring 2025 RICAS administration, using its experienced human scoring staff and the Project Essay Grade (PEG) automated scoring engine. Two approaches were employed depending on the content area. Responses to mathematics and ELA constructed-response (CR) items were scored exclusively by trained human raters, ensuring consistency with the constructs measured in those subjects. ELA-writing responses, by contrast, were scored through a hybrid model that combined automated scoring with human review.

This design aligned scoring practices with the unique demands of each content area while meeting DESE's expectations for validity, reliability, and timely reporting. In ELA–writing, the hybrid process was applied to all items. As part of pre-processing, student responses were screened to identify non-scorable

or special cases (e.g., off-topic responses, alert/crisis content). These responses bypassed the automated system and were routed directly to human raters for scoring. The engine score serves as the score of record for the 90% highest-confidence responses. The remaining 10% lowest confidence-responses are routed to expert human scorers who assign the score of record.

Table 4-1 provides an overview of the items scored during the Spring 2025 administration, including the content area, score point range, and the scoring method applied.

Table 4-1 Items Scored During the Spring 2025 Administration

Content Area	Item ID	Grade	Number of Traits	Score Point Range	Scoring Method
ELA	EL305885122	3	1	0-4	human
ELA	EL311063921	3	2	0-3/0-4	hybrid
ELA	EL307375298	4	1	0-4	human
ELA	EL208631509	4	2	0-3/0-4	hybrid
ELA	EL736478825	5	2	0-3/0-4	hybrid
ELA	EL307267042	5	2	0-3/0-4	hybrid
ELA	EL315605273	6	2	0-3/0-5	hybrid
ELA	EL314484549	6	2	0-3/0-5	hybrid
ELA	EL912346409	7	2	0-3/0-5	hybrid
ELA	EL306400886	7	2	0-3/0-5	hybrid
ELA	EL206813674	8	2	0-3/0-5	hybrid
ELA	EL306640647	8	2	0-3/0-5	hybrid
Mathematics	MA260584A	3	1	0-3	human
Mathematics	MA735951978	3	1	0-3	human
Mathematics	MA261857A	3	1	0-3	human
Mathematics	MA134930940	3	1	0-3	human
Mathematics	MA800780887	4	1	0-4	human
Mathematics	MA303335	4	1	0-4	human
Mathematics	MA231875780	4	1	0-4	human
Mathematics	MA803957909	4	1	0-4	human
Mathematics	MA232642946	5	1	0-4	human
Mathematics	MA715102462	5	1	0-4	human
Mathematics	MA903746975	5	1	0-4	human
Mathematics	MA802371654	5	1	0-4	human
Mathematics	MA900378821	6	1	0-4	human
Mathematics	MA800259417	6	1	0-4	human
Mathematics	MA233752432	6	1	0-4	human
Mathematics	MA287883	6	1	0-4	human
Mathematics	MA802907874	7	1	0-4	human
Mathematics	MA235900389	7	1	0-4	human
Mathematics	MA235905181	7	1	0-4	human
Mathematics	MA717248260	7	1	0-4	human
Mathematics	MA203085405	8	1	0-4	human
Mathematics	MA236558384	8	1	0-4	human
Mathematics	MA311433	8	1	0-4	human
Mathematics	MA800473031	8	1	0-4	human

The 2025 RICAS included a range of constructed-response and essay items, in addition to selected-response and short-answer questions. These open-ended items were scored either entirely by human raters or through the hybrid model, depending on content area and item type.

Constructed-response items in ELA (grades 3 and 4) were scored on a 0–3 scale. In mathematics, constructed-response items were scored on a 0–3 scale in grade 3 and a 0–4 scale in grades 4–8. Essay items were scored for two traits: Idea Development and Language Conventions. For Idea Development, the score range was 0–4 in grades 3–5 and 0–5 in grades 6–8. For Language Conventions, the score range was 0–3 across all grades.

For responses that fell outside the parameters of the scoring rubric, the following condition codes were reported instead of a numerical score:

- Blank: Response box is completely free of text and punctuation.

- Non-English: The response is written in a language other than English or is a mix of English and another language but lacks sufficient English to provide a score.

- Off Topic (OT): A response that is not related to the task/prompt administered or is not a valid attempt at responding to any task/prompt on the assessment.

- Repeat the Prompt (RTP): The response copies all or a portion of the prompt/stimuli materials verbatim and offers no attempt to respond to the task/prompt.

4.5.1 Scoring Plan

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

- All scoring was conducted, applying a virtual/synchronous scoring model maintaining stringent quality control measures applied in a center-based, regional scoring environment.

- Prior to the start of scoring, scorers receive a system requirement document.

- 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 and/or via pre-recorded training module or a recording of live training.

- Scoring directors brought raters back into their team training meetings for 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.

The scheduling and communication used to coordinate MI's scoring services for the 2025 RICAS was managed by the Cognia Scoring Project Manager. The Cognia Scoring Group Manager also facilitated the transition by providing training materials and consultation, as needed, to ensure a successful transition to MI and a consistent approach to scoring.

MI provided an experienced leadership team for the scoring of performance-assessment items for the 2025 RICAS.

The following MI staff members were involved with the scoring of the 2024–2025 RICAS responses:

- The *Director, Performance Assessment Scoring* provided guidance, direction, and leadership to Scoring Services.

- The *Project Manager* was responsible for the communication and coordination between Scoring Services and Cognia.

- The *Scoring Manager* served as the liaison between Scoring Services and the Remote Staffing Team to facilitate staffing assignments and handle all matters of logistics.

- The *Scoring Content Lead* facilitated the training of scoring leadership and oversaw all scorer training sessions. They guaranteed the quality and consistency of scoring across all test administrations and items and the timely completion of scoring.

- The *Scoring Director* was responsible for supervising Scorers and Scoring Team Leaders, and for ensuring quality scoring across all test administrations and items.

- The *Team Leaders (TLs)* provided support and direction to scorers on quality and accuracy.

4.5.2 Rater Recruitment and Qualifications

MI's Remote Staffing Team served as the recruiting organization for temporary associates who were onboarded for scoring RICAS. To ensure that scorers consistently produce accurate and consistent scores, MI recruited raters with experience successfully scoring other large-scale assessments. MI assigned those raters to the grade level, subject area, and item type for which they were most qualified based on their performance on similar projects. Returning raters were selected based on experience and performance, as well as attendance, punctuality, and cooperation with work procedures and MI policies. MI maintains evaluations and performance data for all staff who work on each scoring project to determine employment eligibility for future projects. Finally, MI targeted recruitment of new raters as needed, to continue to identify talent across the country that will best fulfill the hand-scoring requirements.

All raters possessed, at a minimum, a four-year college degree. MI collected proof of degree for all raters as a condition of employment. Those scoring high school students' responses must have at least a 4-year degree and must also have either 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. A summary of scorers' backgrounds is provided in Table 4-2.

Table 4-2 Summary of Rater and Scoring Leadership Backgrounds (Operational Scoring)

Education	# Raters	% Raters	# Leadership	% Leadership
Bachelor's degree or higher	239	62	14	58
Master's degree or higher	149	38	10	42
Teaching Experience	**# Raters**	**% Raters**	**# Leadership**	**% Leadership**
Certified or experienced	190	49	14	58
None	198	51	10	42
Scoring Experience	**# Raters**	**% Raters**	**# Leadership**	**% Leadership**
Less than a year	62	16	0	0
1–3 years	180	46	0	0
3+ years	146	38	24	0

All raters resided in the United States and properly completed Form I-9 to verify their identity and employment authorization. Raters' I-9 forms are retained on file as required by law and made available for inspection by authorized government officers as needed. MI is an equal opportunity employer and believes that a diverse workforce is of the utmost importance. When hiring, MI strives to ensure the work force is diverse across age, ethnicity, gender, and other demographics.

In selecting team leaders who will monitor the raters, MI scoring leadership reviewed records of all returning staff. They looked for experienced team leaders with a record of good performance on previous projects and considered raters recommended for promotion to the team leader position.

MI requires all hand-scoring project staff (scoring directors, team leaders, raters, and clerical staff) to sign a confidentiality/nondisclosure agreement before receiving any training or viewing any secure project materials. The employment agreement indicates that no participant in training and/or scoring may reveal information about the test, the scoring criteria, or the scoring methods to any person.

4.5.3 Scoring Platforms

The hand scoring of open-ended response items for RICAS took place in a locked-down kiosk version (MIRA) of MI's Virtual Scoring Center (VSC). VSC is a secure, centrally administered, password-protected, and HTML-based web application. VSC provides a proven, online, user-friendly interface to train and qualify scorers and allows them to accurately score open-ended response items.

VSC has the following capabilities, which can be configured to meet DESE's requirements.

- Defining scorer roles and qualifications based on training, security requirements, or prior history

- Managing and randomly routing responses to scorers, including double-blind second readings

- Allowing project leaders to spot-check scorers, monitor reliability, and offer feedback

- Allowing scorers to flag responses for a variety of reasons (nonscorable responses, alert responses)

- Generating status reports at any time during the project (progress reports, individual scorer performance reports, score distribution reports, and interrater reliability reports)

- Importing and exporting data to test delivery and reporting systems

Prior to the start of scoring, the Scoring Manager worked with MI's scoring technology team to configure VSC to score open-ended response items for RICAS. A Project Configuration Checklist (PCC) was used to identify open-ended response items and define the scoring rules for VSC. The PCC also identified data thresholds, rules for automated quality assurance processes, and who should be granted access to reports. In addition, Cognia and MI completed a user-acceptance process to ensure that student responses were being routed correctly and that VSC was operating as expected. As a result, our scoring team was able to accurately score the items and deliver the results on time.

Responses were sent from Cognia to MI and loaded into VSC using integrated web services. Scores were then exported from VSC and sent to Cognia using the same pipeline. VSC runs import and export routines several times a day. VSC prioritizes the oldest available responses in the queue so that these responses are scored first. This process facilitates scoring responses such that MI can provide results in a timely manner.

4.5.4 Leadership & Rater Training

Operational training materials were prepared by Cognia, with MI participating in all discussions to finalize materials. Table 4-3 outlines the training materials produced.

Table 4-3 Materials Developed for the RICAS

Material	Description
Anchor sets	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; it 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	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.
Qualification sets	Qualification 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 MCAS-approved scoring standards.

Leadership training was conducted prior to scorer training to ensure Scoring Directors were fully prepared for their responsibilities during the operational scoring window. Scoring Content Leads and the Scoring Manager trained the Scoring Directors on their assigned items, content areas, and grades.

Training was delivered at the item level and followed a structured process:

- Presenting an overview of the assessment program and training process

- Reviewing the student prompt, passage, and scoring rubric

- Conducting a guided review of anchor responses

- Analyzing each anchor response, its score, and the rationale

- Providing instruction on use of the VSC/MIRA scoring platform

- Completing an independent scoring activity using a practice set to simulate operational scoring

- Reviewing practice responses with discussion of scores and rationales

- Reinforcing scoring criteria with attention to differences between adjacent score points

- Completing two qualification sets, each containing 10 responses

- Receiving guidance on applying condition codes and identifying crisis or alert responses

- Accessing a resource library to support ongoing reference throughout scoring

All Team Leader and rater training sessions were conducted live via Microsoft Teams. Following the live training, raters completed practice and qualification activities in VSC, which also housed the training materials—such as rubrics, anchor sets, and other reference documents—used throughout the project. Each grade had two operational writing items, with one Scoring Director and two Team Leaders assigned per grade; raters and Team Leaders trained on the item for their grade. For reading, each grade had two operational items, with one Scoring Director and one Team Leader assigned per grade; raters and Team Leaders trained on both items.

Rater training began with a general orientation to RICAS, covering the goals and purposes of the program as well as any unique features of the test and population. Training for a specific item always began with a review of the scoring guide, which included the task, rubric, and scoring notes. All scoring guides had been approved by DESE during field-test benchmarking meetings and were used without modification.

As part of training, prospective raters reviewed three sets of student responses, some of which had been introduced during field-test training. Anchor sets were DESE-approved collections of typical responses at each score point. Practice sets included atypical or discussion-provoking responses to broaden raters' understanding. Qualification sets consisted of 10 clear, representative responses used to confirm raters' ability to apply the scoring criteria accurately.

Meeting the standard on an item's qualification set was a prerequisite for scoring. Raters were required to achieve at least 70% exact agreement with no non-adjacent scores on either of two qualification sets. For multi-trait ELA items, the accuracy threshold is applied separately to each trait.

These measures are quantified using specific metrics, such as percent exact agreement, quadratic weighted kappa (QWK), and standard mean difference (SMD), among others. The resulting data are used to generate real-time scorer performance statistics, which are integrated into an automated monitoring system that supports ongoing evaluation, retraining, and—when necessary—removal of raters to ensure scoring accuracy and consistency across the scoring pool.

Scorers and Scoring leadership received daily automated messages that summarized scorers' performance to ensure all stakeholders are aware of current performance and any issues that required attention. Scoring leadership could also access specialized reports that they used to identify scorers who require retraining and/or removal because of issues with accuracy or productivity.

Validity Responses

Validity responses selected by scoring leadership were automatically routed to each rater at a rate of 5% of the number of responses scored. All validity responses were deployed blindly—that is, they were indistinguishable from regular student responses and scored without the scorer's awareness of their status.

If validity data showed that a scorer had drifted in their application of the scoring criteria, that scorer was sent a message informing that their accuracy was either below average or unsatisfactory, that their Team Leader would be working with them to help improve their accuracy, and that they should proactively take some time to pause scoring and review their training materials. Any scorer whose validity results were judged "unsatisfactory" had their scores reset (voided) until such time as the scorer improved their accuracy. Scorers who failed to improve despite targeted feedback and one-on-one assistance from leadership were removed from the item(s) on which they were struggling, though they were allowed to continue scoring other items where their validity results were satisfactory.

Double-Blind Scoring

MI's scoring approach ensured that target agreement rates were maintained and the scoring performance for every hand-scored item met or exceeded expectations. Double-blind scoring is the practice that refers to a method whereby the same response is routed to two scorers. The response is independently and anonymously reviewed by each scorer. In double-blind scoring, scorers do not know which response will be (or already has been) scored by another randomly selected scorer. For this assessment, 10% of responses that were exclusively hand-scored were randomly selected and routed to a second scorer for independent review and scoring.

Read-Behind Scoring

Scoring leadership used results of interrater reliability (IRR) derived from double-blind scoring and validity results data as monitoring tools to determine which scorers needed retraining. This retraining was achieved using additional data collected from read-behind scoring. Scorers whose IRR and/or validity results indicated they were drifting from the criteria had their scored sets of responses captured and sent

to a read-behinds queue. Leadership reviewed the responses and the scores given by these scorers. Leadership looked for examples of mis-scored responses and returned these responses and the correct scores for the scorer to review. Leadership was also able to change incorrect scores during the read-behinds process without returning these responses to the scorers. This was done when there were many mis-scored responses that contained the same pattern of drift on the scorer's part; in this case, leadership sent the scorer a representative sample of responses to address the issue and changed scores on the remainder.

Voiding (Resetting)

The scoring system allowed MI to invalidate a rater's work. At the discretion of scoring leadership, if a scorer failed to maintain accuracy standards, his or her work for the impacted time frame might be invalidated. When scores were reset, the affected student responses were routed to other qualified raters for re-scoring.

Interrater Reliability

Interrater reliability (IRR) is a statistical measure used to evaluate the consistency of scores assigned by different human raters to the same student responses. High IRR indicates that scorers are applying the rubric in a consistent and replicable manner, which is essential for ensuring both fairness and validity in score-based decisions. For the Spring 2025 RICAS administration, MI evaluated IRR using four complementary metrics: Quadratic Weighted Kappa (QWK), which assesses the degree of agreement between raters while accounting for the magnitude of scoring differences; exact agreement, which reflects the percentage of instances where raters assigned the identical score; adjacent agreement, which reflects the percentage of instances where raters' scores differed by no more than one point; and Standardized Mean Difference (SMD), which evaluates the presence of systematic scoring severity or leniency across raters.

Summaries of the IRR results are presented in Tables 4-4 and 4-5 for ELA constructed-response, and mathematics responses, respectively[1]. These tables report the reliability statistics for operational items scored as part of the MCAS and RICAS assessments. Scoring is conducted jointly, with responses from RI and MA combined in a single pool rather than handled separately by state. Accordingly, responses from both states are included in the interrater reliability analyses. This joint approach provides a larger pool of double-scored responses, which yields more stable and precise estimates of interrater consistency (e.g., QWK, percent agreement, SMD).

Table 4-4 Summary of Interrater Reliability Statistics—ELA Constructed-Response

Item ID	Grade	Score Points	Double-Scored Responses	QWK	Percent Exact	Percent Adjacent	SMD
EL305885122	3	0–3	7367	0.6661	61.2	35.6	-0.0448
EL307375298	4	0–3	7131	0.6106	57.1	37.0	0.1853

---

[1] *Note IRR is not reported for ELA-writing responses because this administration did not include double human scoring of ELA-writing, and the hybrid scoring design yields no paired human–machine scores; without such pairs, interrater reliability cannot be estimated.*

Table 4-5 Summary of Interrater Reliability Statistics—Mathematics

Item ID	Grade	Score Points	Double-Scored Responses	QWK	Percent Exact	Percent Adjacent	SMD
MA260584A	3	3	7371	0.9342	88.0	11.6	0.0015
MA735951978	3	3	7462	0.9113	81.3	17.8	0.0012
MA261857A	3	3	7489	0.9383	86.8	12.8	0.0070
MA134930940	3	3	7432	0.9416	88.8	10.8	0.0041
MA800780887	4	3	7298	0.9229	77.8	21.0	0.0104
MA303335	4	3	7299	0.8757	70.6	26.4	-0.0037
MA231875780	4	3	7299	0.9439	85.1	14.4	0.0014
MA803957909	4	3	7266	0.8545	70.3	27.6	-0.0001
MA232642946	5	4	7429	0.9581	84.7	14.2	0.0154
MA715102462	5	4	7423	0.9157	77.0	20.6	0.0249
MA903746975	5	4	7429	0.8887	70.9	25.2	-0.0223
MA802371654	5	4	7445	0.9299	82.1	16.2	0.0063
MA900378821	6	4	7459	0.9402	86.7	11.8	0.0080
MA800259417	6	4	7430	0.9364	83.5	14.4	-0.0060
MA233752432	6	4	7491	0.9767	93.0	6.6	0.0010
MA287883	6	4	7450	0.9475	80.3	18.0	-0.0196
MA802907874	7	4	7324	0.9551	87.0	12.2	0.0007
MA235900389	7	4	7473	0.9539	88.9	10.8	0.0058
MA235905181	7	4	7404	0.9528	88.5	10.6	0.0031
MA717248260	7	4	7408	0.9407	86.3	12.8	0.0032
MA203085405	8	4	7549	0.8497	72.2	25.0	-0.0154
MA236558384	8	4	7425	0.9413	81.6	15.8	-0.0414
MA311433	8	4	7361	0.892	72.9	25.0	0.0229
MA800473031	8	4	7450	0.9439	84.5	14.0	-0.0207

4.5.5 Automated Scoring

MI's PEG automated scoring technology was used to score responses to writing items as part of a hybrid human-machine scoring process. This section describes PEG, the training and validation sample and process, and the hybrid human-machine scoring process.

4.5.5.1 Project Essay Grade (PEG)

Figure 4-1 presents the architecture of MI's PEG engine. During engine training, this architecture allows PEG to generate hundreds of custom linguistic (rule-based) features, which are determined by codified English linguistic rules such as syntax and semantics and extracted from representative student responses. In addition to rule-based features, the engine also includes features extracted by Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA) procedures.

Figure 4-1 Engine Architecture
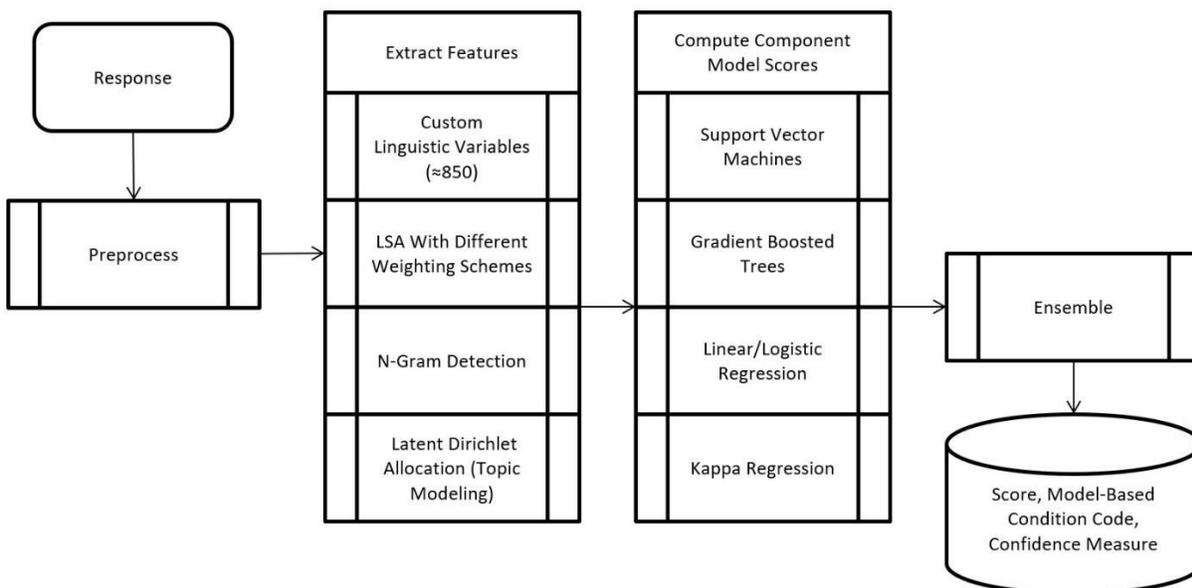
The scoring engine's item- and trait-specific scoring models use computed features from the training responses along with the scores assigned to them by human raters. Using hundreds of parameterizations across several machine-learning algorithms, via cross-validation and optimization, the engine determines which algorithms best predict the human-assigned scores. These algorithms draw on many of the latest advances in the field of machine learning to generate linear and non-linear classification and regression models. These approaches typically result in 100 candidate models for a single item or trait. The engine then uses an ensembling procedure to combine the best models into a robust final model. The ensembling procedure utilizes linear regression, where the objective is to maximize a continuous relaxation of QWK, thus maximizing the engine's agreement with the human raters.

4.5.5.2 Model Training and Validation

Training Data

The training data used to build automated scoring models for the ELA-writing items were drawn from field-test student responses. A target of 2,000 total responses per item, with at least 200 responses per score point, was established to ensure adequate training coverage and model stability.

To support model development, each item's response set was divided into the following:

- Approximately 85% of responses for the training set were used to develop the scoring model.

- Approximately 15% were used for the validation set, reserved to evaluate the model's accuracy.

Model Training

Component model training requires inputs of response "features." For items that assess writing quality (e.g., essays), the engine processes the responses and calculates approximately 850 linguistic variables that describe the responses in mathematical terms. These variables range in complexity from simple to highly complex. Examples of simple variables are measures such as word count or sentence length, word

choice and spelling errors, and the number and severity of grammatical errors. The most complex variables measure patterns that represent style, fluidity, smoothness of transitions, clarity of communication, and other sophisticated concepts.

To build an essay-scoring model, the engine examines the variables and text features of responses, correlates them with the human scores previously assigned, and identifies those features that have high predictive value. The engine then sends the features to hundreds of different algorithms that compete to see which algorithms best associate the features with the human-assigned scores. These algorithms draw on many of the latest advances in the field of machine learning to generate both linear and non-linear models. Examples of approaches used include Support Vector Machines, Gradient Boosted Trees, and various regression approaches.

Note that building component models for each item and trait prevents variables from being generalized across items or traits, allowing the engine to faithfully reproduce humans' application of the scoring rubrics. This means that the resultant models are reasonably robust to gaming attempts, as each represents a unique valuation of the item- or trait-specific text features similarly valued by professional raters.

The approaches described typically result in 100 models for a single item or trait. Ensembling is the process of selecting the "best of the best" models, to result in a small set of strong, yet dissimilar component models. A linear-kappa regression is used to determine the model ensembling weights. The more accurate a given model is, the more weight it carries in the final score decision.

Scoring a response involves first preprocessing the response. The purpose of preprocessing is twofold: (1) create raw and canonical representations of the response from which features can be extracted, and (2) filter out responses for which the scoring model does not apply (e.g., blank or insufficient responses). The response is then scored with the associated component models. A final score is produced performing a weighted sum using the ensembling weights.

Model Validation

Model validation was conducted at two levels: (1) a global evaluation of overall model performance against human scores, and (2) a subgroup-level evaluation to ensure the models did not unfairly advantage or disadvantage particular student groups.

For global validation, each model was applied to an independent validation set of student responses that had not been used for training. Because even two professional raters will not always agree on the exact score for a given response, model performance was judged relative to the level of agreement observed among human raters. A model was considered valid when its agreement with human scores was at least as strong as the agreement observed between independent human raters.

Evaluation followed the framework proposed by Williamson, Xi, and Breyer (2012)[2] and relied on several key metrics. Quadratic Weighted Kappa (QWK) measured agreement on numerical scores, with larger disagreements weighted more heavily. Standardized Mean Difference (SMD) measured differences in average scores (human – machine) relative to pooled standard deviation, capturing systematic tendencies for the model to over- or under-score. Evaluation criteria and thresholds are summarized in Table 4-6.

---

[2] *Williamson, D. M., Xi, X., Breyer, F. J. (2012). A framework for evaluation and use of automated scoring. Educational Measurement: Issues and Practice, 31(1), 2–13. https://doi.org/10.1111/j.1745-3992.2011.00223.x*

Table 4-6 Model Evaluation Criteria

Criterion	Threshold
Human – machine agreement	$QWK_{HM} \geq 0.65$
Degradation	$QWK_{HH} - QWK_{HM} < 0.10$
Human – machine standardized mean score difference	$|SMD_{HM}| < 0.15$

Table 4-7 presents model validation results, by model/trait, for models that passed all evaluation criteria and were deployed operationally.

Table 4-7 Model Validation Results

Grade	Item	Trait	HM *N*	HM QWK	HH QWK	Degradation	HM SMD
3	EL311063921	DEV	282	0.818	0.837	0.020	0.023
3	EL311063921	CONV	280	0.814	0.785	-0.029	0.008
4	EL208631509	DEV	292	0.827	0.819	-0.008	0.030
4	EL208631509	CONV	288	0.786	0.766	-0.019	0.024
5	EL307267042	DEV	290	0.806	0.831	0.025	0.010
5	EL307267042	CONV	283	0.835	0.820	-0.015	0.015
5	EL736478825	DEV	313	0.829	0.821	-0.008	0.035
5	EL736478825	CONV	303	0.808	0.801	-0.007	0.055
6	EL314484549	DEV	288	0.836	0.880	0.044	0.023
6	EL314484549	CONV	284	0.853	0.861	0.008	0.020
6	EL315605273	DEV	286	0.888	0.903	0.015	0.064
6	EL315605273	CONV	283	0.884	0.891	0.006	0.027
7	EL306400886	DEV	286	0.914	0.909	-0.005	0.003
7	EL306400886	CONV	284	0.894	0.907	0.012	0.042
7	EL310944009	DEV	300	0.904	0.956	0.052	0.050
7	EL310944009	CONV	296	0.935	0.957	0.022	0.025
8	EL206813674	DEV	289	0.918	0.910	-0.008	0.017
8	EL206813674	CONV	286	0.910	0.902	-0.008	0.053
8	EL306640647	DEV	290	0.797	0.891	0.094	0.116
8	EL306640647	CONV	284	0.873	0.854	-0.019	0.004

Note: DEV = Idea Development, CONV = Language Conventions.

In addition to global validation, MI also conducted local (subgroup-level) validation to evaluate potential bias across student groups. Subgroup analyses were conducted when at least 10 paired human–machine scores were available for a given subgroup within an item. A subgroup was flagged for potential bias if the absolute SMD was ≥ 0.125 and statistically significant at the 95% level, with a Bonferroni correction applied to adjust for multiple comparisons. An item model was flagged for further review if any subgroup comparison indicated potential bias.

This two-tiered validation approach—global model validation combined with subgroup-level bias analysis—ensured that models met both psychometric quality standards and fairness expectations before being approved for operational use.

Hybrid Scoring Process

Once models associated with a given item passed validation, subsequent student responses were scored using a hybrid human-automated scoring approach. Figure 4-2 shows the response routing rules under the hybrid scoring process. Upon initial receipt and validation of each response, MI routed responses for those items eligible for automated scoring to the engine and the remainder of the responses directly to the VSC hand-scoring system. After automated scoring, certain responses were then routed to the hand-scoring system.

Figure 4-2 Response Routing Rules
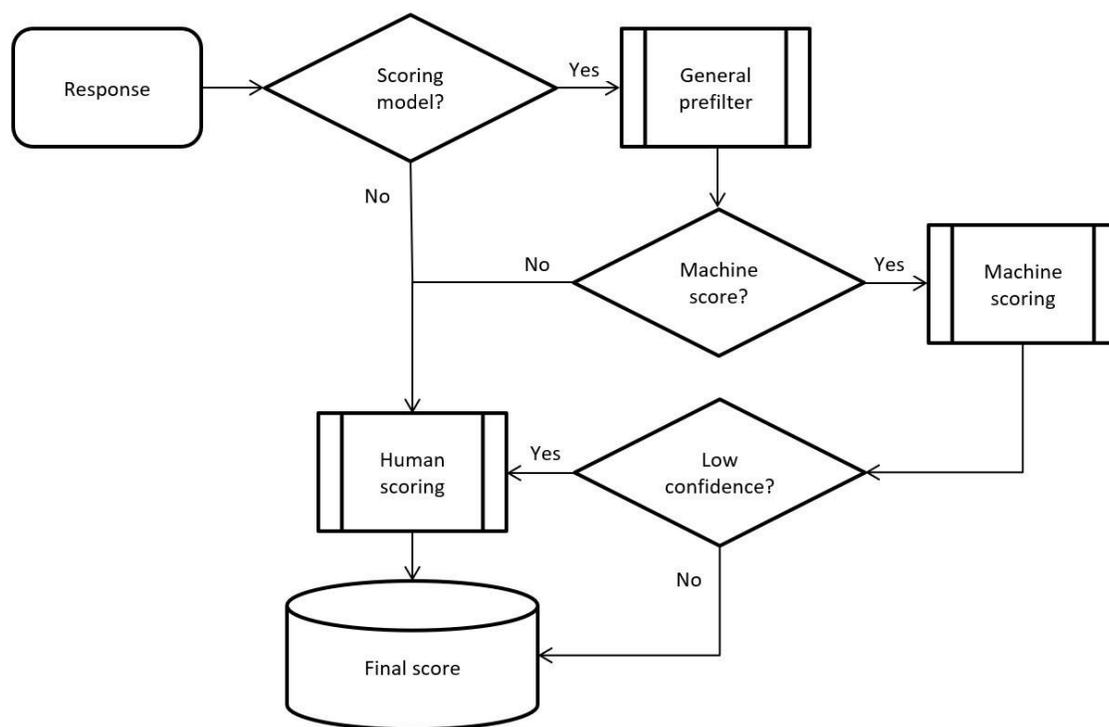
Prior to being assigned a score by the engine, ELA-writing responses were pre-processed to filter out certain non-scorable cases (e.g., off-topic text or high proportion of copied prompt text), and to identify potentially worrying responses that might indicate a personal student crisis. Flags were used to indicate responses with characteristics that meet the definition of condition codes as defined in the hand-scoring criteria (described above). For example, the engine flagged responses that copied the prompt or stimuli (Repeat the prompt; RTP), were written in a language other than English (Not English; NE), were unrelated to the task (Off-topic; OT), or contained alert words or phrases that indicate that the response should be reviewed by the client. The engine also flagged instances where a numerical score could not be returned, such as responses that were excessively long (more than 30,000 characters), an unexpected character or item ID was sent, the scoring process timed out, or a system error occurred. Table 4-8 presents the pre-processing flags.

Table 4-8 RICAS Flags

Flag	Description	Routed to Hand Scoring
0	Standard scoring	If classified in lowest 10% confidence
240	Too long (i.e., 30,000+ characters submitted)	Always
250	Expected essay fields are null or empty; set when nulls are discovered within the processing pipeline.	No, "Blank" condition code automatically assigned
260	In a multipart essay, one or more parts are null, blank, or whitespace.	Always
300	Unexpected character (i.e., non-utf8 string or other encoding issue)	Always
400	Unexpected item_id (i.e., the item_id is not associated with approved scoring model)	Always
500	Scorable alert (i.e., an essay which contains alert language but is otherwise scorable)	Always
620	Applies when the ratio of copied characters exceeds specified threshold (RTP)	Always to expert rater
660	Insufficient English language to score (NE)	Always to expert rater
670	Off-Topic (OT)	Always to expert rater
900	Timeout (i.e., unable to complete essay score prediction within time limits). Not client configurable.	Always
950	System error processing essay (i.e., internal error). Not client configurable.	Always

Following pre-processing, remaining responses were then sent to the automated scoring engine, where text features were extracted, the scoring model(s) applied, and responses assigned a score and measure of score confidence. As shown in Table 4-8, the following types of responses were then sent to human raters for scoring:

- Responses flagged for any reason other than being blank.

- Responses flagged as low-confidence. Low-confidence responses straddle the lines between score point values on a rubric and are difficult to score accurately because they exhibit characteristics of multiple score points. The 10% lowest-confidence responses flagged were scored by the highest-performing raters who were deemed experts.

- Responses with characteristics that meet the definition of condition codes as defined in the hand-scoring criteria (e.g., RTP, NE, OT) were routed to expert raters for confirmation.

4.5.5.3 Quality Assurance

MI's hybrid scoring approach incorporated multiple layers of quality assurance designed to ensure accuracy, fairness, and reliability throughout the scoring process.

First, automated scoring models were trained exclusively on gold-standard human scores, ensuring that the models were calibrated to reflect expert scoring judgments. Each model was then validated against industry-standard criteria, as described above, including agreement with human scores, stability across samples, and the absence of systematic bias. As part of this process, fairness was explicitly evaluated by examining model performance across student subgroups to minimize the risk of differential impact.

Second, quality assurance extended to the human scoring process. Embedded validity responses were used to confirm scorer accuracy and monitor consistency over time. This evaluation was used to identify the most accurate raters, who were then designated as experts and tasked with scoring responses flagged by the system as low-confidence. These checks provided ongoing evidence that both human and automated scoring remained aligned with established scoring standards.

Finally, safeguards were built into the operational workflow to ensure that all responses were scored with the highest degree of accuracy. Responses flagged as non-scorable or requiring special handling (e.g., alert/crisis content) were routed to scoring leadership or expert raters for review and final scoring.

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, training and validation of the model, and monitoring of scoring quality.

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 C.

RICAS results are preliminarily reported in the following deliverables: data files in csv format at the school, district, and state level; a roster PDF produced at the school level; a released essay report at the school and district levels; and a released mathematics constructed-response item report at the school and district levels. The data files and roster contain the results at the item level as well as for reporting categories for ELA and mathematics. The released essay provides a look at the student's essay response as well as the score given to the released essay. The released mathematics constructed-response report shows the student's responses to the released mathematics constructed items. Both paper and online responses are reported for both reports. The preliminary reports and data files are available for school and district access in the Download Center in LightHouse™. Scaled scores, achievement levels, and student growth percentile are not reported in preliminary reporting.

Parent/Guardian Reports are produced in final reporting. 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 the Download Center in LightHouse™. Individual Parent/Guardian Report PDFs were run and collated by grade and school as well as district level and posted online for school, district, and state access.

The Family Portal in LightHouse™ was used for the first time in 2025. Parents or legal guardians may access the Family Portal using the student's SASID. The parent or legal guardian can see results for 2025 as well as results from 2024 and 2023 if they are available. A version of the printed Parent/Guardian Report is available in the Family Portal.

5.2 PARENT/GUARDIAN REPORT

The Parent/Guardian Report (based on the MCAS report design) or "Individual Student Report" was generated for each student who participated in the RICAS tests. The report is designed as a stand-alone 11" x 17" color report that is folded to create four pages; see Appendix D for a sample report. Two full-color copies of each student's report were printed: one for the parent/guardian and one for the school's records. The report is designed to present parents/guardians with a detailed summary of their child's

RICAS performance on ELA and mathematics and to enable comparisons with other students at the school, district, and state levels.

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

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

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

The inside pages of each report contain the achievement level, scaled score, and standard error of the scaled score for each content area. If the student does not receive a scaled score, the reason is displayed after "Your Child's Achievement Level." Each achievement level has its own distinct color, and that color is used throughout the report to highlight important report elements based on the student's achievement level and score. These report elements include the student's earned achievement level, scaled score, the visual scale's achievement-level title and achievement-level cut scores, and the comparison of the student's scaled score to the average scaled score at the student's school, district, and 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, are displayed in a table. A student growth percentile (SGP) for each content area tested is displayed, if available, with a comparison to the average SGP for the student's school and district. An SGP describes the student's learning over time compared to his or her academic peers (peers are other students with similar scores on previous state tests).

The student's performance in each content area's reporting categories is also displayed using pictographs and text that indicates the points earned by the student versus the total points possible in that reporting category. The student's performance on individual test questions is reported at the bottom of the results page in a simplified item response grid. The grid indicates the points earned and points possible for each test question included on the grid. Essay questions are indicated on the grid.

5.3 REPORTING BUSINESS REQUIREMENTS

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

5.4 QUALITY ASSURANCE

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

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

The third aspect of quality control involves procedures to check the accuracy of reported data. Using a sample of schools and districts, the quality assurance group verifies that the reported information is correct. There are two sets of samples selected that may not be mutually exclusive. The first set includes samples that satisfy all the following criteria:

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

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

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

5.5 ADDITIONAL RESOURCES

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

5.5.1 Students and Their Families

For students and their families, a "Resources for Families" page (https://www.ride.ri.gov/InstructionAssessment/Assessment/ResourcesforFamilies.aspx) provides general information about Rhode Island's State Assessment Program. Links to more detailed resources are found throughout the page and include:

- An Assessment Results page which provides access to the Rhode Island Assessment Data Portal and guides its use

- A Family Guide to Understanding RICAS which provides guidance not only for properly interpreting RICAS results but also for appropriate use of the results

- A RICAS Individual Student Report (ISR) Guide which explains the different components of the ISR each student receives, how to interpret them, and how to use them to work with teachers to help their child succeed

- A RICAS Family Portal Guide which explains what information is available in the Family Portal and how to create an account.

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 2025 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 2025. Note that the information presented in this section is based only on the operational items, since student scores are calculated on those items.

6.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

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

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

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

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

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

Content Area	Grade	Item Type	Number of Items	Difficulty Mean	Difficulty Standard Deviation	Discrimination Mean	Discrimination Standard Deviation
ELA	3	All	32	0.54	0.14	0.49	0.11
ELA	3	MC	24	0.59	0.11	0.45	0.08
ELA	3	OR	6	0.45	0.10	0.56	0.09
ELA	3	ES	2	0.23	0.05	0.74	0.01
ELA	4	All	32	0.60	0.15	0.51	0.10
ELA	4	MC	24	0.62	0.12	0.48	0.08
ELA	4	OR	6	0.57	0.19	0.58	0.08
ELA	4	ES	2	0.32	0.06	0.72	0.01
ELA	5	All	33	0.56	0.17	0.48	0.15
ELA	5	MC	24	0.62	0.16	0.42	0.09
ELA	5	OR	5	0.44	0.16	0.50	0.09
ELA	5	ES	4	0.40	0.07	0.78	0.02
ELA	6	All	33	0.55	0.12	0.50	0.16
ELA	6	MC	24	0.56	0.10	0.43	0.09
ELA	6	OR	5	0.63	0.13	0.59	0.08
ELA	6	ES	4	0.41	0.11	0.82	0.01
ELA	7	All	33	0.57	0.12	0.51	0.14
ELA	7	MC	24	0.61	0.11	0.46	0.08
ELA	7	OR	5	0.48	0.08	0.51	0.08
ELA	7	ES	4	0.44	0.13	0.82	0.02
ELA	8	All	33	0.59	0.13	0.51	0.15
ELA	8	MC	24	0.61	0.13	0.44	0.08
ELA	8	OR	5	0.61	0.10	0.57	0.04
ELA	8	ES	4	0.44	0.11	0.83	0.01
Mathematics	3	All	40	0.54	0.14	0.54	0.11
Mathematics	3	MC	18	0.57	0.15	0.47	0.06
Mathematics	3	OR	22	0.51	0.13	0.59	0.12
Mathematics	4	All	40	0.53	0.14	0.54	0.11
Mathematics	4	MC	13	0.59	0.13	0.47	0.09
Mathematics	4	OR	27	0.50	0.13	0.57	0.10
Mathematics	5	All	40	0.48	0.16	0.50	0.13
Mathematics	5	MC	19	0.57	0.14	0.42	0.09
Mathematics	5	OR	21	0.41	0.13	0.56	0.13
Mathematics	6	All	40	0.46	0.15	0.50	0.14
Mathematics	6	MC	16	0.51	0.15	0.40	0.08
Mathematics	6	OR	24	0.43	0.14	0.57	0.12
Mathematics	7	All	40	0.40	0.16	0.50	0.16
Mathematics	7	MC	21	0.44	0.15	0.41	0.10
Mathematics	7	OR	19	0.35	0.16	0.61	0.13
Mathematics	8	All	40	0.44	0.16	0.52	0.13
Mathematics	8	MC	17	0.54	0.12	0.45	0.11
Mathematics	8	OR	23	0.36	0.14	0.58	0.12

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

Difficulty indices for selected-response items tend to be higher (indicating that students performed better on these items) than the difficulty indices for open-response items because selected-response items can be answered correctly by simply identifying rather than providing the correct answer, or by guessing. Similarly, discrimination indices for those open-response items with more than two points tend to be larger than those for dichotomous items because of the greater variability in score points of the former (i.e., the partial credit these items allow). The restriction of range (i.e., only two score categories) in dichotomous

items tends to make the discrimination indices lower. Note that these patterns are more consistent within item type, so when interpreting classical item statistics, comparisons should be emphasized among items of the same type.

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

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 2025 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 2025 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 2025 RICAS, the standardization DIF procedure (Dorans & Kulick, 1986) was employed to evaluate subgroup differences, which denote significant group-level differences in performance for examinees with equivalent achievement levels on the test. The standardization DIF procedure is designed to identify items for which subgroups of interest perform differently and beyond the impact of differences in overall achievement. The DIF procedure calculates the difference in item performance for two groups of students (at a time) matched for achievement on the total test. Specifically, average item performance is calculated for students at every total score. Then an overall average is calculated, weighting the total score distribution so that it is the same for the two groups. DIF statistics were calculated for all subgroups with at least 75 students.

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

The main DIF index produced under the standardization procedure has a theoretical range from -1.0 to 1.0 for selected-response items and open-response items; the latter uses an adjusted index. Dorans and Holland (1993) suggested that index values between -0.05 and 0.05 denote either a negligible amount of DIF or the absence of DIF. The majority of 2025 RICAS items fell within this range. Dorans and Holland further stated that items with values between -0.10 and -0.05 and between 0.05 and 0.10 (i.e., "low" DIF)

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

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

- male compared with female

- not ELL compared with ELL[3]

- 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 2025 spring administration, DIF analyses were conducted again as a post-hoc quality check based on the operational data. The tables in Appendix G present the number of items classified as either "low" or "high" DIF, in total and by group favored. Following Dorans and Holland's recommendation, items with DIF indices between -0.10 and -0.05 and between 0.05 and 0.10 were categorized as "low" DIF, and values outside the -0.10 to 0.10 range were categorized as "high" DIF. Very few items exhibited high DIF in the operational data, which suggested that the item review that occurred after the MCAS field-testing effectively ruled out items displaying large DIF for the RICAS operational spring tests.

6.3 DIMENSIONALITY ANALYSIS

Because tests are constructed with multiple content area subcategories and their associated knowledge and skills, the potential exists for the invocation of multiple dimensions beyond the common primary dimension. Generally, the subcategories are highly correlated with each other; therefore, a primary dimension typically explains the majority of variance in test scores. The presence of one dominant primary dimension is the primary psychometric assumption to support the use of the unidimensional IRT models that are used for calibrating and scaling the items administered on the 2025 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 2025 administration. A total of 12 computer-based test forms were analyzed; the results for these analyses are reported below.

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

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

---

[3] *ELL = English Language Learner (includes current and former English Language Learners).*

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

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

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

DIMTEST and DETECT were applied to the operational items of the RICAS tests administered during spring 2025. The data for each grade were split into a training sample and a cross-validation sample. For all grades, there were over 9,400 student examinees per test form in both ELA and mathematics, so every training sample and cross-validation sample had at least 4,700 students. After randomly splitting the data into training and cross-validation samples, DIMTEST was applied to each data set to see if the null hypothesis of unidimensionality would be rejected. DETECT was then applied to each data set for which the DIMTEST null hypothesis was rejected to estimate the effect size of the multidimensionality.

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

DETECT was then used to estimate the effect size for the violations of local independence for the 2025 tests. Table 6-2 displays the multidimensionality effect-size estimates from DETECT across four operational years.

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

Content Area	Grade	Multidimensionality Effect Size 2022	Multidimensionality Effect Size 2023	Multidimensionality Effect Size 2024	Multidimensionality Effect Size 2025
ELA	3	0.15	0.14	0.18	0.20
ELA	4	0.21	0.17	0.21	0.18
ELA	5	0.21	0.17	0.37	0.35
ELA	6	0.25	0.22	0.34	0.38
ELA	7	0.30	0.34	0.43	0.31
ELA	8	0.30	0.32	0.44	0.41
ELA	**Average**	**0.24**	**0.23**	**0.33**	**0.30**
Mathematics	3	0.20	0.17	0.19	0.13
Mathematics	4	0.18	0.14	0.13	0.08
Mathematics	5	0.18	0.10	0.09	0.12
Mathematics	6	0.13	0.15	0.15	0.15
Mathematics	7	0.13	0.13	0.11	0.13
Mathematics	8	0.18	0.14	0.17	0.18
Mathematics	**Average**	**0.17**	**0.14**	**0.14**	**0.13**

The DETECT values indicate weak ($0.2 <$ DETECT < 0.4) or very weak (DETECT < 0.2) multidimensionality for all the 2025 RICAS test forms, except for ELA grade 8 (DETECT = 0.41), indicating a low-end moderate multidimensionality with more dimensionality distinctiveness of writing prompt items relative to the rest of the test. The patterns are consistent with the previous year's results.

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

In summary, for the 2025 dimensionality analyses, the violations of local independence, as evidenced by the DETECT effect sizes, were either weak or very weak in all test forms. The patterns with respect to the selected-response and open-response items were consistent with those in the previous year, with ELA tending to display more separation than mathematics. However, this separation did not result in an effect size that would suggest using a unidimensional IRT model is inappropriate.

6.4 QUALITY CHECK

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

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

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

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

6.5 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING CLASSICAL ITEM ANALYSES

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

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

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

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

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

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

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

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

Chapter 7. Item Response Theory Analysis

7.1 OVERVIEW

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

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

- evaluations of the calibration processes (e.g., checking the number of cycles required for convergence for reasonableness);

- checking item parameters and their standard errors for reasonableness;

- examination of test characteristic curves (TCCs) and test information function (TIF) curves for reasonableness;

- evaluation of model fit (e.g., test level, item-level, and person-level);

- evaluation of equating items (e.g., delta analyses, *b-b* analyses, beta analyses);

- examination of a-plots and b-plots for reasonableness; and

- evaluation of the scaling results (e.g., comparing look-up tables to the previous year's).

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

7.2 IRT

All RICAS items were calibrated applying IRT on the MCAS data. IRT uses mathematical models to define a relationship between an unobserved measure of student performance, usually referred to as theta (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 $\bar{\theta}$ 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 2025 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:

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

(Equation 1)

where
U represents the scored response on an item,
i indexes the items,
j indexes students,
a 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:

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

(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:

$$P_{ik}^*(\theta_j) = P(U_i \geq k|\theta_j) = \frac{exp[Da_i(\theta_j - b_i + d_{ik})]}{1 + exp[Da_i(\theta_j - b_i + d_{ik})]},$$

(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:

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

(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:

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

(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:

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

(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, $E(U_i|\theta_i) = P_i(\theta_i).$ 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:

$$E(X|\theta_j) = \sum_{i=1}^{n} E(U_i|\theta_j),$$ (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:

$$SEM(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}.$$ (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 H provide the IRT item parameters and associated standard errors of all operational scoring items on the 2025 RICAS tests. The MCAS equating report in Appendix H contains graphs of the TCCs and TIFs, which are defined in the previous section. While the information provided in Appendix H 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 C.

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	34	7
ELA	4	45	7
ELA	5	49	11
ELA	6	31	11
ELA	7	53	10
ELA	8	43	10
Mathematics	3	38	NA
Mathematics	4	63	NA
Mathematics	5	58	NA
Mathematics	6	54	NA
Mathematics	7	105	NA
Mathematics	8	50	NA

7.4 EQUATING

Section 7.4 summarizes the equating procedure and results to place the 2025 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 2025 MCAS equating (Appendix H).

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 2025 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 2025 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]"[4]) 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, it's 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

---

[4] *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

density function for θ in the focal group. The following threshold is used to categorize an item into negligible, moderate or large drift:

- o $|\hat{\beta}| < 0.05,$ negligible drift

- o $0.05 \le |\hat{\beta}| < 0.1,$ moderate drift

- o $|\hat{\beta}| \ge 0.1,$ large drift

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

Content Review

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

- changes to item administration mode

- word/graphic changes to any part of the item

- change to option order

- change in position (e.g., beginning of test vs. end of test)

- whether an item experiences "clueing" in one administration but not in the other

- whether there are test security risks associated with the flagged items

- any other difference that may affect the testing experience

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

The equating items that remained following these evaluation procedures were then employed in the SL method, and the linking relationship obtained from the SL method was used to transform the item parameters for all items in the 2025 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 2025 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	1.15	-0.26
ELA	4	1.16	-0.33
ELA	5	1.17	-0.27
ELA	6	1.50	-0.35
ELA	7	1.34	-0.38
ELA	8	1.42	-0.26
Mathematics	3	1.22	-0.16
Mathematics	4	1.15	-0.15
Mathematics	5	1.14	-0.24
Mathematics	6	1.16	-0.22
Mathematics	7	1.21	-0.28
Mathematics	8	1.14	-0.22

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, 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 2025 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 2025 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 scaled scores 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:

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

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
ELA	4	18.846	499.421
ELA	5	17.686	499.335
ELA	6	18.984	500.202
ELA	7	19.098	499.791
ELA	8	19.900	498.981
Mathematics	3	21.357	499.413
Mathematics	4	20.938	498.869
Mathematics	5	19.039	499.525
Mathematics	6	19.870	500.165
Mathematics	7	20.758	499.353
Mathematics	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 Cut 1	Theta Cut 2	Theta Cut 3	Scale Score Min	Scale Score Cut 1	Scale Score Cut 2	Scale Score Cut 3	Scale Score Max
ELA	3	-1.581	0.011	1.604	440	470	500	530	560
ELA	4	-1.561	0.031	1.623	440	470	500	530	560
ELA	5	-1.659	0.038	1.734	440	470	500	530	560
ELA	6	-1.591	-0.011	1.570	440	470	500	530	560
ELA	7	-1.560	0.011	1.582	440	470	500	530	560
ELA	8	-1.456	0.051	1.559	440	470	500	530	560
Mathematics	3	-1.377	0.027	1.432	440	470	500	530	560
Mathematics	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
Mathematics	6	-1.518	-0.008	1.502	440	470	500	530	560
Mathematics	7	-1.414	0.031	1.476	440	470	500	530	560
Mathematics	8	-1.496	-0.008	1.479	440	470	500	530	560

7.6 QUALITY CHECK

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

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

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

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

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

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

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

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

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

Evidence: Section 7.2 describes the underlying 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 H) 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 H. Section 7.5 describes the processes of applying equating and scaling results to place raw scores onto RICAS score scales. These sections demonstrate a high level of rigor in selection, application, and interpretation of equating results, placing scores from this year's forms on the same scales as forms from prior years.

Chapter 8. Reliability

8.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT

Although an individual item's performance is an important factor in evaluating an assessment, a complete evaluation must also address the way items grouped in a set function as a set and complement one another. Tests that function well provide a dependable assessment of a student's level of ability. Just like the measurement of physical properties such as temperature, any measurement tool contains some amount of measurement error, which leads to different results if the measurements were taken multiple times. As the tools to measure latent ability, the items' quality 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 2025 RICAS tests is the α coefficient of Cronbach (1951). This approach is most easily understood as an extension of a related procedure, split-half reliability. In the split-half approach, a test is split in half, and students' scores on the two half-tests are correlated. To estimate the correlation between two full-length tests, the Spearman-Brown correction (Spearman, 1910; Brown, 1910) is applied. If the correlation is high, this is evidence that the items complement one another and function well as a group, suggesting that measurement error is minimal. The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation since each different possible split of the test into halves will result in a different correlation.

Cronbach's 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:

$$a = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma_{(Y_i)}^2}{\sigma_x^2}\right],$$

(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 0.90 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 Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha (α)	SEM
ELA	3	9,849	44	21.49	9.22	0.90	2.98
ELA	4	9,495	44	24.46	9.25	0.91	2.84
ELA	5	9,749	48	24.66	9.75	0.90	3.13
ELA	6	9,583	50	26.06	11.11	0.91	3.27
ELA	7	9,680	50	25.92	10.80	0.91	3.15
ELA	8	9,800	50	27.39	11.26	0.91	3.30
Mathematics	3	9,797	48	24.95	11.94	0.93	3.05
Mathematics	4	9,403	54	27.52	13.03	0.94	3.32
Mathematics	5	9,716	54	25.38	12.34	0.92	3.50
Mathematics	6	9,515	54	23.98	12.54	0.92	3.54
Mathematics	7	9,560	54	20.20	11.69	0.93	3.19
Mathematics	8	9,664	54	23.26	12.64	0.93	3.41

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 that 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 2025 RICAS online forms. Appendix I 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 0.80 to 0.95 across the tests, with a median of 0.91 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 2025 RICAS content areas. Results and reporting category descriptions are presented in Appendix I. The reliability coefficients for the reporting subcategories range from 0.49 to 0.90, with a median of 0.73 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 2025 RICAS tests, students were classified into one of four achievement levels: *Not Meeting Expectations*, *Partially Meeting Expectations*, *Meeting Expectations*, or *Exceeding Expectations*.

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

Accuracy refers to the extent to which achievement classifications based on test scores match the classifications that would have been assigned if the scores did not contain any measurement error. Accuracy must be assessed because errorless test scores do not exist. Consistency measures the extent to which classifications based on test scores match the classifications based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are administered to the same group of students. In operational testing programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and the consistency of classifications based on a single administration of a test. The Livingston and Lewis (1995) technique was used for the 2025 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 2025 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 2025 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:

$$\kappa = \frac{(\text{Observed agreement}) - (\text{Chance agreement})}{1 - (\text{Chance agreement})} = \frac{\sum_i c_{ii} - \sum_i c_{i.} c_{.i}}{1 - \sum_i c_{i.} c_{.i}},$$

(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 2025 RICAS tests.

Table 8-2 includes overall accuracy indices with consistency indices displayed in parentheses next to the accuracy values, as well as overall kappa values. Overall ranges for accuracy (0.80–0.85), consistency (0.72–0.78), and kappa (0.59–0.67) 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 0.85 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, 85% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.77 indicates that 77% 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 2025 RICAS tests. A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.

In Table 8-3, the accuracy and consistency indices at the *Partially Meeting Expectations/Meeting Expectations* threshold range from 0.91–0.93 and 0.87–0.90, respectively. The false positive and false negative decision rates at the *Partially Meeting Expectations/Meeting Expectations* threshold range from 4%–5% and 3%–4%, respectively. 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	0.80 (0.72)	0.59	0.85 (0.77)	0.80 (0.73)	0.79 (0.72)	0.71 (0.51)
ELA	4	0.82 (0.75)	0.62	0.86 (0.78)	0.82 (0.76)	0.82 (0.75)	0.65 (0.41)
ELA	5	0.84 (0.77)	0.63	0.84 (0.75)	0.86 (0.82)	0.79 (0.72)	0.75 (0.52)
ELA	6	0.80 (0.72)	0.60	0.87 (0.80)	0.82 (0.75)	0.73 (0.66)	0.62 (0.43)
ELA	7	0.81 (0.74)	0.62	0.88 (0.82)	0.81 (0.74)	0.77 (0.70)	0.63 (0.43)
ELA	8	0.81 (0.73)	0.61	0.88 (0.81)	0.80 (0.73)	0.77 (0.69)	0.71 (0.54)
Mathematics	3	0.83 (0.76)	0.65	0.88 (0.80)	0.82 (0.76)	0.82 (0.75)	0.78 (0.64)
Mathematics	4	0.84 (0.78)	0.67	0.86 (0.79)	0.85 (0.80)	0.83 (0.77)	0.77 (0.61)
Mathematics	5	0.85 (0.78)	0.66	0.86 (0.79)	0.86 (0.81)	0.82 (0.75)	0.76 (0.57)
Mathematics	6	0.84 (0.78)	0.66	0.84 (0.78)	0.85 (0.81)	0.83 (0.76)	0.78 (0.60)
Mathematics	7	0.84 (0.77)	0.66	0.87 (0.80)	0.85 (0.80)	0.80 (0.71)	0.83 (0.68)
Mathematics	8	0.84 (0.78)	0.66	0.86 (0.81)	0.84 (0.79)	0.82 (0.73)	0.80 (0.63)

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

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

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

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

8.6 QUALITY CHECK

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

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

8.7 EVIDENCE IN SUPPORT OF VALIDITY ARGUMENTS REGARDING RELIABILITY

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

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

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

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

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

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

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

9.1 RATIONALE FOR VALIDITY ARGUMENT-CENTERED TECHNICAL REPORTING

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

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

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

Figure 9-1 Validity Argument Logic Model
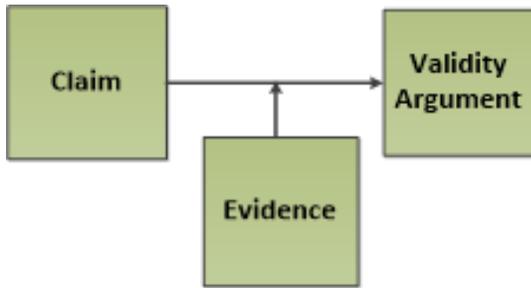
9.2 VALIDITY ARGUMENT FOR INTERPRETATION AND USE OF RICAS TEST SCORES

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

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

1) **Description Inference**: Items sample from the target domain appropriately such that high quality forms can be produced. (Domain to Item)

2) **Evaluation Inference**: Forms sample from items appropriately such that observed scores reflective of the domain can be produced. (Item to Form)

3) **Generalization Inference**: Observed scores on individual forms are reliable such that they are reflective of expected scores across forms. (Form to Score)

4) **Explanation Inference**: Expected scores are associated with classification cuts such that classification decisions are interpretable. (Score to Interpretation)

5) **Utilization Inferences:** Interpretations of scores and classifications are used as intended and only in ways considered appropriate and fair. (Interpretation to Use)

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

Figure 9-2 Chapelle's (2020) Framework: The Arguments and the Inferential Steps
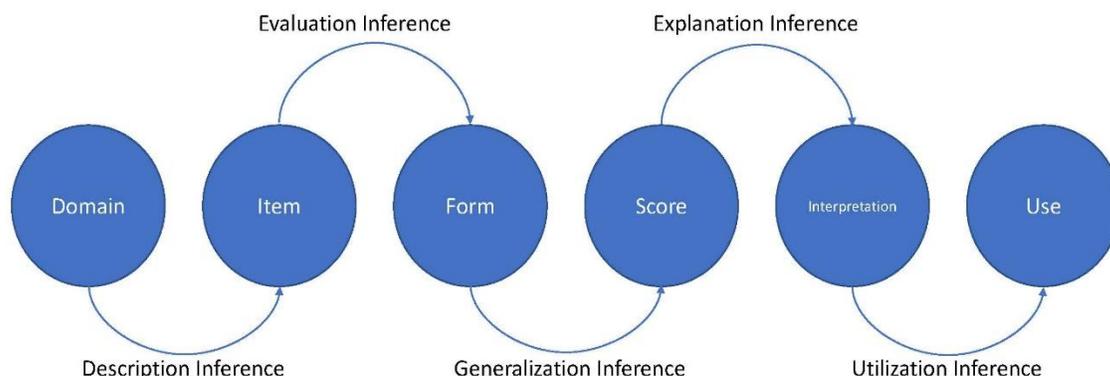
It is important for the gathering of information in support of the Generalization Inference (3) to define what is meant by the term "form" in this context. A test form is not just the set of items on which the score is based, but the structure of the exam in terms of all elements that can affect an individual's performance. This can include, among other things, the raters scoring an exam, the occasion on which the exam is administered, and the setting in which it is administered. Generalization from observed to expected score is optimized when all sources of potential variability of test scores are identified and accounted for such that observed scores maximally reflect a student's ability and not the influence of unwanted sources of variance.

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

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

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

9.2.1 Claims Supporting Intended Interpretations of RICAS Test Scores

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

9.2.2 Claims Supporting Intended Uses of RICAS Test Scores

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

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

    1)   understands the meaning of scores and classifications, appropriate uses and interpretations of those scores and classifications, and any limits on their interpretability, as applied to the intended use, and

    2)   finds the scores and classifications genuinely useful for that intended use.

The evidence described in this section pertains to the aspects that relate to activities performed by Cognia/eMetric 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 and RIDE provided materials and other implementation supports (e.g., town halls, professional development/educational sessions) that put all stakeholders in a strong position to be able to understand the *intended* meanings and uses of the RICAS scores.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

9.3 VALIDATION SUMMARY

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

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

- Following the Chapelle (2020) framework, we have provided a chain of inferences, each building on the previous, to preserve the interpretations as defined by the content standards such that they are realized within the resulting test scores and applicable to the intended uses of the RICAS program. By establishing the description inference, providing evidence that the items used in the assessment target the domain as defined by the standards, we argue that we can create individual forms that produce a test score reflective of achievement on that domain.

- We present evidence we have created such individual forms that these forms do elicit test scores reflecting achievement on the intended domain, which in turn is able to provide a classification for a student achievement level and that such classification decisions are interpretable.

- Finally, we provide evidence to support that the test score and classification interpretations are clearly enough explained as to be used as intended and only in ways considered appropriate and fair. We use this chain of evidence to assert the scores and classifications resulting from RICAS tests are interpretable and used in a way that is intended and fair.

References

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory.* Belmont, CA: Wadsworth, Inc.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association. Available from https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf

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	Number of Students Tested Without Accommodations
ELA	3	1,110	9,852
ELA	4	1,330	9,497
ELA	5	1,473	9,753
ELA	6	1,361	9,589
ELA	7	1,343	9,681
ELA	8	1,352	9,803
Mathematics	3	2,433	10,000
Mathematics	4	2,385	9,609
Mathematics	5	2,296	9,912
Mathematics	6	1,768	9,733
Mathematics	7	1,769	9,802
Mathematics	8	1,756	9,898

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
Mouse Pointer	25	22	7	5	10	12
Large Print Test Edition	2	1	0	1	1	0
Screen Reader Edition	3	0	0	2	0	1
Assistive Technology	4	1	3	4	1	2
Braille Test Edition	0	0	1	0	0	0
Human Read Aloud as a Non-Standard Accommodation	40	25	33	27	9	12
Human Signer as a Non-Standard Accommodation	0	0	0	0	0	1
Human Scribe as a Non-Standard Accommodation	70	52	43	32	18	11
Text-to-Speech	177	153	175	159	152	111
Speech-to-Text as a Non-Standard Accommodation	99	117	136	81	62	41
Typed Responses	10	7	4	8	2	7
Spell-checker	35	72	86	68	63	50
Word Prediction Special	47	86	72	76	72	38
Graphic Organizer/Reference Sheet	832	1,073	1,190	1,031	994	1,016
Bilingual Dictionary and Glossary	191	187	213	272	307	305

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
Mouse Pointer	8	1	3	0	1	1
Large Print Test Edition	1	0	0	1	0	0
Screen Reader Edition	1	0	2	1	11	2
Assistive Technology	2	1	2	3	0	1
Braille Test Edition	0	0	1	0	0	0
Human Read Aloud Standard Accommodation	59	69	53	26	6	9
Human Signer Standard Accommodation	3	5	1	2	2	5
Human Scribe Standard Accommodation	42	31	27	21	11	8
Text-to-Speech	2,030	1,895	1,764	1,169	1,066	904
Speech-to-Text Standard	65	57	62	34	38	19
Typed Responses	9	7	5	8	1	6
Spanish	196	206	203	230	246	252
Calculation Device	103	103	124	145	181	251
Word Prediction Standard	1	23	16	17	13	6
Graphic Organizer/Reference Sheet	712	956	1,061	887	879	915
Bilingual Dictionary and Glossary	205	194	225	190	207	224

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	58,175	100.00
ELL	10,548	18.13
Economically Disadvantaged	33,431	57.47
African American	5,328	9.16
Asian	1,870	3.21
Hispanic	17,969	30.89
Native American/Alaska Native	449	0.77
White	29,276	50.32
Pacific Islander/Hawaiian	105	0.18
Multiracial	3,178	5.46
Male	29,695	51.04
Female	28,472	48.94
Special Education	10,656	18.32
Plan 504	4,104	7.05

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

Description	Number Tested	Percent Tested
All Students	58,954	100.00
ELL	11,412	19.36
Economically Disadvantaged	34,046	57.75
African American	5,456	9.25
Asian	1,932	3.28
Hispanic	18,513	31.40
Native American/Alaska Native	446	0.76
White	29,318	49.73
Pacific Islander/Hawaiian	110	0.19
Multiracial	3,179	5.39
Male	30,090	51.04
Female	28,856	48.95
Special Education	10,635	18.04
Plan 504	4,094	6.94

APPENDIX C

ACHIEVEMENT LEVEL DISTRIBUTIONS

Table C-1. Achievement-Level Distributions by Grade—ELA

Grade	Achievement Level	Percent in Level 2025	Percent in Level 2024	Percent in Level 2023	Percent in Level 2022
3	Not Meeting Expectations	21.49	19.43	21.13	19.41
3	Partially Meeting Expectations	41.83	42.76	42.03	44.03
3	Meeting Expectations	31.81	32.28	31.54	31.57
3	Exceeding Expectations	4.87	5.54	5.30	4.98
4	Not Meeting Expectations	21.13	21.74	19.91	21.67
4	Partially Meeting Expectations	44.57	48.18	46.79	49.30
4	Meeting Expectations	31.38	27.80	29.71	26.49
4	Exceeding Expectations	2.92	2.27	3.59	2.55
5	Not Meeting Expectations	18.12	20.28	20.35	17.74
5	Partially Meeting Expectations	51.06	50.01	44.55	50.51
5	Meeting Expectations	27.72	26.56	32.17	28.25
5	Exceeding Expectations	3.10	3.15	2.94	3.50
6	Not Meeting Expectations	28.28	33.73	31.12	32.10
6	Partially Meeting Expectations	38.59	37.87	36.94	36.72
6	Meeting Expectations	27.25	22.18	27.16	25.72
6	Exceeding Expectations	5.88	6.21	4.78	5.47
7	Not Meeting Expectations	29.07	29.33	27.50	28.80
7	Partially Meeting Expectations	38.38	43.86	43.45	42.01
7	Meeting Expectations	27.81	23.17	24.41	25.74
7	Exceeding Expectations	4.74	3.64	4.63	3.46
8	Not Meeting Expectations	27.83	34.19	31.50	27.54
8	Partially Meeting Expectations	37.33	33.89	36.28	43.45
8	Meeting Expectations	28.29	24.83	25.60	24.87
8	Exceeding Expectations	6.56	7.09	6.62	4.14

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

Grade	Achievement Level	Percent in Level 2025	Percent in Level 2024	Percent in Level 2023	Percent in Level 2022
3	Not Meeting Expectations	22.65	22.56	25.20	24.84
3	Partially Meeting Expectations	38.92	38.30	40.33	40.14
3	Meeting Expectations	31.14	31.85	28.99	30.99
3	Exceeding Expectations	7.29	7.29	5.47	4.04
4	Not Meeting Expectations	20.31	20.23	22.32	23.68
4	Partially Meeting Expectations	44.33	44.36	41.67	46.13
4	Meeting Expectations	30.43	30.94	30.66	27.02
4	Exceeding Expectations	4.92	4.46	5.36	3.17
5	Not Meeting Expectations	20.82	17.27	17.90	23.90
5	Partially Meeting Expectations	47.57	52.66	52.13	50.16
5	Meeting Expectations	28.16	27.24	27.49	24.31
5	Exceeding Expectations	3.45	2.84	2.48	1.64
6	Not Meeting Expectations	22.56	24.76	24.09	23.77
6	Partially Meeting Expectations	47.47	49.32	47.18	48.87
6	Meeting Expectations	26.65	22.67	25.99	25.65
6	Exceeding Expectations	3.32	3.26	2.74	1.72
7	Not Meeting Expectations	27.65	27.06	31.83	29.61
7	Partially Meeting Expectations	45.78	47.63	42.48	47.67
7	Meeting Expectations	22.27	22.29	22.34	20.14
7	Exceeding Expectations	4.31	3.01	3.36	2.57
8	Not Meeting Expectations	29.06	29.83	30.86	28.65
8	Partially Meeting Expectations	44.53	45.30	46.16	50.59
8	Meeting Expectations	22.91	21.47	19.92	18.40
8	Exceeding Expectations	3.50	3.40	3.06	2.36

APPENDIX D
SAMPLE REPORTS

Building Brighter Futures

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

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

How many days of instruction has your child missed?

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

2024 Percent Proficient: Chronically Absent Students vs. Not Chronically Absent Students

Legend: ● Chronically Absent ● Not Chronically Absent

Categories: RICAS - ELA (Gap: 18.0%), RICAS - Math (Gap: 21.2%), PSAT - ELA (Gap: 22.6%), PSAT - Math (Gap: 21.5%), SAT - ELA (Gap: 24.8%), SAT - Math (Gap: 20.8%)
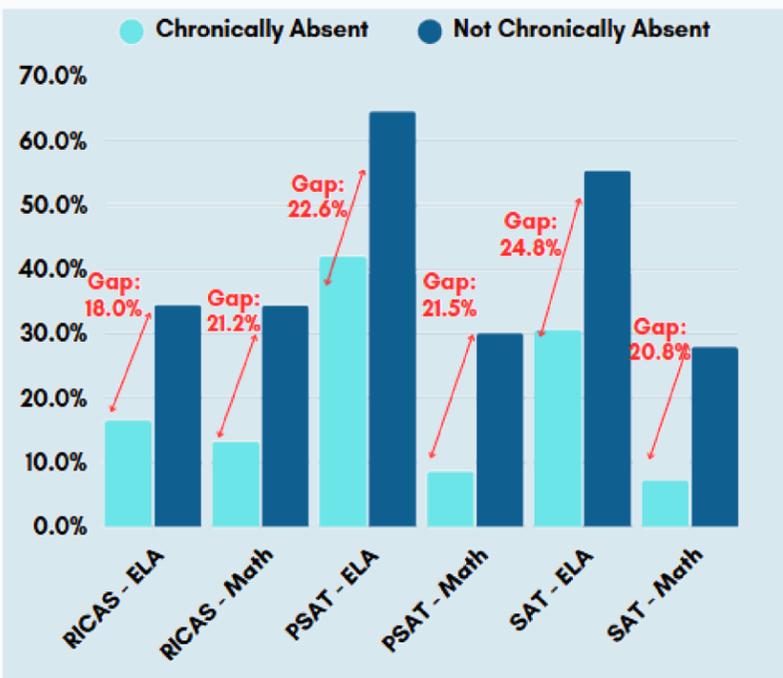
Why does it matter?

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

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

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

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

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

www.ride.ri.gov | www.attendancemattersri.org

Spring 2025 RICAS
Individual Student Report

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

This report provides your child's results from the 2025 Rhode Island Comprehensive Assessment System (RICAS) tests in English Language Arts (reading and writing) and mathematics.

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

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

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

Your Child's Overall Results

English Language Arts	**Mathematics**
Achievement Level	Achievement Level
Partially Meeting Expectations	**Partially Meeting Expectations**
Score	Score
482	**485**
(Score range: 440-560)	(Score range: 440-560)
Growth Percentile	Growth Percentile
77	**89**
Details on page 2	Details on page 3

Did you know that establishing family routines can help your child succeed?
Make a habit of setting up designated times for homework, reading, mealtimes, family conversations, bedtime, and leaving for school each day.

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

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

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

English Language Arts

Your Child's Achievement Level:	**Partially Meeting Expectations**
Your Child's Score:	**482**

482

440	470	500	530	560
Not Meeting Expectations	**Partially Meeting Expectations**	Meeting Expectations	Exceeding Expectations	

A student who performed at this level did not meet grade-level expectations in this subject.

The school, in consultation with the student's parent/guardian, should determine the coordinated academic assistance and/or additional instruction the student needs to succeed in this subject.

A student who performed at this level partially met grade-level expectations in this subject.

The school, in consultation with the student's parent/guardian, should consider whether the student needs additional academic assistance to succeed in this subject.

A student who performed at this level met grade-level expectations and is academically on track to succeed in the current grade in this subject.

A student who performed at this level exceeded grade-level expectations by demonstrating mastery of the subject matter.

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

Achievement

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

Your Child's Grade	Score	Year	School	District	State
4	482	2025	487	486	490
3	472	2024			

Average Score

2025 Student Growth Percentiles

Your child's score this year is the same as or better than 77 percent of Rhode Island students who had a similar score to your child on the assessment(s) in a previous year(s).

Your Child — 77
School — 57
District — 48
State Avg. — 50

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

Reporting Category	Points Earned by Your Child	Total Possible Points	School	District	State
Reading	18	31	18.6	18.1	18.9
Language†	3	9	4.0	4.2	4.5
Writing ‡	0	4	0.9	0.9	1.1

Average Points

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

Individual Test Questions

Question Number	1	2	3	4	5	6	7	8	9	10	11	12 CV	12 ID	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Points Earned	1/1	0/1	1/1	1/1	1/1	0/1	1/1	2/2	1/1	1/1	1/2	0/3	0/4	0/1	1/1	1/1	1/1	0/1	1/1	0/2	1/1	1/1	0/1	1/1	0/1	0/1	0/1	1/2	1/2	0/1	1/3	

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

Mathematics

Your Child's Achievement Level:	**Partially Meeting Expectations**
Your Child's Score:	**485**

485

440	470	500	530	560
Not Meeting Expectations	**Partially Meeting Expectations**	Meeting Expectations	Exceeding Expectations	

A student who performed at this level did not meet grade-level expectations in this subject.

The school, in consultation with the student's parent/guardian, should determine the coordinated academic assistance and/or additional instruction the student needs to succeed in this subject.

A student who performed at this level partially met grade-level expectations in this subject.

The school, in consultation with the student's parent/guardian, should consider whether the student needs additional academic assistance to succeed in this subject.

A student who performed at this level met grade-level expectations and is academically on track to succeed in the current grade in this subject.

A student who performed at this level exceeded grade-level expectations by demonstrating mastery of the subject matter.

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

Achievement

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

Your Child's Grade	Score	Year	School	District	State
4	485	2025	489	484	491
3	464	2024			

Average Score

2025 Student Growth Percentiles

Your child's score this year is the same as or better than 89 percent of Rhode Island students who had a similar score to your child on the assessment(s) in a previous year(s).

Your Child — 89
School — 65
District — 51
State Avg. — 50

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

Reporting Category	Points Earned by Your Child	Total Possible Points	School	District	State
Operations and Algebraic Thinking	7	11	4.8	4.4	5.1
Number and Operations in Base Ten	3	10	5.9	4.9	5.7
Number and Operations-Fractions	8	16	8.8	7.6	8.2
Measurement and Data	5	11	4.8	4.6	5.6
Geometry	0	6	1.6	2.0	2.6

Average Points

Individual Test Questions

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

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

APPENDIX E

ITEM-LEVEL CLASSICAL STATISTICS

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

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
EL305848619	MC	9,849	0.74	0.49	0
EL305848846	MC	9,849	0.47	0.41	0
EL305849356	MC	9,849	0.38	0.19	0
EL305855968	MC	9,849	0.68	0.41	0
EL305856828	MC	9,849	0.53	0.35	0
EL305857179	MC	9,849	0.67	0.45	0
EL305857557	MC	9,849	0.50	0.37	0
EL305867369	MC	9,849	0.55	0.54	0
EL305871854	MC	9,849	0.59	0.52	0
EL306536295	MC	9,849	0.71	0.57	0
EL306538230	MC	9,849	0.68	0.52	0
EL306542014	MC	9,849	0.41	0.43	0
EL306542408	MC	9,849	0.62	0.48	0
EL306542695	MC	9,849	0.42	0.50	0
EL306551827	MC	9,849	0.69	0.47	0
EL311051383	MC	9,849	0.67	0.41	0
EL311056839	MC	9,849	0.61	0.48	0
EL311058588	MC	9,849	0.46	0.40	0
EL311059365	MC	9,849	0.66	0.47	0
EL311132236	MC	9,849	0.76	0.49	0
EL311137156	MC	9,849	0.57	0.54	0
EL311155968	MC	9,849	0.66	0.46	0
EL311158115	MC	9,849	0.50	0.36	0
EL311160399	MC	9,849	0.61	0.49	0
EL305868518	OR	9,849	0.37	0.64	0
EL305871312	OR	9,849	0.58	0.44	0
EL305885122	OR	9,849	0.36	0.65	1
EL306539868	OR	9,849	0.36	0.55	0
EL311057786	OR	9,849	0.57	0.46	0
EL311078713	OR	9,849	0.47	0.59	0
EL311063921#SCORE_TRAIT_CONV	ES	9,849	0.26	0.75	1
EL311063921#SCORE_TRAIT_IDEADEV	ES	9,849	0.19	0.74	1

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

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

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
EL208629059	MC	9,495	0.63	0.58	0
EL208629200	MC	9,495	0.60	0.53	0
EL208629667	MC	9,495	0.59	0.46	0
EL208629781	MC	9,495	0.65	0.47	0
EL208630729	MC	9,495	0.62	0.46	0
EL208631197	MC	9,495	0.62	0.38	0
EL208631698	MC	9,495	0.77	0.51	0
EL208632061	MC	9,495	0.83	0.55	0
EL208632192	MC	9,495	0.77	0.57	0
EL291009	MC	9,495	0.74	0.55	0
EL291015	MC	9,495	0.60	0.35	0
EL307283777	MC	9,495	0.44	0.39	0
EL307365872	MC	9,495	0.73	0.58	0
EL307370948	MC	9,495	0.60	0.42	0
EL307385093	MC	9,495	0.63	0.55	0
EL307400674	MC	9,495	0.68	0.54	0
EL307470135	MC	9,495	0.56	0.47	0
EL307500095	MC	9,495	0.59	0.44	0
EL307581198	MC	9,495	0.74	0.56	0
EL307583130	MC	9,495	0.54	0.29	0
EL909549145	MC	9,495	0.44	0.45	0
EL909550937	MC	9,495	0.40	0.48	0
EL912955945	MC	9,495	0.43	0.34	0
EL916346163	MC	9,495	0.80	0.52	0
EL208630869	OR	9,495	0.77	0.63	0
EL208631896	OR	9,495	0.75	0.47	0
EL291024	OR	9,495	0.36	0.58	0
EL307375298	OR	9,495	0.37	0.62	0
EL307485646	OR	9,495	0.49	0.66	0
EL307584265	OR	9,495	0.70	0.49	0
EL208631509#SCORE_TRAIT_CONV	ES	9,495	0.36	0.73	1
EL208631509#SCORE_TRAIT_IDEADEV	ES	9,495	0.28	0.71	1

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

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

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
EL207164820	MC	9,749	0.52	0.49	0
EL207274050	MC	9,749	0.45	0.44	0
EL207278877	MC	9,749	0.87	0.48	0
EL207341702	MC	9,749	0.59	0.30	0
EL207342693	MC	9,749	0.58	0.32	0
EL207360230	MC	9,749	0.72	0.52	0
EL307061151	MC	9,749	0.65	0.49	0
EL307063923	MC	9,749	0.60	0.49	0
EL307077265	MC	9,749	0.58	0.40	0
EL307137436	MC	9,749	0.27	0.21	0
EL307138100	MC	9,749	0.89	0.42	0
EL307153223	MC	9,749	0.70	0.48	0
EL307175118	MC	9,749	0.80	0.57	0
EL307249807	MC	9,749	0.74	0.52	0
EL307253066	MC	9,749	0.80	0.56	0
EL736467737	MC	9,749	0.76	0.36	0
EL736469533	MC	9,749	0.42	0.34	0
EL736469872	MC	9,749	0.71	0.54	0
EL736470482	MC	9,749	0.46	0.39	0
EL736471910	MC	9,749	0.67	0.41	0
EL736472217	MC	9,749	0.62	0.35	0
EL736473790	MC	9,749	0.45	0.44	0
EL736474369	MC	9,749	0.60	0.36	0
EL827627427	MC	9,749	0.35	0.30	0
EL207277420	OR	9,749	0.35	0.44	0
EL307156922	OR	9,749	0.70	0.57	0
EL307176181	OR	9,749	0.49	0.60	0
EL736475157	OR	9,749	0.35	0.37	0
EL736477190	OR	9,749	0.33	0.51	0
EL307267042#SCORE_TRAIT_CONV	ES	9,749	0.49	0.80	0
EL307267042#SCORE_TRAIT_IDEADEV	ES	9,749	0.39	0.78	0
EL736478825#SCORE_TRAIT_CONV	ES	9,749	0.40	0.76	0
EL736478825#SCORE_TRAIT_IDEADEV	ES	9,749	0.31	0.78	0

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

Table E-4. Item-Level Classical Test Theory Statistics—ELA Grade 6

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
EL308558989	MC	9,583	0.57	0.43	0
EL308559785	MC	9,583	0.64	0.49	0
EL308576679	MC	9,583	0.52	0.47	0
EL308630437	MC	9,583	0.42	0.29	0
EL308672079	MC	9,583	0.69	0.58	0
EL312142482	MC	9,583	0.55	0.45	0
EL312978480	MC	9,583	0.62	0.49	0
EL312980576	MC	9,583	0.51	0.45	0
EL314063698	MC	9,583	0.53	0.44	0
EL314148707	MC	9,583	0.62	0.42	0
EL314158725	MC	9,583	0.51	0.27	0
EL314164847	MC	9,583	0.59	0.50	0
EL314429542	MC	9,583	0.54	0.38	0
EL314478387	MC	9,583	0.66	0.47	0
EL314482379	MC	9,583	0.35	0.31	0
EL315274742	MC	9,583	0.44	0.29	0
EL315278077	MC	9,583	0.75	0.47	0
EL315281344	MC	9,583	0.71	0.29	0
EL315301385	MC	9,583	0.55	0.49	0
EL315381843	MC	9,583	0.68	0.52	0
EL315462998	MC	9,583	0.59	0.48	0
EL315482249	MC	9,583	0.60	0.39	0
EL315486389	MC	9,583	0.48	0.52	0
EL319246845	MC	9,583	0.38	0.38	0
EL308553441	OR	9,583	0.72	0.66	0
EL314140609	OR	9,583	0.59	0.62	0
EL314379145	OR	9,583	0.44	0.46	0
EL315276310	OR	9,583	0.77	0.61	0
EL315300123	OR	9,583	0.64	0.59	0
EL314484549#SCORE_TRAIT_CONV	ES	9,583	0.49	0.83	0
EL314484549#SCORE_TRAIT_IDEADEV	ES	9,583	0.31	0.82	0
EL315605273#SCORE_TRAIT_CONV	ES	9,583	0.52	0.83	1
EL315605273#SCORE_TRAIT_IDEADEV	ES	9,583	0.33	0.82	1

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

Table E-5. Item-Level Classical Test Theory Statistics—ELA Grade 7

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
EL206078126	MC	9,680	0.56	0.49	0
EL206078710	MC	9,680	0.44	0.33	0
EL206079400	MC	9,680	0.48	0.48	0
EL213939721	MC	9,680	0.58	0.53	0
EL305230669	MC	9,680	0.58	0.45	0
EL305244758	MC	9,680	0.51	0.47	0
EL306347661	MC	9,680	0.74	0.45	0
EL306348894	MC	9,680	0.49	0.35	0
EL306443148	MC	9,680	0.74	0.57	0
EL306569606	MC	9,680	0.71	0.41	0
EL306571635	MC	9,680	0.71	0.48	0
EL306633578	MC	9,680	0.72	0.48	0
EL306636364	MC	9,680	0.71	0.50	0
EL306637041	MC	9,680	0.62	0.56	0
EL306637866	MC	9,680	0.57	0.54	0
EL912338456	MC	9,680	0.41	0.31	0
EL912344338	MC	9,680	0.64	0.43	0
EL912500351	MC	9,680	0.63	0.42	0
EL912501540	MC	9,680	0.71	0.55	0
EL912582905	MC	9,680	0.73	0.52	0
EL912620913	MC	9,680	0.50	0.36	0
EL912621697	MC	9,680	0.63	0.55	0
EL912622150	MC	9,680	0.70	0.48	0
EL912623426	MC	9,680	0.43	0.34	0
EL305250618	OR	9,680	0.47	0.55	0
EL306463000	OR	9,680	0.58	0.62	0
EL306568260	OR	9,680	0.37	0.46	0
EL912583971	OR	9,680	0.53	0.41	0
EL917655389	OR	9,680	0.45	0.51	0
EL306400886#SCORE_TRAIT_CONV	ES	9,680	0.55	0.84	1
EL306400886#SCORE_TRAIT_IDEADEV	ES	9,680	0.31	0.83	1
EL912346409#SCORE_TRAIT_CONV	ES	9,680	0.54	0.81	1
EL912346409#SCORE_TRAIT_IDEADEV	ES	9,680	0.35	0.81	1

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

Table E-6. Item-Level Classical Test Theory Statistics—ELA Grade 8

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
EL206801492	MC	9,800	0.79	0.49	0
EL206802113	MC	9,800	0.79	0.45	0
EL206804483	MC	9,800	0.77	0.54	0
EL206804808	MC	9,800	0.66	0.41	0
EL206805580	MC	9,800	0.64	0.47	0
EL206806025	MC	9,800	0.75	0.39	0
EL206808101	MC	9,800	0.56	0.42	0
EL208447958	MC	9,800	0.52	0.35	0
EL208451039	MC	9,800	0.77	0.56	0
EL209839730	MC	9,800	0.48	0.31	0
EL209877427	MC	9,800	0.52	0.39	0
EL209878098	MC	9,800	0.67	0.49	0
EL209878664	MC	9,800	0.68	0.57	0
EL213257312	MC	9,800	0.50	0.54	0
EL215725121	MC	9,800	0.69	0.52	0
EL306531694	MC	9,800	0.63	0.32	0
EL306532662	MC	9,800	0.65	0.38	0
EL306541779	MC	9,800	0.33	0.32	0
EL306542147	MC	9,800	0.68	0.47	0
EL306543569	MC	9,800	0.53	0.46	0
EL306550178	MC	9,800	0.40	0.34	0
EL306557022	MC	9,800	0.65	0.39	0
EL306619088	MC	9,800	0.50	0.52	0
EL306620254	MC	9,800	0.47	0.51	0
EL206809440	OR	9,800	0.56	0.61	0
EL208451735	OR	9,800	0.69	0.59	0
EL209978403	OR	9,800	0.50	0.50	0
EL306633886	OR	9,800	0.73	0.58	0
EL306634908	OR	9,800	0.58	0.58	0
EL206813674#SCORE_TRAIT_CONV	ES	9,800	0.55	0.82	1
EL206813674#SCORE_TRAIT_IDEADEV	ES	9,800	0.36	0.82	1
EL306640647#SCORE_TRAIT_CONV	ES	9,800	0.53	0.83	1
EL306640647#SCORE_TRAIT_IDEADEV	ES	9,800	0.33	0.84	1

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

Table E-7. Item-Level Classical Test Theory Statistics—Mathematics Grade 3

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
MA001034206	MC	9,797	0.75	0.44	0
MA001336857	MC	9,797	0.68	0.46	0
MA134676677	MC	9,797	0.33	0.45	0
MA134830909	MC	9,797	0.23	0.43	0
MA134867959	MC	9,797	0.55	0.47	0
MA231152219	MC	9,797	0.47	0.43	0
MA231157499	MC	9,797	0.46	0.41	0
MA293459	MC	9,797	0.51	0.55	0
MA297412	MC	9,797	0.66	0.46	0
MA310834	MC	9,797	0.57	0.54	0
MA310842	MC	9,797	0.60	0.53	0
MA311274	MC	9,797	0.77	0.49	0
MA735665944	MC	9,797	0.55	0.54	0
MA735732140	MC	9,797	0.81	0.51	0
MA735734501	MC	9,797	0.72	0.45	0
MA735755736	MC	9,797	0.63	0.45	0
MA902374248	MC	9,797	0.55	0.51	0
MA902654429	MC	9,797	0.41	0.33	0
MA001137079	OR	9,797	0.60	0.61	0
MA001330803	OR	9,797	0.51	0.64	0
MA001631751	OR	9,797	0.72	0.57	0
MA002751340	OR	9,797	0.31	0.27	0
MA134872363	OR	9,797	0.52	0.64	0
MA134930940	OR	9,797	0.44	0.75	0
MA134940265	OR	9,797	0.71	0.47	0
MA134942387	OR	9,797	0.55	0.63	0
MA231230636	OR	9,797	0.53	0.54	0
MA260584A	OR	9,797	0.53	0.72	0
MA261857A	OR	9,797	0.40	0.75	1
MA297504	OR	9,797	0.63	0.48	0
MA310884	OR	9,797	0.52	0.65	0
MA703073138	OR	9,797	0.33	0.56	0
MA735754028	OR	9,797	0.72	0.64	0
MA735848960	OR	9,797	0.37	0.52	0
MA735951978	OR	9,797	0.35	0.75	1
MA736028646	OR	9,797	0.55	0.64	0
MA900371312	OR	9,797	0.62	0.59	0
MA900371363	OR	9,797	0.36	0.40	0
MA900573685	OR	9,797	0.54	0.62	0
MA935137496	OR	9,797	0.43	0.56	1

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

Table E-8. Item-Level Classical Test Theory Statistics—Mathematics Grade 4

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
MA136447421	MC	9,403	0.56	0.57	0
MA202029218	MC	9,403	0.59	0.33	0
MA232239522	MC	9,403	0.54	0.35	0
MA233049856	MC	9,403	0.46	0.45	0
MA233051799	MC	9,403	0.79	0.35	0
MA297984	MC	9,403	0.53	0.60	0
MA307075	MC	9,403	0.87	0.42	0
MA307314	MC	9,403	0.66	0.50	0
MA311543	MC	9,403	0.59	0.49	0
MA311554	MC	9,403	0.61	0.47	0
MA800763627	MC	9,403	0.44	0.61	0
MA900750340	MC	9,403	0.63	0.48	0
MA900754381	MC	9,403	0.38	0.55	0
MA000732007	OR	9,403	0.48	0.54	0
MA002162929	OR	9,403	0.40	0.45	0
MA010534486	OR	9,403	0.38	0.50	0
MA136448521	OR	9,403	0.44	0.59	0
MA200344609	OR	9,403	0.65	0.52	0
MA231836735	OR	9,403	0.48	0.61	0
MA231843918	OR	9,403	0.48	0.59	0
MA231875780	OR	9,403	0.52	0.77	0
MA232254177	OR	9,403	0.48	0.52	0
MA232261850	OR	9,403	0.71	0.47	0
MA232725235	OR	9,403	0.64	0.62	0
MA232909389	OR	9,403	0.42	0.55	0
MA232979874	OR	9,403	0.83	0.48	0
MA250533	OR	9,403	0.51	0.54	0
MA258228	OR	9,403	0.71	0.49	0
MA303317	OR	9,403	0.39	0.56	0
MA303335	OR	9,403	0.40	0.70	0
MA713677363	OR	9,403	0.69	0.59	0
MA800652951	OR	9,403	0.23	0.46	0
MA800750938	OR	9,403	0.45	0.75	0
MA800780887	OR	9,403	0.49	0.81	0
MA803957909	OR	9,403	0.35	0.70	0
MA900740880	OR	9,403	0.41	0.58	0
MA900741771	OR	9,403	0.53	0.59	0
MA900750085	OR	9,403	0.45	0.60	0
MA900751271	OR	9,403	0.58	0.47	0
MA900776517	OR	9,403	0.36	0.46	0

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

Table E-9. Item-Level Classical Test Theory Statistics—Mathematics Grade 5

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
MA000859040	MC	9,716	0.64	0.35	0
MA000952312	MC	9,716	0.49	0.41	0
MA000963503	MC	9,716	0.62	0.35	0
MA201552031	MC	9,716	0.62	0.31	0
MA202000453	MC	9,716	0.77	0.45	0
MA233233270	MC	9,716	0.45	0.38	0
MA282124	MC	9,716	0.53	0.27	0
MA298013	MC	9,716	0.75	0.52	0
MA298014	MC	9,716	0.77	0.51	0
MA306430	MC	9,716	0.45	0.59	0
MA311279	MC	9,716	0.61	0.55	0
MA704359445A	MC	9,716	0.42	0.37	0
MA801230425	MC	9,716	0.30	0.34	0
MA801654509	MC	9,716	0.47	0.52	0
MA900660071	MC	9,716	0.40	0.52	0
MA900662425	MC	9,716	0.63	0.41	0
MA900980248	MC	9,716	0.56	0.42	0
MA901137741	MC	9,716	0.49	0.39	0
MA904353319	MC	9,716	0.79	0.40	0
MA000846693	OR	9,716	0.41	0.61	0
MA000965213	OR	9,716	0.38	0.54	0
MA001060832	OR	9,716	0.50	0.52	0
MA001145679	OR	9,716	0.21	0.43	1
MA232642946	OR	9,716	0.30	0.79	0
MA233369406	OR	9,716	0.46	0.63	0
MA233458877	OR	9,716	0.27	0.52	0
MA311283	OR	9,716	0.62	0.54	0
MA704359624	OR	9,716	0.69	0.34	0
MA715102228	OR	9,716	0.35	0.62	0
MA715102462	OR	9,716	0.42	0.76	0
MA800662477	OR	9,716	0.26	0.47	0
MA801673695	OR	9,716	0.20	0.38	0
MA802302084	OR	9,716	0.37	0.52	0
MA802306160	OR	9,716	0.59	0.58	0
MA802371654	OR	9,716	0.46	0.78	0
MA900835182	OR	9,716	0.45	0.42	0
MA901364051	OR	9,716	0.25	0.42	0
MA903746975	OR	9,716	0.48	0.77	0
MA904453014	OR	9,716	0.52	0.53	0
MA935150419	OR	9,716	0.45	0.61	0

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

Table E-10. Item-Level Classical Test Theory Statistics—Mathematics Grade 6

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
MA135360248	MC	9,515	0.83	0.37	0
MA272290	MC	9,515	0.54	0.45	0
MA282127	MC	9,515	0.49	0.52	0
MA296377	MC	9,515	0.70	0.42	0
MA296382	MC	9,515	0.47	0.31	0
MA303705	MC	9,515	0.37	0.39	0
MA307250	MC	9,515	0.34	0.47	0
MA307363	MC	9,515	0.76	0.34	0
MA311652	MC	9,515	0.40	0.48	0
MA736452061	MC	9,515	0.31	0.49	0
MA736453104	MC	9,515	0.59	0.37	0
MA736454897	MC	9,515	0.46	0.29	0
MA805104566	MC	9,515	0.42	0.42	0
MA805111429	MC	9,515	0.50	0.31	0
MA900281418	MC	9,515	0.63	0.49	0
MA908142878	MC	9,515	0.37	0.29	0
MA001554177	OR	9,515	0.51	0.50	0
MA001556929	OR	9,515	0.72	0.55	0
MA001577731	OR	9,515	0.49	0.56	0
MA001603294	OR	9,515	0.72	0.57	0
MA002526793	OR	9,515	0.59	0.45	0
MA003477341	OR	9,515	0.45	0.63	0
MA233752432	OR	9,515	0.39	0.79	0
MA233849848	OR	9,515	0.34	0.61	0
MA287883	OR	9,515	0.29	0.76	1
MA307362	OR	9,515	0.60	0.62	0
MA703149512	OR	9,515	0.56	0.50	0
MA713677108	OR	9,515	0.33	0.54	0
MA736449649	OR	9,515	0.24	0.58	0
MA800203270	OR	9,515	0.29	0.74	0
MA800259417	OR	9,515	0.49	0.76	1
MA805109765	OR	9,515	0.42	0.47	0
MA805280133	OR	9,515	0.24	0.34	0
MA805280170	OR	9,515	0.55	0.55	0
MA805301630	OR	9,515	0.44	0.41	0
MA900378821	OR	9,515	0.41	0.77	1
MA900432715	OR	9,515	0.26	0.54	0
MA901737396	OR	9,515	0.35	0.55	0
MA902348152	OR	9,515	0.33	0.48	0
MA902761648	OR	9,515	0.32	0.39	0

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

Table E-11. Item-Level Classical Test Theory Statistics—Mathematics Grade 7

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
MA000971342	MC	9,560	0.39	0.52	0
MA001684984	MC	9,560	0.28	0.33	0
MA005200422	MC	9,560	0.48	0.22	0
MA005204778	MC	9,560	0.41	0.51	0
MA200144334	MC	9,560	0.65	0.43	0
MA200256305	MC	9,560	0.58	0.47	0
MA201106705	MC	9,560	0.46	0.43	0
MA234550897	MC	9,560	0.43	0.50	0
MA272764	MC	9,560	0.53	0.31	0
MA281422	MC	9,560	0.36	0.58	0
MA298180	MC	9,560	0.75	0.38	0
MA301838	MC	9,560	0.36	0.44	0
MA306595	MC	9,560	0.69	0.46	0
MA306624	MC	9,560	0.23	0.16	0
MA311074	MC	9,560	0.39	0.36	0
MA900552355	MC	9,560	0.51	0.46	0
MA900556478	MC	9,560	0.30	0.37	0
MA900566128	MC	9,560	0.23	0.40	0
MA900665080	MC	9,560	0.29	0.43	0
MA903153837	MC	9,560	0.36	0.28	0
MA904159889	MC	9,560	0.52	0.47	0
MA001178265	OR	9,560	0.22	0.56	0
MA001678587	OR	9,560	0.27	0.65	1
MA005077768	OR	9,560	0.43	0.59	0
MA201203484	OR	9,560	0.25	0.70	0
MA234551702	OR	9,560	0.26	0.48	0
MA234553959	OR	9,560	0.71	0.51	0
MA234555649	OR	9,560	0.50	0.55	0
MA234561626	OR	9,560	0.23	0.58	0
MA235900389	OR	9,560	0.35	0.75	0
MA235905181	OR	9,560	0.36	0.82	0
MA236233939	OR	9,560	0.68	0.51	0
MA703876323	OR	9,560	0.12	0.68	0
MA717248260	OR	9,560	0.29	0.77	1
MA801105441	OR	9,560	0.21	0.53	0
MA802907874	OR	9,560	0.24	0.81	1
MA900567252	OR	9,560	0.33	0.61	0
MA900739359	OR	9,560	0.40	0.62	0
MA900765087	OR	9,560	0.51	0.66	0
MA904001752	OR	9,560	0.28	0.27	0

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

Table E-12. Item-Level Classical Test Theory Statistics—Mathematics Grade 8

Item Number	Item Type	Number	Difficulty	Discrimination	Percent Omitted
MA001736920	MC	9,664	0.53	0.41	0
MA002159196	MC	9,664	0.35	0.52	0
MA002178533	MC	9,664	0.44	0.46	0
MA002181298	MC	9,664	0.58	0.49	0
MA203023828	MC	9,664	0.79	0.44	0
MA228311	MC	9,664	0.72	0.52	0
MA235755845	MC	9,664	0.38	0.30	0
MA236476077	MC	9,664	0.78	0.50	0
MA264730	MC	9,664	0.53	0.52	0
MA301682	MC	9,664	0.54	0.37	0
MA301698	MC	9,664	0.51	0.52	0
MA306554	MC	9,664	0.56	0.54	0
MA307586	MC	9,664	0.47	0.49	0
MA307600	MC	9,664	0.47	0.29	0
MA311391	MC	9,664	0.60	0.20	0
MA800750287	MC	9,664	0.50	0.65	0
MA901143033	MC	9,664	0.49	0.41	0
MA001737758	OR	9,664	0.44	0.56	0
MA002170014	OR	9,664	0.51	0.61	0
MA003128642	OR	9,664	0.43	0.55	0
MA202726454	OR	9,664	0.48	0.52	0
MA202828823	OR	9,664	0.36	0.70	0
MA203016132	OR	9,664	0.55	0.57	0
MA203085405	OR	9,664	0.63	0.66	0
MA235753974	OR	9,664	0.31	0.58	0
MA236473688	OR	9,664	0.37	0.47	0
MA236558384	OR	9,664	0.37	0.84	1
MA301481	OR	9,664	0.21	0.66	1
MA311433	OR	9,664	0.32	0.78	2
MA704839576	OR	9,664	0.13	0.54	0
MA715919560	OR	9,664	0.17	0.55	0
MA715919661	OR	9,664	0.63	0.34	0
MA715919788	OR	9,664	0.30	0.63	0
MA800473031	OR	9,664	0.24	0.77	1
MA800659905	OR	9,664	0.25	0.42	0
MA800756470	OR	9,664	0.28	0.56	0
MA800880893	OR	9,664	0.47	0.53	0
MA804043870	OR	9,664	0.29	0.65	0
MA901253257	OR	9,664	0.35	0.38	0
MA902265010	OR	9,664	0.26	0.48	0

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

APPENDIX F

SCORE DISTRIBUTIONS

Table F-1. Item-Level Score Distributions for SR and OR Items and ESs—ELA

Grade	Item Number	Number	Total Possible Points	Percent of Students at Score Point 0	Percent of Students at Score Point 1	Percent of Students at Score Point 2	Percent of Students at Score Point 3	Percent of Students at Score Point 4	Percent of Students at Score Point 5
3	EL305868518	9,849	2	55	15	30	N/A	N/A	N/A
3	EL305871312	9,849	2	30	23	47	N/A	N/A	N/A
3	EL305885122	9,849	3	24	46	25	4	N/A	N/A
3	EL306539868	9,849	2	54	21	25	N/A	N/A	N/A
3	EL311057786	9,849	2	28	29	43	N/A	N/A	N/A
3	EL311063921#SCORE_TRAIT_Conv	9,849	3	44	36	15	4	N/A	N/A
3	EL311063921#SCORE_TRAIT_Ideadev	9,849	4	48	32	14	5	0	N/A
3	EL311078713	9,849	2	33	40	27	N/A	N/A	N/A
3	EL408927809	9,802	2	16	34	50	N/A	N/A	N/A
3	EL409231916	9,802	2	36	25	39	N/A	N/A	N/A
3	EL409345080	9,802	2	34	17	49	N/A	N/A	N/A
4	EL208630869	9,495	2	7	31	62	N/A	N/A	N/A
4	EL208631509#SCORE_TRAIT_Conv	9,495	3	27	44	22	6	N/A	N/A
4	EL208631509#SCORE_TRAIT_Ideadev	9,495	4	23	49	20	5	2	N/A
4	EL208631896	9,495	2	10	30	60	N/A	N/A	N/A
4	EL291024	9,495	2	56	16	29	N/A	N/A	N/A
4	EL307375298	9,495	3	23	47	25	5	N/A	N/A
4	EL307485646	9,495	2	36	30	34	N/A	N/A	N/A
4	EL307584265	9,495	2	22	15	63	N/A	N/A	N/A
4	EL408472889	9,458	2	28	33	39	N/A	N/A	N/A
4	EL408484798	9,458	2	7	37	55	N/A	N/A	N/A
4	EL408564605	9,458	2	16	22	61	N/A	N/A	N/A
5	EL207277420	9,749	2	59	11	30	N/A	N/A	N/A
5	EL301646707	9,707	2	4	48	49	N/A	N/A	N/A
5	EL301775613	9,707	2	20	29	51	N/A	N/A	N/A
5	EL307156922	9,749	2	12	35	52	N/A	N/A	N/A
5	EL307176181	9,749	2	48	6	46	N/A	N/A	N/A
5	EL307267042#SCORE_TRAIT_Conv	9,749	3	19	32	29	19	N/A	N/A
5	EL307267042#SCORE_TRAIT_Ideadev	9,749	4	14	39	29	16	3	N/A
5	EL405956937	9,707	2	19	15	65	N/A	N/A	N/A
5	EL736475157	9,749	2	49	33	18	N/A	N/A	N/A
5	EL736477190	9,749	2	57	20	23	N/A	N/A	N/A
5	EL736478825#SCORE_TRAIT_Conv	9,749	3	22	42	28	7	N/A	N/A
5	EL736478825#SCORE_TRAIT_Ideadev	9,749	4	28	34	28	8	2	N/A
6	EL308553441	9,583	2	19	19	62	N/A	N/A	N/A
6	EL314140609	9,583	2	29	24	47	N/A	N/A	N/A
6	EL314379145	9,583	2	34	43	23	N/A	N/A	N/A
6	EL314484549#SCORE_TRAIT_Conv	9,583	3	22	30	27	20	N/A	N/A
6	EL314484549#SCORE_TRAIT_Ideadev	9,583	5	22	29	27	14	6	1
6	EL314851900	9,551	2	38	45	16	N/A	N/A	N/A
6	EL314858443	9,551	2	24	18	57	N/A	N/A	N/A
6	EL315276310	9,583	2	19	7	74	N/A	N/A	N/A
6	EL315300123	9,583	2	19	33	48	N/A	N/A	N/A
6	EL315605273#SCORE_TRAIT_Conv	9,583	3	20	28	26	25	N/A	N/A
6	EL315605273#SCORE_TRAIT_Ideadev	9,583	5	20	28	24	17	9	1
6	EL405961354	9,551	2	23	32	45	N/A	N/A	N/A
7	EL305250618	9,680	2	27	52	21	N/A	N/A	N/A
7	EL306400886#SCORE_TRAIT_Conv	9,680	3	17	28	26	28	N/A	N/A
7	EL306400886#SCORE_TRAIT_Ideadev	9,680	5	18	32	28	19	3	0
7	EL306463000	9,680	2	31	21	47	N/A	N/A	N/A
7	EL306568260	9,680	2	45	38	18	N/A	N/A	N/A
7	EL332401361	9,669	2	31	31	38	N/A	N/A	N/A

Grade	Item Number	Number	Total Possible Points	Percent of Students at Score Point 0	Percent of Students at Score Point 1	Percent of Students at Score Point 2	Percent of Students at Score Point 3	Percent of Students at Score Point 4	Percent of Students at Score Point 5
7	EL334250934	9,669	2	59	15	26	N/A	N/A	N/A
7	EL403134059	9,669	2	20	19	61	N/A	N/A	N/A
7	EL912346409#SCORE_TRAIT_Conv	9,680	3	15	32	29	24	N/A	N/A
7	EL912346409#SCORE_TRAIT_Ideadev	9,680	5	12	35	25	21	6	0
7	EL912583971	9,680	2	18	57	25	N/A	N/A	N/A
7	EL917655389	9,680	2	52	5	43	N/A	N/A	N/A
8	EL206809440	9,800	2	40	8	52	N/A	N/A	N/A
8	EL206813674#SCORE_TRAIT_Conv	9,800	3	17	30	23	30	N/A	N/A
8	EL206813674#SCORE_TRAIT_Ideadev	9,800	5	17	29	22	19	12	1
8	EL208451735	9,800	2	15	33	52	N/A	N/A	N/A
8	EL209978403	9,800	2	40	20	40	N/A	N/A	N/A
8	EL306633886	9,800	2	19	17	64	N/A	N/A	N/A
8	EL306634908	9,800	2	25	34	41	N/A	N/A	N/A
8	EL306640647#SCORE_TRAIT_Conv	9,800	3	21	23	29	26	N/A	N/A
8	EL306640647#SCORE_TRAIT_Ideadev	9,800	5	22	23	27	19	7	0
8	EL406435295	9,785	2	22	53	25	N/A	N/A	N/A
8	EL406943951	9,785	2	22	37	41	N/A	N/A	N/A
8	EL406945342	9,785	2	12	37	51	N/A	N/A	N/A

Table F-2 Item-Level Score Distributions for SR and OR Items—Mathematics

Grade	Item Number	Number	Total Possible Points	Percent of Students at Score Point 0	Percent of Students at Score Point 1	Percent of Students at Score Point 2	Percent of Students at Score Point 3	Percent of Students at Score Point 4	Percent of Students at Score Point 5
3	MA134930940	9,797	3	28	25	34	13	N/A	N/A
3	MA260584A	9,797	3	17	27	37	19	N/A	N/A
3	MA261857A	9,797	3	31	32	20	16	N/A	N/A
3	MA735951978	9,797	3	43	22	19	16	N/A	N/A
4	MA231875780	9,403	4	12	23	27	25	14	N/A
4	MA303335	9,403	4	20	33	21	15	10	N/A
4	MA713677363	9,403	2	15	31	54	N/A	N/A	N/A
4	MA800750938	9,403	2	41	27	31	N/A	N/A	N/A
4	MA800780887	9,403	4	14	26	24	20	15	N/A
4	MA803957909	9,403	4	21	39	22	12	5	N/A
5	MA000965213	9,716	2	39	45	16	N/A	N/A	N/A
5	MA232642946	9,716	4	45	23	11	9	12	N/A
5	MA715102462	9,716	4	26	18	31	12	13	N/A
5	MA802306160	9,716	2	14	54	32	N/A	N/A	N/A
5	MA802371654	9,716	4	20	20	28	18	14	N/A
5	MA903746975	9,716	4	16	27	21	19	17	N/A
6	MA233752432	9,515	4	27	24	21	18	9	N/A
6	MA287883	9,515	4	50	15	11	12	11	N/A
6	MA302246376	9,490	2	52	25	24	N/A	N/A	N/A
6	MA800203270	9,515	2	59	24	17	N/A	N/A	N/A
6	MA800259417	9,515	4	20	17	22	26	15	N/A
6	MA805301630	9,515	2	48	16	36	N/A	N/A	N/A
6	MA900378821	9,515	4	14	43	18	13	11	N/A
7	MA235900389	9,560	4	25	33	24	12	5	N/A
7	MA235905181	9,560	4	16	48	18	10	8	N/A
7	MA703876323	9,560	2	83	9	8	N/A	N/A	N/A
7	MA717248260	9,560	4	30	39	15	10	4	N/A
7	MA802907874	9,560	4	49	19	18	6	6	N/A
7	MA900765087	9,560	2	19	59	21	N/A	N/A	N/A
8	MA203016132	9,664	2	26	38	36	N/A	N/A	N/A
8	MA203085405	9,664	4	6	14	22	35	22	N/A
8	MA203126627	9,656	2	29	57	13	N/A	N/A	N/A
8	MA236558384	9,664	4	32	30	11	11	15	N/A
8	MA311433	9,664	4	32	29	20	11	6	N/A
8	MA800473031	9,664	4	55	14	15	10	5	N/A
8	MA800880893	9,664	2	40	25	34	N/A	N/A	N/A

APPENDIX G

DIFFERENTIAL ITEM FUNCTIONING RESULTS

Table G-1. Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored—ELA

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

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

Grade	Group Reference	Group Focal	Item Type	Number of Items	Number "Low" Total	Number "Low" Favoring Reference	Number "Low" Favoring Focal	Number "High" Total	Number "High" Favoring Reference	Number "High" Favoring Focal
8	Students Without Disabilities	Students with Disabilities	OR	5	0	0	0	0	0	0
8	Students Without Disabilities	Students with Disabilities	ES	3	0	0	0	0	0	0

Table G-2. Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored—Mathematics

Grade	Group Reference	Group Focal	Item Type	Number of Items	Number "Low" Total	Number "Low" Favoring Reference	Number "Low" Favoring Focal	Number "High" Total	Number "High" Favoring Reference	Number "High" Favoring Focal
3	Male	Female	MC	18	3	0	3	0	0	0
3	Male	Female	OR	22	2	2	0	1	1	0
3	Not ELL	ELL	MC	18	0	0	0	0	0	0
3	Not ELL	ELL	OR	22	4	2	2	0	0	0
3	Not Economically Disadvantaged	Economically Disadvantaged	MC	18	0	0	0	0	0	0
3	Not Economically Disadvantaged	Economically Disadvantaged	OR	22	0	0	0	0	0	0
3	White	African American	MC	18	2	1	1	0	0	0
3	White	African American	OR	22	6	4	2	0	0	0
3	White	Hispanic / Latino	MC	18	0	0	0	0	0	0
3	White	Hispanic / Latino	OR	22	1	0	1	0	0	0
3	Students Without Disabilities	Students with Disabilities	MC	18	0	0	0	0	0	0
3	Students Without Disabilities	Students with Disabilities	OR	22	0	0	0	0	0	0
4	Male	Female	MC	13	2	2	0	0	0	0
4	Male	Female	OR	27	6	4	2	0	0	0
4	Not ELL	ELL	MC	13	0	0	0	0	0	0
4	Not ELL	ELL	OR	27	2	1	1	0	0	0
4	Not Economically Disadvantaged	Economically Disadvantaged	MC	13	0	0	0	0	0	0
4	Not Economically Disadvantaged	Economically Disadvantaged	OR	27	0	0	0	0	0	0
4	White	African American	MC	13	0	0	0	0	0	0
4	White	African American	OR	27	2	1	1	0	0	0
4	White	Hispanic / Latino	MC	13	0	0	0	0	0	0
4	White	Hispanic / Latino	OR	27	2	1	1	0	0	0
4	Students Without Disabilities	Students with Disabilities	MC	13	1	0	1	0	0	0
4	Students Without Disabilities	Students with Disabilities	OR	27	0	0	0	0	0	0
5	Male	Female	MC	19	2	2	0	0	0	0
5	Male	Female	OR	21	0	0	0	0	0	0
5	Not ELL	ELL	MC	19	1	1	0	0	0	0
5	Not ELL	ELL	OR	21	0	0	0	0	0	0
5	Not Economically Disadvantaged	Economically Disadvantaged	MC	19	0	0	0	0	0	0
5	Not Economically Disadvantaged	Economically Disadvantaged	OR	21	0	0	0	0	0	0
5	White	African American	MC	19	1	1	0	0	0	0
5	White	African American	OR	21	0	0	0	0	0	0
5	White	Hispanic / Latino	MC	19	0	0	0	0	0	0
5	White	Hispanic / Latino	OR	21	0	0	0	0	0	0
5	Students Without Disabilities	Students with Disabilities	MC	19	0	0	0	0	0	0
5	Students Without Disabilities	Students with Disabilities	OR	21	0	0	0	0	0	0
6	Male	Female	MC	16	2	2	0	0	0	0
6	Male	Female	OR	24	3	1	2	0	0	0
6	Not ELL	ELL	MC	16	3	0	3	0	0	0
6	Not ELL	ELL	OR	24	2	2	0	0	0	0
6	Not Economically Disadvantaged	Economically Disadvantaged	MC	16	0	0	0	0	0	0
6	Not Economically Disadvantaged	Economically Disadvantaged	OR	24	0	0	0	0	0	0
6	White	African American	MC	16	2	0	2	0	0	0
6	White	African American	OR	24	0	0	0	0	0	0

Grade	Group Reference	Group Focal	Item Type	Number of Items	Number "Low" Total	Number "Low" Favoring Reference	Number "Low" Favoring Focal	Number "High" Total	Number "High" Favoring Reference	Number "High" Favoring Focal
6	White	Hispanic / Latino	MC	16	0	0	0	0	0	0
6	White	Hispanic / Latino	OR	24	0	0	0	0	0	0
6	Students Without Disabilities	Students with Disabilities	MC	16	0	0	0	0	0	0
6	Students Without Disabilities	Students with Disabilities	OR	24	1	1	0	0	0	0
7	Male	Female	MC	21	3	2	1	0	0	0
7	Male	Female	OR	19	0	0	0	0	0	0
7	Not ELL	ELL	MC	21	4	1	3	0	0	0
7	Not ELL	ELL	OR	19	1	1	0	0	0	0
7	Not Economically Disadvantaged	Economically Disadvantaged	MC	21	2	0	2	0	0	0
7	Not Economically Disadvantaged	Economically Disadvantaged	OR	19	0	0	0	0	0	0
7	White	African American	MC	21	3	2	1	0	0	0
7	White	African American	OR	19	0	0	0	0	0	0
7	White	Hispanic / Latino	MC	21	3	1	2	0	0	0
7	White	Hispanic / Latino	OR	19	0	0	0	0	0	0
7	Students Without Disabilities	Students with Disabilities	MC	21	1	0	1	0	0	0
7	Students Without Disabilities	Students with Disabilities	OR	19	1	1	0	0	0	0
8	Male	Female	MC	17	0	0	0	0	0	0
8	Male	Female	OR	23	0	0	0	0	0	0
8	Not ELL	ELL	MC	17	0	0	0	0	0	0
8	Not ELL	ELL	OR	23	0	0	0	0	0	0
8	Not Economically Disadvantaged	Economically Disadvantaged	MC	17	0	0	0	0	0	0
8	Not Economically Disadvantaged	Economically Disadvantaged	OR	23	0	0	0	0	0	0
8	White	African American	MC	17	3	1	2	0	0	0
8	White	African American	OR	23	1	0	1	0	0	0
8	White	Hispanic / Latino	MC	17	0	0	0	0	0	0
8	White	Hispanic / Latino	OR	23	0	0	0	0	0	0
8	Students Without Disabilities	Students with Disabilities	MC	17	1	1	0	0	0	0
8	Students Without Disabilities	Students with Disabilities	OR	23	0	0	0	0	0	0

APPENDIX H
2024–25 MCAS EQUATING REPORT

Massachusetts Comprehensive Assessment System

2024–2025 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. Tabled Delta Analysis Results
    5. Tabled B/B Analysis Results
    6. Tabled Beta Analysis Results
    7. Final Item Parameters
    8. Decision Accuracy and Consistency (DAC)
    9. Fit Plots and Beta Plots of Watchlist Items

The final results of this equating will be included as part of the 2024 - 2025 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[1]

---

[1] *Note: 2025 is the first year that students with text-to-speech accommodations were included in the equating sample. Due to this change in the business rules, results in this section should be interpreted with caution.*

Table 1.1.1

Percentage of Students by Achievement Levels Categories

English Language Arts

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS	SD SS
3	2025	63649	17	39	38	6	44	-0.5	494.2	23.8
3	2024	59965	15	41	38	7	44	-1.8	495.7	22.9
3	2023	60542	14	40	39	7	46	0.3	496.0	23.4
3	2022	61648	13	41	39	7	46	-6.9	497.1	22.8
3	2021	50011	8	39	43	10	53	-5.7	500.1	22.4
3	2019	63602	6	36	48	11	58	5.8	504.8	21.8
3	2018	43046	6	41	43	10	53	1.0	501.8	20.6
3	2017	26459	7	41	43	8	52		500.1	21.1
4	2025	62216	17	41	37	5	42	3.0	493.5	23.7
4	2024	61103	16	46	34	5	39	-3.2	493.5	22.2
4	2023	61836	14	44	36	6	42	1.7	494.7	22.2
4	2022	62100	14	46	36	4	40	-11.2	493.8	21.1
4	2021	50867	11	38	45	6	51	-2.7	498.9	22.8
4	2019	65450	7	39	44	10	54	-0.5	502.6	21.1
4	2018	69078	7	38	44	10	55	3.5	502.2	21.6
4	2017	63918	8	41	43	8	51		500.1	21.2
5	2025	63664	13	47	34	5	39	-0.6	494.7	22.1
5	2024	62601	13	47	34	6	40	-6.0	494.6	21.9
5	2023	62316	13	41	41	5	46	3.8	496.0	21.8
5	2022	63620	11	47	37	5	42	-7.1	495.6	20.4
5	2021	51362	10	41	41	8	49	-5.0	497.9	22.1
5	2019	67933	6	39	47	8	54	-1.5	501.9	21.0
5	2018	69390	6	38	49	7	56	4.4	502.3	20.0
5	2017	28547	7	42	46	5	51		499.9	19.8
6	2025	64470	20	37	33	10	43	1.3	494.1	28.9
6	2024	62688	22	36	30	11	42	-2.5	494.6	28.9
6	2023	63574	21	35	36	8	44	1.3	494.0	28.1
6	2022	63887	20	37	35	8	43	-7.2	494.0	26.0
6	2021	51319	19	31	37	13	50	-5.7	498.4	28.3
6	2019	67612	11	33	42	13	56	3.4	502.5	25.6
6	2018	53988	10	38	42	11	52	-0.7	501.3	23.5
6	2017	29369	8	39	47	6	53		500.3	20.7
7	2025	64737	21	37	35	7	42	4.6	493.0	26.6
7	2024	63626	19	43	31	7	38	-4.4	492.6	25.6
7	2023	63711	17	41	34	8	42	-0.3	494.2	25.1
7	2022	65584	17	40	37	6	42	-3.2	493.7	23.6
7	2021	51120	17	37	39	7	46	-4.8	495.6	24.8
7	2019	67462	11	39	42	9	50	3.6	499.8	23.0
7	2018	66410	13	40	39	8	47	-6.5	497.4	23.4
7	2017	30209	8	38	48	6	53		500.2	20.8

Table 1.1.1 (continued)
Percentage of Students by Achievement Levels Categories
English Language Arts

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS	SD SS
8	2025	65482	21	34	34	11	45	0.5	494.5	28.7
8	2024	64139	21	34	33	12	44	-0.9	494.9	28.5
8	2023	65553	20	35	35	10	45	2.2	495.4	28.5
8	2022	67919	16	40	36	7	43	-0.3	494.8	24.8
8	2021	50822	15	41	37	7	44	-10.0	496.2	25.5
8	2019	67350	11	35	42	12	54	1.1	500.6	24.5
8	2018	69486	13	34	42	10	52	1.1	499.6	24.9
8	2017	65314	9	40	43	8	51		499.5	21.8

Table 1.1.2
Percentage of Students by Achievement Levels Categories
Mathematics

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS	SD SS
3	2025	63570	19	36	35	10	45	-8.3	496.3	26.8
3	2024	50046	12	35	41	12	53	4.6	500.6	24.5
3	2023	51707	12	39	39	10	49	1.2	499.1	24.3
3	2022	53433	13	39	40	8	48	9.8	497.5	24.7
3	2021	45242	20	42	32	6	38	-18.0	491.2	25.7
3	2019	56176	7	37	45	11	56	7.2	503.0	22.4
3	2018	43501	11	40	40	9	49	-3.2	499.1	23.2
3	2017	26659	11	37	44	8	52		499.2	22.7
4	2025	61946	15	40	36	8	44	-10.9	496.0	25.1
4	2024	50882	8	37	45	10	55	1.5	502.5	23.0
4	2023	52554	10	37	43	10	54	4.0	501.1	22.9
4	2022	53577	10	40	43	7	50	10.9	498.8	23.0
4	2021	45553	17	44	34	4	39	-17.8	491.7	23.9
4	2019	57629	6	37	47	10	57	7.7	503.0	21.2
4	2018	69779	11	40	42	7	49	-1.9	498.0	22.6
4	2017	64473	10	39	44	6	51		498.7	22.0
5	2025	63371	15	44	35	7	41	-6.5	494.9	23.2
5	2024	53399	7	45	41	7	48	-0.3	499.6	20.2
5	2023	54159	7	45	42	6	48	5.7	499.3	20.6
5	2022	55635	10	48	38	5	42	3.1	496.5	20.6
5	2021	46011	13	47	35	5	39	-15.5	493.5	21.3
5	2019	60444	5	40	48	6	55	8.0	501.7	19.8
5	2018	70083	9	45	42	5	47	-2.7	497.7	19.7
5	2017	29285	8	42	42	8	49		499.4	20.7
6	2025	64578	15	43	35	7	42	-4.9	495.7	24.1
6	2024	54954	10	43	39	9	47	-0.1	499.4	22.5
6	2023	56389	11	42	40	8	47	-0.3	498.1	21.9
6	2022	56939	9	43	42	6	48	8.9	498.2	21.2
6	2021	46699	16	45	34	5	39	-18.9	493.4	22.7
6	2019	61719	6	37	46	12	58	9.6	504.0	22.0
6	2018	54582	9	43	42	6	48	-4.2	498.4	21.9
6	2017	29704	9	39	46	6	52		499.7	21.2
7	2025	64726	19	41	31	9	40	-3.6	493.3	25.9
7	2024	56448	12	45	34	9	43	-0.1	497.0	24.2
7	2023	57234	16	41	35	9	44	1.7	496.1	24.5
7	2022	59311	13	45	34	8	42	2.7	495.5	23.5
7	2021	46839	13	48	32	7	39	-13.4	494.9	23.5
7	2019	62495	9	39	41	12	53	5.3	501.0	23.7
7	2018	66925	12	40	40	8	47	-0.9	497.7	23.4
7	2017	30144	9	43	40	8	48		498.9	22.0

Table 1.1.2 (continued)
Percentage of Students by Achievement Levels Categories
Mathematics

Grade	Year	Count	NM	PM	ME	EE	ME+EE	Delta	Ave. SS	SD SS
8	2025	65585	20	41	31	8	39	-4.9	493.6	25.4
8	2024	57441	13	43	35	9	44	1.3	497.4	23.6
8	2023	59572	14	43	34	9	43	2.3	496.4	23.8
8	2022	62311	12	48	32	8	40	4.4	495.8	22.7
8	2021	47150	16	48	32	4	36	-15.1	492.0	22.9
8	2019	62817	8	41	40	11	51	0.4	501.5	22.5
8	2018	70044	11	39	42	8	51	1.1	498.9	22.7
8	2017	66077	9	42	40	9	49		500.3	21.6

Section 1.2

Raw Scores Associated with Cutpoints

Table 1.2.1
Raw Scores Associated with Cutpoints

Subject	Grade	Cut Point	2024 Actual	2025 Actual	2025 Pred
English Language Arts	3	NM–PM	14	13	12
English Language Arts	3	PM–ME	27	26	25
English Language Arts	3	ME–EE	37	37	36
English Language Arts	4	NM–PM	15	16	14
English Language Arts	4	PM–ME	29	30	29
English Language Arts	4	ME–EE	39	40	39
English Language Arts	5	NM–PM	16	15	15
English Language Arts	5	PM–ME	34	31	30
English Language Arts	5	ME–EE	44	42	41
English Language Arts	6	NM–PM	19	19	17
English Language Arts	6	PM–ME	32	33	30
English Language Arts	6	ME–EE	42	43	41
English Language Arts	7	NM–PM	19	19	17
English Language Arts	7	PM–ME	34	33	31
English Language Arts	7	ME–EE	44	43	42
English Language Arts	8	NM–PM	20	20	20
English Language Arts	8	PM–ME	33	34	33
English Language Arts	8	ME–EE	43	44	43
Mathematics	3	NM–PM	14	14	14
Mathematics	3	PM–ME	29	30	31
Mathematics	3	ME–EE	42	43	44
Mathematics	4	NM–PM	16	15	15
Mathematics	4	PM–ME	34	34	34
Mathematics	4	ME–EE	49	49	49
Mathematics	5	NM–PM	13	14	13
Mathematics	5	PM–ME	32	32	32
Mathematics	5	ME–EE	49	49	49
Mathematics	6	NM–PM	12	13	13
Mathematics	6	PM–ME	30	31	30
Mathematics	6	ME–EE	48	49	48
Mathematics	7	NM–PM	11	12	12
Mathematics	7	PM–ME	28	26	26
Mathematics	7	ME–EE	47	46	46
Mathematics	8	NM–PM	14	14	13
Mathematics	8	PM–ME	32	32	31
Mathematics	8	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 beta(5,17), and 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:

Table 1.3.1.a
Number of Cycles to Convergence

Subject	Grade	Initial Cycles
English Language Arts	Grade 3	34
English Language Arts	Grade 4	45
English Language Arts	Grade 5	49
English Language Arts	Grade 6	31
English Language Arts	Grade 7	53
English Language Arts	Grade 8	43
Mathematics	Grade 3	38
Mathematics	Grade 4	63
Mathematics	Grade 5	58
Mathematics	Grade 6	54
Mathematics	Grade 7	105
Mathematics	Grade 8	50

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 are listed in the following table:

Table 1.3.1.b
Stocking and Lord Constants

Subject	Grade	Slope	Intercept	Num Eq Items	Num Eq Items Rem
English Language Arts	3	1.15	-0.26	18	1
English Language Arts	4	1.16	-0.33	18	0
English Language Arts	5	1.17	-0.27	16	0
English Language Arts	6	1.50	-0.35	16	0
English Language Arts	7	1.34	-0.38	16	0
English Language Arts	8	1.42	-0.26	16	0
Mathematics	3	1.22	-0.16	20	0
Mathematics	4	1.15	-0.15	20	1
Mathematics	5	1.14	-0.24	20	0
Mathematics	6	1.16	-0.22	20	0
Mathematics	7	1.21	-0.28	20	1
Mathematics	8	1.14	-0.22	20	0

The third table shows the number of cycles to achieve convergence in the equating (FCIP) calibration runs for ELA Stage 2:

Table 1.3.1.c
Number of Cycles to Convergence

Subject	Grade	Initial Cycles	Equating Cycles
English Language Arts	Grade 3	119	7
English Language Arts	Grade 4	44	7
English Language Arts	Grade 5	26	11
English Language Arts	Grade 6	35	11
English Language Arts	Grade 7	52	10
English Language Arts	Grade 8	48	10

Three methods were used to evaluate the suitability of the equating items: delta analysis, b/b analysis, and beta analysis. The b/b analysis and beta analysis were calculated for descriptive purposes. As a result of the beta analysis, flagged items were reviewed by content personnel, and one item each from grade 3 ELA, grade 4 math, and grade 7 math was removed from the equating analysis. Results from these analyses are included in Section II of this report.

Items flagged by the beta analysis 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	IA00280 (EL308824)	beta analysis	retained for equating
English Language Arts	3	IA00445 (EL626043435)	beta analysis	removed from equating
English Language Arts	4	IA00408 (EL624647580)	beta analysis	retained for equating
English Language Arts	6	IA00175 (EL303504)	beta analysis	retained for equating
Mathematics	4	IA01055 (MA311572)	beta analysis	retained for equating
Mathematics	4	IA01093 (MA623879088)	beta analysis	retained for equating
Mathematics	4	IA02819 (MA713583365)	beta analysis	removed from equating
Mathematics	4	IA04965 (MA800867144)	beta analysis	retained for equating
Mathematics	5	IA02552 (MA311324)	beta analysis	retained for equating
Mathematics	5	IA05002 (MA801652356)	beta analysis	retained for equating
Mathematics	7	IA10268 (MA005205470)	beta analysis	removed from equating
Mathematics	8	IA04665 (MA307399)	beta analysis	retained for equating
Mathematics	8	IA04957 (MA800770988)	beta analysis	retained for equating

Section 1.4

Equating Item Summary Statistics

Table 1.4.1

Equating Item Summary Statistics

(2025 a and b values are derived from unscaled parameters before equating)

Subject	Grade	Year	P-Value Mean	P-Value Std Dev	Point Biserial Mean	Point Biserial Std Dev	a Mean	a Std Dev	b Mean	b Std Dev
English Language Arts	03	2025	0.55	0.13	0.47	0.10	1.07	0.37	0.21	0.68
English Language Arts	03	Previous	0.57	0.12	0.47	0.09	0.94	0.28	-0.20	0.55
English Language Arts	04	2025	0.62	0.15	0.44	0.09	0.86	0.20	-0.27	0.74
English Language Arts	04	Previous	0.63	0.16	0.42	0.09	0.78	0.19	-0.66	0.85
English Language Arts	05	2025	0.66	0.14	0.41	0.07	0.90	0.32	-0.43	0.62
English Language Arts	05	Previous	0.67	0.14	0.41	0.08	0.74	0.24	-0.83	0.74
English Language Arts	06	2025	0.67	0.10	0.42	0.09	0.92	0.25	-0.41	0.68
English Language Arts	06	Previous	0.68	0.11	0.41	0.09	0.60	0.16	-0.97	1.03
English Language Arts	07	2025	0.69	0.10	0.45	0.09	0.98	0.28	-0.54	0.60
English Language Arts	07	Previous	0.69	0.11	0.43	0.09	0.71	0.19	-1.12	0.83
English Language Arts	08	2025	0.63	0.13	0.42	0.08	0.89	0.30	-0.26	0.60
English Language Arts	08	Previous	0.63	0.13	0.41	0.07	0.59	0.18	-0.68	0.81
Mathematics	03	2025	0.60	0.14	0.50	0.14	1.06	0.32	-0.19	0.65
Mathematics	03	Previous	0.63	0.15	0.48	0.12	0.87	0.24	-0.35	0.79
Mathematics	04	2025	0.58	0.12	0.52	0.14	1.11	0.36	-0.08	0.55
Mathematics	04	Previous	0.64	0.12	0.50	0.13	1.01	0.28	-0.18	0.55
Mathematics	05	2025	0.55	0.18	0.51	0.12	1.12	0.32	-0.08	0.78
Mathematics	05	Previous	0.59	0.19	0.49	0.11	1.00	0.26	-0.28	0.85
Mathematics	06	2025	0.52	0.14	0.52	0.11	1.12	0.35	0.08	0.64
Mathematics	06	Previous	0.56	0.16	0.50	0.10	0.94	0.29	-0.19	0.77
Mathematics	07	2025	0.51	0.18	0.52	0.13	1.32	0.38	0.19	0.70
Mathematics	07	Previous	0.54	0.18	0.50	0.13	1.09	0.33	-0.04	0.86
Mathematics	08	2025	0.52	0.13	0.52	0.13	1.28	0.41	0.15	0.52
Mathematics	08	Previous	0.55	0.14	0.51	0.12	1.10	0.35	-0.08	0.62

Section 2.1

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

Note: "CR" in this section refers to all non-MC items in the equating set, different from typical use of CR in the MCAS context.

A/A Plot: English Language Arts Grade 3
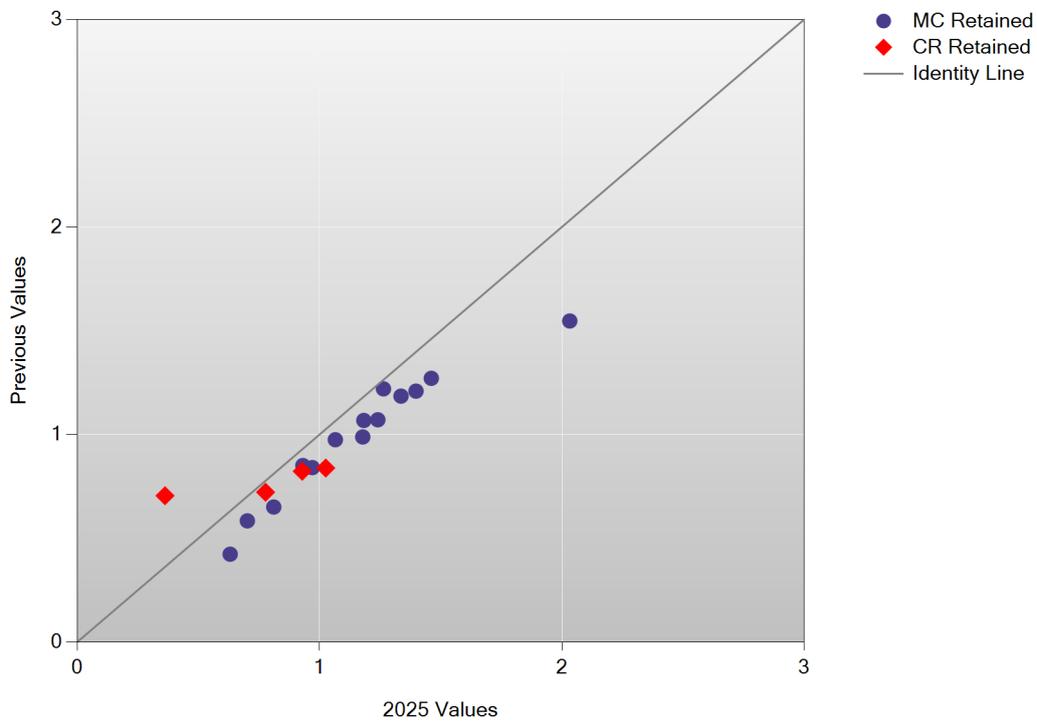
B/B Plot: English Language Arts Grade 3
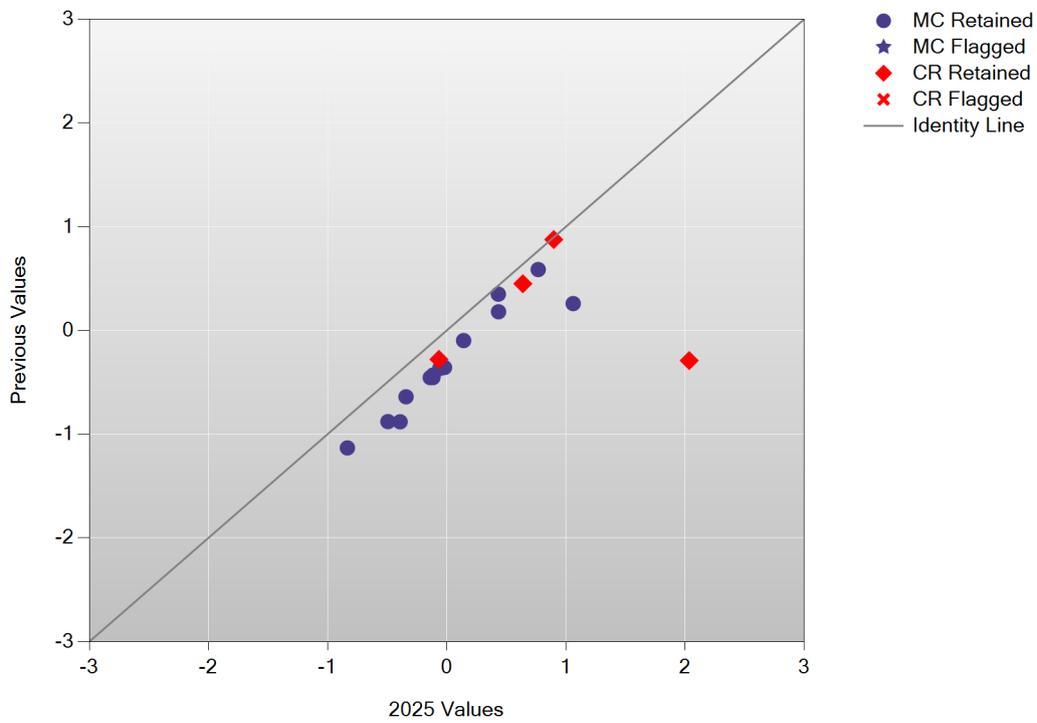
Delta Plot: English Language Arts Grade 3
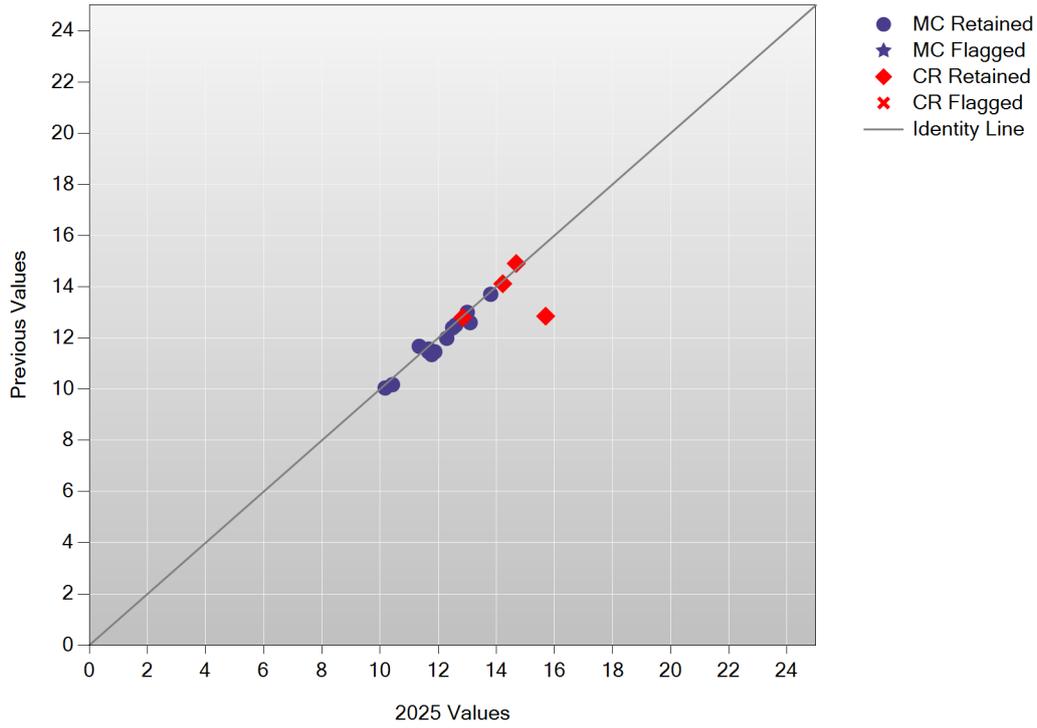
Test Characteristic Curve:English Language Arts Grade 3
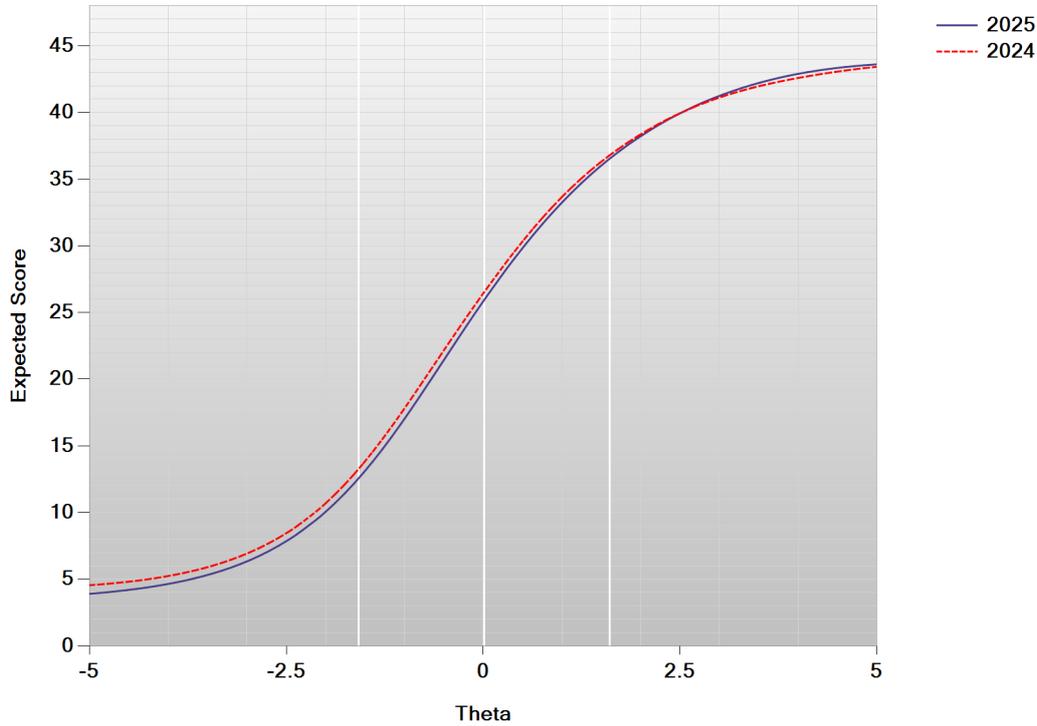
Test Information Function: English Language Arts Grade 3
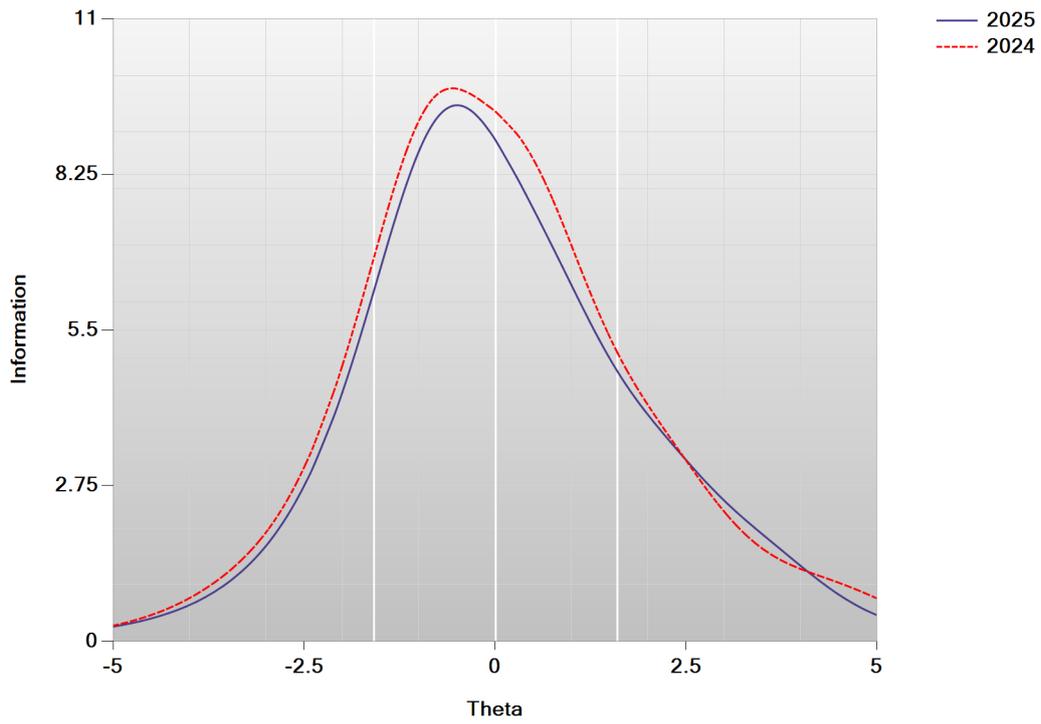
Cumulative Scale Score Distributions: English Language Arts Grade 3
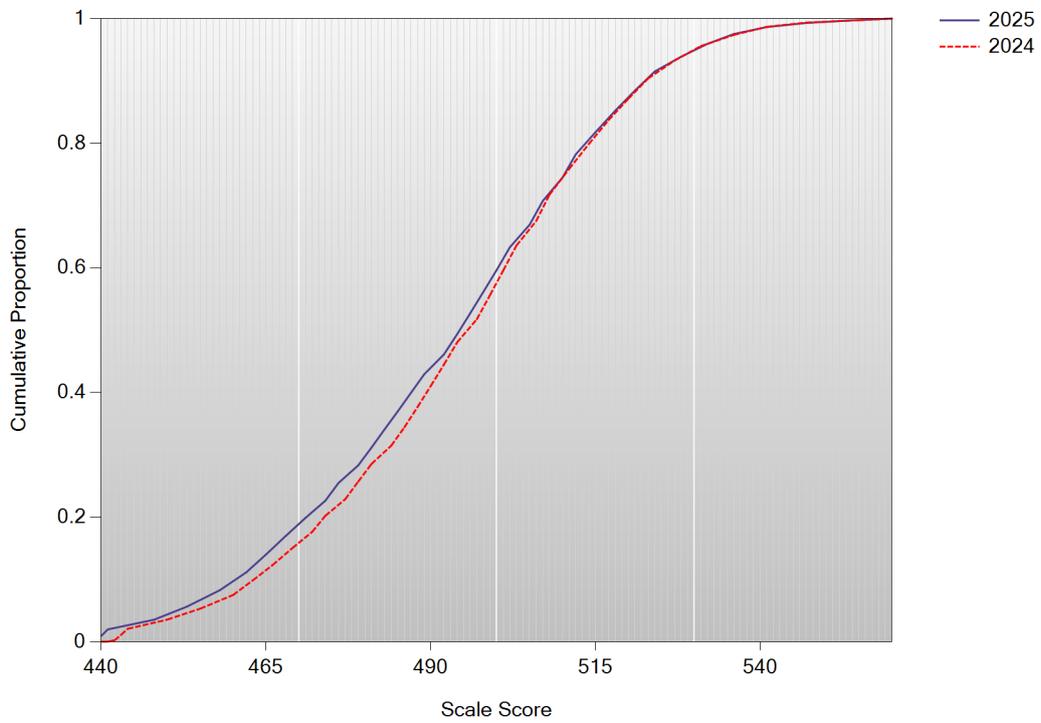
A/A Plot: English Language Arts Grade 4
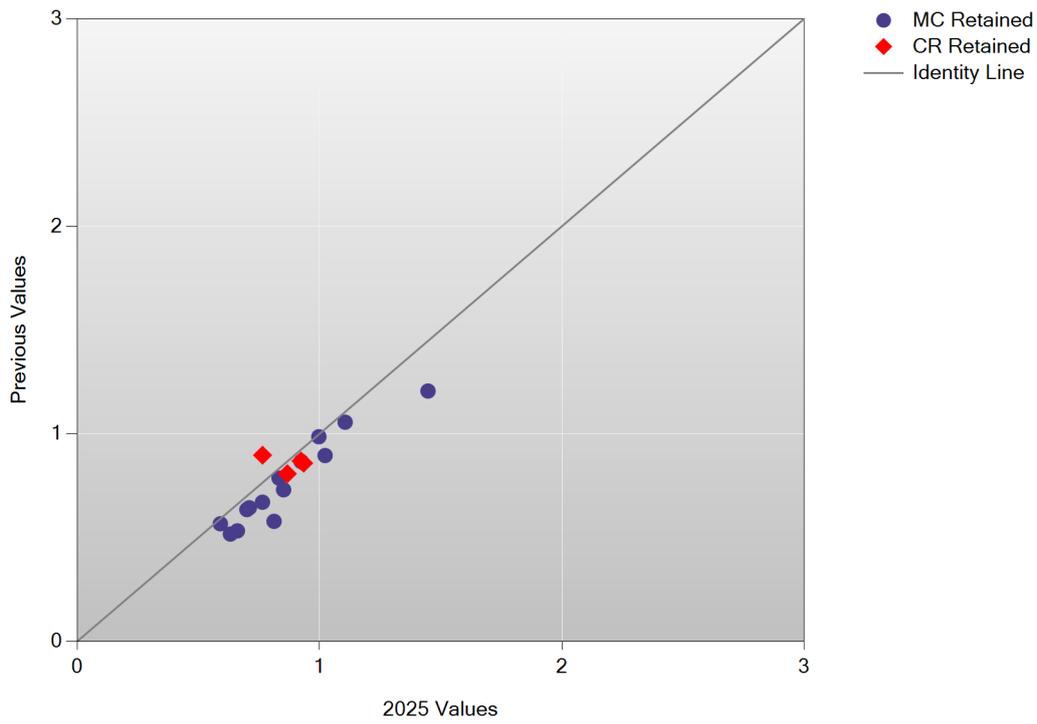
B/B Plot: English Language Arts Grade 4
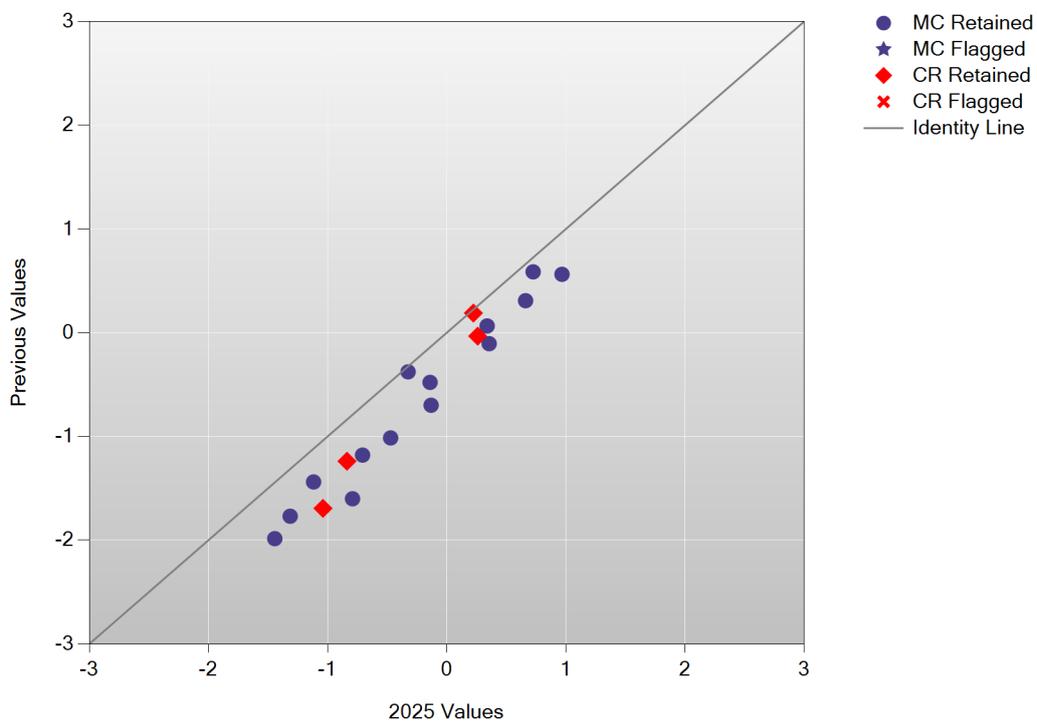
Delta Plot: English Language Arts Grade 4
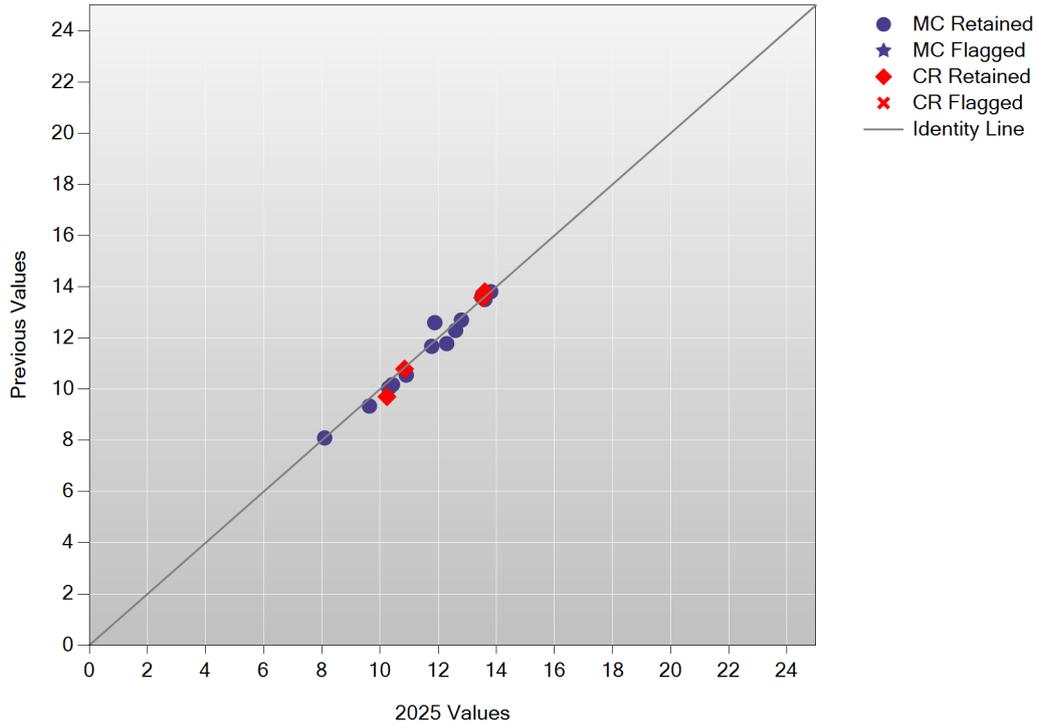
Test Characteristic Curve:English Language Arts Grade 4
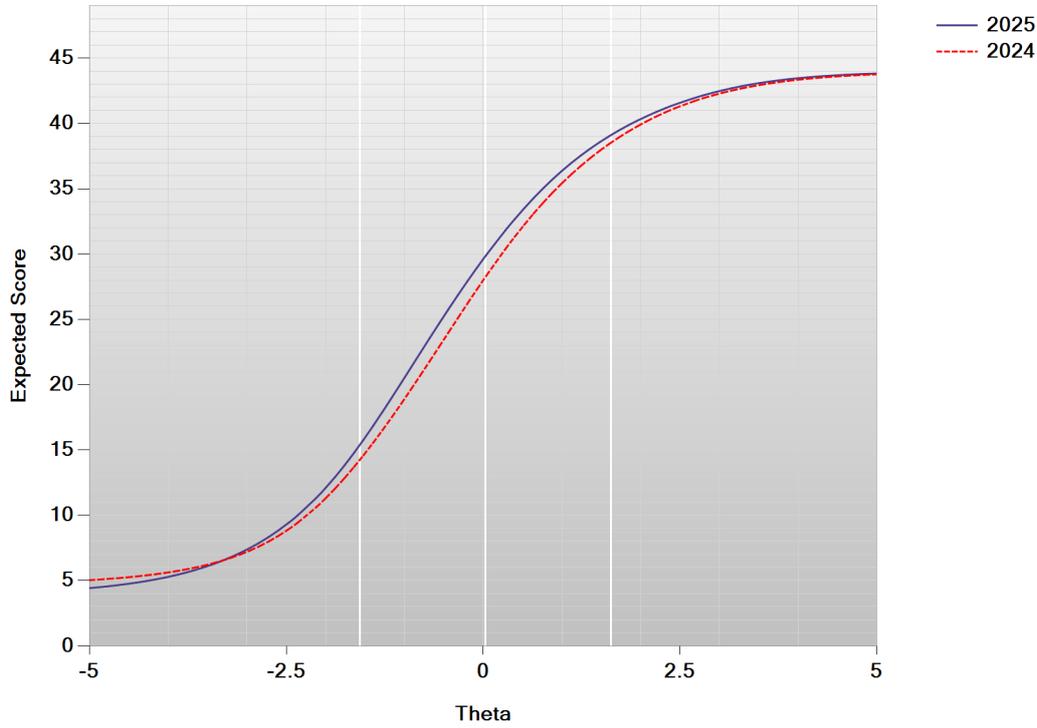
Test Information Function: English Language Arts Grade 4
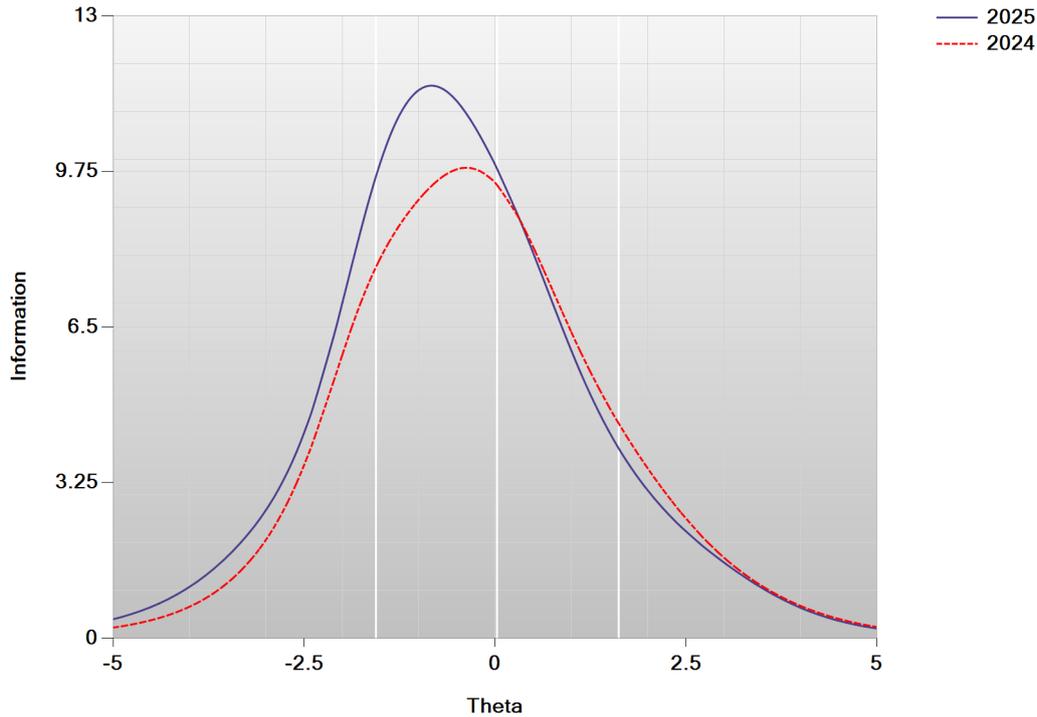
Cumulative Scale Score Distributions: English Language Arts Grade 4
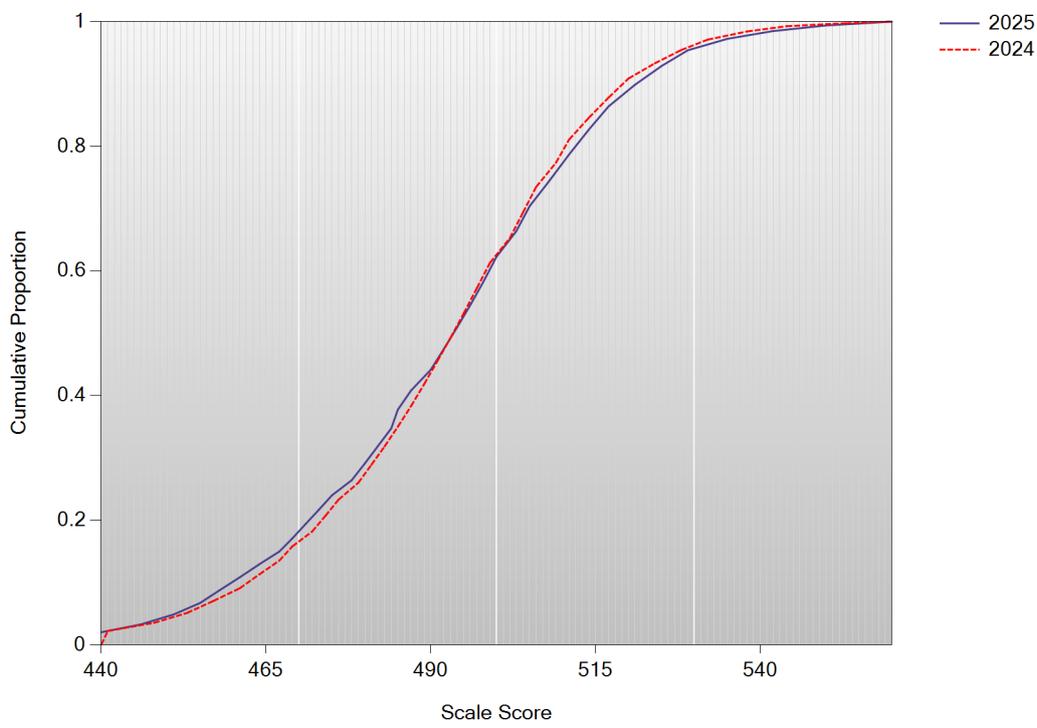
A/A Plot: English Language Arts Grade 5
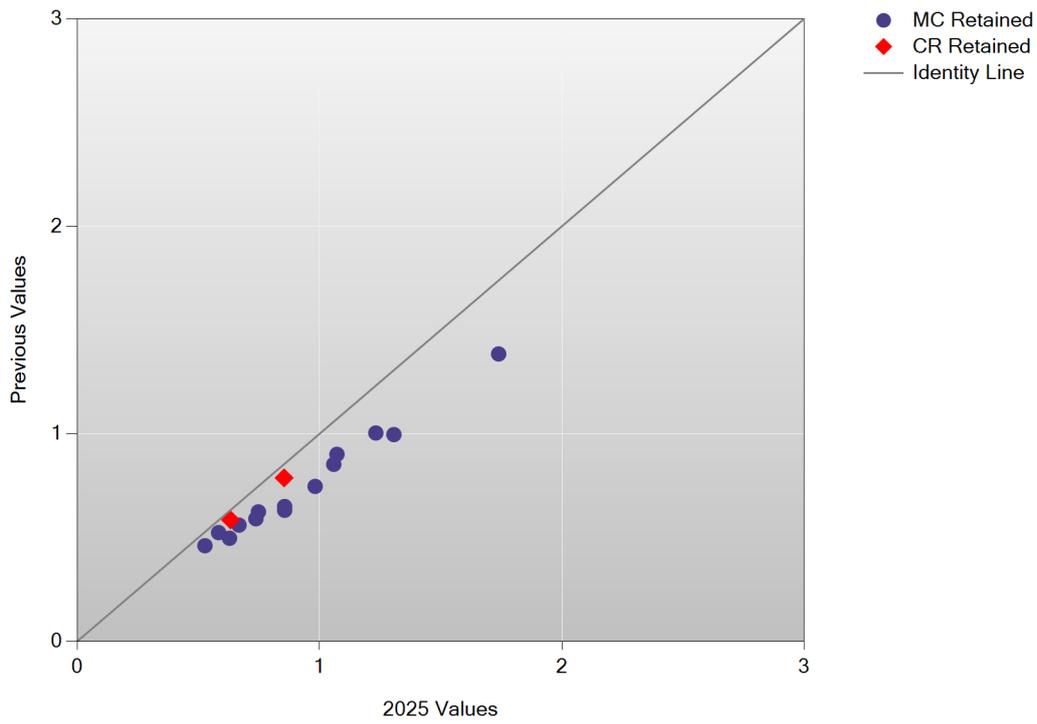
B/B Plot: English Language Arts Grade 5
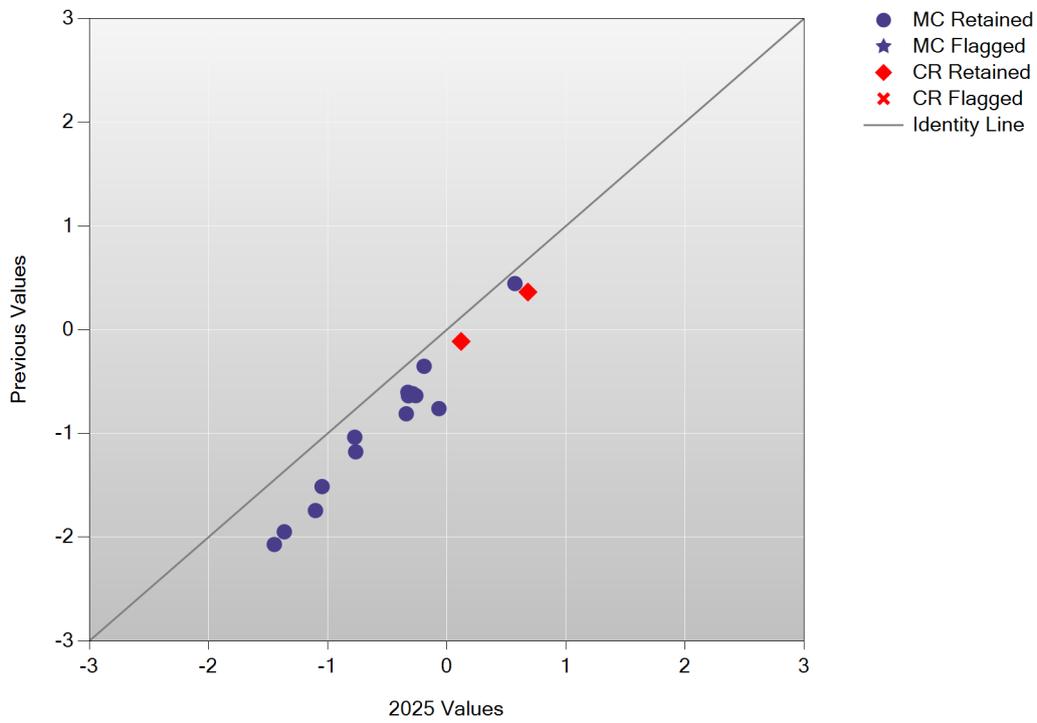
Delta Plot: English Language Arts Grade 5
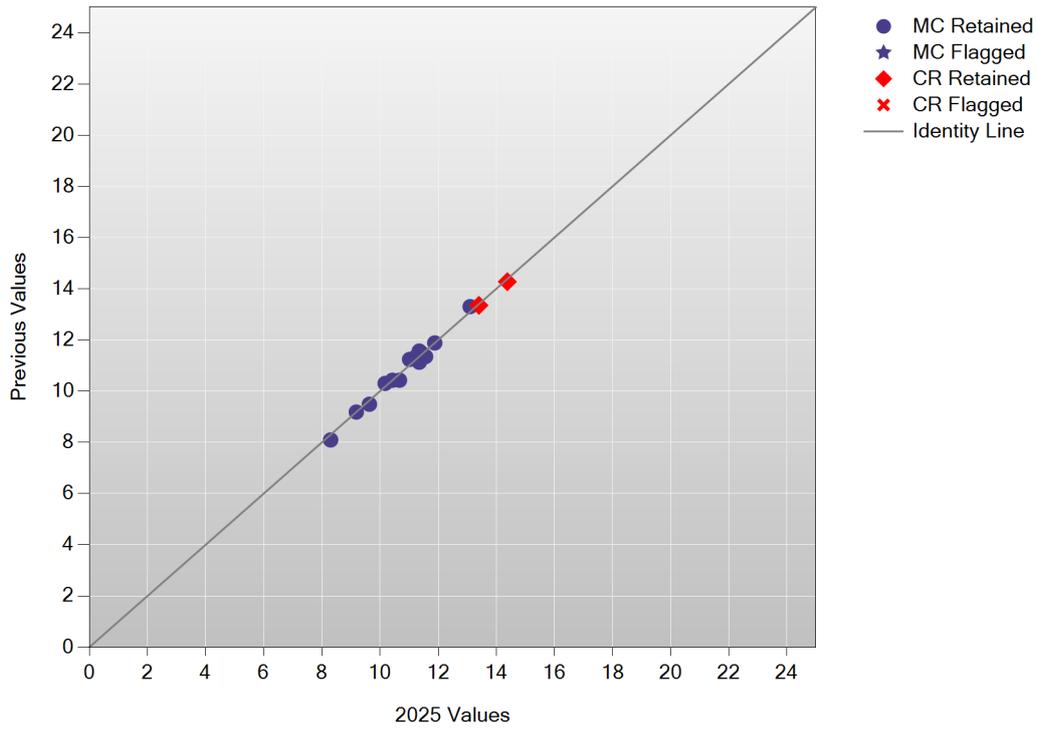
Test Characteristic Curve:English Language Arts Grade 5
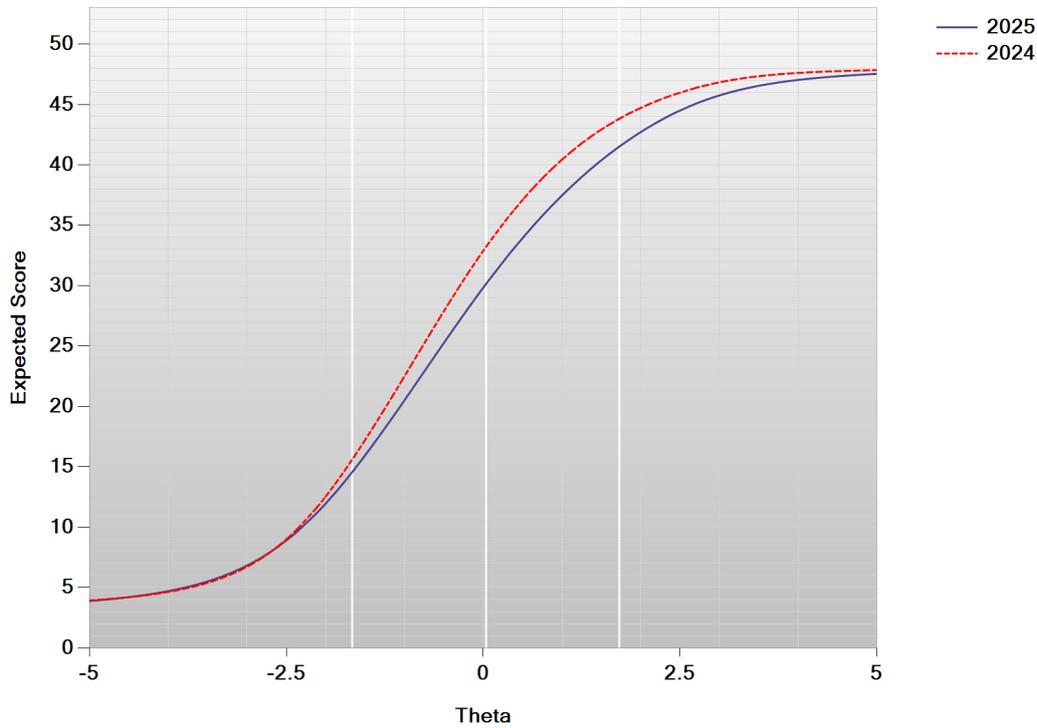
Test Information Function: English Language Arts Grade 5
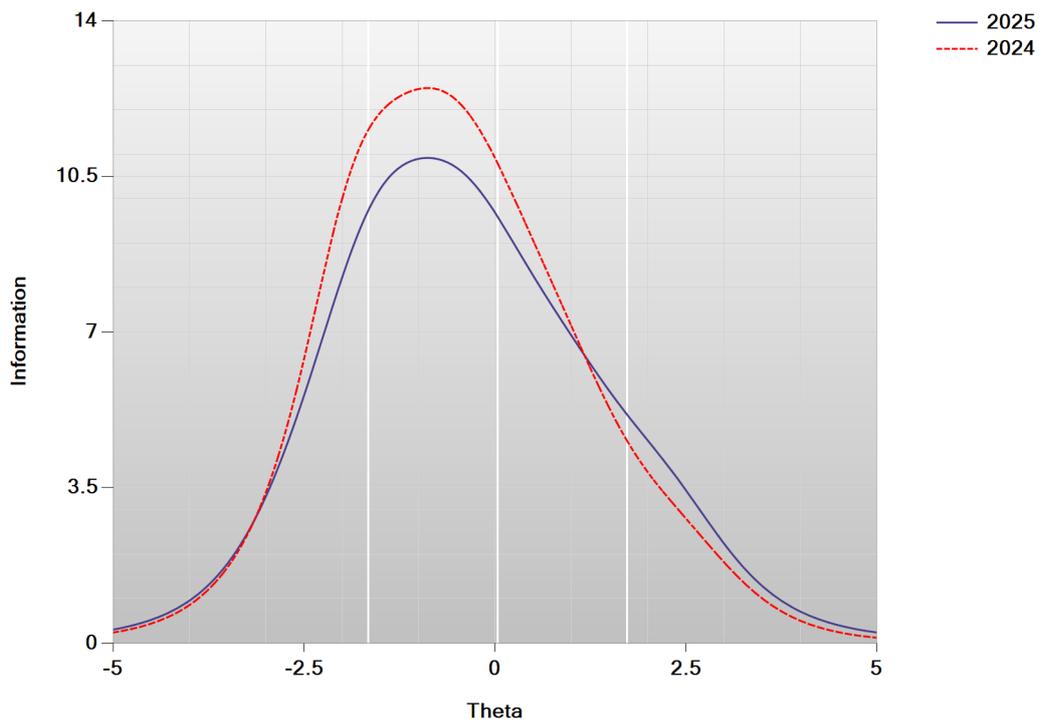
Cumulative Scale Score Distributions: English Language Arts Grade 5
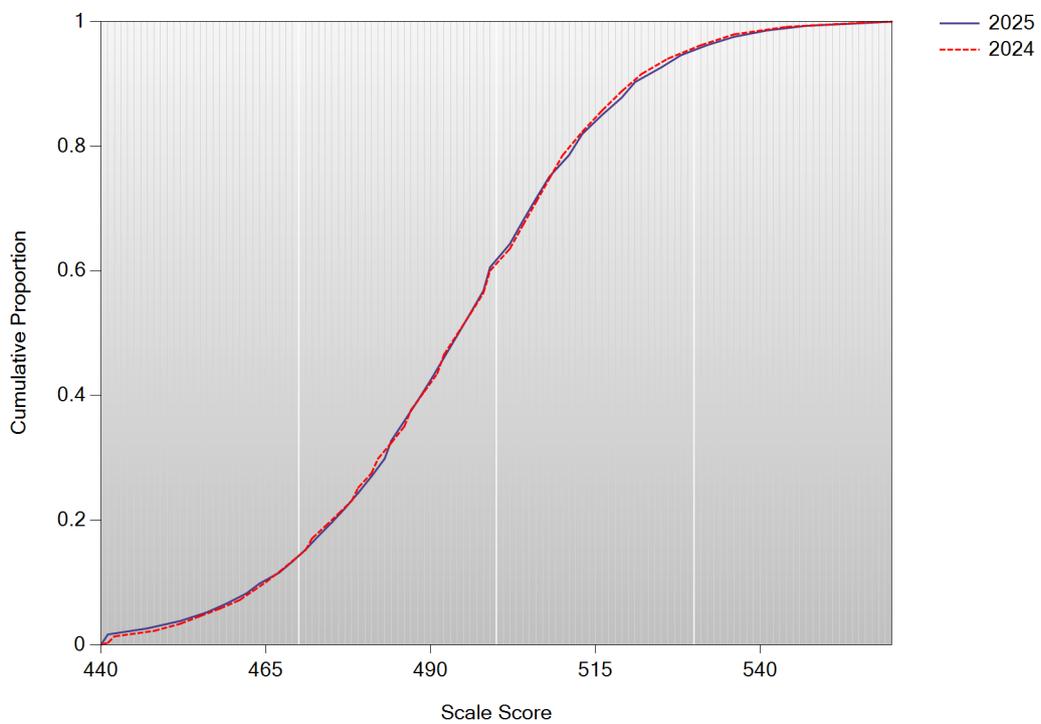
A/A Plot: English Language Arts Grade 6
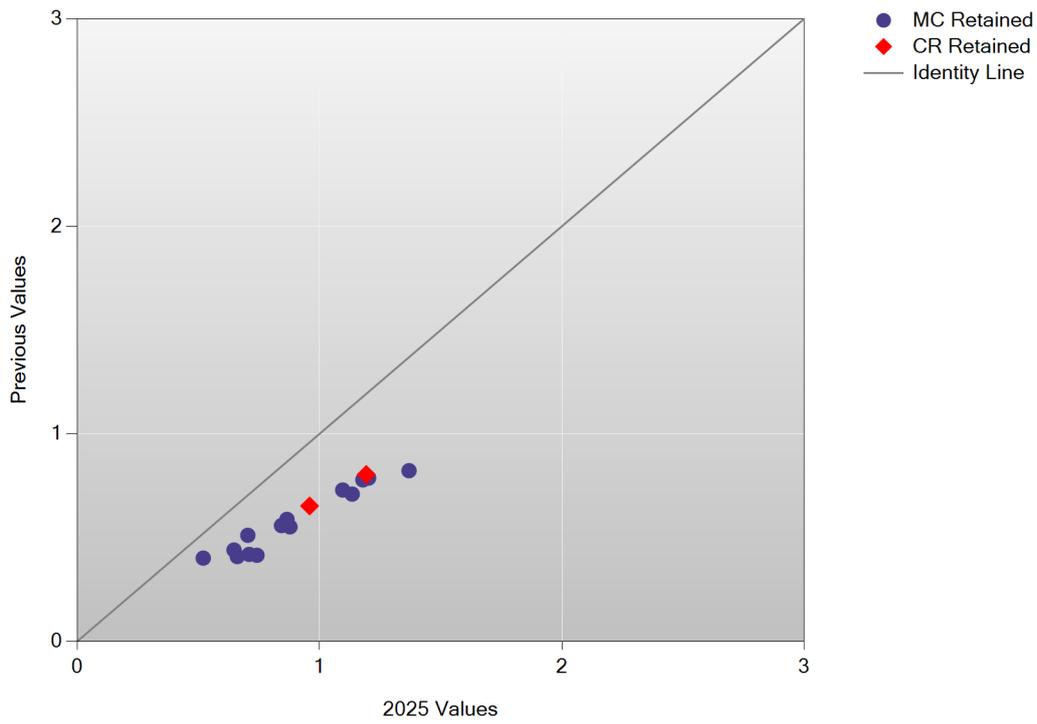
B/B Plot: English Language Arts Grade 6
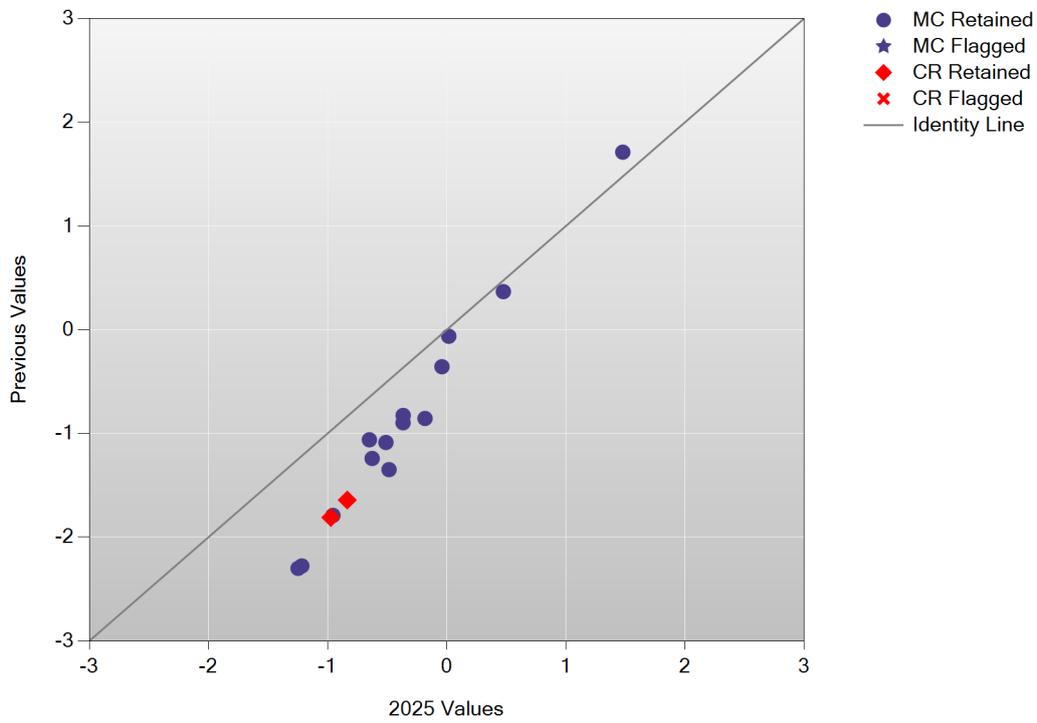
Delta Plot: English Language Arts Grade 6
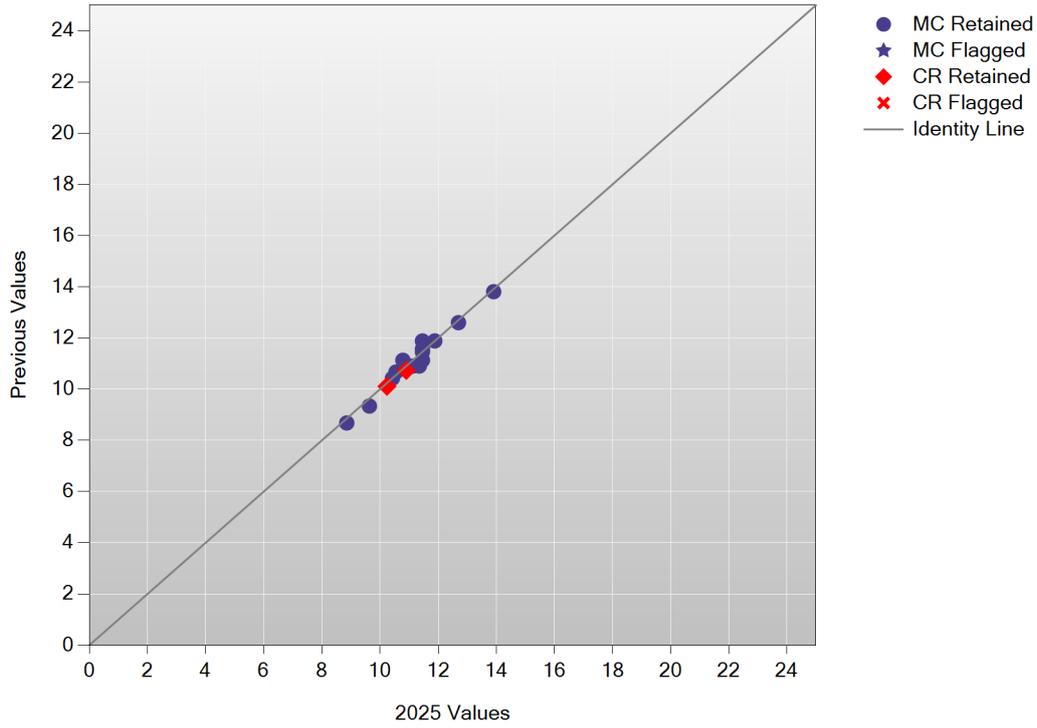
Test Characteristic Curve:English Language Arts Grade 6
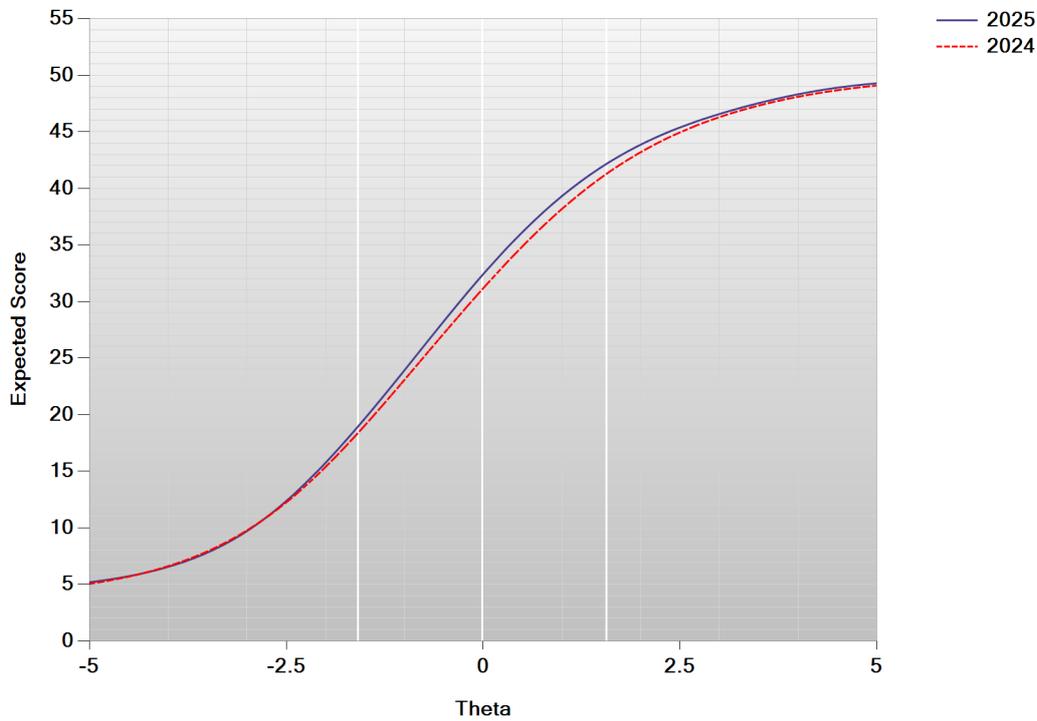
Test Information Function: English Language Arts Grade 6
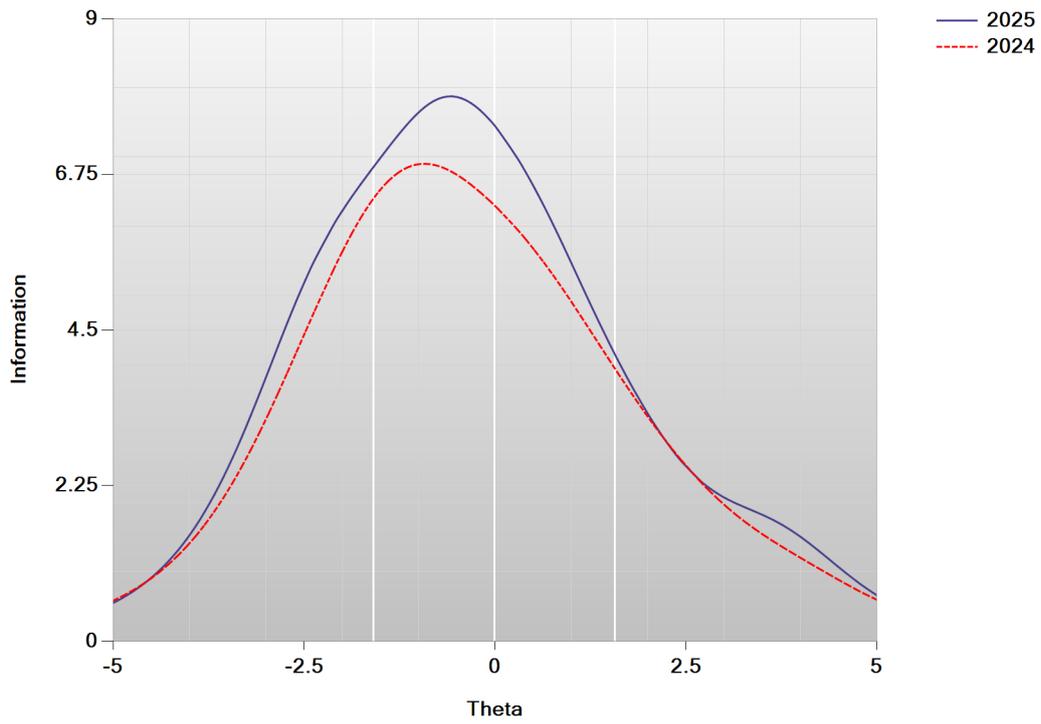
Cumulative Scale Score Distributions: English Language Arts Grade 6
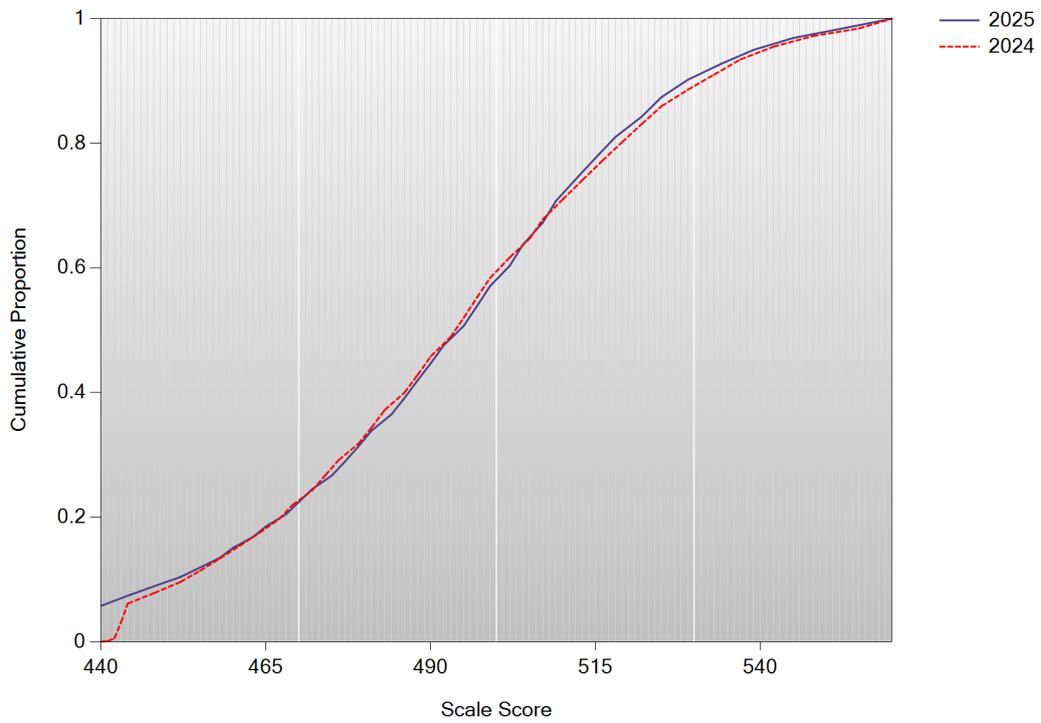
A/A Plot: English Language Arts Grade 7
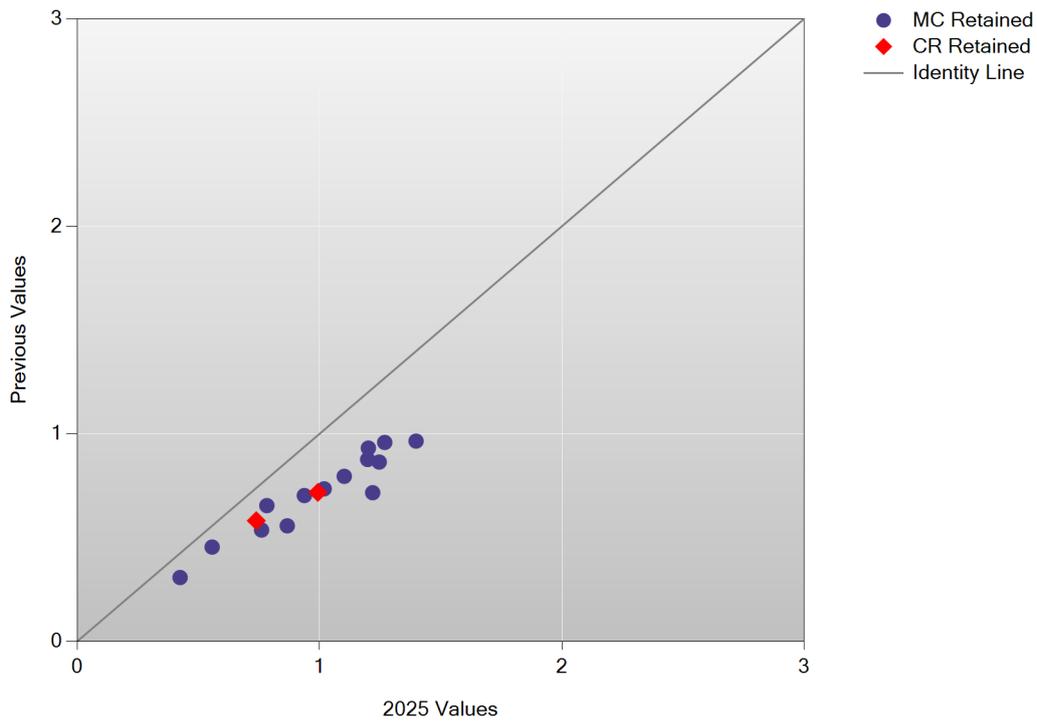
B/B Plot: English Language Arts Grade 7
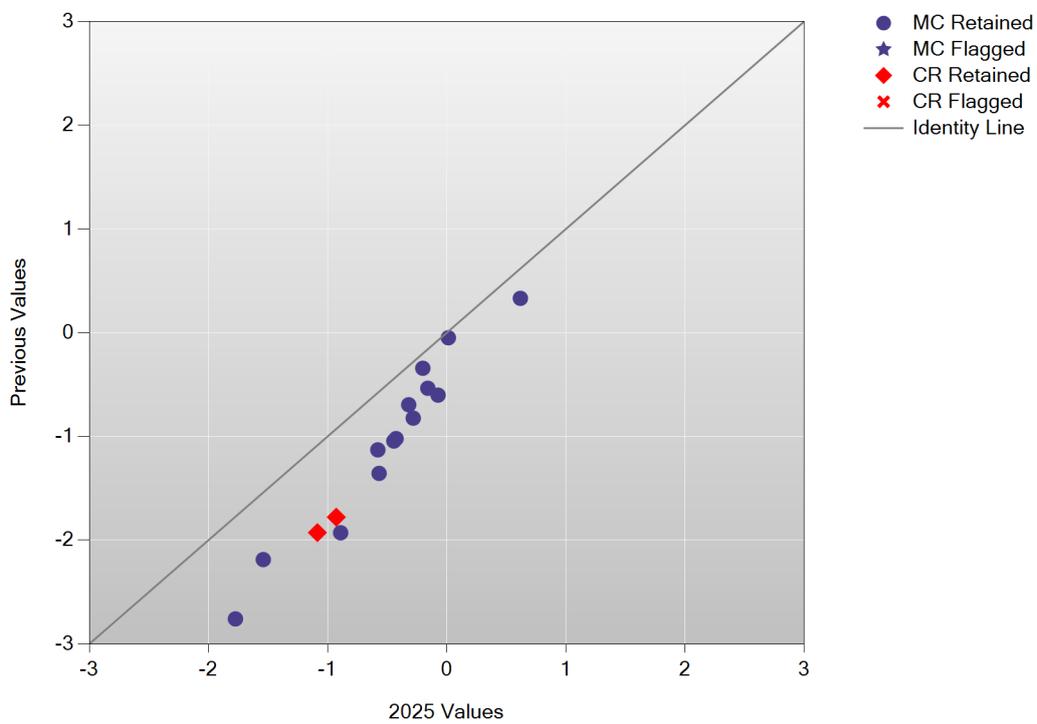
Delta Plot: English Language Arts Grade 7
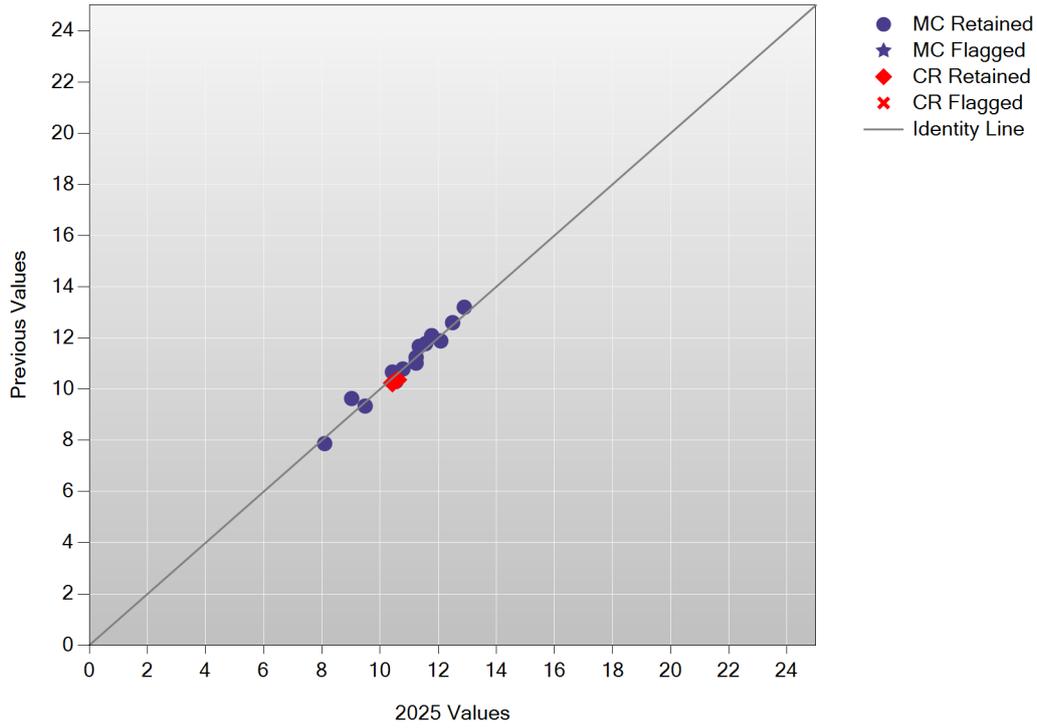
Test Characteristic Curve:English Language Arts Grade 7
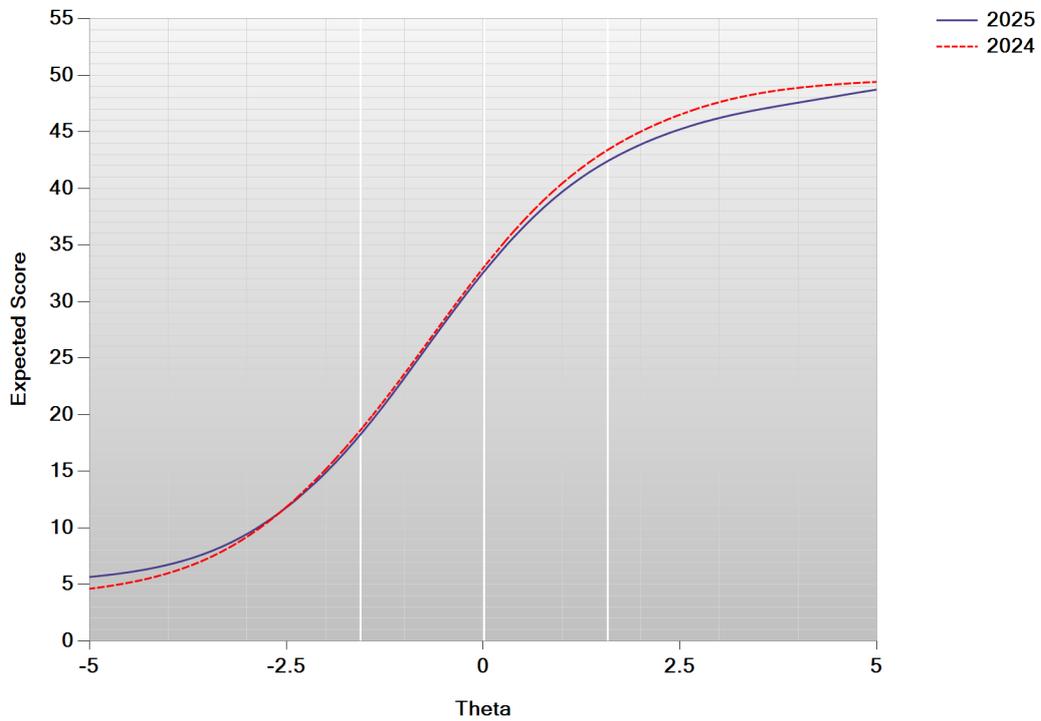
Test Information Function: English Language Arts Grade 7
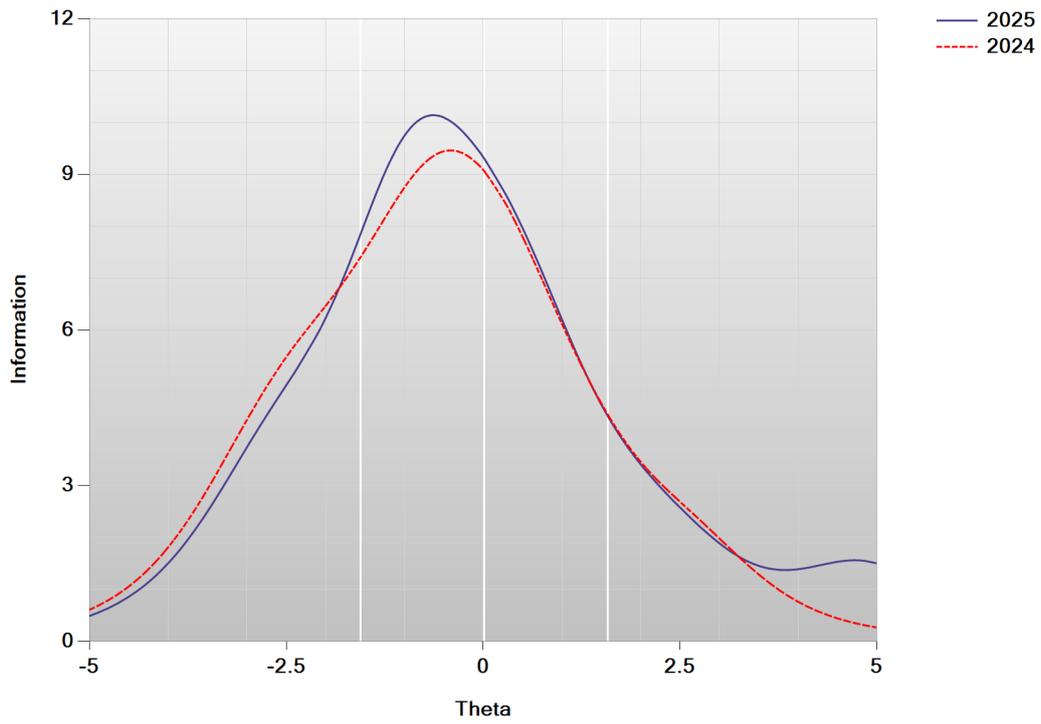
Cumulative Scale Score Distributions: English Language Arts Grade 7
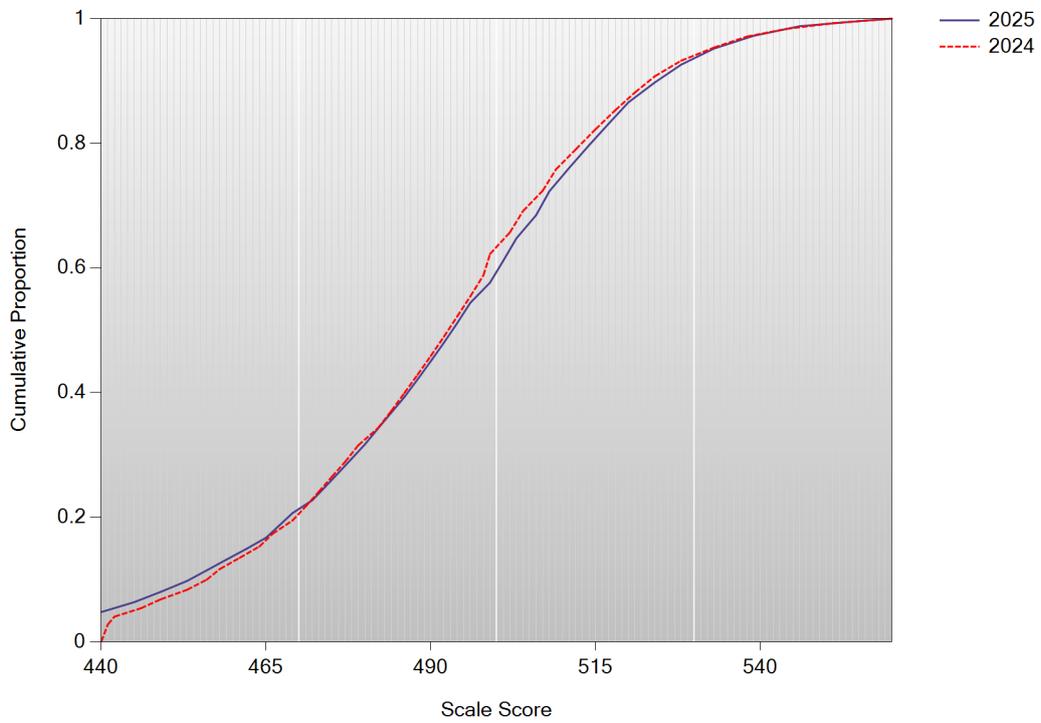
A/A Plot: English Language Arts Grade 8
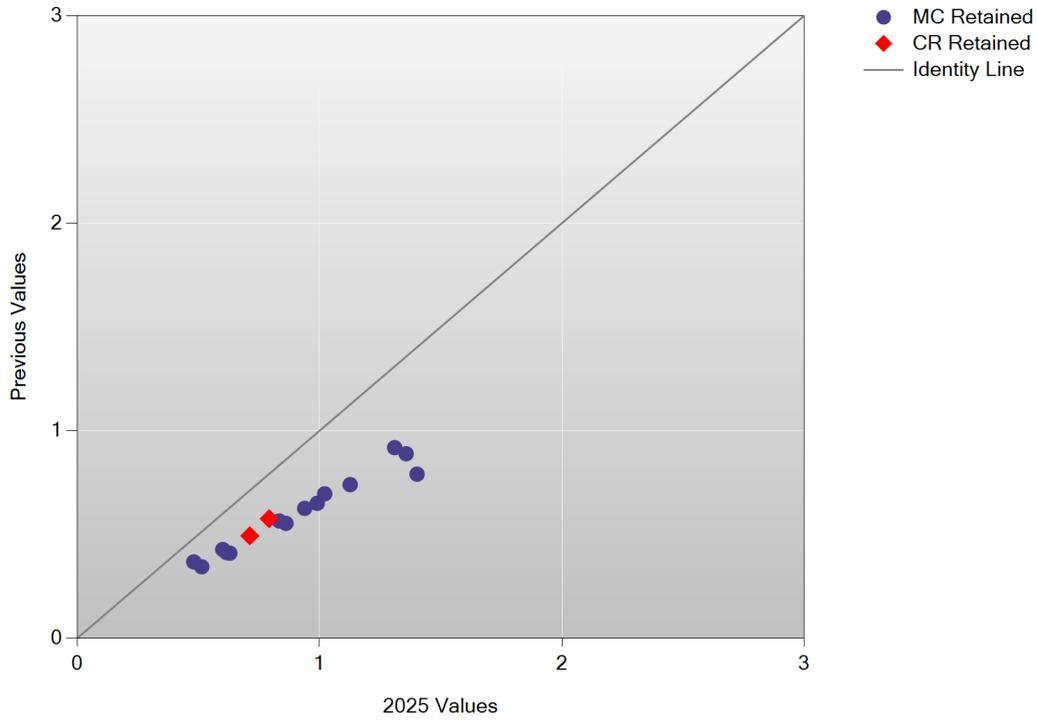
B/B Plot: English Language Arts Grade 8
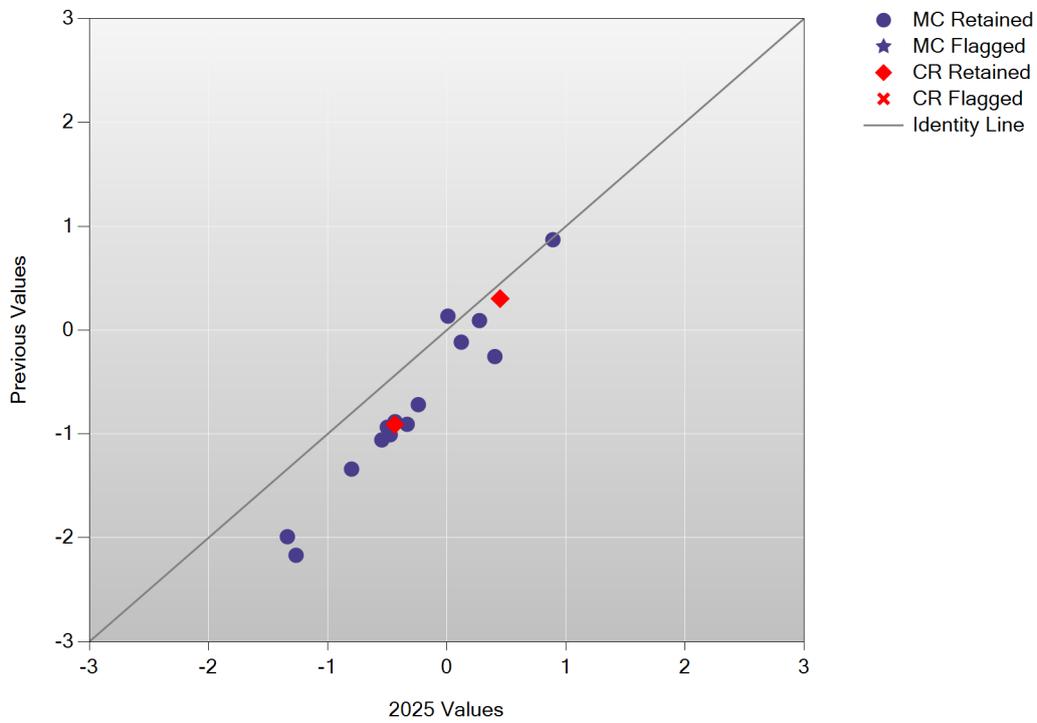
Delta Plot: English Language Arts Grade 8
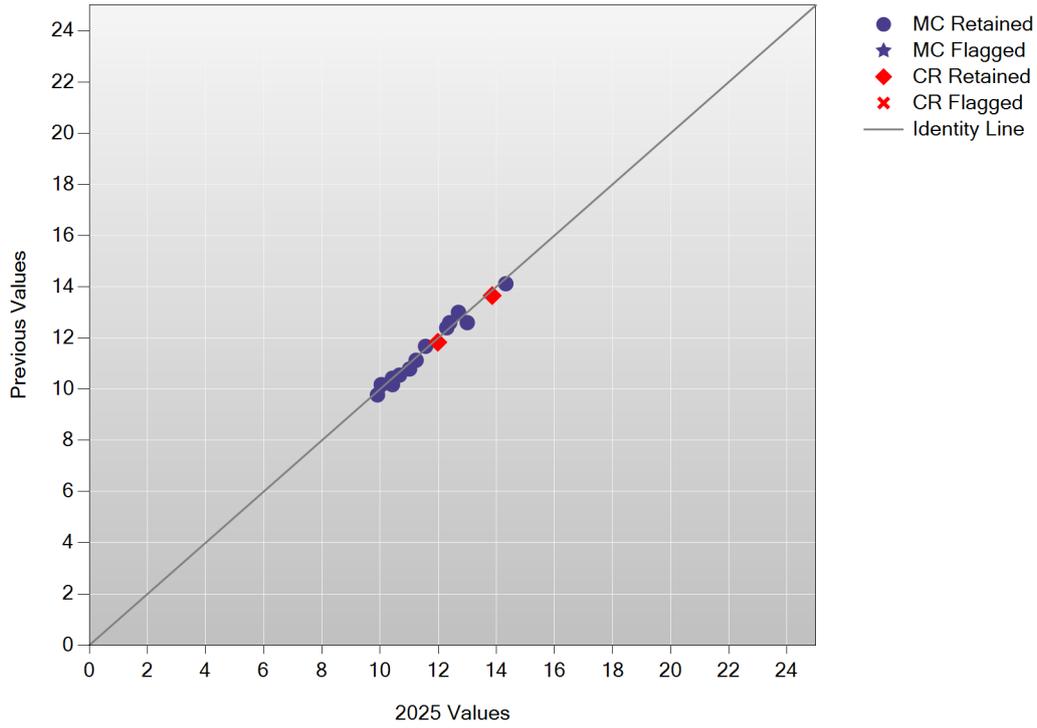
Test Characteristic Curve:English Language Arts Grade 8
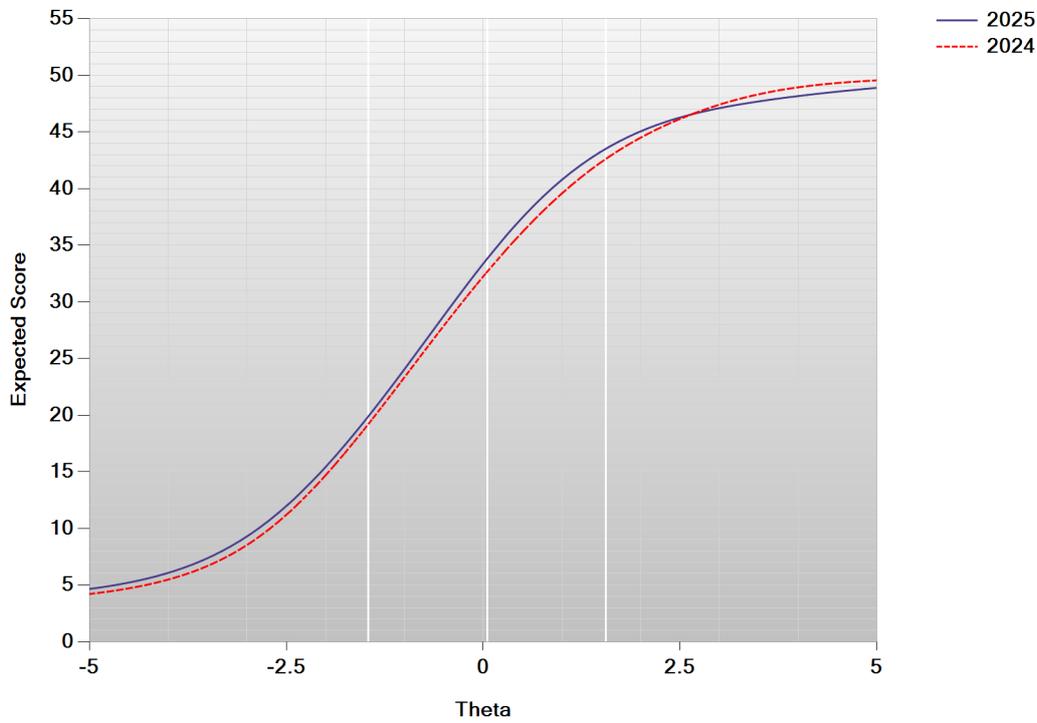
Test Information Function: English Language Arts Grade 8
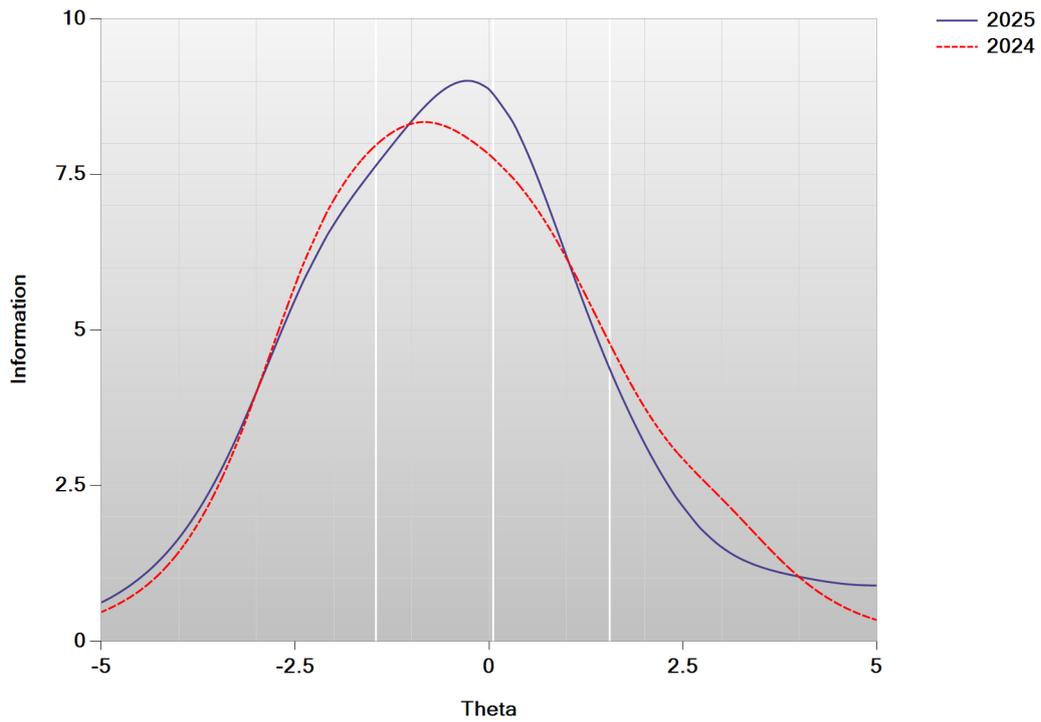
Cumulative Scale Score Distributions: English Language Arts Grade 8
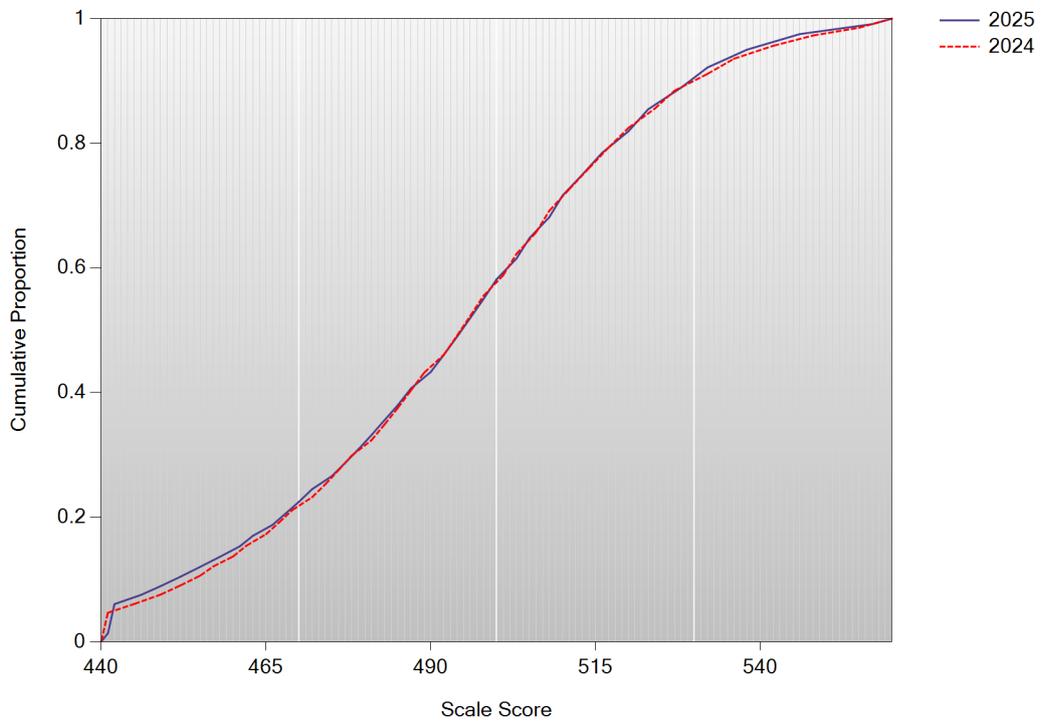
A/A Plot: Mathematics Grade 3

B/B Plot: Mathematics Grade 3

Delta Plot: Mathematics Grade 3
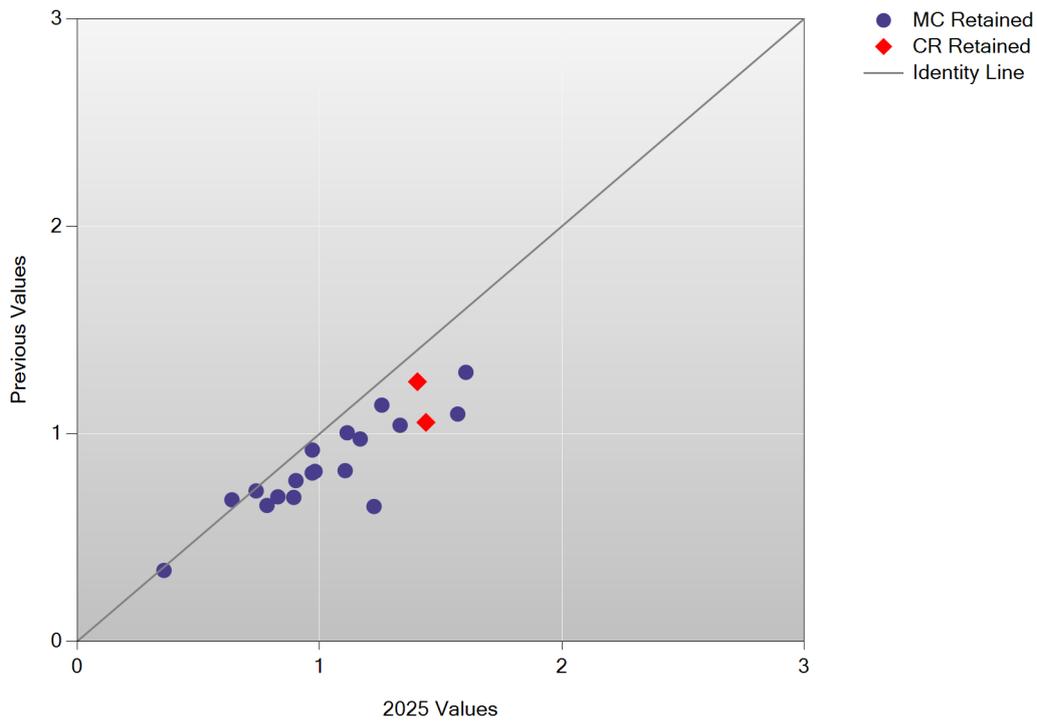
Test Characteristic Curve:Mathematics Grade 3
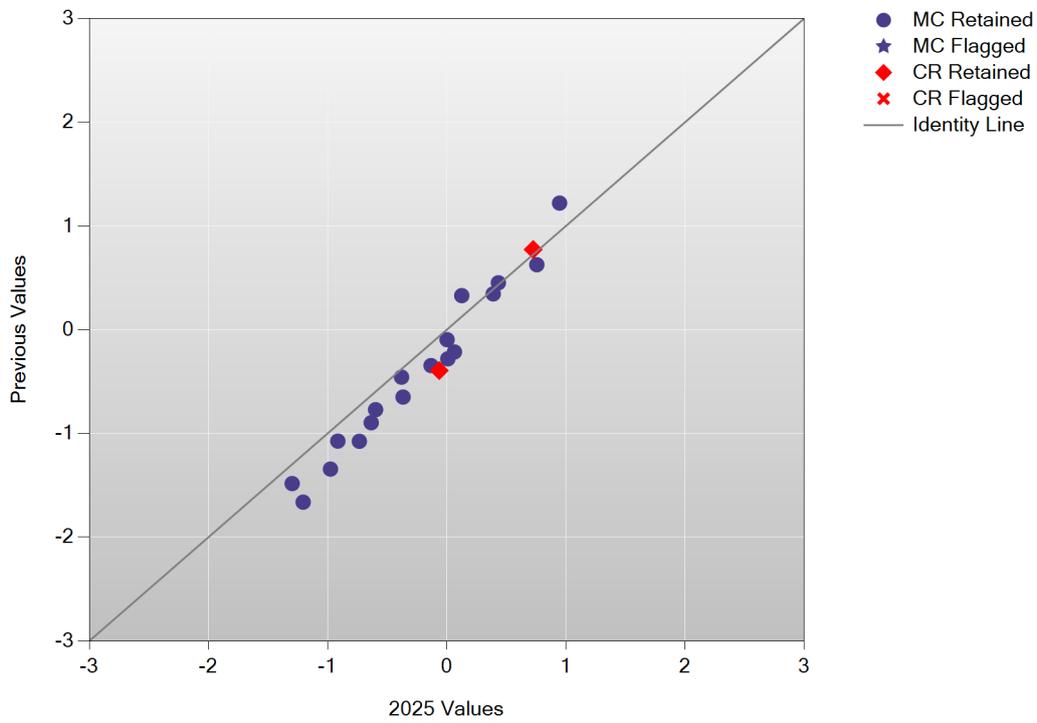
Test Information Function: Mathematics Grade 3
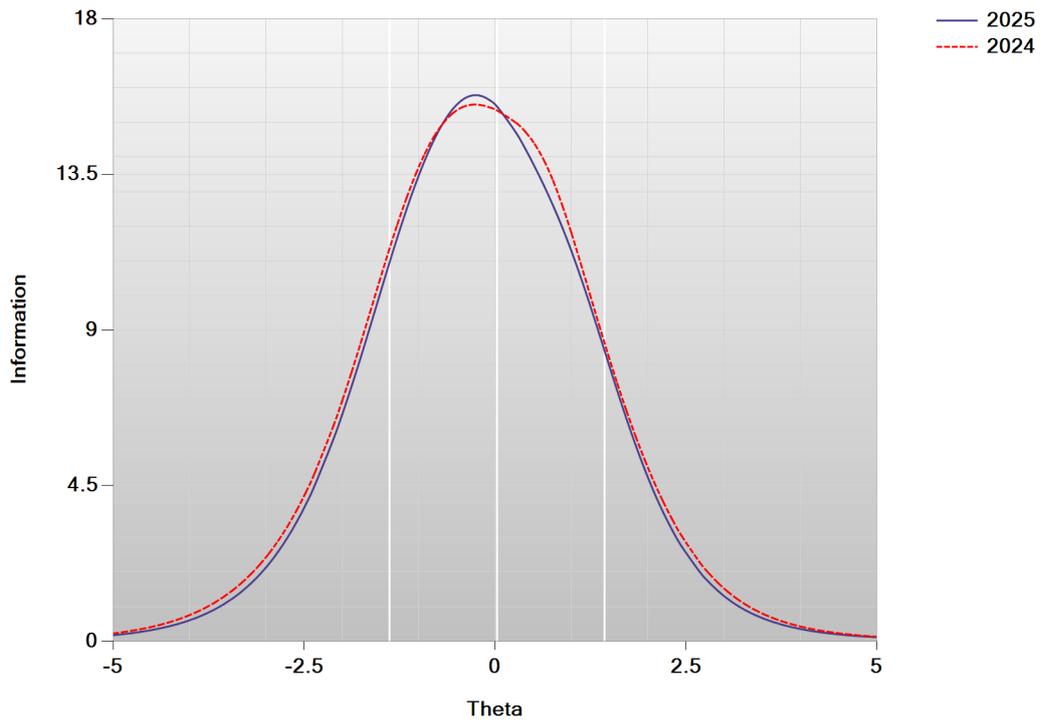
Cumulative Scale Score Distributions: Mathematics Grade 3
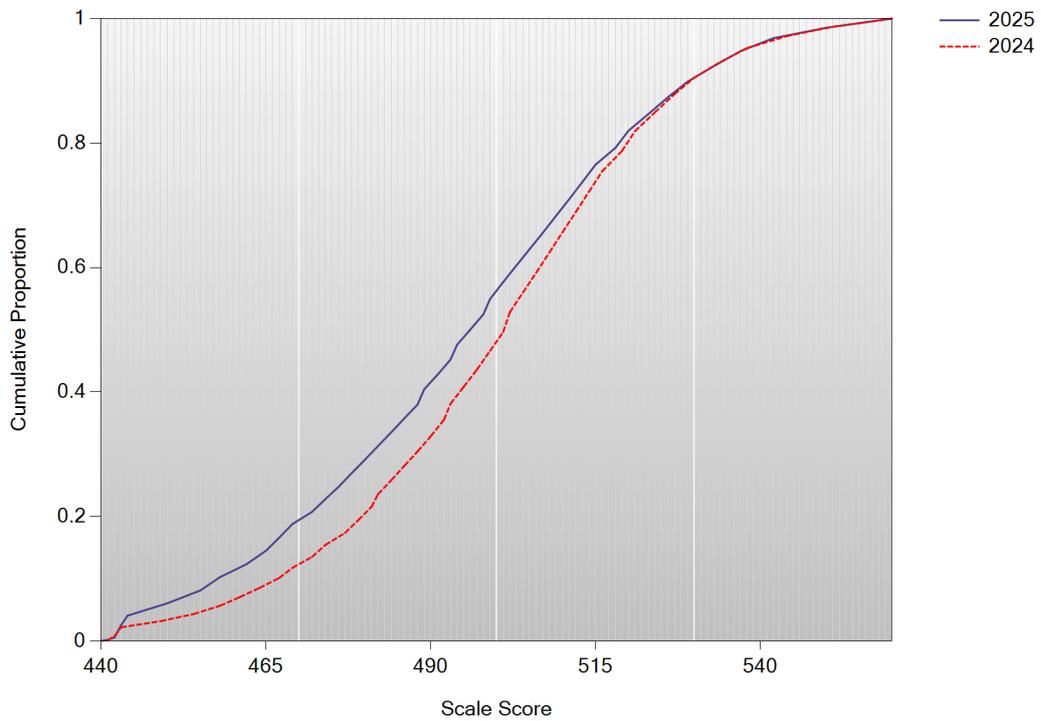
A/A Plot: Mathematics Grade 4
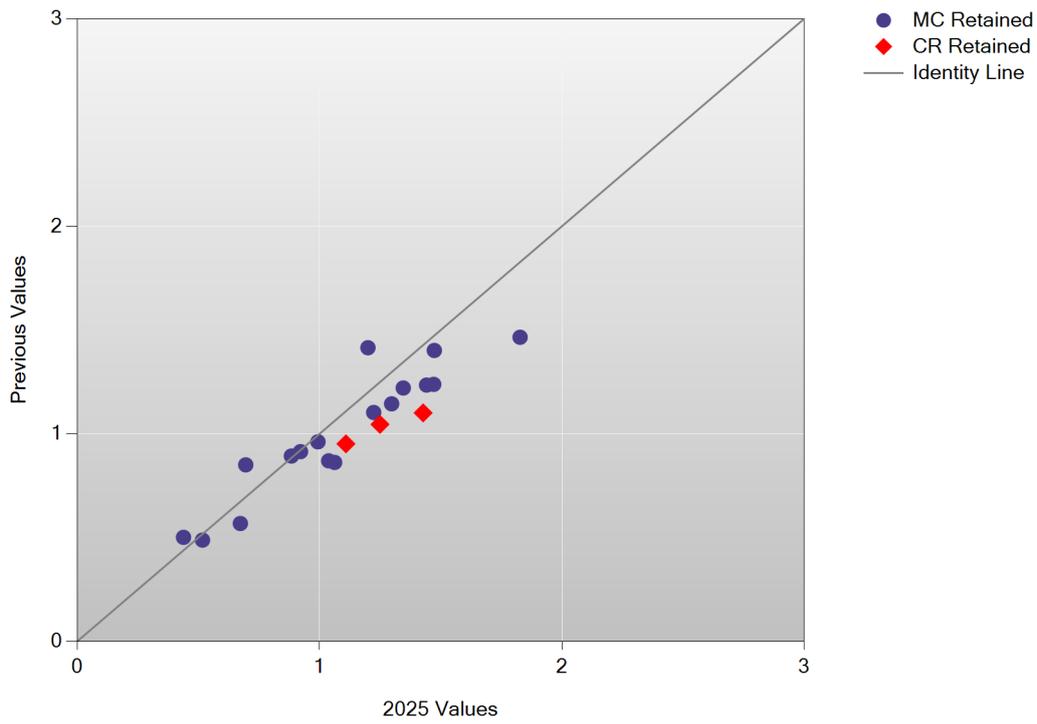
B/B Plot: Mathematics Grade 4
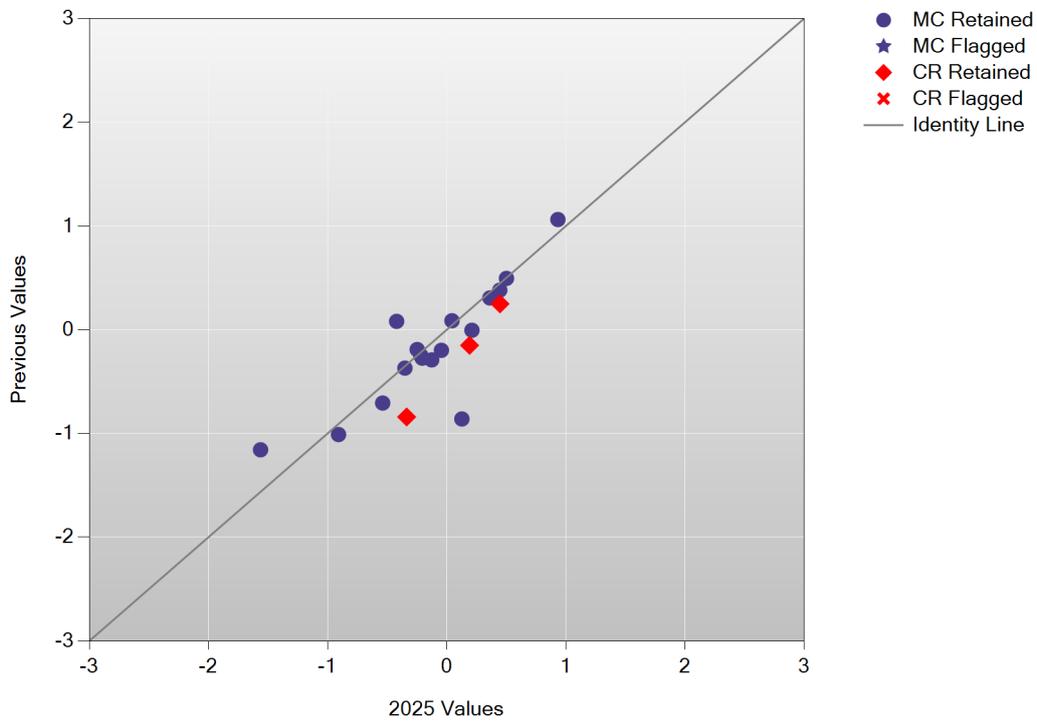
Delta Plot: Mathematics Grade 4
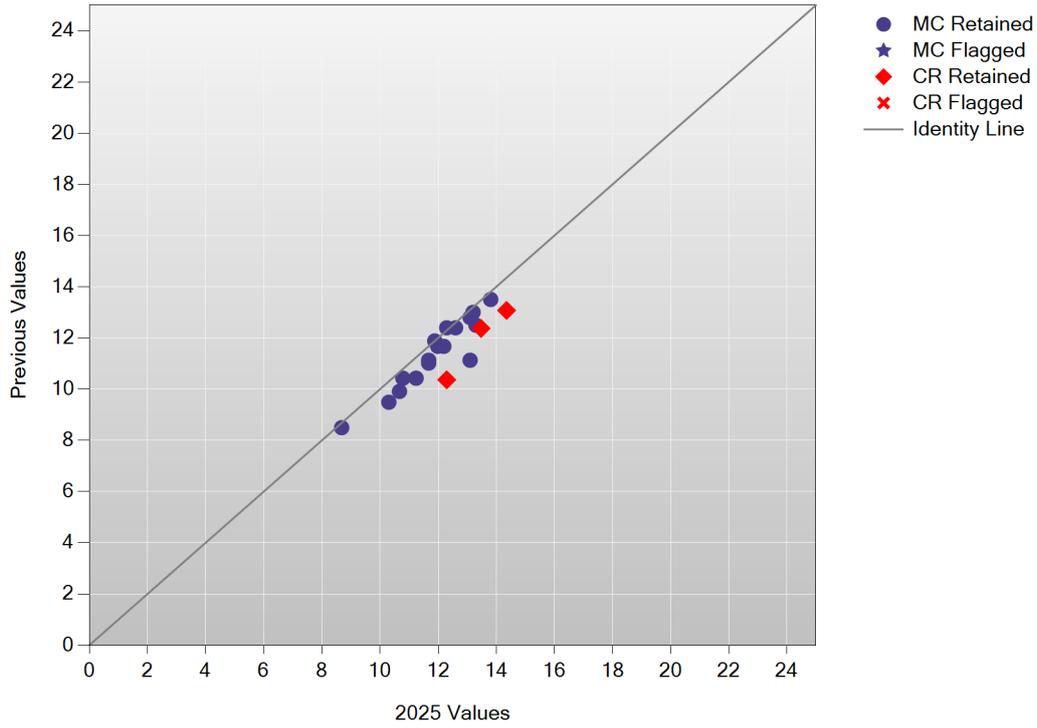
Test Characteristic Curve:Mathematics Grade 4
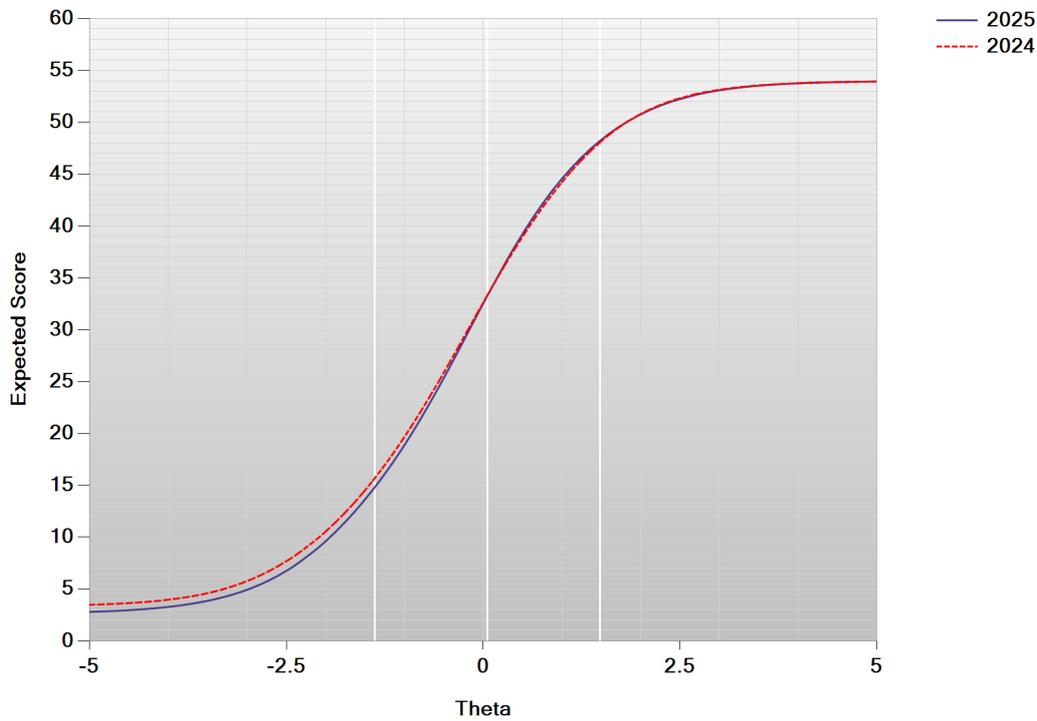
Test Information Function: Mathematics Grade 4
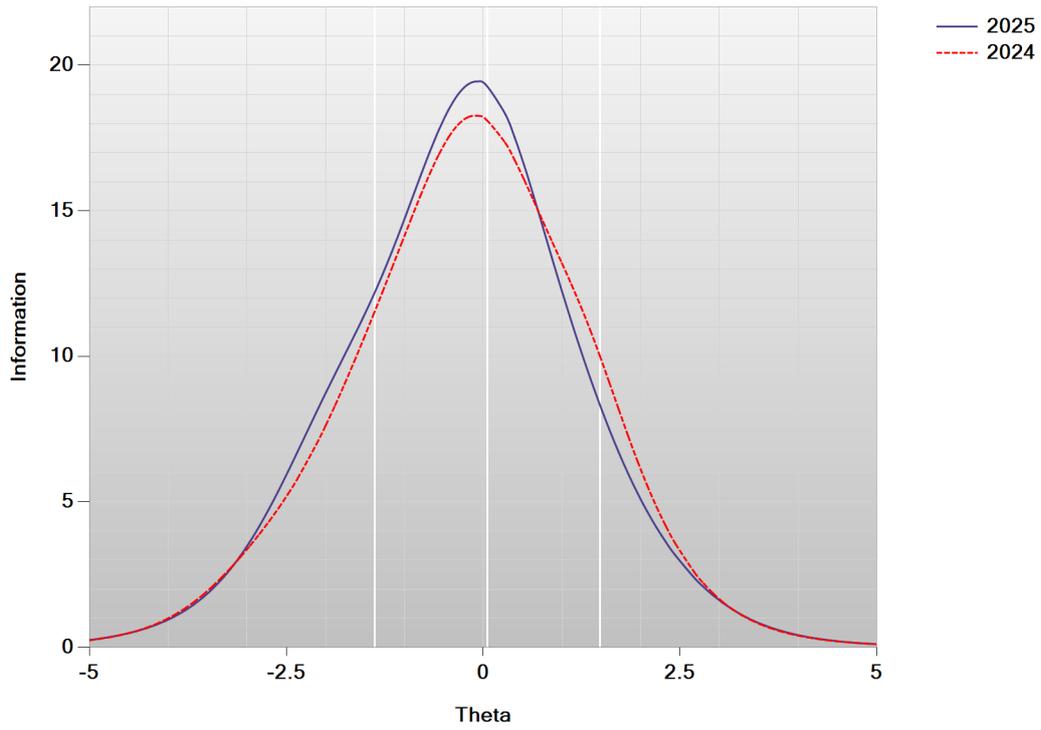
Cumulative Scale Score Distributions: Mathematics Grade 4
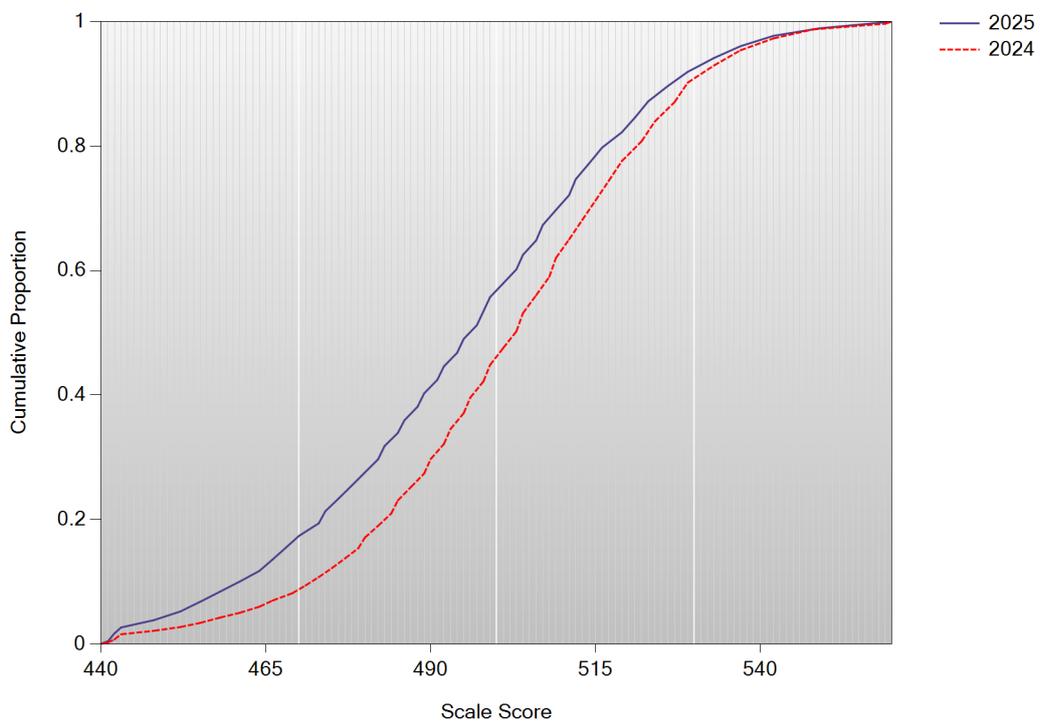
A/A Plot: Mathematics Grade 5
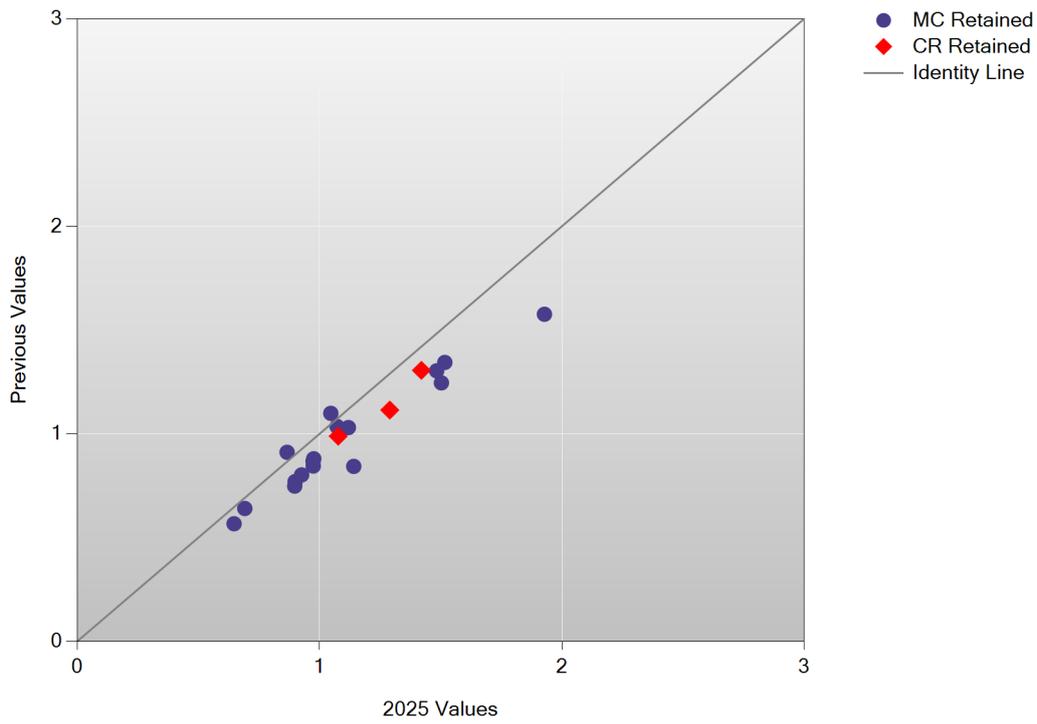
B/B Plot: Mathematics Grade 5
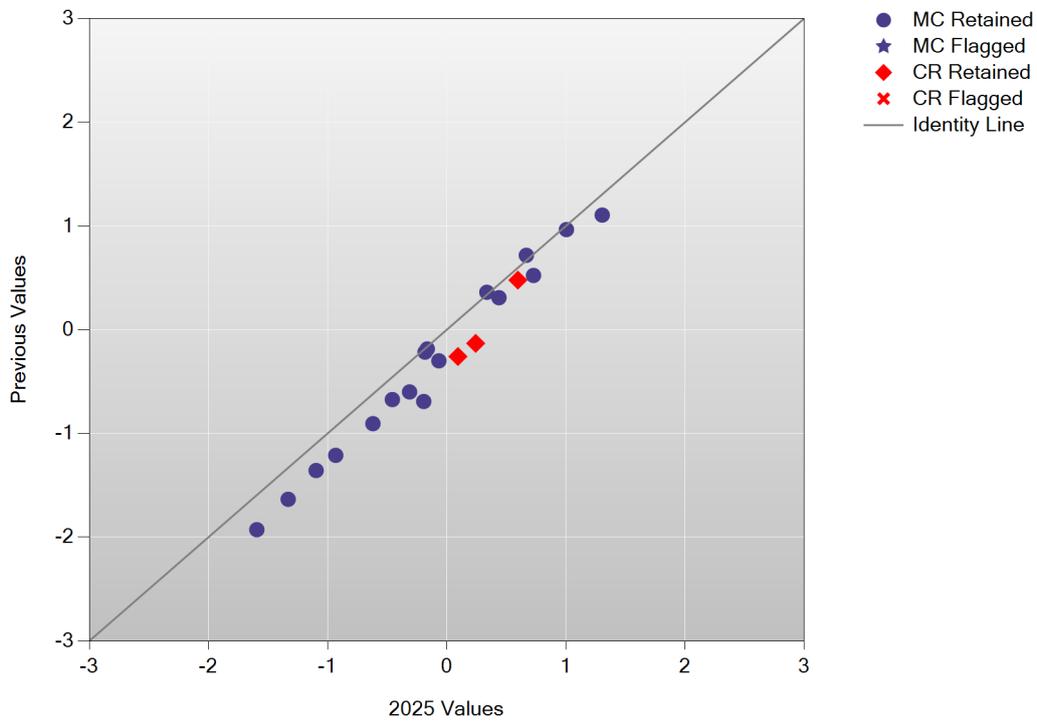
Delta Plot: Mathematics Grade 5
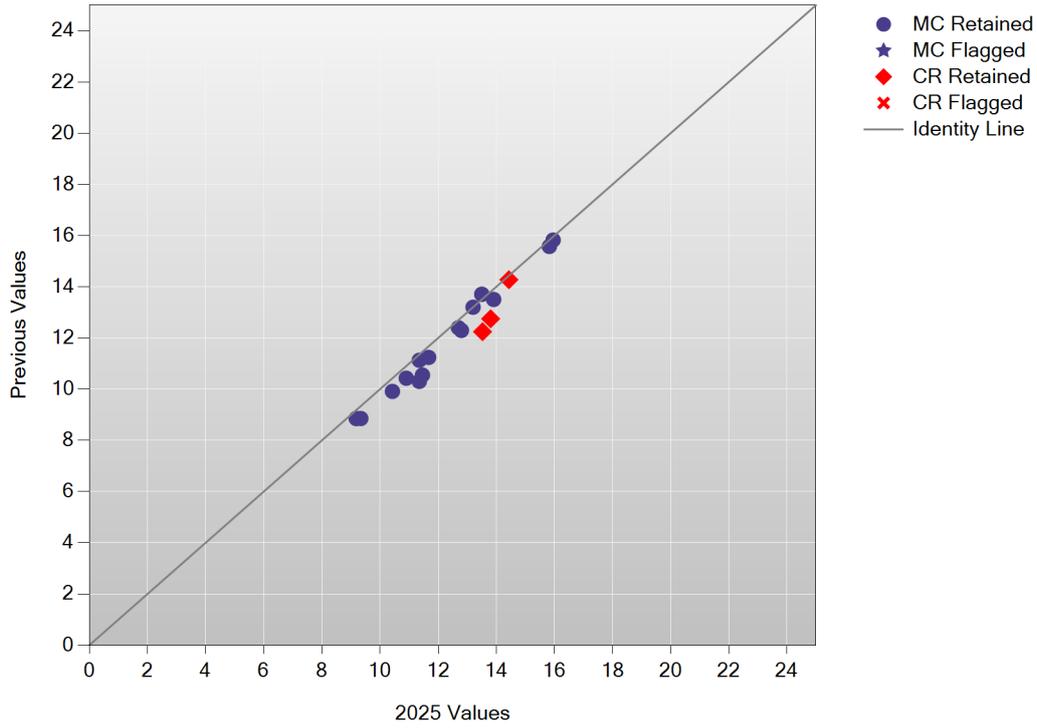
Test Characteristic Curve:Mathematics Grade 5
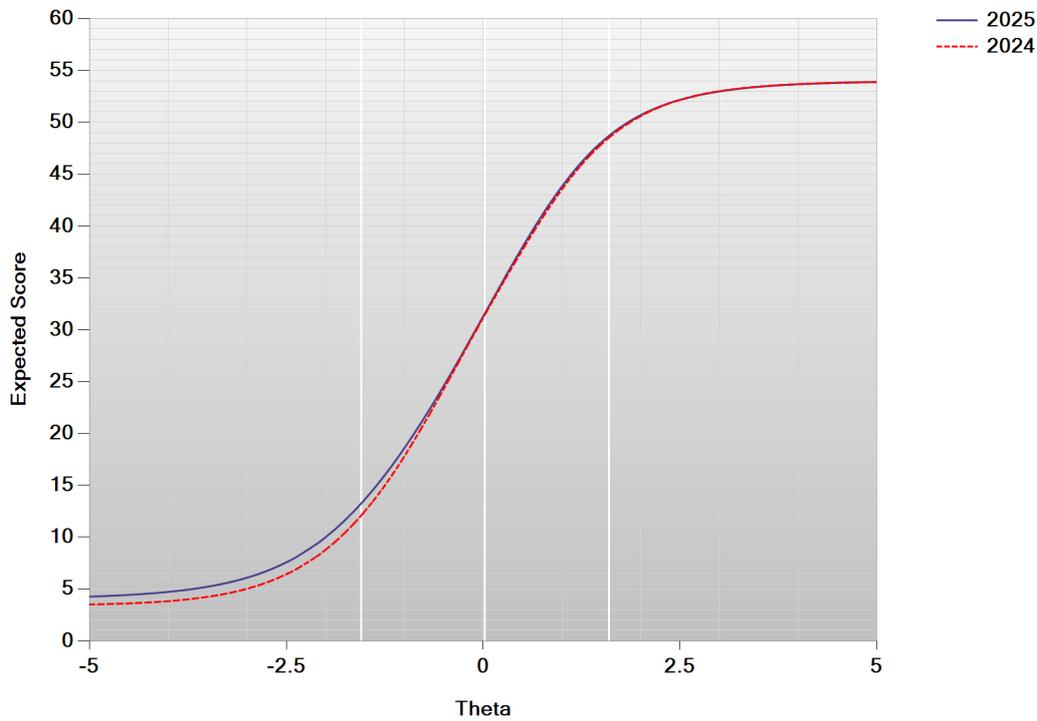
Test Information Function: Mathematics Grade 5
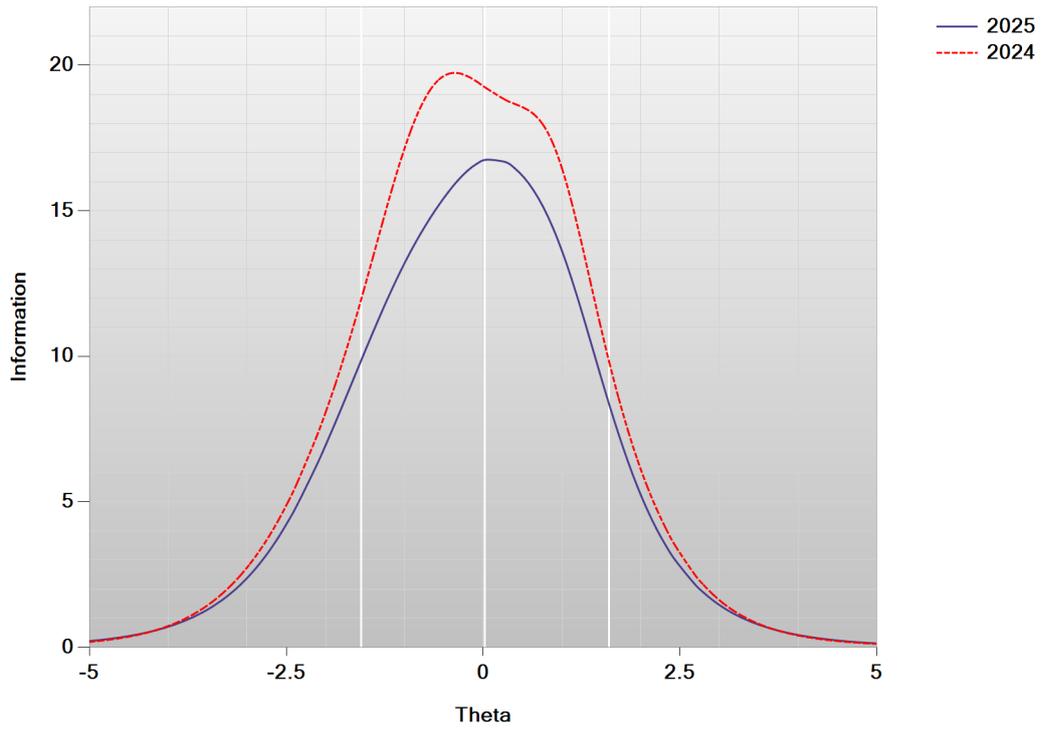
Cumulative Scale Score Distributions: Mathematics Grade 5
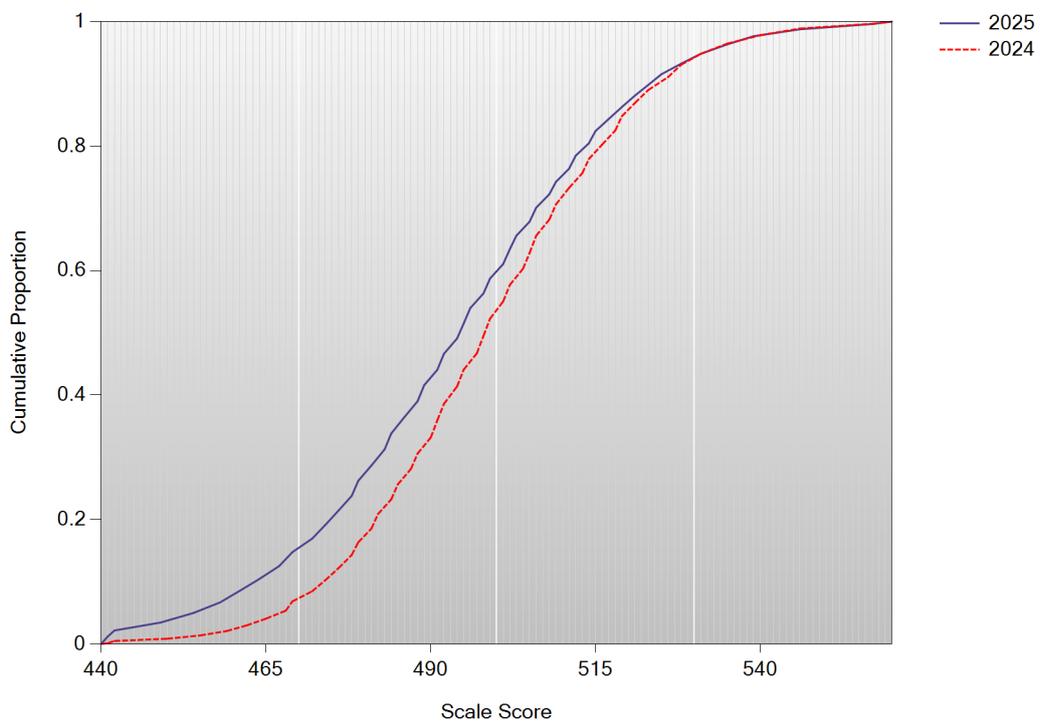
A/A Plot: Mathematics Grade 6
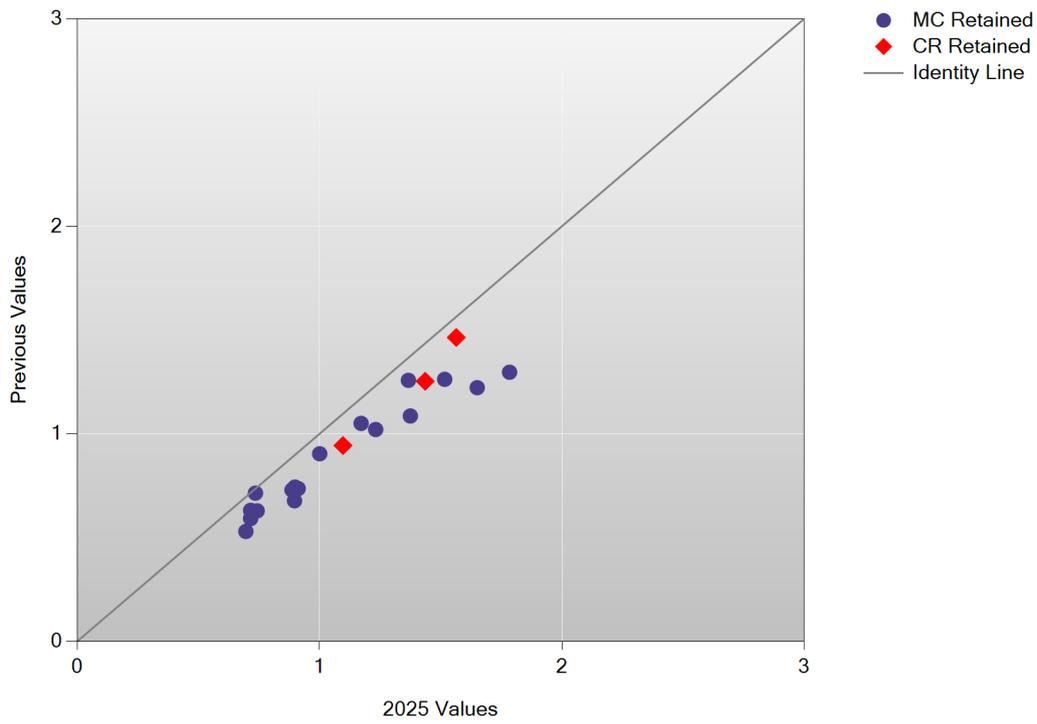
B/B Plot: Mathematics Grade 6
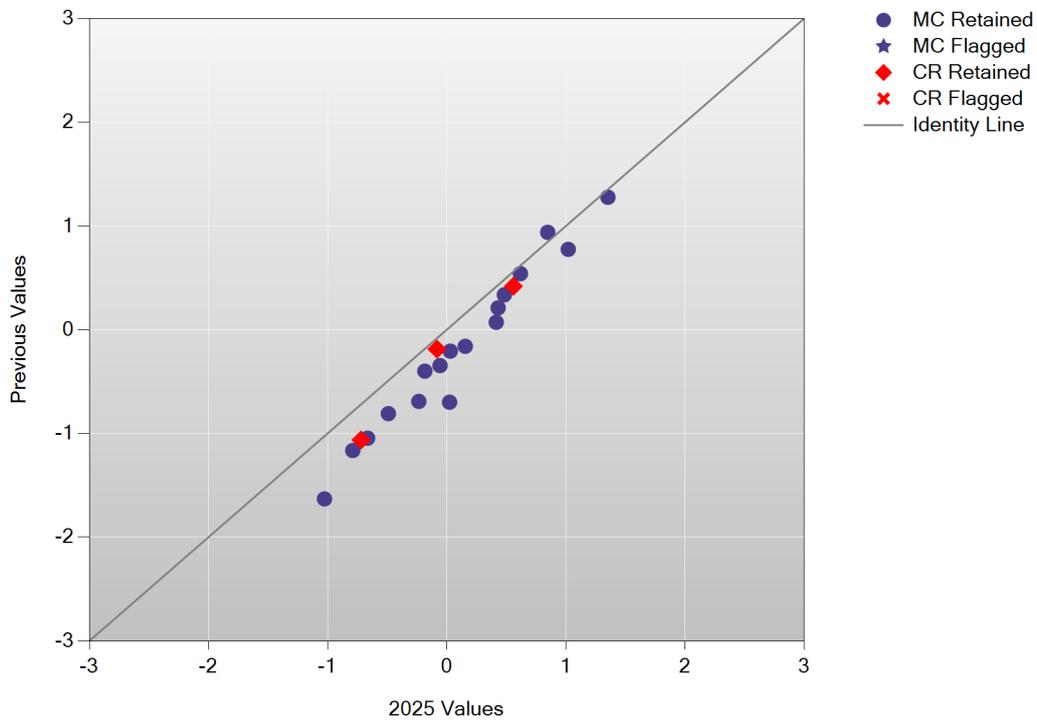
Delta Plot: Mathematics Grade 6
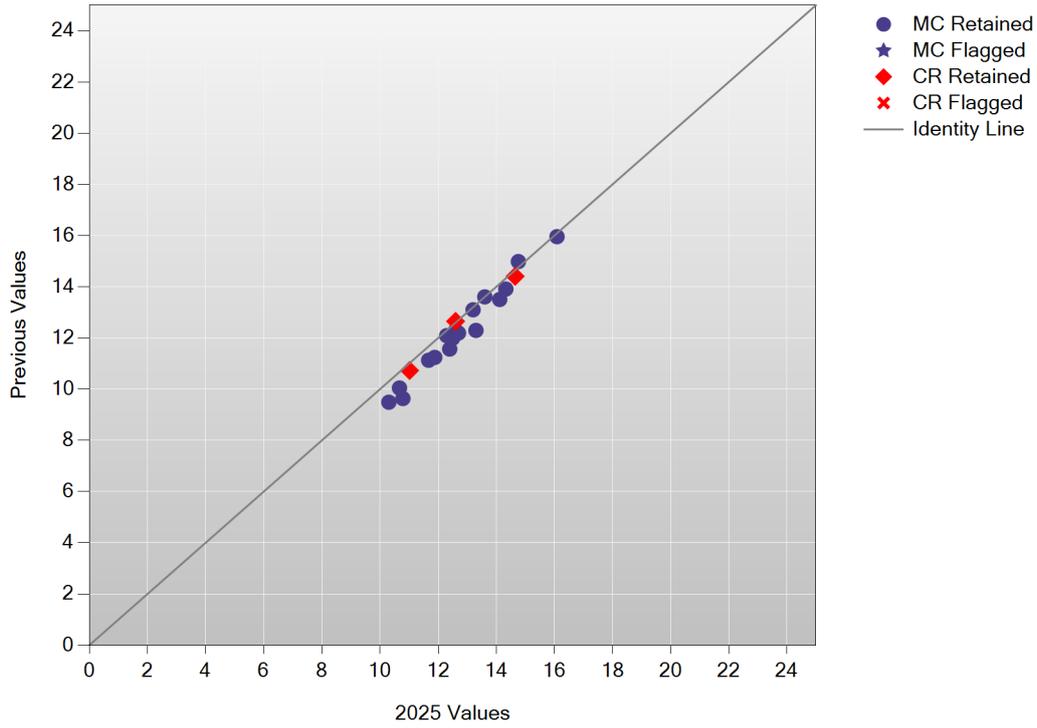
Test Characteristic Curve:Mathematics Grade 6
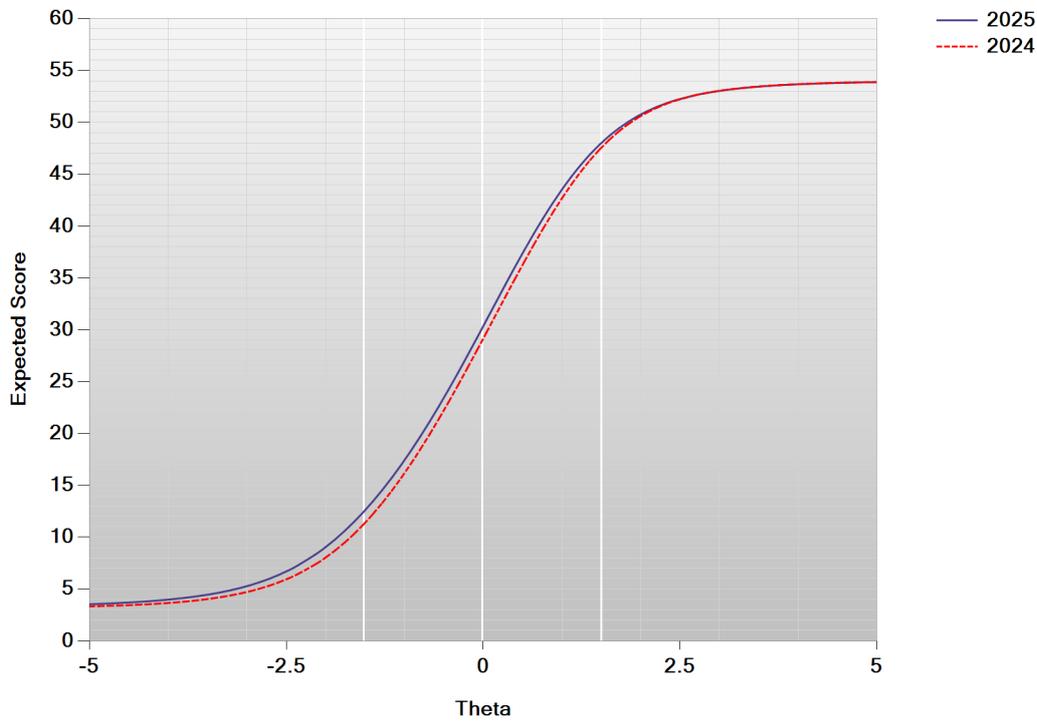
Test Information Function: Mathematics Grade 6
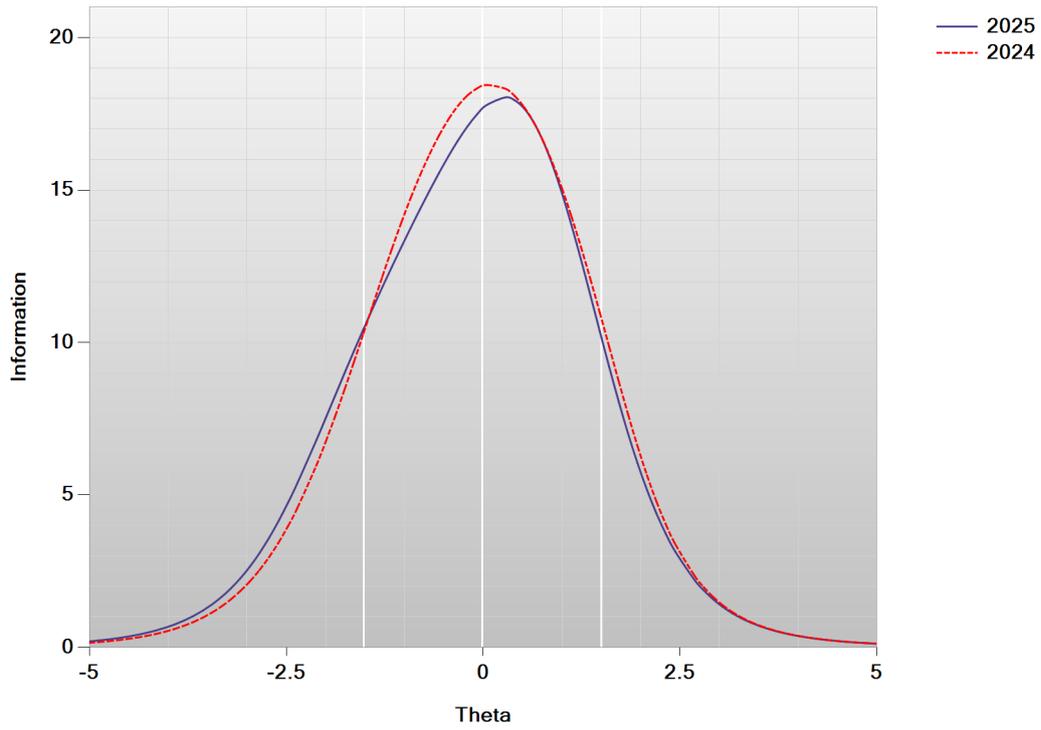
Cumulative Scale Score Distributions: Mathematics Grade 6
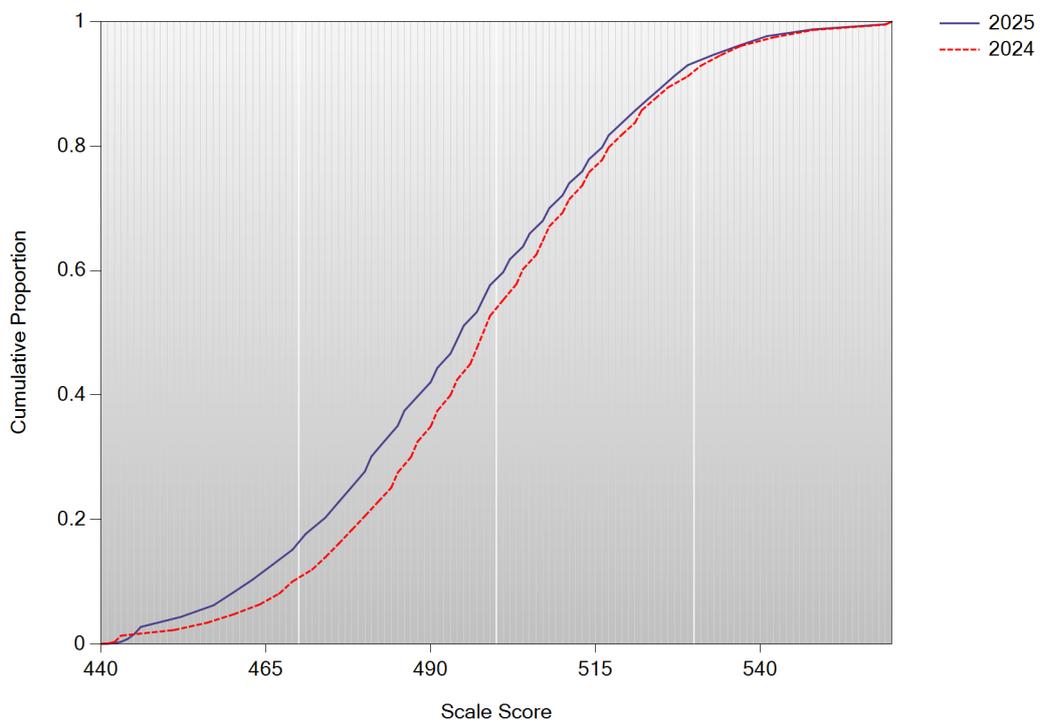
A/A Plot: Mathematics Grade 7
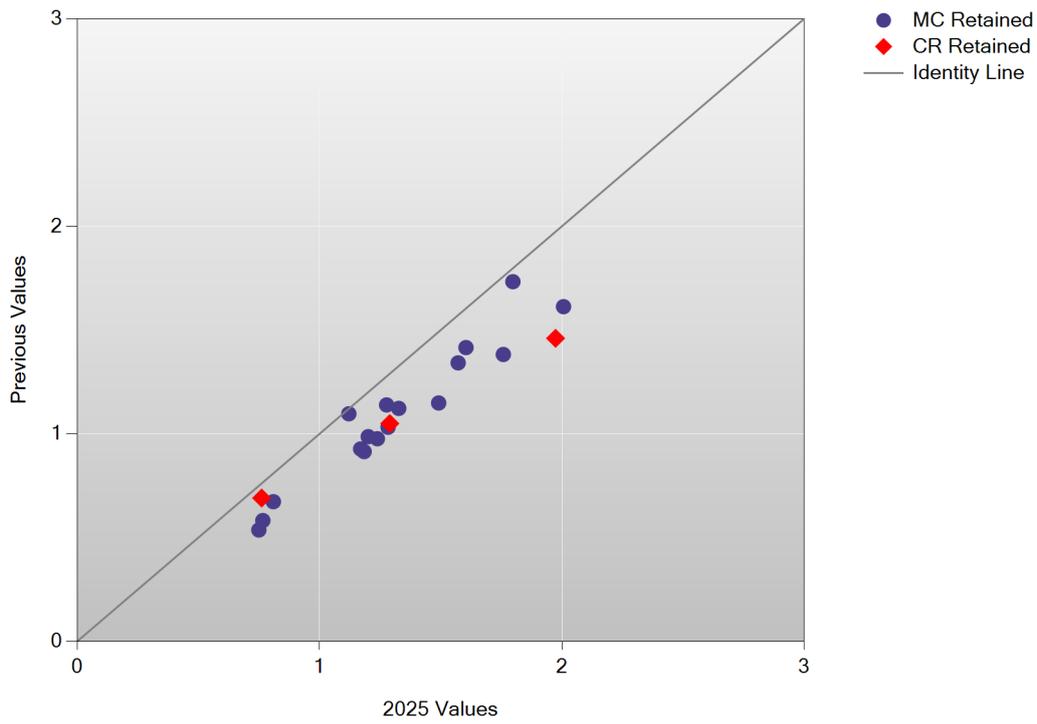
B/B Plot: Mathematics Grade 7
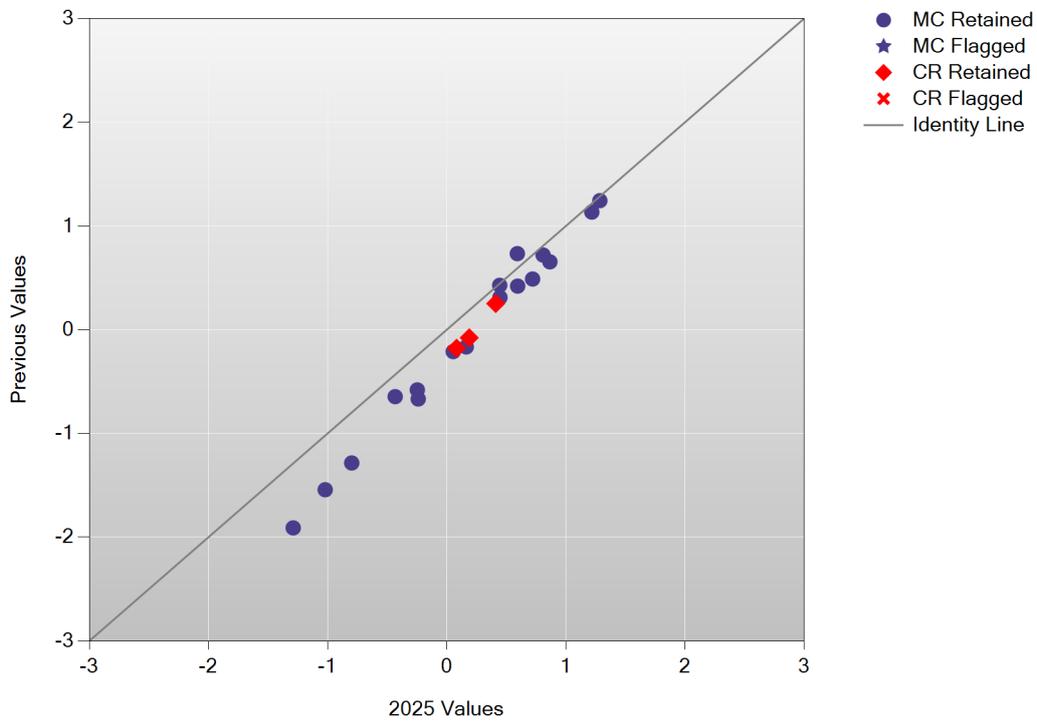
Delta Plot: Mathematics Grade 7
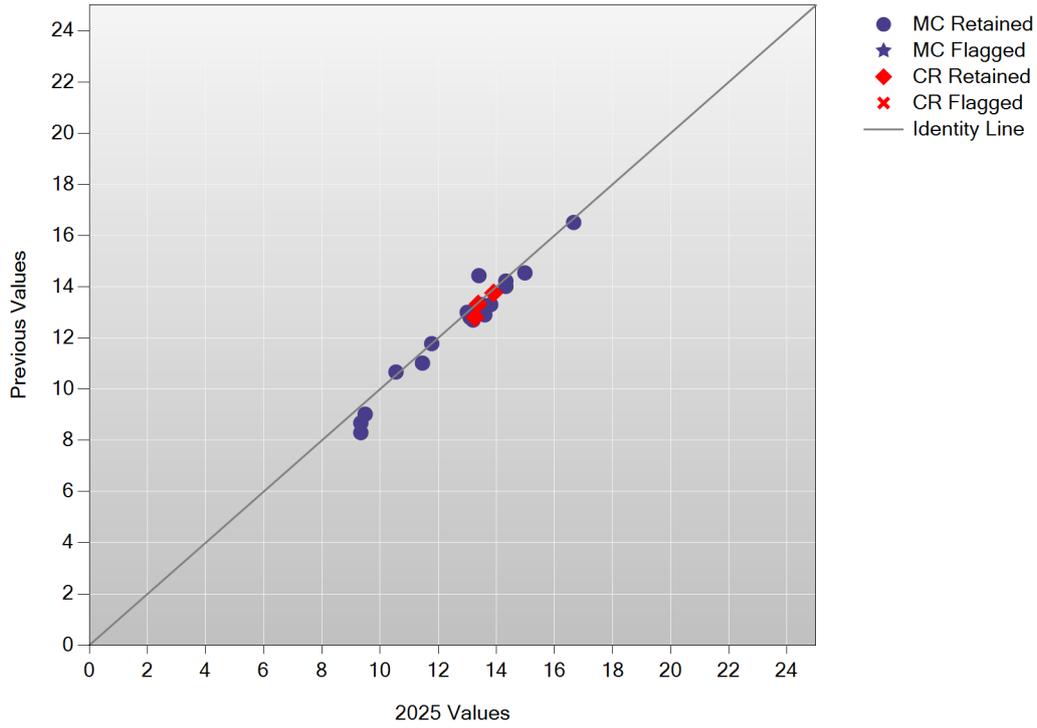
Test Characteristic Curve:Mathematics Grade 7
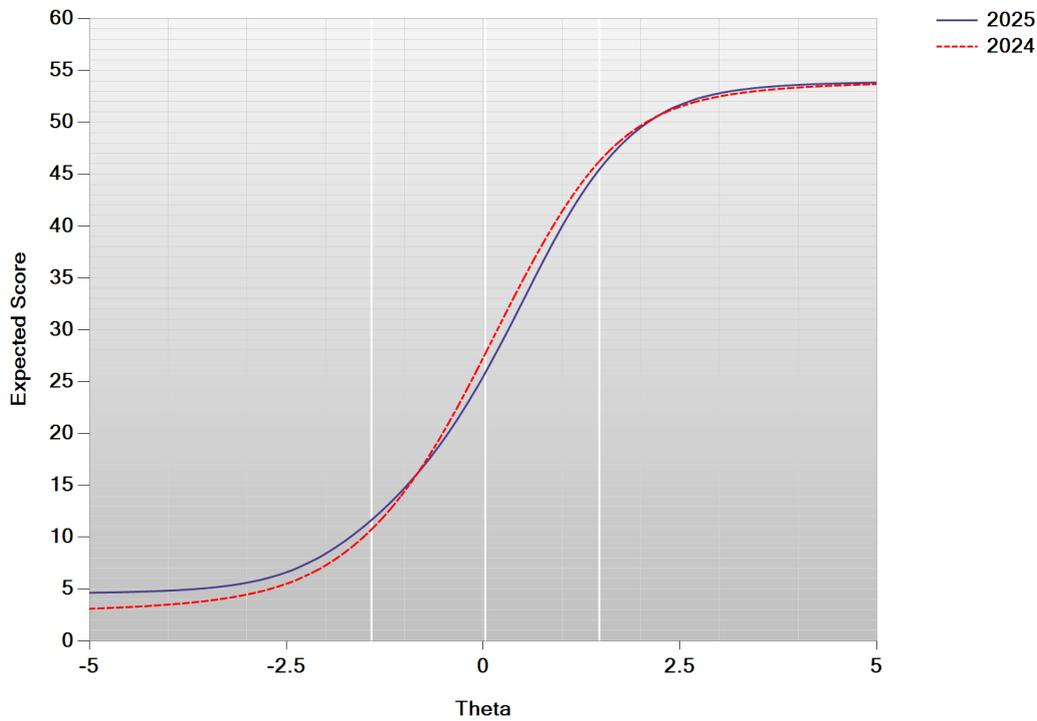
Test Information Function: Mathematics Grade 7
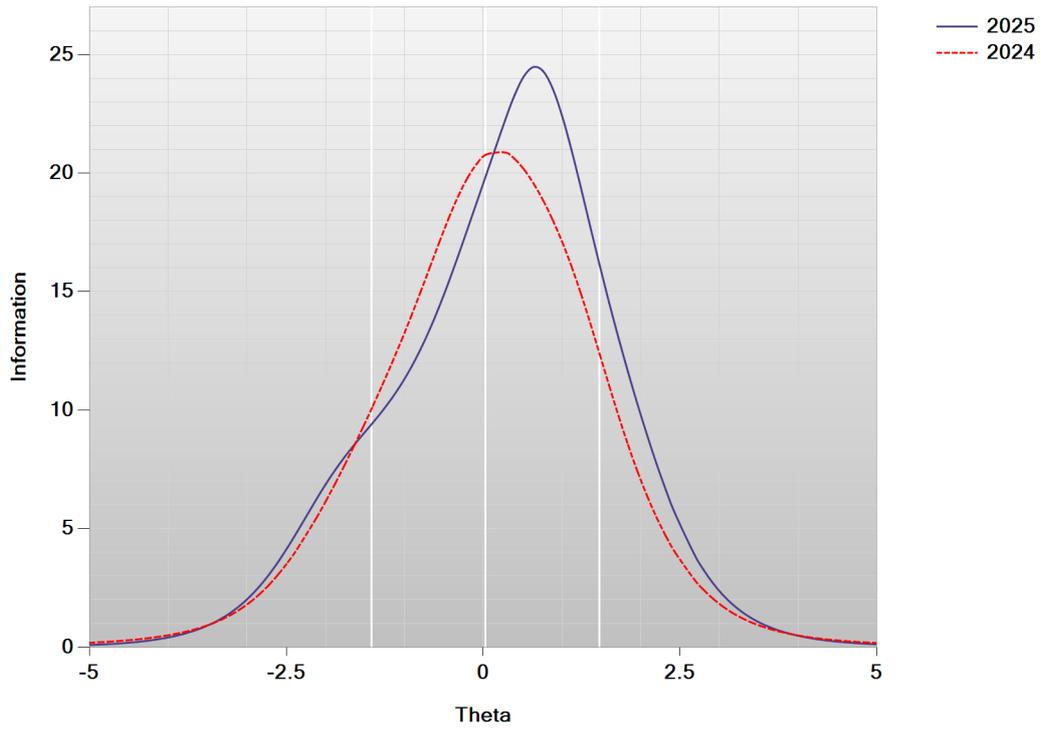
Cumulative Scale Score Distributions: Mathematics Grade 7
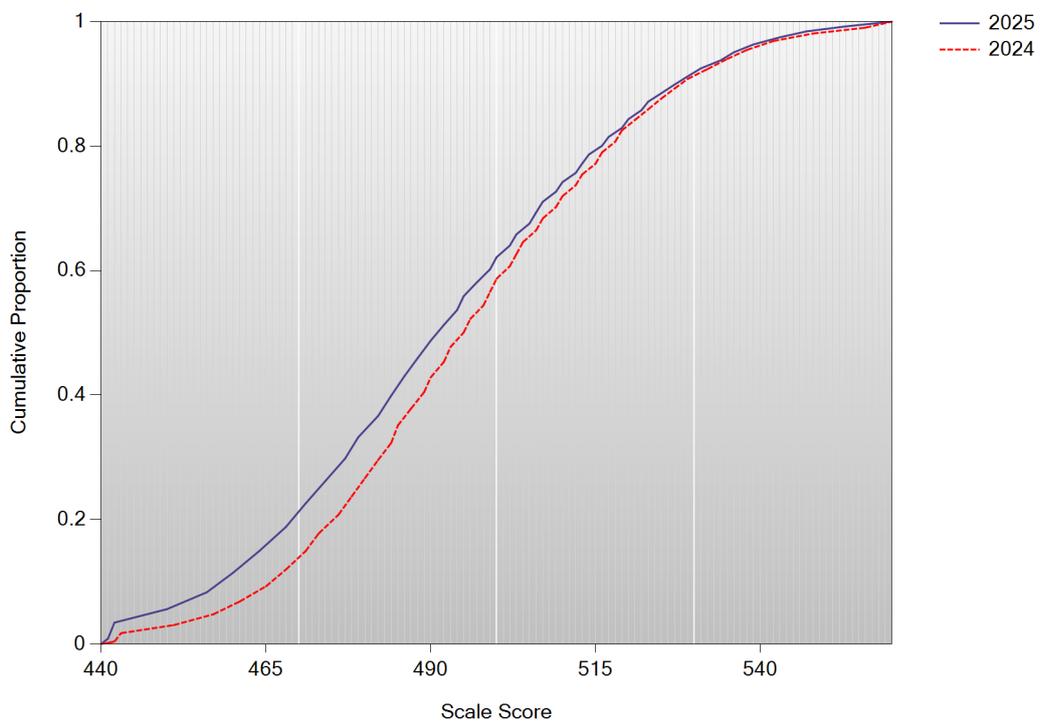
A/A Plot: Mathematics Grade 8
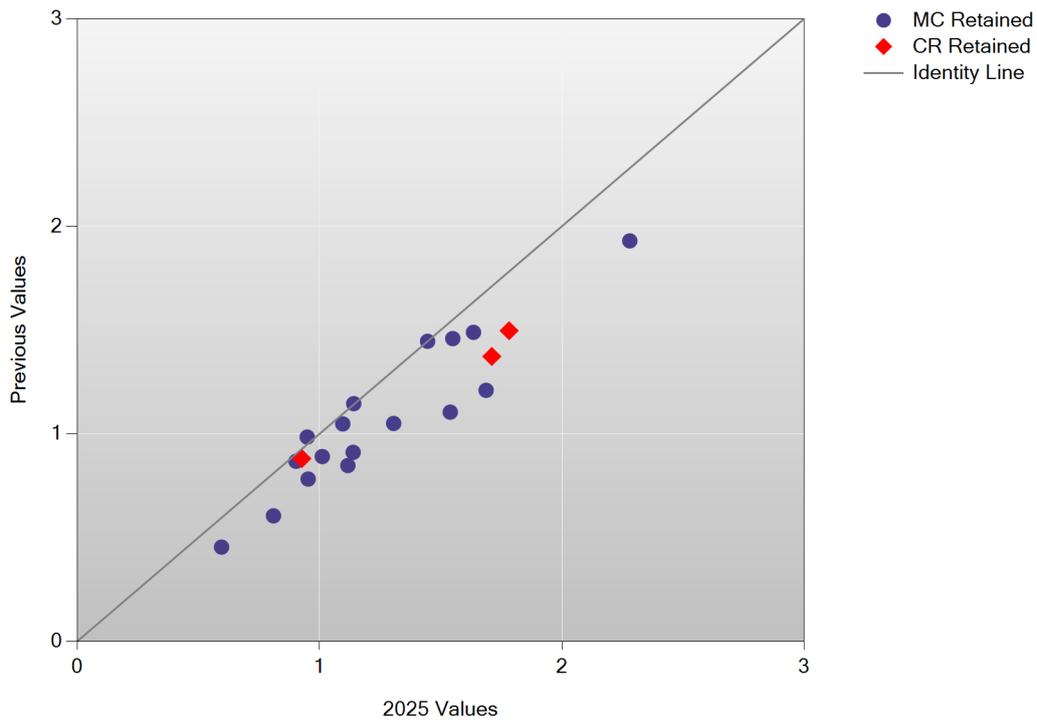
B/B Plot: Mathematics Grade 8
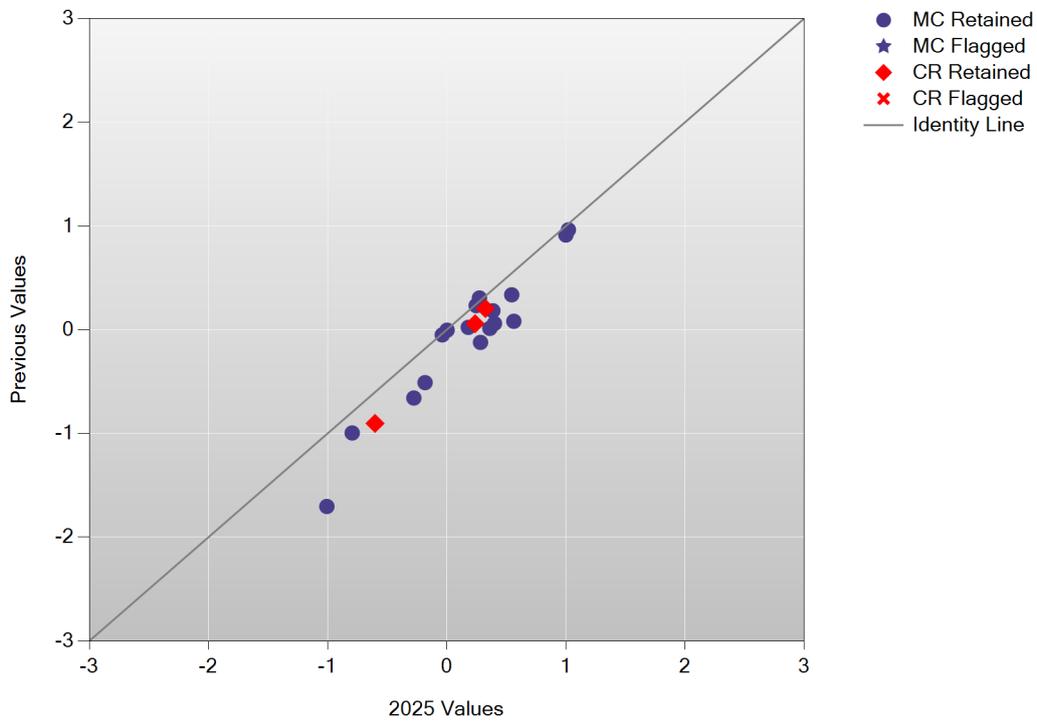
Delta Plot: Mathematics Grade 8
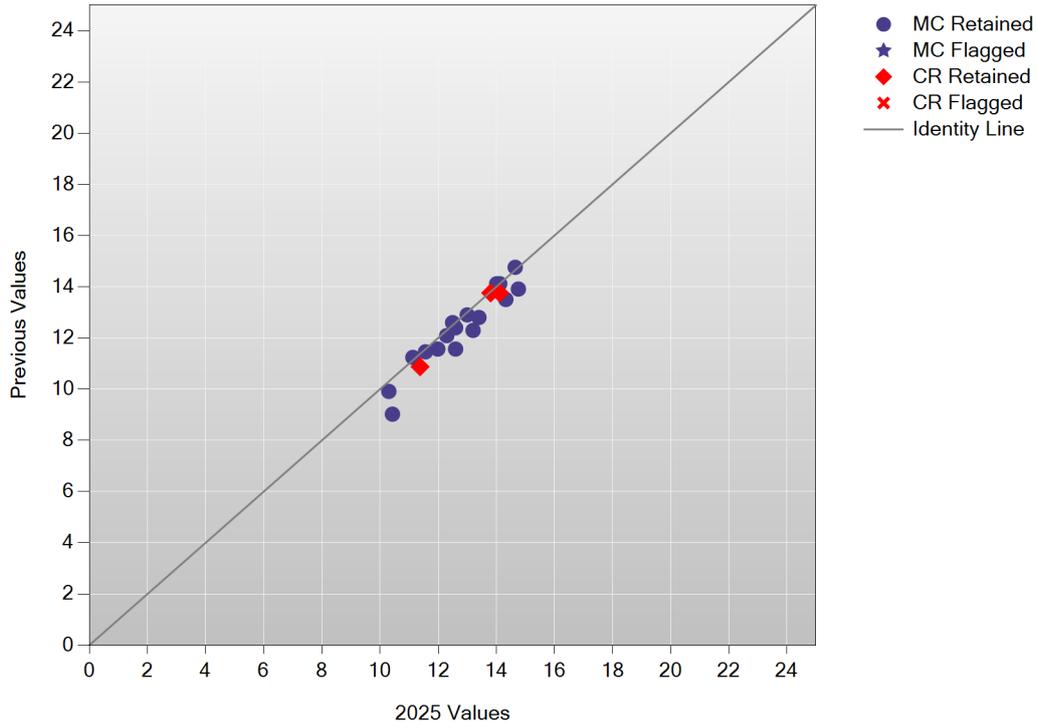
Test Characteristic Curve:Mathematics Grade 8
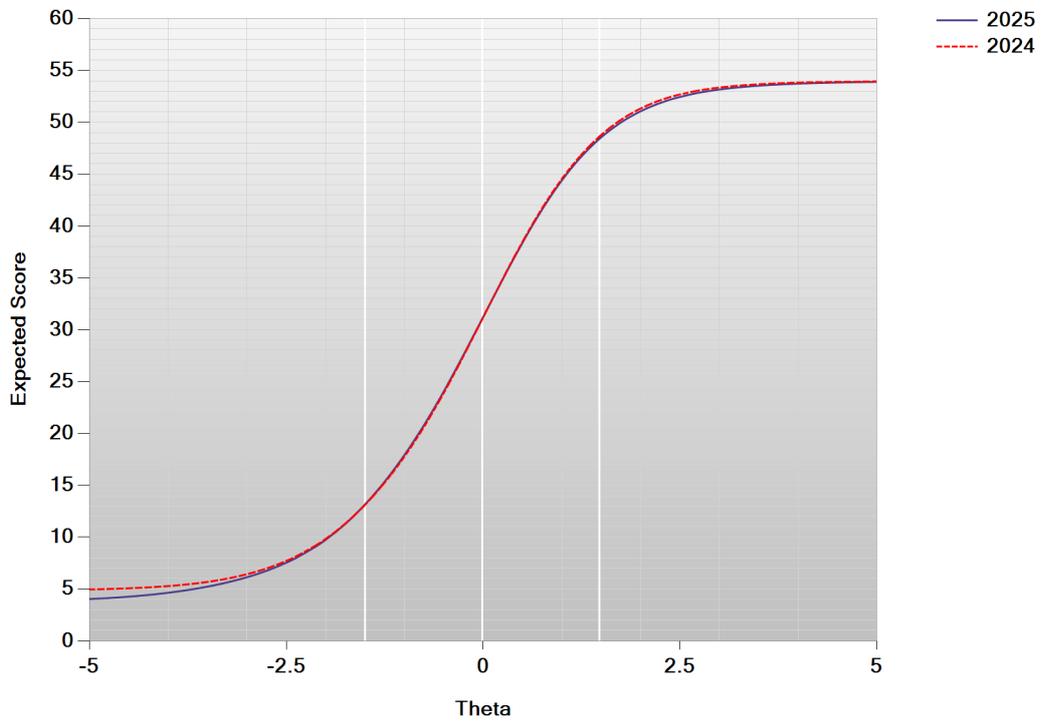
Test Information Function: Mathematics Grade 8
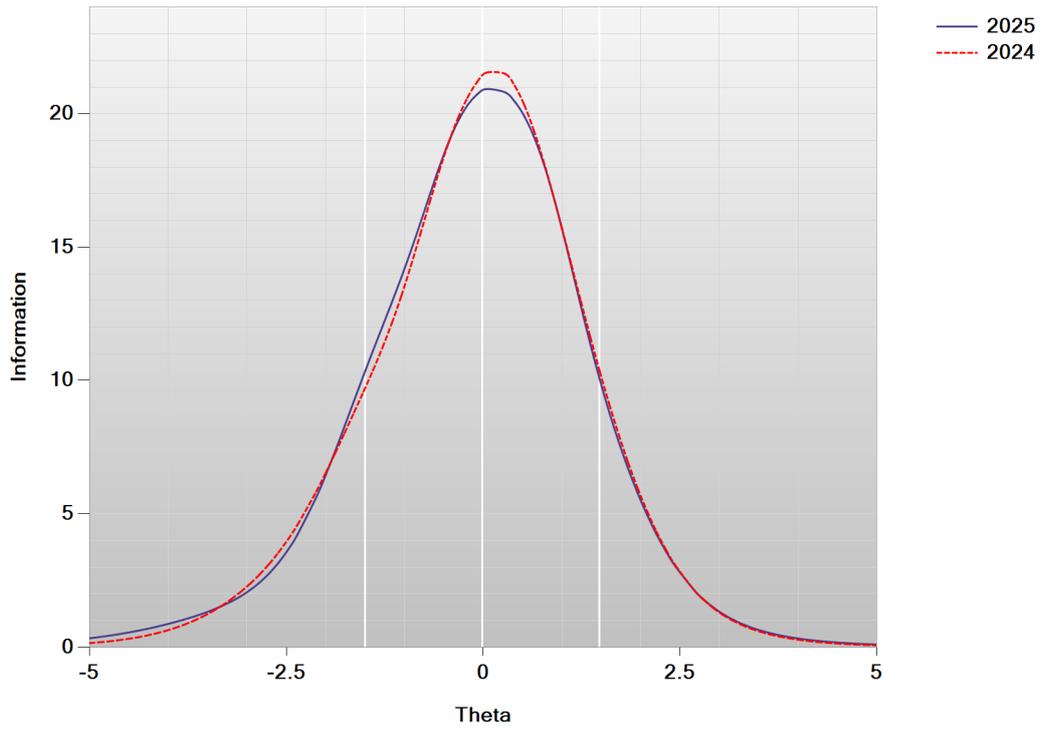
Cumulative Scale Score Distributions: Mathematics Grade 8
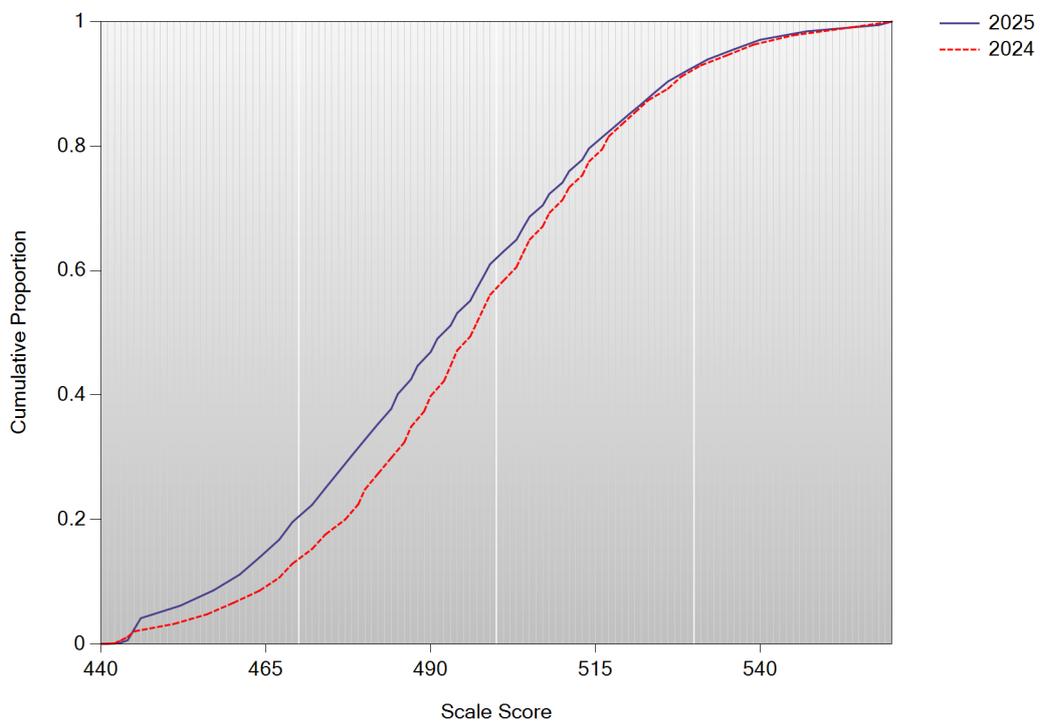
Section 2.2

Lookup Tables

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

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.173	1.42	10.0	440	1	440	1
1	-3.168	1.43	10.0	440	1	441	1
2	-3.163	1.43	10.0	440	1	441	1
3	-3.158	1.44	10.0	440	1	442	1
4	-3.153	1.45	10.0	440	1	442	1
5	-3.148	1.45	10.0	440	1	443	1
6	-3.143	1.46	10.0	441	1	443	1
7	-2.756	2.14	10.0	448	1	444	1
8	-2.460	2.85	10.0	453	1	450	1
9	-2.218	3.60	9.9	458	1	455	1
10	-2.012	4.35	9.0	462	1	460	1
11	-1.830	5.10	8.3	465	1	463	1
12	-1.667	5.82	7.8	468	1	466	1
13	-1.518	6.51	7.4	471	2	469	1
14	-1.378	7.15	7.0	474	2	472	2
15	-1.247	7.72	6.8	476	2	474	2
16	-1.123	8.22	6.6	479	2	477	2
17	-1.002	8.64	6.4	481	2	479	2
18	-0.886	8.98	6.3	483	2	481	2
19	-0.772	9.23	6.2	485	2	484	2
20	-0.659	9.39	6.1	487	2	486	2
21	-0.548	9.47	6.1	489	2	488	2
22	-0.436	9.46	6.1	492	2	490	2
23	-0.325	9.39	6.1	494	2	492	2
24	-0.212	9.25	6.2	496	2	494	2
25	-0.097	9.06	6.3	498	2	497	2
26	0.020	8.82	6.3	500	3	499	2
27	0.139	8.55	6.4	502	3	501	3
28	0.263	8.25	6.6	505	3	503	3
29	0.390	7.93	6.7	507	3	506	3
30	0.522	7.59	6.8	510	3	508	3
31	0.661	7.22	7.0	512	3	511	3
32	0.806	6.83	7.2	515	3	514	3
33	0.960	6.40	7.4	518	3	517	3
34	1.124	5.96	7.7	521	3	520	3
35	1.301	5.49	8.0	524	3	523	3
36	1.494	5.01	8.4	528	3	527	3
37	1.706	4.55	8.8	532	4	531	4

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	1.942	4.11	9.3	536	4	536	4
39	2.209	3.66	9.8	541	4	541	4
40	2.518	3.18	10.0	547	4	547	4
41	2.891	2.63	10.0	554	4	555	4
42	3.196	2.24	10.0	560	4	560	4
43	3.196	2.24	10.0	560	4	560	4
44	3.196	2.24	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.153	2.34	10.0	440	1	440	1
1	-3.149	2.34	10.0	440	1	440	1
2	-3.146	2.35	10.0	440	1	440	1
3	-3.142	2.36	10.0	440	1	441	1
4	-3.138	2.37	10.0	440	1	441	1
5	-3.135	2.37	10.0	440	1	441	1
6	-3.131	2.38	10.0	440	1	441	1
7	-3.127	2.39	10.0	440	1	441	1
8	-2.814	3.17	10.0	446	1	448	1
9	-2.564	4.03	9.4	451	1	453	1
10	-2.355	4.97	8.5	455	1	457	1
11	-2.175	5.94	7.7	458	1	461	1
12	-2.015	6.91	7.2	461	1	464	1
13	-1.871	7.82	6.7	464	1	467	1
14	-1.737	8.65	6.4	467	1	469	1
15	-1.612	9.37	6.2	469	1	472	2
16	-1.494	9.98	6.0	471	2	474	2
17	-1.380	10.48	5.8	473	2	476	2
18	-1.269	10.87	5.7	475	2	479	2
19	-1.161	11.17	5.6	478	2	481	2
20	-1.055	11.37	5.6	480	2	483	2
21	-0.950	11.50	5.6	482	2	485	2
22	-0.844	11.54	5.5	484	2	487	2
23	-0.739	11.52	5.6	485	2	489	2
24	-0.633	11.42	5.6	487	2	491	2
25	-0.525	11.26	5.6	490	2	493	2
26	-0.416	11.05	5.7	492	2	495	2
27	-0.305	10.79	5.7	494	2	497	2
28	-0.190	10.48	5.8	496	2	499	2
29	-0.073	10.12	5.9	498	2	502	3
30	0.049	9.73	6.0	500	3	504	3
31	0.175	9.29	6.2	503	3	506	3
32	0.308	8.80	6.4	505	3	509	3
33	0.448	8.26	6.6	508	3	511	3
34	0.596	7.65	6.8	511	3	514	3
35	0.756	6.99	7.1	514	3	517	3
36	0.931	6.28	7.5	517	3	520	3
37	1.125	5.53	8.0	521	3	524	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	1.343	4.78	8.6	525	3	528	3
39	1.593	4.03	9.4	529	3	532	4
40	1.887	3.32	10.0	535	4	538	4
41	2.246	2.63	10.0	542	4	544	4
42	2.710	1.94	10.0	550	4	553	4
43	3.215	1.33	10.0	560	4	560	4
44	3.215	1.33	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.360	2.14	10.0	440	1	440	1
1	-3.347	2.18	10.0	440	1	440	1
2	-3.333	2.22	10.0	440	1	441	1
3	-3.320	2.25	10.0	441	1	441	1
4	-3.307	2.29	10.0	441	1	441	1
5	-3.294	2.33	10.0	441	1	442	1
6	-3.281	2.36	10.0	441	1	442	1
7	-2.944	3.52	9.4	447	1	448	1
8	-2.689	4.64	8.2	452	1	452	1
9	-2.479	5.69	7.4	456	1	455	1
10	-2.299	6.65	6.9	459	1	458	1
11	-2.139	7.51	6.5	462	1	461	1
12	-1.993	8.27	6.1	464	1	463	1
13	-1.858	8.93	5.8	467	1	465	1
14	-1.730	9.47	5.7	469	1	467	1
15	-1.609	9.91	5.6	471	2	469	1
16	-1.492	10.25	5.5	473	2	471	2
17	-1.379	10.50	5.5	475	2	472	2
18	-1.269	10.68	5.4	477	2	474	2
19	-1.160	10.80	5.4	479	2	476	2
20	-1.053	10.88	5.4	481	2	478	2
21	-0.947	10.91	5.4	483	2	479	2
22	-0.841	10.91	5.4	484	2	481	2
23	-0.736	10.89	5.4	486	2	482	2
24	-0.631	10.82	5.4	488	2	484	2
25	-0.526	10.72	5.4	490	2	486	2
26	-0.419	10.58	5.4	492	2	487	2
27	-0.312	10.40	5.5	494	2	489	2
28	-0.203	10.18	5.5	496	2	491	2
29	-0.093	9.92	5.6	498	2	492	2
30	0.021	9.63	5.7	499	2	494	2
31	0.137	9.31	5.8	502	3	496	2
32	0.256	8.96	5.9	504	3	498	2
33	0.380	8.61	6.0	506	3	499	2
34	0.509	8.24	6.2	508	3	502	3
35	0.642	7.87	6.3	511	3	504	3
36	0.782	7.49	6.5	513	3	506	3
37	0.930	7.10	6.6	516	3	508	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	1.085	6.70	6.8	519	3	510	3
39	1.251	6.27	7.1	521	3	513	3
40	1.429	5.83	7.3	525	3	516	3
41	1.622	5.38	7.6	528	3	519	3
42	1.834	4.91	8.0	532	4	522	3
43	2.071	4.41	8.4	536	4	526	3
44	2.343	3.81	9.1	541	4	531	4
45	2.675	3.01	10.0	547	4	536	4
46	3.131	1.95	10.0	555	4	544	4
47	3.430	1.39	10.0	560	4	555	4
48	3.430	1.39	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.171	3.34	10.0	440	1	440	1
1	-3.170	3.34	10.0	440	1	440	1
2	-3.169	3.34	10.0	440	1	441	1
3	-3.169	3.35	10.0	440	1	441	1
4	-3.168	3.35	10.0	440	1	442	1
5	-3.167	3.35	10.0	440	1	442	1
6	-3.166	3.35	10.0	440	1	443	1
7	-3.165	3.36	10.0	440	1	443	1
8	-3.164	3.36	10.0	440	1	443	1
9	-3.163	3.36	10.0	440	1	444	1
10	-2.935	4.00	9.5	444	1	444	1
11	-2.737	4.56	8.9	448	1	448	1
12	-2.560	5.03	8.5	452	1	452	1
13	-2.398	5.43	8.1	455	1	455	1
14	-2.247	5.76	7.9	458	1	458	1
15	-2.104	6.04	7.7	460	1	461	1
16	-1.967	6.28	7.6	463	1	464	1
17	-1.835	6.49	7.5	465	1	467	1
18	-1.707	6.69	7.3	468	1	469	1
19	-1.583	6.87	7.2	470	2	472	2
20	-1.461	7.05	7.2	472	2	474	2
21	-1.341	7.22	7.1	475	2	476	2
22	-1.223	7.38	7.0	477	2	479	2
23	-1.106	7.53	6.9	479	2	481	2
24	-0.990	7.66	6.9	481	2	483	2
25	-0.875	7.76	6.8	484	2	486	2
26	-0.760	7.84	6.8	486	2	488	2
27	-0.645	7.87	6.8	488	2	490	2
28	-0.529	7.88	6.8	490	2	493	2
29	-0.413	7.84	6.8	492	2	495	2
30	-0.294	7.77	6.8	495	2	497	2
31	-0.174	7.66	6.9	497	2	499	2
32	-0.051	7.51	6.9	499	2	502	3
33	0.075	7.34	7.0	502	3	505	3
34	0.205	7.14	7.1	504	3	507	3
35	0.340	6.90	7.2	507	3	510	3
36	0.480	6.63	7.4	509	3	513	3
37	0.627	6.32	7.5	512	3	516	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.782	5.98	7.8	515	3	519	3
39	0.946	5.59	8.0	518	3	522	3
40	1.124	5.17	8.3	522	3	525	3
41	1.317	4.72	8.7	525	3	529	3
42	1.529	4.23	9.2	529	3	533	4
43	1.768	3.73	9.8	534	4	537	4
44	2.041	3.22	10.0	539	4	542	4
45	2.361	2.71	10.0	545	4	548	4
46	2.746	2.27	10.0	552	4	555	4
47	3.150	1.99	10.0	560	4	560	4
48	3.150	1.99	10.0	560	4	560	4
49	3.150	1.99	10.0	560	4	560	4
50	3.150	1.99	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.131	3.41	10.0	440	1	440	1
1	-3.129	3.41	10.0	440	1	440	1
2	-3.128	3.42	10.0	440	1	440	1
3	-3.126	3.42	10.0	440	1	441	1
4	-3.124	3.42	10.0	440	1	441	1
5	-3.123	3.43	10.0	440	1	441	1
6	-3.121	3.43	10.0	440	1	441	1
7	-3.120	3.44	10.0	440	1	441	1
8	-3.118	3.44	10.0	440	1	441	1
9	-3.116	3.44	10.0	440	1	442	1
10	-2.866	4.07	9.5	445	1	446	1
11	-2.651	4.59	8.9	449	1	449	1
12	-2.461	5.03	8.5	453	1	453	1
13	-2.288	5.44	8.2	456	1	456	1
14	-2.130	5.85	7.9	459	1	458	1
15	-1.982	6.29	7.6	462	1	461	1
16	-1.843	6.76	7.3	465	1	464	1
17	-1.712	7.24	7.1	467	1	466	1
18	-1.587	7.73	6.9	469	1	469	1
19	-1.468	8.21	6.7	472	2	471	2
20	-1.353	8.66	6.5	474	2	473	2
21	-1.241	9.06	6.3	476	2	475	2
22	-1.132	9.41	6.2	478	2	477	2
23	-1.026	9.69	6.1	480	2	479	2
24	-0.920	9.90	6.1	482	2	482	2
25	-0.816	10.05	6.0	484	2	484	2
26	-0.711	10.13	6.0	486	2	486	2
27	-0.607	10.14	6.0	488	2	488	2
28	-0.502	10.10	6.0	490	2	490	2
29	-0.395	10.01	6.0	492	2	492	2
30	-0.287	9.87	6.1	494	2	494	2
31	-0.177	9.69	6.1	496	2	496	2
32	-0.063	9.47	6.2	499	2	498	2
33	0.053	9.21	6.3	501	3	499	2
34	0.174	8.92	6.4	503	3	502	3
35	0.300	8.58	6.5	506	3	504	3
36	0.432	8.18	6.7	508	3	507	3
37	0.572	7.73	6.9	511	3	509	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.722	7.21	7.1	514	3	512	3
39	0.883	6.63	7.4	517	3	515	3
40	1.060	5.99	7.8	520	3	518	3
41	1.258	5.31	8.3	524	3	521	3
42	1.481	4.62	8.9	528	3	524	3
43	1.739	3.95	9.6	533	4	528	3
44	2.042	3.33	10.0	539	4	533	4
45	2.409	2.72	10.0	546	4	538	4
46	2.878	2.04	10.0	555	4	544	4
47	3.153	1.72	10.0	560	4	551	4
48	3.153	1.72	10.0	560	4	560	4
49	3.153	1.72	10.0	560	4	560	4
50	3.153	1.72	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-2.964	4.11	9.8	440	1	440	1
1	-2.953	4.15	9.8	440	1	440	1
2	-2.941	4.18	9.7	440	1	440	1
3	-2.930	4.22	9.7	441	1	440	1
4	-2.919	4.25	9.7	441	1	441	1
5	-2.908	4.28	9.6	441	1	441	1
6	-2.896	4.32	9.6	441	1	441	1
7	-2.885	4.35	9.5	442	1	441	1
8	-2.874	4.39	9.5	442	1	441	1
9	-2.863	4.42	9.5	442	1	441	1
10	-2.851	4.45	9.4	442	1	445	1
11	-2.664	5.02	8.9	446	1	449	1
12	-2.495	5.51	8.5	449	1	452	1
13	-2.339	5.93	8.2	452	1	455	1
14	-2.194	6.28	7.9	455	1	457	1
15	-2.057	6.59	7.8	458	1	460	1
16	-1.926	6.85	7.6	461	1	462	1
17	-1.800	7.08	7.5	463	1	465	1
18	-1.678	7.28	7.4	466	1	467	1
19	-1.560	7.48	7.3	468	1	469	1
20	-1.445	7.66	7.2	470	2	472	2
21	-1.332	7.84	7.1	472	2	474	2
22	-1.222	8.02	7.0	475	2	476	2
23	-1.113	8.19	7.0	477	2	478	2
24	-1.005	8.35	6.9	479	2	481	2
25	-0.898	8.50	6.8	481	2	483	2
26	-0.792	8.64	6.8	483	2	485	2
27	-0.687	8.76	6.7	485	2	487	2
28	-0.581	8.87	6.7	487	2	489	2
29	-0.475	8.94	6.7	490	2	492	2
30	-0.369	8.99	6.6	492	2	494	2
31	-0.261	9.01	6.6	494	2	496	2
32	-0.152	8.98	6.6	496	2	498	2
33	-0.041	8.90	6.7	498	2	501	3
34	0.073	8.77	6.7	500	3	503	3
35	0.191	8.58	6.8	503	3	506	3
36	0.313	8.32	6.9	505	3	508	3
37	0.440	8.00	7.0	508	3	511	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.574	7.61	7.2	510	3	514	3
39	0.717	7.15	7.4	513	3	517	3
40	0.871	6.63	7.7	516	3	520	3
41	1.039	6.05	8.1	520	3	524	3
42	1.225	5.41	8.6	523	3	527	3
43	1.437	4.73	9.1	528	3	532	4
44	1.683	4.01	9.9	532	4	536	4
45	1.982	3.23	10.0	538	4	542	4
46	2.370	2.39	10.0	546	4	548	4
47	2.929	1.58	10.0	557	4	555	4
48	3.066	1.45	10.0	560	4	560	4
49	3.066	1.45	10.0	560	4	560	4
50	3.066	1.45	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-2.782	2.78	10.0	440	1	440	1
1	-2.751	2.88	10.0	441	1	441	1
2	-2.719	2.99	10.0	441	1	441	1
3	-2.688	3.10	10.0	442	1	442	1
4	-2.657	3.22	10.0	443	1	442	1
5	-2.626	3.34	10.0	443	1	443	1
6	-2.595	3.46	10.0	444	1	443	1
7	-2.318	4.72	9.8	450	1	449	1
8	-2.100	5.94	8.8	455	1	454	1
9	-1.920	7.09	8.0	458	1	458	1
10	-1.763	8.17	7.5	462	1	461	1
11	-1.625	9.17	7.1	465	1	464	1
12	-1.499	10.10	6.7	467	1	467	1
13	-1.383	10.94	6.5	469	1	469	1
14	-1.275	11.71	6.2	472	2	472	2
15	-1.173	12.40	6.1	474	2	474	2
16	-1.076	13.01	5.9	476	2	477	2
17	-0.983	13.56	5.8	478	2	479	2
18	-0.893	14.05	5.7	480	2	481	2
19	-0.806	14.47	5.6	482	2	482	2
20	-0.721	14.83	5.5	484	2	484	2
21	-0.638	15.14	5.5	486	2	486	2
22	-0.556	15.38	5.4	488	2	488	2
23	-0.475	15.57	5.4	489	2	490	2
24	-0.395	15.71	5.4	491	2	492	2
25	-0.314	15.78	5.4	493	2	493	2
26	-0.234	15.79	5.4	494	2	495	2
27	-0.154	15.75	5.4	496	2	497	2
28	-0.072	15.65	5.4	498	2	499	2
29	0.010	15.50	5.4	499	2	501	3
30	0.093	15.29	5.5	501	3	502	3
31	0.178	15.04	5.5	503	3	504	3
32	0.265	14.74	5.6	505	3	506	3
33	0.355	14.41	5.6	507	3	508	3
34	0.447	14.03	5.7	509	3	510	3
35	0.542	13.62	5.8	511	3	512	3
36	0.642	13.17	5.9	513	3	514	3
37	0.745	12.67	6.0	515	3	516	3

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

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.855	12.11	6.1	518	3	519	3
39	0.970	11.46	6.3	520	3	521	3
40	1.095	10.70	6.5	523	3	524	3
41	1.230	9.81	6.8	526	3	527	3
42	1.380	8.78	7.2	529	3	530	4
43	1.552	7.58	7.8	533	4	534	4
44	1.755	6.21	8.6	537	4	538	4
45	2.013	4.67	9.9	542	4	544	4
46	2.376	2.99	10.0	550	4	551	4
47	2.837	1.62	10.0	560	4	560	4
48	2.837	1.62	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-2.812	4.30	10.0	440	1	440	1
1	-2.789	4.41	10.0	440	1	440	1
2	-2.766	4.52	9.8	441	1	441	1
3	-2.743	4.63	9.7	441	1	441	1
4	-2.721	4.75	9.6	442	1	442	1
5	-2.698	4.86	9.5	442	1	442	1
6	-2.675	4.98	9.4	443	1	443	1
7	-2.447	6.21	8.4	448	1	443	1
8	-2.257	7.29	7.8	452	1	448	1
9	-2.092	8.22	7.3	455	1	452	1
10	-1.945	9.04	7.0	458	1	455	1
11	-1.810	9.77	6.7	461	1	458	1
12	-1.686	10.44	6.5	464	1	461	1
13	-1.570	11.08	6.3	466	1	464	1
14	-1.460	11.70	6.1	468	1	466	1
15	-1.357	12.31	6.0	470	2	469	1
16	-1.259	12.93	5.8	473	2	471	2
17	-1.165	13.56	5.7	474	2	473	2
18	-1.075	14.18	5.6	476	2	475	2
19	-0.989	14.79	5.4	478	2	477	2
20	-0.906	15.40	5.3	480	2	479	2
21	-0.825	15.98	5.2	482	2	480	2
22	-0.747	16.54	5.1	483	2	482	2
23	-0.671	17.07	5.1	485	2	484	2
24	-0.597	17.56	5.0	486	2	485	2
25	-0.525	18.01	4.9	488	2	487	2
26	-0.453	18.40	4.9	489	2	489	2
27	-0.383	18.74	4.8	491	2	490	2
28	-0.313	19.02	4.8	492	2	492	2
29	-0.244	19.23	4.8	494	2	493	2
30	-0.175	19.38	4.8	495	2	495	2
31	-0.106	19.44	4.7	497	2	496	2
32	-0.036	19.43	4.8	498	2	498	2
33	0.033	19.33	4.8	499	2	499	2
34	0.104	19.14	4.8	501	3	501	3
35	0.176	18.87	4.8	503	3	503	3
36	0.249	18.51	4.9	504	3	504	3
37	0.324	18.06	4.9	506	3	506	3

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

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.401	17.52	5.0	507	3	508	3
39	0.480	16.90	5.1	509	3	509	3
40	0.563	16.20	5.2	511	3	511	3
41	0.650	15.43	5.3	512	3	513	3
42	0.741	14.60	5.5	514	3	515	3
43	0.837	13.72	5.7	516	3	517	3
44	0.939	12.79	5.9	519	3	519	3
45	1.050	11.81	6.1	521	3	522	3
46	1.169	10.78	6.4	523	3	524	3
47	1.301	9.71	6.7	526	3	527	3
48	1.449	8.58	7.1	529	3	529	3
49	1.617	7.38	7.7	533	4	533	4
50	1.816	6.11	8.5	537	4	537	4
51	2.062	4.77	9.6	542	4	542	4
52	2.394	3.33	10.0	549	4	548	4
53	2.920	1.78	10.0	560	4	559	4
54	2.920	1.78	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.126	2.03	10.0	440	1	440	1
1	-3.112	2.07	10.0	440	1	440	1
2	-3.097	2.11	10.0	441	1	441	1
3	-3.082	2.14	10.0	441	1	441	1
4	-3.068	2.18	10.0	441	1	442	1
5	-3.053	2.22	10.0	441	1	442	1
6	-3.038	2.26	10.0	442	1	450	1
7	-2.662	3.52	10.0	449	1	455	1
8	-2.393	4.73	8.8	454	1	459	1
9	-2.179	5.88	7.9	458	1	462	1
10	-2.001	6.94	7.2	461	1	465	1
11	-1.845	7.93	6.8	464	1	468	1
12	-1.705	8.83	6.4	467	1	469	1
13	-1.578	9.66	6.1	469	1	472	2
14	-1.461	10.43	5.9	472	2	474	2
15	-1.351	11.13	5.7	474	2	476	2
16	-1.247	11.78	5.5	476	2	478	2
17	-1.148	12.36	5.4	478	2	479	2
18	-1.054	12.90	5.3	479	2	481	2
19	-0.963	13.39	5.2	481	2	482	2
20	-0.875	13.83	5.1	483	2	484	2
21	-0.790	14.24	5.0	484	2	485	2
22	-0.707	14.61	5.0	486	2	487	2
23	-0.626	14.95	4.9	488	2	488	2
24	-0.546	15.26	4.9	489	2	490	2
25	-0.468	15.55	4.8	491	2	491	2
26	-0.392	15.82	4.8	492	2	492	2
27	-0.316	16.06	4.8	494	2	494	2
28	-0.241	16.27	4.7	495	2	495	2
29	-0.167	16.45	4.7	496	2	497	2
30	-0.094	16.59	4.7	498	2	498	2
31	-0.020	16.70	4.7	499	2	499	2
32	0.053	16.76	4.7	501	3	501	3
33	0.126	16.78	4.6	502	3	502	3
34	0.200	16.76	4.7	503	3	504	3
35	0.274	16.69	4.7	505	3	505	3
36	0.348	16.57	4.7	506	3	506	3
37	0.424	16.41	4.7	508	3	508	3

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

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.501	16.20	4.7	509	3	509	3
39	0.580	15.94	4.8	511	3	511	3
40	0.661	15.62	4.8	512	3	513	3
41	0.745	15.22	4.9	514	3	514	3
42	0.831	14.75	5.0	515	3	516	3
43	0.922	14.18	5.1	517	3	518	3
44	1.018	13.50	5.2	519	3	519	3
45	1.120	12.68	5.3	521	3	521	3
46	1.231	11.72	5.6	523	3	523	3
47	1.353	10.61	5.8	525	3	526	3
48	1.490	9.35	6.2	528	3	528	3
49	1.648	7.93	6.8	531	4	531	4
50	1.838	6.38	7.5	535	4	535	4
51	2.081	4.74	8.7	539	4	540	4
52	2.424	3.05	10.0	546	4	546	4
53	3.017	1.42	10.0	557	4	557	4
54	3.176	1.16	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-3.028	2.43	10.0	440	1	440	1
1	-2.975	2.60	10.0	441	1	441	1
2	-2.923	2.78	10.0	442	1	441	1
3	-2.870	2.97	10.0	443	1	442	1
4	-2.817	3.18	10.0	444	1	443	1
5	-2.764	3.39	10.0	445	1	443	1
6	-2.712	3.62	10.0	446	1	451	1
7	-2.420	5.05	8.8	452	1	456	1
8	-2.195	6.33	7.9	457	1	460	1
9	-2.008	7.46	7.3	460	1	464	1
10	-1.846	8.46	6.8	463	1	467	1
11	-1.703	9.34	6.5	466	1	469	1
12	-1.572	10.13	6.2	469	1	472	2
13	-1.452	10.83	6.0	471	2	474	2
14	-1.340	11.48	5.9	474	2	476	2
15	-1.234	12.07	5.7	476	2	478	2
16	-1.134	12.62	5.6	478	2	480	2
17	-1.038	13.13	5.5	480	2	482	2
18	-0.947	13.62	5.4	481	2	484	2
19	-0.859	14.08	5.3	483	2	485	2
20	-0.774	14.51	5.2	485	2	487	2
21	-0.691	14.93	5.1	486	2	488	2
22	-0.611	15.32	5.1	488	2	490	2
23	-0.533	15.69	5.0	490	2	491	2
24	-0.457	16.04	5.0	491	2	493	2
25	-0.382	16.37	4.9	493	2	494	2
26	-0.308	16.67	4.9	494	2	496	2
27	-0.236	16.95	4.8	495	2	497	2
28	-0.164	17.21	4.8	497	2	499	2
29	-0.094	17.44	4.8	498	2	499	2
30	-0.023	17.63	4.7	499	2	501	3
31	0.047	17.80	4.7	501	3	503	3
32	0.117	17.92	4.7	502	3	504	3
33	0.187	18.00	4.7	504	3	506	3
34	0.257	18.04	4.7	505	3	507	3
35	0.328	18.02	4.7	507	3	508	3
36	0.400	17.94	4.7	508	3	510	3
37	0.473	17.80	4.7	510	3	511	3

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

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.547	17.59	4.7	511	3	513	3
39	0.623	17.30	4.8	513	3	514	3
40	0.701	16.94	4.8	514	3	516	3
41	0.781	16.48	4.9	516	3	517	3
42	0.865	15.93	5.0	517	3	519	3
43	0.953	15.28	5.1	519	3	521	3
44	1.046	14.52	5.2	521	3	522	3
45	1.145	13.64	5.4	523	3	524	3
46	1.252	12.63	5.6	525	3	526	3
47	1.369	11.48	5.9	527	3	529	3
48	1.500	10.18	6.2	529	3	531	4
49	1.651	8.71	6.7	533	4	534	4
50	1.832	7.08	7.5	537	4	537	4
51	2.063	5.28	8.6	541	4	542	4
52	2.390	3.36	10.0	548	4	548	4
53	2.974	1.46	10.0	559	4	559	4
54	3.011	1.38	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-2.859	2.49	10.0	440	1	440	1
1	-2.843	2.56	10.0	440	1	441	1
2	-2.827	2.62	10.0	441	1	441	1
3	-2.811	2.68	10.0	441	1	442	1
4	-2.795	2.74	10.0	441	1	443	1
5	-2.779	2.81	10.0	442	1	443	1
6	-2.762	2.88	10.0	442	1	451	1
7	-2.367	4.84	9.4	450	1	457	1
8	-2.095	6.36	8.2	456	1	461	1
9	-1.877	7.48	7.6	460	1	465	1
10	-1.688	8.32	7.2	464	1	468	1
11	-1.520	9.00	6.9	468	1	471	2
12	-1.366	9.61	6.7	471	2	473	2
13	-1.223	10.21	6.5	474	2	476	2
14	-1.090	10.83	6.3	477	2	478	2
15	-0.966	11.50	6.1	479	2	480	2
16	-0.850	12.20	5.9	482	2	482	2
17	-0.741	12.95	5.8	484	2	484	2
18	-0.637	13.72	5.6	486	2	485	2
19	-0.540	14.52	5.4	488	2	487	2
20	-0.447	15.33	5.3	490	2	489	2
21	-0.358	16.13	5.2	492	2	490	2
22	-0.273	16.93	5.0	494	2	492	2
23	-0.191	17.71	4.9	495	2	493	2
24	-0.112	18.47	4.8	497	2	495	2
25	-0.036	19.20	4.7	499	2	496	2
26	0.038	19.92	4.7	500	3	498	2
27	0.111	20.62	4.6	502	3	499	2
28	0.181	21.30	4.5	503	3	500	3
29	0.251	21.97	4.4	505	3	502	3
30	0.319	22.60	4.4	506	3	503	3
31	0.387	23.18	4.3	507	3	504	3
32	0.454	23.69	4.3	509	3	506	3
33	0.520	24.10	4.2	510	3	507	3
34	0.587	24.37	4.2	512	3	509	3
35	0.654	24.49	4.2	513	3	510	3
36	0.721	24.42	4.2	514	3	512	3
37	0.789	24.18	4.2	516	3	513	3

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

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.859	23.75	4.3	517	3	515	3
39	0.931	23.15	4.3	519	3	516	3
40	1.004	22.39	4.4	520	3	518	3
41	1.081	21.50	4.5	522	3	519	3
42	1.161	20.47	4.6	523	3	521	3
43	1.245	19.34	4.7	525	3	523	3
44	1.334	18.11	4.9	527	3	525	3
45	1.430	16.79	5.1	529	3	527	3
46	1.533	15.40	5.3	531	4	529	3
47	1.647	13.94	5.6	534	4	532	4
48	1.774	12.39	5.9	536	4	535	4
49	1.919	10.71	6.3	539	4	538	4
50	2.091	8.85	7.0	543	4	542	4
51	2.307	6.75	8.0	547	4	548	4
52	2.608	4.39	9.9	553	4	556	4
53	2.922	2.68	10.0	560	4	560	4
54	2.922	2.68	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)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
0	-2.983	2.09	10.0	440	1	440	1
1	-2.938	2.19	10.0	441	1	441	1
2	-2.892	2.29	10.0	442	1	441	1
3	-2.847	2.40	10.0	443	1	442	1
4	-2.802	2.52	10.0	444	1	443	1
5	-2.756	2.65	10.0	445	1	443	1
6	-2.711	2.78	10.0	445	1	444	1
7	-2.666	2.93	10.0	446	1	445	1
8	-2.374	4.14	9.9	452	1	451	1
9	-2.145	5.46	8.6	457	1	456	1
10	-1.955	6.78	7.7	461	1	460	1
11	-1.792	8.03	7.1	464	1	464	1
12	-1.648	9.17	6.7	467	1	467	1
13	-1.518	10.20	6.3	469	1	469	1
14	-1.399	11.12	6.1	472	2	472	2
15	-1.288	11.96	5.8	474	2	474	2
16	-1.184	12.74	5.7	476	2	477	2
17	-1.086	13.50	5.5	478	2	479	2
18	-0.992	14.25	5.3	480	2	480	2
19	-0.903	14.99	5.2	482	2	482	2
20	-0.817	15.73	5.1	484	2	484	2
21	-0.735	16.46	5.0	485	2	486	2
22	-0.656	17.16	4.9	487	2	487	2
23	-0.579	17.83	4.8	488	2	489	2
24	-0.504	18.44	4.7	490	2	490	2
25	-0.431	18.99	4.6	491	2	492	2
26	-0.359	19.48	4.6	493	2	493	2
27	-0.289	19.89	4.5	494	2	494	2
28	-0.220	20.23	4.5	496	2	496	2
29	-0.151	20.50	4.5	497	2	497	2
30	-0.083	20.71	4.4	498	2	499	2
31	-0.015	20.86	4.4	499	2	499	2
32	0.053	20.94	4.4	501	3	501	3
33	0.121	20.97	4.4	503	3	503	3
34	0.189	20.94	4.4	504	3	504	3
35	0.258	20.85	4.4	505	3	505	3
36	0.328	20.69	4.4	507	3	507	3
37	0.398	20.47	4.5	508	3	508	3

Raw Score	Theta	Information	SE (Scale Score)	2025 Scale Score	2025 Achievement Levels	2024 Scale Score	2024 Achievement Levels
38	0.471	20.17	4.5	510	3	510	3
39	0.545	19.79	4.5	511	3	511	3
40	0.621	19.31	4.6	513	3	513	3
41	0.700	18.72	4.7	514	3	514	3
42	0.782	18.01	4.8	516	3	516	3
43	0.869	17.16	4.9	518	3	517	3
44	0.960	16.16	5.0	520	3	519	3
45	1.059	15.01	5.2	522	3	521	3
46	1.166	13.70	5.4	524	3	523	3
47	1.284	12.27	5.8	526	3	526	3
48	1.416	10.74	6.2	529	3	528	3
49	1.569	9.12	6.7	532	4	531	4
50	1.752	7.43	7.4	536	4	535	4
51	1.981	5.64	8.5	540	4	539	4
52	2.298	3.71	10.0	547	4	545	4
53	2.850	1.66	10.0	558	4	554	4
54	2.966	1.40	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	594	0.00933	0.00933
441	NM	659	0.01035	0.01969
448	NM	984	0.01546	0.03515
453	NM	1342	0.02108	0.05623
458	NM	1679	0.02638	0.08261
462	NM	1820	0.02859	0.11120
465	NM	1827	0.02870	0.13991
468	NM	1896	0.02979	0.16970
471	PM	1846	0.02900	0.19870
474	PM	1747	0.02745	0.22615
476	PM	1817	0.02855	0.25469
479	PM	1801	0.02830	0.28299
481	PM	1810	0.02844	0.31143
483	PM	1871	0.02940	0.34082
485	PM	1842	0.02894	0.36976
487	PM	1881	0.02955	0.39931
489	PM	1891	0.02971	0.42902
492	PM	2036	0.03199	0.46101
494	PM	2059	0.03235	0.49336
496	PM	2156	0.03387	0.52724
498	PM	2194	0.03447	0.56171
500	ME	2212	0.03475	0.59646
502	ME	2321	0.03647	0.63292
505	ME	2303	0.03618	0.66911
507	ME	2429	0.03816	0.70727
510	ME	2394	0.03761	0.74488
512	ME	2367	0.03719	0.78207
515	ME	2267	0.03562	0.81769
518	ME	2198	0.03453	0.85222
521	ME	2061	0.03238	0.88460
524	ME	1948	0.03061	0.91521
528	ME	1509	0.02371	0.93891
532	EE	1311	0.02060	0.95951
536	EE	1008	0.01584	0.97535
541	EE	709	0.01114	0.98649
547	EE	418	0.00657	0.99306
554	EE	265	0.00416	0.99722
560	EE	177	0.00278	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	1242	0.01996	0.01996
446	NM	778	0.01250	0.03247
451	NM	1001	0.01609	0.04856
455	NM	1143	0.01837	0.06693
458	NM	1288	0.02070	0.08763
461	NM	1265	0.02033	0.10796
464	NM	1315	0.02114	0.12910
467	NM	1254	0.02016	0.14925
469	NM	1328	0.02134	0.17060
471	PM	1425	0.02290	0.19350
473	PM	1416	0.02276	0.21626
475	PM	1451	0.02332	0.23958
478	PM	1511	0.02429	0.26387
480	PM	1679	0.02699	0.29086
482	PM	1749	0.02811	0.31897
484	PM	1743	0.02802	0.34698
485	PM	1867	0.03001	0.37699
487	PM	1916	0.03080	0.40779
490	PM	2079	0.03342	0.44120
492	PM	2090	0.03359	0.47480
494	PM	2149	0.03454	0.50934
496	PM	2218	0.03565	0.54499
498	PM	2343	0.03766	0.58265
500	ME	2511	0.04036	0.62301
503	ME	2544	0.04089	0.66390
505	ME	2475	0.03978	0.70368
508	ME	2541	0.04084	0.74452
511	ME	2634	0.04234	0.78686
514	ME	2486	0.03996	0.82681
517	ME	2319	0.03727	0.86409
521	ME	2157	0.03467	0.89876
525	ME	1854	0.02980	0.92856
529	ME	1580	0.02540	0.95395
535	EE	1145	0.01840	0.97235
542	EE	783	0.01259	0.98494
550	EE	559	0.00898	0.99392
560	EE	378	0.00608	1.00000

Table 2.3.3
Cumulative Scale Score Distribution
English Language Arts Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	33	0.00052	0.00052
441	NM	1008	0.01583	0.01635
447	NM	621	0.00975	0.02611
452	NM	759	0.01192	0.03803
456	NM	879	0.01381	0.05183
459	NM	896	0.01407	0.06591
462	NM	1027	0.01613	0.08204
464	NM	1014	0.01593	0.09797
467	NM	1118	0.01756	0.11553
469	NM	1131	0.01777	0.13329
471	PM	1222	0.01919	0.15249
473	PM	1422	0.02234	0.17482
475	PM	1363	0.02141	0.19623
477	PM	1450	0.02278	0.21901
479	PM	1565	0.02458	0.24359
481	PM	1679	0.02637	0.26996
483	PM	1809	0.02841	0.29838
484	PM	1831	0.02876	0.32714
486	PM	2034	0.03195	0.35909
488	PM	2043	0.03209	0.39118
490	PM	2143	0.03366	0.42484
492	PM	2268	0.03562	0.46046
494	PM	2252	0.03537	0.49584
496	PM	2287	0.03592	0.53176
498	PM	2320	0.03644	0.56820
499	PM	2386	0.03748	0.60568
502	ME	2375	0.03731	0.64299
504	ME	2386	0.03748	0.68046
506	ME	2286	0.03591	0.71637
508	ME	2210	0.03471	0.75108
511	ME	2214	0.03478	0.78586
513	ME	2156	0.03387	0.81973
516	ME	1938	0.03044	0.85017
519	ME	1784	0.02802	0.87819
521	ME	1574	0.02472	0.90291
525	ME	1508	0.02369	0.92660
528	ME	1252	0.01967	0.94626
532	EE	1030	0.01618	0.96244
536	EE	832	0.01307	0.97551
541	EE	656	0.01030	0.98582

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

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
547	EE	474	0.00745	0.99326
555	EE	257	0.00404	0.99730
560	EE	172	0.00270	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	3713	0.05759	0.05759
444	NM	1039	0.01612	0.07371
448	NM	974	0.01511	0.08882
452	NM	963	0.01494	0.10375
455	NM	982	0.01523	0.11899
458	NM	1001	0.01553	0.13451
460	NM	1040	0.01613	0.15064
463	NM	1108	0.01719	0.16783
465	NM	1107	0.01717	0.18500
468	NM	1232	0.01911	0.20411
470	PM	1301	0.02018	0.22429
472	PM	1319	0.02046	0.24475
475	PM	1416	0.02196	0.26671
477	PM	1452	0.02252	0.28924
479	PM	1537	0.02384	0.31308
481	PM	1638	0.02541	0.33848
484	PM	1668	0.02587	0.36436
486	PM	1704	0.02643	0.39079
488	PM	1833	0.02843	0.41922
490	PM	1771	0.02747	0.44669
492	PM	1886	0.02925	0.47594
495	PM	1981	0.03073	0.50667
497	PM	2037	0.03160	0.53827
499	PM	2085	0.03234	0.57061
502	ME	2121	0.03290	0.60351
504	ME	2182	0.03385	0.63735
507	ME	2267	0.03516	0.67251
509	ME	2243	0.03479	0.70731
512	ME	2248	0.03487	0.74217
515	ME	2201	0.03414	0.77631
518	ME	2157	0.03346	0.80977
522	ME	2113	0.03277	0.84255
525	ME	2015	0.03125	0.87380
529	ME	1805	0.02800	0.90180
534	EE	1632	0.02531	0.92711
539	EE	1467	0.02275	0.94987
545	EE	1230	0.01908	0.96895
552	EE	919	0.01425	0.98320
560	EE	1083	0.01680	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	3076	0.04752	0.04752
445	NM	1026	0.01585	0.06336
449	NM	1066	0.01647	0.07983
453	NM	1133	0.01750	0.09733
456	NM	1107	0.01710	0.11443
459	NM	1119	0.01729	0.13172
462	NM	1101	0.01701	0.14872
465	NM	1169	0.01806	0.16678
467	NM	1270	0.01962	0.18640
469	NM	1277	0.01973	0.20613
472	PM	1335	0.02062	0.22675
474	PM	1414	0.02184	0.24859
476	PM	1451	0.02241	0.27100
478	PM	1467	0.02266	0.29367
480	PM	1480	0.02286	0.31653
482	PM	1663	0.02569	0.34222
484	PM	1634	0.02524	0.36746
486	PM	1627	0.02513	0.39259
488	PM	1833	0.02831	0.42090
490	PM	1882	0.02907	0.44997
492	PM	1936	0.02991	0.47988
494	PM	2017	0.03116	0.51104
496	PM	2121	0.03276	0.54380
499	PM	2105	0.03252	0.57632
501	ME	2265	0.03499	0.61130
503	ME	2330	0.03599	0.64730
506	ME	2416	0.03732	0.68462
508	ME	2475	0.03823	0.72285
511	ME	2426	0.03747	0.76032
514	ME	2339	0.03613	0.79645
517	ME	2261	0.03493	0.83138
520	ME	2215	0.03422	0.86559
524	ME	2066	0.03191	0.89751
528	ME	1851	0.02859	0.92610
533	EE	1677	0.02590	0.95201
539	EE	1323	0.02044	0.97244
546	EE	996	0.01539	0.98783
555	EE	529	0.00817	0.99600
560	EE	259	0.00400	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	19	0.00029	0.00029
441	NM	846	0.01292	0.01321
442	NM	3079	0.04702	0.06023
446	NM	962	0.01469	0.07492
449	NM	928	0.01417	0.08909
452	NM	986	0.01506	0.10415
455	NM	1037	0.01584	0.11999
458	NM	1060	0.01619	0.13617
461	NM	1093	0.01669	0.15287
463	NM	1109	0.01694	0.16980
466	NM	1158	0.01768	0.18749
468	NM	1208	0.01845	0.20593
470	PM	1218	0.01860	0.22453
472	PM	1324	0.02022	0.24475
475	PM	1373	0.02097	0.26572
477	PM	1375	0.02100	0.28672
479	PM	1427	0.02179	0.30851
481	PM	1506	0.02300	0.33151
483	PM	1585	0.02421	0.35572
485	PM	1557	0.02378	0.37949
487	PM	1767	0.02698	0.40648
490	PM	1726	0.02636	0.43284
492	PM	1821	0.02781	0.46065
494	PM	1895	0.02894	0.48958
496	PM	1959	0.02992	0.51950
498	PM	1985	0.03031	0.54982
500	ME	2089	0.03190	0.58172
503	ME	2161	0.03300	0.61472
505	ME	2178	0.03326	0.64798
508	ME	2187	0.03340	0.68138
510	ME	2272	0.03470	0.71607
513	ME	2191	0.03346	0.74953
516	ME	2289	0.03496	0.78449
520	ME	2266	0.03460	0.81910
523	ME	2316	0.03537	0.85446
528	ME	2273	0.03471	0.88918
532	EE	2120	0.03238	0.92155
538	EE	1875	0.02863	0.95018
546	EE	1637	0.02500	0.97518
557	EE	1047	0.01599	0.99117
560	EE	578	0.00883	1.00000

Table 2.3.7
Cumulative Scale Score Distribution
Mathematics Grade 3

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	7	0.00011	0.00011
441	NM	112	0.00176	0.00187
442	NM	225	0.00354	0.00541
443	NM	1261	0.01984	0.02525
444	NM	975	0.01534	0.04059
450	NM	1255	0.01974	0.06033
455	NM	1306	0.02054	0.08087
458	NM	1363	0.02144	0.10231
462	NM	1314	0.02067	0.12298
465	NM	1394	0.02193	0.14491
467	NM	1333	0.02097	0.16588
469	NM	1382	0.02174	0.18762
472	PM	1259	0.01980	0.20742
474	PM	1293	0.02034	0.22776
476	PM	1250	0.01966	0.24743
478	PM	1409	0.02216	0.26959
480	PM	1363	0.02144	0.29103
482	PM	1406	0.02212	0.31315
484	PM	1395	0.02194	0.33510
486	PM	1431	0.02251	0.35761
488	PM	1411	0.02220	0.37980
489	PM	1551	0.02440	0.40420
491	PM	1474	0.02319	0.42739
493	PM	1565	0.02462	0.45201
494	PM	1514	0.02382	0.47582
496	PM	1549	0.02437	0.50019
498	PM	1570	0.02470	0.52489
499	PM	1553	0.02443	0.54932
501	ME	1740	0.02737	0.57669
503	ME	1688	0.02655	0.60324
505	ME	1664	0.02618	0.62942
507	ME	1668	0.02624	0.65566
509	ME	1722	0.02709	0.68274
511	ME	1710	0.02690	0.70964
513	ME	1757	0.02764	0.73728
515	ME	1774	0.02791	0.76519
518	ME	1708	0.02687	0.79206
520	ME	1756	0.02762	0.81968
523	ME	1699	0.02673	0.84641
526	ME	1726	0.02715	0.87356

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

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
529	ME	1604	0.02523	0.89879
533	EE	1597	0.02512	0.92391
537	EE	1509	0.02374	0.94765
542	EE	1331	0.02094	0.96859
550	EE	1060	0.01667	0.98526
560	EE	937	0.01474	1.00000

Table 2.3.8
Cumulative Scale Score Distribution
Mathematics Grade 4

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	28	0.00045	0.00045
441	NM	231	0.00373	0.00418
442	NM	788	0.01272	0.01690
443	NM	585	0.00944	0.02635
448	NM	734	0.01185	0.03819
452	NM	861	0.01390	0.05209
455	NM	962	0.01553	0.06762
458	NM	1001	0.01616	0.08378
461	NM	1010	0.01630	0.10009
464	NM	1065	0.01719	0.11728
466	NM	1135	0.01832	0.13560
468	NM	1171	0.01890	0.15451
470	PM	1177	0.01900	0.17351
473	PM	1250	0.02018	0.19368
474	PM	1206	0.01947	0.21315
476	PM	1256	0.02028	0.23343
478	PM	1288	0.02079	0.25422
480	PM	1314	0.02121	0.27543
482	PM	1304	0.02105	0.29648
483	PM	1338	0.02160	0.31808
485	PM	1291	0.02084	0.33892
486	PM	1256	0.02028	0.35920
488	PM	1360	0.02195	0.38115
489	PM	1328	0.02144	0.40259
491	PM	1344	0.02170	0.42429
492	PM	1342	0.02166	0.44595
494	PM	1345	0.02171	0.46767
495	PM	1386	0.02237	0.49004
497	PM	1368	0.02208	0.51212
498	PM	1390	0.02244	0.53456
499	PM	1402	0.02263	0.55719
501	ME	1376	0.02221	0.57941
503	ME	1387	0.02239	0.60180
504	ME	1453	0.02346	0.62525
506	ME	1432	0.02312	0.64837
507	ME	1530	0.02470	0.67307
509	ME	1496	0.02415	0.69722
511	ME	1480	0.02389	0.72111
512	ME	1580	0.02551	0.74662
514	ME	1555	0.02510	0.77172

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
516	ME	1594	0.02573	0.79745
519	ME	1529	0.02468	0.82214
521	ME	1471	0.02375	0.84588
523	ME	1603	0.02588	0.87176
526	ME	1537	0.02481	0.89657
529	ME	1405	0.02268	0.91925
533	EE	1382	0.02231	0.94156
537	EE	1170	0.01889	0.96045
542	EE	1033	0.01668	0.97713
549	EE	749	0.01209	0.98922
560	EE	668	0.01078	1.00000

Table 2.3.9
Cumulative Scale Score Distribution
Mathematics Grade 5

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	12	0.00019	0.00019
441	NM	753	0.01188	0.01207
442	NM	599	0.00945	0.02152
449	NM	807	0.01273	0.03426
454	NM	979	0.01545	0.04971
458	NM	1062	0.01676	0.06647
461	NM	1185	0.01870	0.08517
464	NM	1218	0.01922	0.10439
467	NM	1310	0.02067	0.12506
469	NM	1415	0.02233	0.14739
472	PM	1376	0.02171	0.16910
474	PM	1397	0.02204	0.19114
476	PM	1464	0.02310	0.21425
478	PM	1490	0.02351	0.23776
479	PM	1538	0.02427	0.26203
481	PM	1565	0.02470	0.28672
483	PM	1635	0.02580	0.31252
484	PM	1616	0.02550	0.33803
486	PM	1672	0.02638	0.36441
488	PM	1613	0.02545	0.38986
489	PM	1635	0.02580	0.41566
491	PM	1572	0.02481	0.44047
492	PM	1625	0.02564	0.46611
494	PM	1567	0.02473	0.49084
495	PM	1527	0.02410	0.51494
496	PM	1566	0.02471	0.53965
498	PM	1491	0.02353	0.56318
499	PM	1518	0.02395	0.58713
501	ME	1480	0.02335	0.61048
502	ME	1508	0.02380	0.63428
503	ME	1375	0.02170	0.65598
505	ME	1429	0.02255	0.67853
506	ME	1422	0.02244	0.70097
508	ME	1372	0.02165	0.72262
509	ME	1255	0.01980	0.74242
511	ME	1347	0.02126	0.76368
512	ME	1326	0.02092	0.78460
514	ME	1259	0.01987	0.80447
515	ME	1243	0.01961	0.82408
517	ME	1225	0.01933	0.84341

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

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
519	ME	1219	0.01924	0.86265
521	ME	1175	0.01854	0.88119
523	ME	1076	0.01698	0.89817
525	ME	1090	0.01720	0.91537
528	ME	1076	0.01698	0.93235
531	EE	1013	0.01599	0.94834
535	EE	954	0.01505	0.96339
539	EE	836	0.01319	0.97658
546	EE	698	0.01101	0.98760
557	EE	537	0.00847	0.99607
560	EE	249	0.00393	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	10	0.00015	0.00020
442	NM	62	0.00096	0.00116
443	NM	139	0.00215	0.00331
444	NM	293	0.00454	0.00785
445	NM	508	0.00787	0.01572
446	NM	744	0.01152	0.02724
452	NM	1024	0.01586	0.04310
457	NM	1207	0.01869	0.06179
460	NM	1326	0.02053	0.08232
463	NM	1367	0.02117	0.10349
466	NM	1540	0.02385	0.12733
469	NM	1541	0.02386	0.15120
471	PM	1630	0.02524	0.17644
474	PM	1704	0.02639	0.20282
476	PM	1604	0.02484	0.22766
478	PM	1580	0.02447	0.25213
480	PM	1602	0.02481	0.27694
481	PM	1572	0.02434	0.30128
483	PM	1606	0.02487	0.32615
485	PM	1580	0.02447	0.35061
486	PM	1552	0.02403	0.37465
488	PM	1491	0.02309	0.39774
490	PM	1482	0.02295	0.42069
491	PM	1474	0.02283	0.44351
493	PM	1487	0.02303	0.46654
494	PM	1421	0.02200	0.48854
495	PM	1466	0.02270	0.51124
497	PM	1430	0.02214	0.53339
498	PM	1406	0.02177	0.55516
499	PM	1376	0.02131	0.57647
501	ME	1348	0.02087	0.59734
502	ME	1331	0.02061	0.61795
504	ME	1324	0.02050	0.63845
505	ME	1342	0.02078	0.65923
507	ME	1328	0.02056	0.67980
508	ME	1319	0.02042	0.70022
510	ME	1328	0.02056	0.72079
511	ME	1252	0.01939	0.74017
513	ME	1240	0.01920	0.75938

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

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
514	ME	1241	0.01922	0.77859
516	ME	1241	0.01922	0.79781
517	ME	1255	0.01943	0.81724
519	ME	1279	0.01981	0.83705
521	ME	1279	0.01981	0.85686
523	ME	1222	0.01892	0.87578
525	ME	1186	0.01837	0.89414
527	ME	1209	0.01872	0.91287
529	ME	1106	0.01713	0.92999
533	EE	1100	0.01703	0.94703
537	EE	985	0.01525	0.96228
541	EE	926	0.01434	0.97662
548	EE	702	0.01087	0.98749
559	EE	528	0.00818	0.99566
560	EE	280	0.00434	1.00000

Table 2.3.11
Cumulative Scale Score Distribution
Mathematics Grade 7

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	12	0.00019	0.00019
441	NM	518	0.00800	0.00819
442	NM	1683	0.02600	0.03419
450	NM	1405	0.02171	0.05590
456	NM	1743	0.02693	0.08283
460	NM	2036	0.03146	0.11428
464	NM	2276	0.03516	0.14945
468	NM	2475	0.03824	0.18768
471	PM	2449	0.03784	0.22552
474	PM	2339	0.03614	0.26166
477	PM	2338	0.03612	0.29778
479	PM	2220	0.03430	0.33208
482	PM	2202	0.03402	0.36610
484	PM	2123	0.03280	0.39890
486	PM	2022	0.03124	0.43014
488	PM	1867	0.02884	0.45898
490	PM	1835	0.02835	0.48733
492	PM	1638	0.02531	0.51264
494	PM	1542	0.02382	0.53646
495	PM	1439	0.02223	0.55869
497	PM	1443	0.02229	0.58099
499	PM	1356	0.02095	0.60194
500	ME	1265	0.01954	0.62148
502	ME	1199	0.01852	0.64001
503	ME	1168	0.01805	0.65805
505	ME	1139	0.01760	0.67565
506	ME	1149	0.01775	0.69340
507	ME	1103	0.01704	0.71044
509	ME	1047	0.01618	0.72662
510	ME	1005	0.01553	0.74214
512	ME	950	0.01468	0.75682
513	ME	989	0.01528	0.77210
514	ME	927	0.01432	0.78642
516	ME	913	0.01411	0.80053
517	ME	917	0.01417	0.81470
519	ME	928	0.01434	0.82903
520	ME	917	0.01417	0.84320
522	ME	945	0.01460	0.85780
523	ME	878	0.01356	0.87137
525	ME	876	0.01353	0.88490

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
527	ME	897	0.01386	0.89876
529	ME	859	0.01327	0.91203
531	EE	831	0.01284	0.92487
534	EE	833	0.01287	0.93774
536	EE	834	0.01289	0.95062
539	EE	826	0.01276	0.96338
543	EE	739	0.01142	0.97480
547	EE	613	0.00947	0.98427
553	EE	520	0.00803	0.99231
560	EE	498	0.00769	1.00000

Table 2.3.12
Cumulative Scale Score Distribution
Mathematics Grade 8

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
440	NM	3	0.00005	0.00005
441	NM	4	0.00006	0.00011
442	NM	22	0.00034	0.00044
443	NM	106	0.00162	0.00206
444	NM	245	0.00374	0.00579
445	NM	1174	0.01790	0.02369
446	NM	1154	0.01760	0.04129
452	NM	1319	0.02011	0.06140
457	NM	1602	0.02443	0.08583
461	NM	1676	0.02555	0.11138
464	NM	1801	0.02746	0.13884
467	NM	1873	0.02856	0.16740
469	NM	1850	0.02821	0.19561
472	PM	1830	0.02790	0.22351
474	PM	1747	0.02664	0.25015
476	PM	1706	0.02601	0.27616
478	PM	1718	0.02620	0.30236
480	PM	1685	0.02569	0.32805
482	PM	1668	0.02543	0.35348
484	PM	1587	0.02420	0.37768
485	PM	1560	0.02379	0.40146
487	PM	1545	0.02356	0.42502
488	PM	1431	0.02182	0.44684
490	PM	1466	0.02235	0.46919
491	PM	1384	0.02110	0.49030
493	PM	1384	0.02110	0.51140
494	PM	1317	0.02008	0.53148
496	PM	1312	0.02000	0.55148
497	PM	1306	0.01991	0.57140
498	PM	1244	0.01897	0.59036
499	PM	1294	0.01973	0.61009
501	ME	1316	0.02007	0.63016
503	ME	1257	0.01917	0.64933
504	ME	1236	0.01885	0.66817
505	ME	1200	0.01830	0.68647
507	ME	1202	0.01833	0.70480
508	ME	1206	0.01839	0.72318
510	ME	1193	0.01819	0.74137
511	ME	1196	0.01824	0.75961
513	ME	1211	0.01846	0.77807

Scale Score	Achievement Levels	N	Proportion	Cumulative Proportion
514	ME	1176	0.01793	0.79601
516	ME	1185	0.01807	0.81407
518	ME	1177	0.01795	0.83202
520	ME	1195	0.01822	0.85024
522	ME	1147	0.01749	0.86773
524	ME	1208	0.01842	0.88615
526	ME	1152	0.01756	0.90371
529	ME	1182	0.01802	0.92174
532	EE	1134	0.01729	0.93903
536	EE	1071	0.01633	0.95536
540	EE	1014	0.01546	0.97082
547	EE	886	0.01351	0.98433
558	EE	668	0.01019	0.99451
560	EE	360	0.00549	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	1.15500	1.27000	0.86296	0.91173	0.13326	False
IA00288	3	0.90000	0.95500	0.82059	0.85241	0.06703	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	1.31000	1.30000	0.73935	0.87970	-0.01353	False
IA00226	3	1.27500	1.45500	0.76308	0.82546	0.23589	False

Table 2.4.3
Rescore Analysis
Mathematics Grade 3

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA10020	3	1.84000	1.86000	0.89913	0.87993	0.02224	False
IA01081	3	0.86500	0.86500	0.94936	0.95989	0.00000	False

Table 2.4.4
Rescore Analysis
Mathematics Grade 4

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01057	4	2.14500	2.25000	1.15788	1.17661	0.09068	False
IA12525	4	1.81500	1.80000	1.17374	1.14742	-0.01278	False

Table 2.4.5
Rescore Analysis
Mathematics Grade 5

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA01159	4	2.21500	2.28000	1.32194	1.31539	0.04917	False
IA02736	4	2.14000	2.23000	1.41436	1.42028	0.06363	False

Table 2.4.6

Rescore Analysis

Mathematics Grade 6

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA02078	4	1.48000	1.55500	1.40337	1.38802	0.05344	False
IA00972	4	2.83000	2.82000	1.12133	1.11976	-0.00892	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	2.12000	2.14000	1.20117	1.16498	0.01665	False
IA02722	4	1.83000	1.77500	1.61684	1.57696	-0.03402	False

Table 2.4.8

Rescore Analysis

Mathematics Grade 8

Item Id	Max	Old Mean	New Mean	Old StDev	New StDev	Effect Size	Discard
IA12109	4	1.68500	1.76000	1.63067	1.62035	0.04599	False
IA00864	4	2.85500	2.81500	1.49168	1.53707	-0.02682	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)	0.77000	0.76000	10.04461	10.17479	1	False	-0.46571
IA00280 (EL308824)	0.63000	0.66000	11.67259	11.35015	1	False	0.23942
IA00281 (EL308826)	0.55000	0.54000	12.49735	12.59827	1	False	-0.30358
IA00282 (EL308827)	0.60000	0.57000	11.98661	12.29450	1	False	-0.66799
IA00283 (EL308835)	0.56000	0.55000	12.39612	12.49735	1	False	-0.33145
IA00284 (EL308837)	0.64000	0.63000	11.56616	11.67259	1	False	-0.56459
IA00285 (EL308838)	0.76000	0.74000	10.17479	10.42662	1	False	-0.28088
IA00286 (EL308842)	0.43000	0.42000	13.70550	13.80757	1	False	0.02002
IA00287 (EL308855)	0.39000	0.38000	14.11728	14.22192	3	False	0.12639
IA00288 (EL308857)	0.31667	0.33667	14.90816	14.68631	3	False	0.92988
IA00443 (EL626042844)	0.66000	0.62000	11.35015	11.77808	1	False	-0.27939
IA00444 (EL626043062)	0.65000	0.63000	11.45872	11.67259	1	False	-0.69565
IA00445 (EL626043435)	0.51500	0.25000	12.84957	15.69796	2	False	3.69225
IA00446 (EL626049849)	0.50000	0.50000	13.00000	13.00000	1	False	0.01436
IA00450 (EL626050679)	0.65000	0.62000	11.45872	11.77808	1	False	-0.50494
IA00451 (EL626050927)	0.54000	0.49000	12.59827	13.10028	1	False	-0.48195
IA00452 (EL626051097)	0.52000	0.51500	12.79939	12.84957	2	False	-0.13044
IA00453 (EL626051328)	0.65000	0.61000	11.45872	11.88272	1	False	-0.31575

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)	0.82000	0.80000	9.33854	9.63352	1	False	-0.67306
IA00219 (EL307709)	0.77000	0.75000	10.04461	10.30204	1	False	-0.63280
IA00220 (EL307710)	0.45000	0.44000	13.50265	13.60388	1	False	-0.29112
IA00221 (EL307713)	0.53000	0.52000	12.69892	12.79939	1	False	-0.56401
IA00222 (EL307714)	0.77000	0.75000	10.04461	10.30204	1	False	-0.63280
IA00223 (EL307719)	0.57000	0.54000	12.29450	12.59827	1	False	0.36174
IA00224 (EL307724)	0.73000	0.70000	10.54875	10.90240	1	False	0.03806
IA00225 (EL307728)	0.44333	0.44667	13.57009	13.53635	3	False	-0.75933
IA00226 (EL307729)	0.42000	0.44000	13.80757	13.60388	3	False	0.04825
IA00289 (EL309792)	0.62000	0.57000	11.77808	12.29450	1	False	1.29892
IA00407 (EL624647403)	0.42000	0.42000	13.80757	13.80757	1	False	-0.71746
IA00408 (EL624647580)	0.54000	0.61000	12.59827	11.88272	1	False	3.12427
IA00411 (EL624652450)	0.76000	0.74000	10.17479	10.42662	1	False	-0.61848
IA00412 (EL624652621)	0.89000	0.89000	8.09389	8.09389	1	False	0.89666
IA00414 (EL624652989)	0.44000	0.45000	13.60388	13.50265	1	False	-0.41840
IA00415 (EL624653348)	0.63000	0.62000	11.67259	11.77808	1	False	-0.85120
IA00416 (EL624653492)	0.71000	0.70500	10.78646	10.84466	2	False	-0.30789
IA00419 (EL624654711)	0.79500	0.75500	9.70443	10.23876	2	False	0.69865

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	0.76000	10.30204	10.17479	1	False	0.45633
IA00497 (EL626304969)	0.67000	0.69000	11.24035	11.01660	1	False	1.43816
IA00500 (EL626332335)	0.64000	0.66000	11.56616	11.35015	1	False	1.24258
IA00501 (EL626332592)	0.83000	0.83000	9.18334	9.18334	1	False	-0.87300
IA00502 (EL626333002)	0.81000	0.80000	9.48841	9.63352	1	False	-0.56102
IA00505 (EL626355215)	0.61000	0.61000	11.88272	11.88272	1	False	-1.66004
IA00506 (EL626355557)	0.65000	0.65000	11.45872	11.45872	1	False	-1.53642
IA00508 (EL626356291)	0.37500	0.36500	14.27456	14.38050	2	False	0.32504
IA00638 (EL627351056)	0.66000	0.67000	11.35015	11.24035	1	False	-0.07629
IA01669 (EL711809263)	0.74000	0.74000	10.42662	10.42662	1	False	-1.23549
IA01670 (EL711809592)	0.74000	0.72000	10.42662	10.66863	1	False	0.97337
IA01671 (EL711827203)	0.89000	0.88000	8.09389	8.30005	1	False	-0.17317
IA01672 (EL711827807)	0.68000	0.66000	11.12920	11.35015	1	False	0.90406
IA01679 (EL711868011)	0.46500	0.46000	13.35138	13.40173	2	False	-0.66733
IA01680 (EL711900602)	0.66000	0.64000	11.35015	11.56616	1	False	0.90440
IA01691 (EL712167015)	0.47000	0.49000	13.30108	13.10028	1	False	0.53882

Table 2.5.4

Delta Analysis

English Language Arts Grade 6

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00173 (EL303496)	0.74000	0.74000	10.42662	10.42662	1	False	-0.63502
IA00174 (EL303500)	0.64000	0.65000	11.56616	11.45872	1	False	-0.14429
IA00175 (EL303504)	0.61000	0.65000	11.88272	11.45872	1	False	2.07964
IA00176 (EL303508)	0.72000	0.73000	10.66863	10.54875	1	False	0.17596
IA00177 (EL303510)	0.86000	0.85000	8.67872	8.85427	1	False	-1.03467
IA00178 (EL303513)	0.68000	0.71000	11.12920	10.78646	1	False	1.68078
IA00179 (EL303514)	0.61000	0.61000	11.88272	11.88272	1	False	-1.00765
IA00180 (EL303518)	0.54000	0.53000	12.59827	12.69892	1	False	-0.57692
IA00515 (EL626864414)	0.82000	0.80000	9.33854	9.63352	1	False	0.00379
IA00517 (EL626864724)	0.70000	0.66000	10.90240	11.35015	1	False	1.51637
IA00518 (EL626865003)	0.68000	0.65000	11.12920	11.45872	1	False	0.71351
IA00520 (EL626865416)	0.42000	0.41000	13.80757	13.91018	1	False	-0.25324
IA00522 (EL626865773)	0.70000	0.68000	10.90240	11.12920	1	False	-0.09236
IA00523 (EL626865942)	0.65000	0.65000	11.45872	11.45872	1	False	-0.89914
IA00528 (EL626867605)	0.76500	0.75500	10.11008	10.23876	2	False	-1.00960
IA00530 (EL626868748)	0.71500	0.70000	10.72779	10.90240	2	False	-0.51715

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)	0.71000	0.71000	10.78646	10.78646	1	False	-1.12203
IA00066 (EL292163)	0.59000	0.62000	12.08982	11.77808	1	False	0.08470
IA00067 (EL292168)	0.48000	0.51000	13.20061	12.89972	1	False	-0.39729
IA00068 (EL292170)	0.69000	0.67000	11.01660	11.24035	1	False	0.48186
IA00069 (EL292172)	0.62000	0.64000	11.77808	11.56616	1	False	-0.47859
IA00070 (EL292176)	0.61000	0.59000	11.88272	12.08982	1	False	0.68715
IA00081 (EL293802)	0.67000	0.67000	11.24035	11.24035	1	False	-0.95519
IA00082 (EL293804)	0.54000	0.55000	12.59827	12.49735	1	False	-1.14126
IA00257 (EL308358)	0.82000	0.81000	9.33854	9.48841	1	False	-0.63656
IA00258 (EL308360)	0.74500	0.72500	10.36465	10.60896	2	False	0.38184
IA00262 (EL308382)	0.63000	0.66000	11.67259	11.35015	1	False	0.31069
IA00265 (EL308389)	0.90000	0.89000	7.87379	8.09389	1	False	-0.69816
IA00269 (EL308397)	0.80000	0.84000	9.63352	9.02217	1	False	3.02192
IA00655 (EL628647210)	0.72000	0.74000	10.66863	10.42662	1	False	0.13362
IA00657 (EL628647689)	0.75000	0.73000	10.30204	10.54875	1	False	0.37510
IA00658 (EL628653398)	0.75500	0.74000	10.23876	10.42662	2	False	-0.04778

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)	0.76000	0.77000	10.17479	10.04461	1	False	-0.03231
IA00057 (EL290798)	0.79000	0.78000	9.77432	9.91123	1	False	-0.62456
IA00058 (EL290799)	0.74000	0.74000	10.42662	10.42662	1	False	-1.38588
IA00059 (EL290800)	0.68000	0.67000	11.12920	11.24035	1	False	-1.10065
IA00060 (EL290801)	0.76000	0.74000	10.17479	10.42662	1	False	0.54394
IA00061 (EL290805)	0.56000	0.57000	12.39612	12.29450	1	False	-0.00831
IA00062 (EL290808)	0.54000	0.56000	12.59827	12.39612	1	False	1.09568
IA00063 (EL290814)	0.39000	0.37000	14.11728	14.32741	1	False	-0.48565
IA00368 (EL623873883)	0.73000	0.72000	10.54875	10.66863	1	False	-0.92122
IA00371 (EL623951471)	0.61500	0.60000	11.83050	11.98661	2	False	-0.72408
IA00373 (EL623952377)	0.43500	0.41500	13.65463	13.85881	2	False	-0.48084
IA00374 (EL623952612)	0.71000	0.69000	10.78646	11.01660	1	False	0.22155
IA00378 (EL623955555)	0.50000	0.53000	13.00000	12.69892	1	False	2.21231
IA00379 (EL623955757)	0.54000	0.50000	12.59827	13.00000	1	False	1.78655
IA00383 (EL623959265)	0.63000	0.64000	11.67259	11.56616	1	False	-0.06420
IA00699 (EL632808123)	0.76000	0.77000	10.17479	10.04461	1	False	-0.03231

Table 2.5.7
Delta Analysis
Mathematics Grade 3

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00834 (MA293457)	0.80000	0.74000	9.63352	10.42662	1	False	-0.01129
IA00838 (MA293524)	0.79000	0.78000	9.77432	9.91123	1	False	0.55036
IA00850 (MA297405)	0.71000	0.68000	10.78646	11.12920	1	False	-0.72480
IA00852 (MA297438)	0.65000	0.60000	11.45872	11.98661	1	False	-0.63709
IA00924 (MA306310)	0.49000	0.50000	13.10028	13.00000	1	False	0.55950
IA00930 (MA306359)	0.68000	0.64000	11.12920	11.56616	1	False	-1.16220
IA00932 (MA306375)	0.51000	0.47000	12.89972	13.30108	1	False	-0.75033
IA01019 (MA311277)	0.81000	0.77000	9.48841	10.04461	1	False	-1.14745
IA01081 (MA623656013)	0.32000	0.29000	14.87080	15.21354	3	False	-0.37904
IA02323 (MA301611A)	0.87000	0.80000	8.49444	9.63352	1	False	1.20879
IA04659 (MA307303)	0.43000	0.44000	13.70550	13.60388	1	False	0.36895
IA04760 (MA713752330)	0.80000	0.79000	9.63352	9.77432	1	False	0.57826
IA04813 (MA735572247)	0.72000	0.69000	10.66863	11.01660	1	False	-0.71050
IA04828 (MA735653938)	0.59000	0.50000	12.08982	13.00000	1	False	1.32533
IA04844 (MA735735757)	0.62000	0.59000	11.77808	12.08982	1	False	-0.90464
IA07855 (MA900579464)	0.47000	0.47000	13.30108	13.30108	1	False	0.03329
IA10020 (MA001639117)	0.58667	0.47000	12.12409	13.30108	3	False	2.56294
IA12584 (MA900373094)	0.60000	0.56000	11.98661	12.39612	1	False	-1.00966
IA14263 (MA231151939)	0.52000	0.50000	12.79939	13.00000	1	False	-0.72580
IA14291 (MA231376908)	0.61000	0.62000	11.88272	11.77808	1	False	0.97538

Table 2.5.8
Delta Analysis
Mathematics Grade 4

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00828 (MA287237)	0.81000	0.75000	9.48841	10.30204	1	False	-0.65533
IA00841 (MA293718)	0.74000	0.67000	10.42662	11.24035	1	False	-0.69909
IA00861 (MA297629)	0.87000	0.86000	8.49444	8.67872	1	False	-0.06084
IA00906 (MA301811)	0.74000	0.71000	10.42662	10.78646	1	False	-0.45816
IA00958 (MA307055)	0.50000	0.48000	13.00000	13.20061	1	False	0.10529
IA00961 (MA307081)	0.56000	0.54000	12.39612	12.59827	1	False	0.07269
IA00963 (MA307085)	0.69000	0.63000	11.01660	11.67259	1	False	-1.16528
IA01049 (MA311537)	0.68000	0.63000	11.12920	11.67259	1	False	-0.93539
IA01055 (MA311572)	0.61000	0.61000	11.88272	11.88272	1	False	0.61052
IA01057 (MA311581)	0.56250	0.45250	12.37076	13.47739	4	False	0.02388
IA01093 (MA623879088)	0.74500	0.57000	10.36465	12.29450	2	False	2.40650
IA02175 (MA286769)	0.78000	0.72000	9.91123	10.66863	1	False	-0.83146
IA02819 (MA713583365)	0.68000	0.49000	11.12920	13.10028	1	False	2.48518
IA02841 (MA713774890)	0.55000	0.47000	12.49735	13.30108	1	False	-0.82411
IA02902 (MA714251321)	0.45000	0.42000	13.50265	13.80757	1	False	-0.16110
IA04945 (MA800744502)	0.63000	0.60000	11.67259	11.98661	1	False	-0.27230
IA04965 (MA800867144)	0.63000	0.58000	11.67259	12.19243	1	False	-0.84443
IA12273 (MA200267672)	0.56000	0.57000	12.39612	12.29450	1	False	0.91711
IA12278 (MA200334941)	0.52000	0.49000	12.79939	13.10028	1	False	-0.18288
IA12525 (MA713956124)	0.49250	0.36750	13.07520	14.35393	4	False	0.46920

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)	0.78000	0.74000	9.91123	10.42662	1	False	-1.01130
IA00776 (MA221207)	0.67000	0.63000	11.24035	11.67259	1	False	-0.95737
IA00803 (MA262207)	0.73000	0.65000	10.54875	11.45872	1	False	0.57503
IA00806 (MA272292)	0.48000	0.48000	13.20061	13.20061	1	False	0.35540
IA00880 (MA298106)	0.24000	0.23000	15.82521	15.95539	1	False	-0.71067
IA00885 (MA299556)	0.75000	0.66000	10.30204	11.35015	1	False	1.07379
IA00943 (MA306466)	0.68000	0.66000	11.12920	11.35015	1	False	-0.09253
IA00989 (MA307638)	0.85000	0.83000	8.85427	9.18334	1	False	-0.04865
IA01020 (MA311280)	0.45000	0.41000	13.50265	13.91018	1	False	-0.81006
IA01029 (MA311337)	0.85000	0.82000	8.85427	9.33854	1	False	-0.66685
IA01149 (MA624347774)	0.43000	0.45000	13.70550	13.50265	1	False	1.05808
IA01155 (MA624357395)	0.37500	0.36000	14.27456	14.43384	2	False	-0.50309
IA01159 (MA624377498)	0.57500	0.44750	12.24353	13.52792	4	False	2.42001
IA02552 (MA311324)	0.43000	0.45000	13.70550	13.50265	1	False	1.05808
IA02725 (MA704359215)	0.57000	0.52000	12.29450	12.79939	1	False	-0.67434
IA02736 (MA704359678)	0.52500	0.42000	12.74917	13.80757	4	False	1.62531
IA02917 (MA715102107)	0.85000	0.83000	8.85427	9.18334	1	False	-0.04865
IA04970 (MA800974344)	0.74000	0.70000	10.42662	10.90240	1	False	-0.96104
IA05002 (MA801652356)	0.56000	0.53000	12.39612	12.69892	1	False	-0.68288
IA05015 (MA801763240)	0.26000	0.24000	15.57338	15.82521	1	False	-0.99828

Table 2.5.10
Delta Analysis
Mathematics Grade 6

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00778 (MA221669)	0.81000	0.75000	9.48841	10.30204	1	False	-1.26049
IA00804 (MA264305)	0.77000	0.72000	10.04461	10.66863	1	False	-0.76720
IA00817 (MA280989)	0.49000	0.48000	13.10028	13.20061	1	False	0.13545
IA00818 (MA282268)	0.41000	0.37000	13.91018	14.32741	1	False	-0.22350
IA00819 (MA282277)	0.45000	0.39000	13.50265	14.11728	1	False	0.72387
IA00827 (MA287186)	0.58000	0.53000	12.19243	12.69892	1	False	-1.04601
IA00845 (MA296349)	0.57000	0.54000	12.29450	12.59827	1	False	-0.52613
IA00972 (MA307339)	0.71500	0.69000	10.72779	11.01660	4	False	0.84885
IA01058 (MA311658)	0.44000	0.44000	13.60388	13.60388	1	False	0.37518
IA02037 (MA217493)	0.68000	0.63000	11.12920	11.67259	1	False	-1.12911
IA02078 (MA251350)	0.36250	0.34000	14.40714	14.64985	4	False	-0.94959
IA02597 (MA311693)	0.80000	0.71000	9.63352	10.78646	1	False	1.05781
IA02698 (MA703179529)	0.23000	0.22000	15.95539	16.08877	1	False	-0.39546
IA02906 (MA714281467)	0.64000	0.56000	11.56616	12.39612	1	False	0.54028
IA04745 (MA703231515)	0.53500	0.54000	12.64862	12.59827	2	False	1.48089
IA04884 (MA736365836)	0.67000	0.61000	11.24035	11.88272	1	False	-0.94167
IA04899 (MA736511626)	0.60000	0.55000	11.98661	12.49735	1	False	-1.18635
IA05126 (MA805103779)	0.57000	0.47000	12.29450	13.30108	1	False	2.27948
IA05134 (MA805167086)	0.31000	0.33000	14.98340	14.75965	1	False	0.70051
IA05135 (MA805171807)	0.59000	0.57000	12.08982	12.29450	1	False	0.28320

Table 2.5.11
Delta Analysis
Mathematics Grade 7

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
IA00831 (MA288414)	0.72000	0.73000	10.66863	10.54875	1	False	1.05131
IA00847 (MA296358)	0.53000	0.48000	12.69892	13.20061	1	False	-0.03849
IA00909 (MA301846)	0.88000	0.82000	8.30005	9.33854	1	False	0.73690
IA00948 (MA306600)	0.84000	0.81000	9.02217	9.48841	1	False	-0.70620
IA00949 (MA306605)	0.51000	0.44000	12.89972	13.60388	1	False	0.79063
IA01006 (MA311093)	0.86000	0.82000	8.67872	9.33854	1	False	-0.59620
IA01011 (MA311109)	0.47000	0.45000	13.30108	13.50265	1	False	-0.93269
IA01018 (MA311140)	0.38000	0.37000	14.22192	14.32741	1	False	-0.83244
IA01069 (MA316886)	0.52000	0.47250	12.79939	13.27595	4	False	-0.10522
IA01097 (MA623950280)	0.35000	0.31000	14.54128	14.98340	1	False	0.26620
IA02722 (MA703943185)	0.46750	0.46250	13.32622	13.37655	4	False	-0.36387
IA04486 (MA227988)	0.69000	0.65000	11.01660	11.45872	1	False	-0.75083
IA04538 (MA282218)	0.62000	0.62000	11.77808	11.77808	1	False	0.27453
IA04591 (MA298183)	0.52000	0.49000	12.79939	13.10028	1	False	-0.77436
IA04593 (MA298192)	0.19000	0.18000	16.51159	16.66146	1	False	-0.27844
IA04629 (MA303730)	0.50000	0.50000	13.00000	13.00000	1	False	-0.07805
IA07614 (MA306636)	0.47000	0.42000	13.30108	13.80757	1	False	0.15355
IA08208 (MA904265644)	0.42500	0.41000	13.75647	13.91018	2	False	-0.88180
IA10259 (MA005177715)	0.40000	0.37000	14.01339	14.32741	1	False	-0.37404
IA10268 (MA005205470)	0.36000	0.46000	14.43384	13.40173	1	False	3.43953

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)	0.64000	0.54000	11.56616	12.59827	1	False	1.18946
IA00858 (MA297513)	0.78000	0.75000	9.91123	10.30204	1	False	-0.54732
IA00864 (MA297652)	0.70250	0.65750	10.87358	11.37740	4	False	-1.54549
IA00865 (MA297656)	0.56000	0.54000	12.39612	12.59827	1	False	-0.76119
IA00905 (MA301702)	0.52000	0.46000	12.79939	13.40173	1	False	-0.35784
IA00979 (MA307472)	0.65000	0.64000	11.45872	11.56616	1	False	0.13545
IA00985 (MA307570)	0.54000	0.55000	12.59827	12.49735	1	False	0.63671
IA01033 (MA311384)	0.59000	0.57000	12.08982	12.29450	1	False	-0.63292
IA01037 (MA311414)	0.41000	0.33000	13.91018	14.75965	1	False	1.36833
IA01042 (MA311448)	0.51000	0.50000	12.89972	13.00000	1	False	-0.49153
IA01125 (MA624247061)	0.42500	0.42000	13.75647	13.80757	2	False	-0.64336
IA02495 (MA309741)	0.45000	0.37000	13.50265	14.32741	1	False	1.05950
IA04665 (MA307399)	0.57000	0.48000	12.29450	13.20061	1	False	0.90441
IA04678 (MA309738)	0.39000	0.39000	14.11728	14.11728	1	False	-0.55780
IA04957 (MA800770988)	0.39000	0.40000	14.11728	14.01339	1	False	-0.04675
IA05059 (MA803856627)	0.67000	0.68000	11.24035	11.12920	1	False	1.31111
IA05070 (MA804042487)	0.33000	0.34000	14.75965	14.64985	1	False	-0.31289
IA10142 (MA002243189)	0.84000	0.74000	9.02217	10.42662	1	False	1.85199
IA12109 (MA002234564)	0.42500	0.39000	13.75647	14.11728	4	False	-1.10622
IA14628 (MA311461)	0.64000	0.60000	11.56616	11.98661	1	False	-1.45366

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)	-1.13290	-0.83870	-0.22240	False
IA00280 (EL308824)	-0.63980	-0.34500	-0.62366	False
IA00281 (EL308826)	0.34970	0.42860	0.42261	False
IA00282 (EL308827)	-0.35740	-0.02300	-0.71217	False
IA00283 (EL308835)	-0.09790	0.13640	-0.28189	False
IA00284 (EL308837)	-0.43050	-0.11830	-0.72341	False
IA00285 (EL308838)	-0.87910	-0.49830	-0.24065	False
IA00286 (EL308842)	0.58720	0.76220	0.40682	False
IA00287 (EL308855)	0.45036	0.63330	0.27778	False
IA00288 (EL308857)	0.87636	0.89317	0.98807	False
IA00443 (EL626042844)	-0.88090	-0.39500	-0.00985	False
IA00444 (EL626043062)	-0.45480	-0.14290	-0.73739	False
IA00445 (EL626043435)	-0.29030	2.03130	3.51336	False
IA00446 (EL626049849)	0.18010	0.42980	-0.08853	False
IA00450 (EL626050679)	-0.36670	-0.06120	-0.65670	False
IA00451 (EL626050927)	0.25900	1.05760	-0.25820	False
IA00452 (EL626051097)	-0.27885	-0.07070	-0.37257	False
IA00453 (EL626051328)	-0.45490	-0.11730	-0.68123	False

Table 2.6.2

b/b Analysis

English Language Arts Grade 4

Item Id	Old b	New b	Std Dist	Flag
IA00218 (EL307705)	-1.76920	-1.31430	-0.35817	False
IA00219 (EL307709)	-1.60100	-0.79100	1.85036	False
IA00220 (EL307710)	0.30890	0.66170	-0.37199	False
IA00221 (EL307713)	0.06570	0.33910	-1.05180	False
IA00222 (EL307714)	-1.43870	-1.11800	0.63092	False
IA00223 (EL307719)	-0.10560	0.35620	0.21760	False
IA00224 (EL307724)	-1.01420	-0.47130	-0.22216	False
IA00225 (EL307728)	-0.03253	0.25973	-1.11505	False
IA00226 (EL307729)	0.18940	0.22353	1.39305	False
IA00289 (EL309792)	-0.69930	-0.13140	0.51309	False
IA00407 (EL624647403)	0.56330	0.96740	0.56336	False
IA00408 (EL624647580)	-0.37720	-0.32470	2.02076	False
IA00411 (EL624652450)	-1.17980	-0.70680	-1.23167	False
IA00412 (EL624652621)	-1.98540	-1.44380	-0.99326	False
IA00414 (EL624652989)	0.58700	0.72510	-0.32988	False
IA00415 (EL624653348)	-0.47880	-0.14000	-0.97262	False
IA00416 (EL624653492)	-1.23880	-0.83720	-0.54951	False
IA00419 (EL624654711)	-1.69335	-1.03900	0.00698	False

Table 2.6.3

b/b Analysis

English Language Arts Grade 5

Item Id	Old b	New b	Std Dist	Flag
IA00495 (EL626304658)	-1.51250	-1.04690	-0.53049	False
IA00497 (EL626304969)	-0.35250	-0.19000	0.92830	False
IA00500 (EL626332335)	-0.60290	-0.32620	-0.01680	False
IA00501 (EL626332592)	-2.07030	-1.44780	-0.73458	False
IA00502 (EL626333002)	-1.74360	-1.10210	0.20310	False
IA00505 (EL626355215)	-0.61680	-0.28540	-0.69415	False
IA00506 (EL626355557)	-0.63560	-0.25970	-0.88131	False
IA00508 (EL626356291)	0.36345	0.68125	0.48464	False
IA00638 (EL627351056)	-0.63730	-0.32110	-0.45431	False
IA01669 (EL711809263)	-1.17720	-0.76440	-0.55864	False
IA01670 (EL711809592)	-0.81030	-0.34000	-0.03166	False
IA01671 (EL711827203)	-1.94850	-1.36350	-0.96108	False
IA01672 (EL711827807)	-1.03620	-0.77250	1.06921	False
IA01679 (EL711868011)	-0.11275	0.12080	-0.49772	False
IA01680 (EL711900602)	-0.76040	-0.06550	2.97620	False
IA01691 (EL712167015)	0.44530	0.57210	-0.30070	False

Table 2.6.4

b/b Analysis

English Language Arts Grade 6

Item Id	Old b	New b	Std Dist	Flag
IA00173 (EL303496)	-1.24100	-0.62630	-0.42978	False
IA00174 (EL303500)	-1.06110	-0.64960	1.71487	False
IA00175 (EL303504)	-0.06260	0.01700	1.64768	False
IA00176 (EL303508)	-1.79060	-0.95550	-0.96001	False
IA00177 (EL303510)	-2.30090	-1.24810	-0.40034	False
IA00178 (EL303513)	-0.89730	-0.36690	-0.90329	False
IA00179 (EL303514)	-0.35670	-0.04010	-0.42672	False
IA00180 (EL303518)	0.36760	0.47530	-0.95458	False
IA00515 (EL626864414)	-2.27770	-1.21670	-0.15947	False
IA00517 (EL626864724)	-1.34900	-0.48480	1.58598	False
IA00518 (EL626865003)	-1.08720	-0.51070	-0.63449	False
IA00520 (EL626865416)	1.71230	1.47800	0.57482	False
IA00522 (EL626865773)	-0.85560	-0.18250	1.21168	False
IA00523 (EL626865942)	-0.82610	-0.36590	-0.20822	False
IA00528 (EL626867605)	-1.81020	-0.97285	-0.89455	False
IA00530 (EL626868748)	-1.64220	-0.83545	-0.76358	False

Table 2.6.5

b/b Analysis

English Language Arts Grade 7

Item Id	Old b	New b	Std Dist	Flag
IA00065 (EL292160)	-1.04420	-0.44410	-0.82624	False
IA00066 (EL292163)	-0.34240	-0.20090	1.38851	False
IA00067 (EL292168)	0.33170	0.61810	0.25078	False
IA00068 (EL292170)	-1.02080	-0.42550	-0.80083	False
IA00069 (EL292172)	-0.69460	-0.31940	-0.36825	False
IA00070 (EL292176)	-0.53550	-0.15890	-0.99195	False
IA00081 (EL293802)	-0.82360	-0.28150	-0.75563	False
IA00082 (EL293804)	-0.04860	0.01360	1.32169	False
IA00257 (EL308358)	-1.92860	-0.89080	1.60913	False
IA00258 (EL308360)	-1.77805	-0.92765	-0.30118	False
IA00262 (EL308382)	-0.60150	-0.07300	-0.09105	False
IA00265 (EL308389)	-2.75910	-1.77420	-0.59390	False
IA00269 (EL308397)	-2.18720	-1.54030	1.70912	False
IA00655 (EL628647210)	-1.12960	-0.57860	-1.04294	False
IA00657 (EL628647689)	-1.35630	-0.56820	0.47755	False
IA00658 (EL628653398)	-1.92730	-1.08565	-0.98480	False

Table 2.6.6

b/b Analysis

English Language Arts Grade 8

Item Id	Old b	New b	Std Dist	Flag
IA00056 (EL290795)	-1.98940	-1.33810	0.28013	False
IA00057 (EL290798)	-2.16830	-1.26490	0.13886	False
IA00058 (EL290799)	-0.90680	-0.33230	0.16809	False
IA00059 (EL290800)	-0.93510	-0.49910	-0.30099	False
IA00060 (EL290801)	-0.88170	-0.43390	-0.54257	False
IA00061 (EL290805)	-0.11560	0.12190	-0.46164	False
IA00062 (EL290808)	0.09220	0.27480	-0.45970	False
IA00063 (EL290814)	0.87100	0.89110	-0.63999	False
IA00368 (EL623873883)	-1.00760	-0.47480	-0.46945	False
IA00371 (EL623951471)	-0.90930	-0.43625	-0.71063	False
IA00373 (EL623952377)	0.30370	0.44835	-0.62525	False
IA00374 (EL623952612)	-1.05740	-0.54510	-0.71315	False
IA00378 (EL623955555)	0.13500	0.00990	2.31015	False
IA00379 (EL623955757)	-0.25470	0.40410	2.55949	False
IA00383 (EL623959265)	-0.71780	-0.23890	-0.26038	False
IA00699 (EL632808123)	-1.33820	-0.79910	-0.27299	False

Table 2.6.7

b/b Analysis

Mathematics Grade 3

Item Id	Old b	New b	Std Dist	Flag
IA00834 (MA293457)	-1.07532	-0.73350	-0.52335	False
IA00838 (MA293524)	-1.07391	-0.91440	0.54837	False
IA00850 (MA297405)	-0.77079	-0.59740	-0.46090	False
IA00852 (MA297438)	-0.34564	-0.13160	-0.48831	False
IA00924 (MA306310)	1.22125	0.94720	0.90811	False
IA00930 (MA306359)	-0.64887	-0.36670	-0.27961	False
IA00932 (MA306375)	0.62678	0.75600	0.82801	False
IA01019 (MA311277)	-1.34368	-0.97610	-0.85167	False
IA01081 (MA623656013)	0.77344	0.72563	-1.26629	False
IA02323 (MA301611A)	-1.66238	-1.20540	-0.36800	False
IA04659 (MA307303)	0.45256	0.43330	-0.84022	False
IA04760 (MA713752330)	-1.48337	-1.29760	1.24309	False
IA04813 (MA735572247)	-0.89674	-0.63420	-1.22792	False
IA04828 (MA735653938)	-0.28038	0.00930	0.80821	False
IA04844 (MA735735757)	-0.21413	0.06450	0.81959	False
IA07855 (MA900579464)	0.32890	0.12600	2.21535	False
IA10020 (MA001639117)	-0.39311	-0.06270	1.11459	False
IA12584 (MA900373094)	-0.09591	0.00280	-1.13955	False
IA14263 (MA231151939)	0.34550	0.39030	-1.17126	False
IA14291 (MA231376908)	-0.45740	-0.37980	0.13176	False

Table 2.6.8

b/b Analysis

Mathematics Grade 4

Item Id	Old b	New b	Std Dist	Flag
IA00828 (MA287237)	-1.01067	-0.90860	-0.84020	False
IA00841 (MA293718)	-0.19829	-0.04490	-0.60209	False
IA00861 (MA297629)	-1.15739	-1.56340	1.35096	False
IA00906 (MA301811)	0.08108	-0.42080	1.74166	False
IA00958 (MA307055)	0.38174	0.44520	-0.71155	False
IA00961 (MA307081)	0.30650	0.36230	-0.67692	False
IA00963 (MA307085)	-0.29072	-0.12640	-0.55659	False
IA01049 (MA311537)	-0.24776	-0.21810	-0.55288	False
IA01055 (MA311572)	-0.19054	-0.24840	-0.17513	False
IA01057 (MA311581)	-0.15074	0.19180	0.21766	False
IA01093 (MA623879088)	-0.84106	-0.33675	0.90440	False
IA02175 (MA286769)	-0.70687	-0.53810	-0.54549	False
IA02819 (MA713583365)	-0.86050	0.12690	2.99528	False
IA02841 (MA713774890)	-0.00494	0.21150	-0.32535	False
IA02902 (MA714251321)	0.49503	0.50130	-0.46621	False
IA04945 (MA800744502)	-0.27406	-0.20610	-0.71816	False
IA04965 (MA800867144)	-0.37026	-0.35130	-0.50415	False
IA12273 (MA200267672)	1.06302	0.93300	0.11263	False
IA12278 (MA200334941)	0.08674	0.04400	-0.24603	False
IA12525 (MA713956124)	0.24926	0.44688	-0.40185	False

Table 2.6.9

b/b Analysis

Mathematics Grade 5

Item Id	Old b	New b	Std Dist	Flag
IA00771 (MA204911)	-1.21101	-0.93240	-1.02442	False
IA00776 (MA221207)	-0.67351	-0.45660	-0.85872	False
IA00803 (MA262207)	-0.59990	-0.31220	-0.46718	False
IA00806 (MA272292)	0.30894	0.43820	-0.88844	False
IA00880 (MA298106)	1.10548	1.30510	0.43247	False
IA00885 (MA299556)	-0.69172	-0.19310	2.21357	False
IA00943 (MA306466)	-0.21607	-0.18190	1.00617	False
IA00989 (MA307638)	-1.35755	-1.09670	-0.61184	False
IA01020 (MA311280)	0.52346	0.72820	-0.20471	False
IA01029 (MA311337)	-1.63471	-1.33110	-0.84216	False
IA01149 (MA624347774)	0.36148	0.33670	1.08696	False
IA01155 (MA624357395)	0.47839	0.59760	-0.96065	False
IA01159 (MA624377498)	-0.25810	0.09415	0.80126	False
IA02552 (MA311324)	0.71760	0.66820	0.98152	False
IA02725 (MA704359215)	-0.29968	-0.06580	-0.81598	False
IA02736 (MA704359678)	-0.13145	0.24328	1.25216	False
IA02917 (MA715102107)	-1.92789	-1.59440	-0.88258	False
IA04970 (MA800974344)	-0.90575	-0.61970	-0.85947	False
IA05002 (MA801652356)	-0.18540	-0.16350	1.13144	False
IA05015 (MA801763240)	0.96557	1.00460	-0.48938	False

Table 2.6.10

b/b Analysis

Mathematics Grade 6

Item Id	Old b	New b	Std Dist	Flag
IA00778 (MA221669)	-1.04570	-0.66440	-0.59495	False
IA00804 (MA264305)	-1.63124	-1.02650	-0.14095	False
IA00817 (MA280989)	0.33658	0.48330	-0.66670	False
IA00818 (MA282268)	0.77507	1.01980	0.60534	False
IA00819 (MA282277)	0.21070	0.43140	-0.96359	False
IA00827 (MA287186)	-0.20649	0.02970	-0.60392	False
IA00845 (MA296349)	0.07174	0.41560	0.30887	False
IA00972 (MA307339)	-1.06280	-0.71965	-0.06650	False
IA01058 (MA311658)	0.54087	0.61850	-0.23496	False
IA02037 (MA217493)	-0.69872	0.02380	3.45478	False
IA02078 (MA251350)	0.41917	0.55955	-0.76963	False
IA02597 (MA311693)	-1.16455	-0.78920	-0.25312	False
IA02698 (MA703179529)	1.27740	1.35310	-0.44527	False
IA02906 (MA714281467)	-0.69116	-0.23450	0.05507	False
IA04745 (MA703231515)	-0.18692	-0.08215	1.04127	False
IA04884 (MA736365836)	-0.80895	-0.49040	-0.31716	False
IA04899 (MA736511626)	-0.16000	0.15540	-0.57430	False
IA05126 (MA805103779)	-0.34567	-0.05640	-0.97529	False
IA05134 (MA805167086)	0.93914	0.84770	1.04852	False
IA05135 (MA805171807)	-0.39930	-0.18380	0.09248	False

Table 2.6.11

b/b Analysis

Mathematics Grade 7

Item Id	Old b	New b	Std Dist	Flag
IA00831 (MA288414)	-0.66739	-0.23950	0.30658	False
IA00847 (MA296358)	-0.21084	0.05440	-0.91710	False
IA00909 (MA301846)	-1.54287	-1.02030	-0.85178	False
IA00948 (MA306600)	-1.28457	-0.79780	-0.64217	False
IA00949 (MA306605)	-0.16589	0.16380	0.23151	False
IA01006 (MA311093)	-1.91073	-1.28900	-0.34094	False
IA01011 (MA311109)	0.42841	0.44470	1.25810	False
IA01018 (MA311140)	0.31130	0.44650	-0.37191	False
IA01069 (MA316886)	-0.17776	0.08295	-0.94486	False
IA01097 (MA623950280)	0.42086	0.59520	-0.53602	False
IA02722 (MA703943185)	-0.07728	0.18883	-0.55851	False
IA04486 (MA227988)	-0.57992	-0.24720	-0.89187	False
IA04538 (MA282218)	-0.64530	-0.43300	1.33652	False
IA04591 (MA298183)	0.65452	0.86580	0.81827	False
IA04593 (MA298192)	1.24472	1.28520	-0.19799	False
IA04629 (MA303730)	0.72024	0.80870	-0.87073	False
IA07614 (MA306636)	0.48891	0.72070	0.64283	False
IA08208 (MA904265644)	0.25219	0.41160	-0.59304	False
IA10259 (MA005177715)	1.13439	1.21760	0.17345	False
IA10268 (MA005205470)	0.73389	0.59240	2.94964	False

Table 2.6.12

b/b Analysis

Mathematics Grade 8

Item Id	Old b	New b	Std Dist	Flag
IA00849 (MA296757)	-0.65826	-0.27640	-0.65842	False
IA00858 (MA297513)	-0.99524	-0.79380	0.78589	False
IA00864 (MA297652)	-0.90338	-0.60175	-0.61746	False
IA00865 (MA297656)	0.23254	0.24710	0.54521	False
IA00905 (MA301702)	0.01325	0.36170	0.30802	False
IA00979 (MA307472)	-0.00438	0.00290	1.11755	False
IA00985 (MA307570)	0.33655	0.54550	-0.72417	False
IA01033 (MA311384)	0.02285	0.18070	-0.76507	False
IA01037 (MA311414)	0.18301	0.38720	-1.09559	False
IA01042 (MA311448)	-0.04950	-0.03750	1.15245	False
IA01125 (MA624247061)	0.20402	0.32530	-0.69145	False
IA02495 (MA309741)	0.06002	0.40020	0.30322	False
IA04665 (MA307399)	0.08175	0.56300	2.05933	False
IA04678 (MA309738)	0.91355	0.99980	-1.03423	False
IA04957 (MA800770988)	0.30600	0.27420	0.95766	False
IA05059 (MA803856627)	-0.51002	-0.18220	-1.01129	False
IA05070 (MA804042487)	0.96363	1.02110	-1.28113	False
IA10142 (MA002243189)	-1.70405	-1.00660	1.03443	False
IA12109 (MA002234564)	0.05829	0.23858	-1.10969	False
IA14628 (MA311461)	-0.12220	0.28340	0.72476	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	0.77000	0.76000	ON03	ON03	0.01634	False
IA00280 (EL308824)	2	0.63000	0.66000	ON03	ON03	0.05297	True
IA00281 (EL308826)	2	0.55000	0.54000	ON03	ON03	0.01716	False
IA00282 (EL308827)	2	0.60000	0.57000	ON03	ON03	0.00201	False
IA00283 (EL308835)	2	0.56000	0.55000	ON03	ON03	0.02700	False
IA00284 (EL308837)	2	0.64000	0.63000	ON03	ON03	0.02227	False
IA00285 (EL308838)	2	0.76000	0.74000	ON03	ON03	0.01464	False
IA00286 (EL308842)	2	0.43000	0.42000	ON03	ON03	0.01527	False
IA00287 (EL308855)	4	1.17000	1.14000	ON03	ON03	0.02159	False
IA00288 (EL308857)	4	0.95000	1.01000	ON03	ON03	0.04763	False
IA00443 (EL626042844)	2	0.66000	0.62000	ON04	ON04	-0.00425	False
IA00444 (EL626043062)	2	0.65000	0.63000	ON04	ON04	0.01934	False
IA00445 (EL626043435)	3	1.03000	0.50000	ON04	ON04	-0.26062	True
IA00446 (EL626049849)	2	0.50000	0.50000	ON04	ON04	0.03783	False
IA00450 (EL626050679)	2	0.65000	0.62000	ON04	ON04	0.01211	False
IA00451 (EL626050927)	2	0.54000	0.49000	ON04	ON04	-0.02607	False
IA00452 (EL626051097)	3	1.04000	1.03000	ON04	ON04	0.03325	False
IA00453 (EL626051328)	2	0.65000	0.61000	ON04	ON04	0.00265	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	0.82000	0.80000	ON03	ON03	-0.00360	False
IA00219 (EL307709)	2	0.77000	0.75000	ON03	ON03	-0.01569	False
IA00220 (EL307710)	2	0.45000	0.44000	ON03	ON03	-0.01246	False
IA00221 (EL307713)	2	0.53000	0.52000	ON03	ON03	-0.00057	False
IA00222 (EL307714)	2	0.77000	0.75000	ON03	ON03	0.00085	False
IA00223 (EL307719)	2	0.57000	0.54000	ON03	ON03	-0.02669	False
IA00224 (EL307724)	2	0.73000	0.70000	ON03	ON03	-0.01806	False
IA00225 (EL307728)	4	1.33000	1.34000	ON03	ON03	0.01038	False
IA00226 (EL307729)	4	1.26000	1.32000	ON03	ON03	0.03814	False
IA00289 (EL309792)	2	0.62000	0.57000	ON03	ON03	-0.04151	False
IA00407 (EL624647403)	2	0.42000	0.42000	ON04	ON04	-0.01350	False
IA00408 (EL624647580)	2	0.54000	0.61000	ON04	ON04	0.07508	True
IA00411 (EL624652450)	2	0.76000	0.74000	ON04	ON04	-0.01330	False
IA00412 (EL624652621)	2	0.89000	0.89000	ON04	ON04	0.00221	False
IA00414 (EL624652989)	2	0.44000	0.45000	ON04	ON04	0.00334	False
IA00415 (EL624653348)	2	0.63000	0.62000	ON04	ON04	-0.01768	False
IA00416 (EL624653492)	3	1.42000	1.41000	ON04	ON04	-0.00061	False
IA00419 (EL624654711)	3	1.59000	1.51000	ON04	ON04	-0.02973	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	0.76000	ON03	ON03	0.00139	False
IA00497 (EL626304969)	2	0.67000	0.69000	ON03	ON03	0.01860	False
IA00500 (EL626332335)	2	0.64000	0.66000	ON03	ON03	0.01981	False
IA00501 (EL626332592)	2	0.83000	0.83000	ON03	ON03	0.00137	False
IA00502 (EL626333002)	2	0.81000	0.80000	ON03	ON03	-0.00915	False
IA00505 (EL626355215)	2	0.61000	0.61000	ON03	ON03	-0.00055	False
IA00506 (EL626355557)	2	0.65000	0.65000	ON03	ON03	-0.00321	False
IA00508 (EL626356291)	3	0.75000	0.73000	ON03	ON03	-0.02072	False
IA00638 (EL627351056)	2	0.66000	0.67000	ON03	ON03	0.01103	False
IA01669 (EL711809263)	2	0.74000	0.74000	ON04	ON04	0.00789	False
IA01670 (EL711809592)	2	0.74000	0.72000	ON04	ON04	-0.01363	False
IA01671 (EL711827203)	2	0.89000	0.88000	ON04	ON04	-0.00290	False
IA01672 (EL711827807)	2	0.68000	0.66000	ON04	ON04	-0.00867	False
IA01679 (EL711868011)	3	0.93000	0.92000	ON04	ON04	0.00149	False
IA01680 (EL711900602)	2	0.66000	0.64000	ON04	ON04	-0.01717	False
IA01691 (EL712167015)	2	0.47000	0.49000	ON04	ON04	0.02929	False

Table 2.7.4

Beta Analysis

English Language Arts Grade 6

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00173 (EL303496)	2	0.74000	0.74000	ON04	ON04	0.00344	False
IA00174 (EL303500)	2	0.64000	0.65000	ON04	ON04	0.00825	False
IA00175 (EL303504)	2	0.61000	0.65000	ON04	ON04	0.05249	True
IA00176 (EL303508)	2	0.72000	0.73000	ON04	ON04	0.01759	False
IA00177 (EL303510)	2	0.86000	0.85000	ON04	ON04	-0.00176	False
IA00178 (EL303513)	2	0.68000	0.71000	ON04	ON04	0.02383	False
IA00179 (EL303514)	2	0.61000	0.61000	ON04	ON04	0.01075	False
IA00180 (EL303518)	2	0.54000	0.53000	ON04	ON04	0.00072	False
IA00515 (EL626864414)	2	0.82000	0.80000	ON03	ON03	-0.01390	False
IA00517 (EL626864724)	2	0.70000	0.66000	ON03	ON03	-0.03596	False
IA00518 (EL626865003)	2	0.68000	0.65000	ON03	ON03	-0.02008	False
IA00520 (EL626865416)	2	0.42000	0.41000	ON03	ON03	-0.00877	False
IA00522 (EL626865773)	2	0.70000	0.68000	ON03	ON03	-0.01581	False
IA00523 (EL626865942)	2	0.65000	0.65000	ON03	ON03	0.00681	False
IA00528 (EL626867605)	3	1.53000	1.51000	ON03	ON03	-0.00240	False
IA00530 (EL626868748)	3	1.43000	1.40000	ON03	ON03	-0.00942	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	0.71000	0.71000	ON04	ON04	0.00047	False
IA00066 (EL292163)	2	0.59000	0.62000	ON04	ON04	0.02439	False
IA00067 (EL292168)	2	0.48000	0.51000	ON04	ON04	0.01915	False
IA00068 (EL292170)	2	0.69000	0.67000	ON04	ON04	-0.01960	False
IA00069 (EL292172)	2	0.62000	0.64000	ON04	ON04	0.01708	False
IA00070 (EL292176)	2	0.61000	0.59000	ON04	ON04	-0.02299	False
IA00081 (EL293802)	2	0.67000	0.67000	ON04	ON04	0.00248	False
IA00082 (EL293804)	2	0.54000	0.55000	ON04	ON04	0.01836	False
IA00257 (EL308358)	2	0.82000	0.81000	ON03	ON03	0.00410	False
IA00258 (EL308360)	3	1.49000	1.45000	ON03	ON03	-0.02245	False
IA00262 (EL308382)	2	0.63000	0.66000	ON03	ON03	0.02148	False
IA00265 (EL308389)	2	0.90000	0.89000	ON03	ON03	-0.00095	False
IA00269 (EL308397)	2	0.80000	0.84000	ON03	ON03	0.02983	False
IA00655 (EL628647210)	2	0.72000	0.74000	ON03	ON03	0.01670	False
IA00657 (EL628647689)	2	0.75000	0.73000	ON03	ON03	-0.02025	False
IA00658 (EL628653398)	3	1.51000	1.48000	ON03	ON03	-0.01867	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	0.76000	0.77000	ON04	ON04	0.02191	False
IA00057 (EL290798)	2	0.79000	0.78000	ON04	ON04	0.00140	False
IA00058 (EL290799)	2	0.74000	0.74000	ON04	ON04	0.00093	False
IA00059 (EL290800)	2	0.68000	0.67000	ON04	ON04	0.00495	False
IA00060 (EL290801)	2	0.76000	0.74000	ON04	ON04	-0.00320	False
IA00061 (EL290805)	2	0.56000	0.57000	ON04	ON04	0.01463	False
IA00062 (EL290808)	2	0.54000	0.56000	ON04	ON04	0.02092	False
IA00063 (EL290814)	2	0.39000	0.37000	ON04	ON04	-0.02239	False
IA00368 (EL623873883)	2	0.73000	0.72000	ON03	ON03	-0.00809	False
IA00371 (EL623951471)	3	1.23000	1.20000	ON03	ON03	-0.00987	False
IA00373 (EL623952377)	3	0.87000	0.83000	ON03	ON03	-0.01229	False
IA00374 (EL623952612)	2	0.71000	0.69000	ON03	ON03	-0.00806	False
IA00378 (EL623955555)	2	0.50000	0.53000	ON03	ON03	0.04093	False
IA00379 (EL623955757)	2	0.54000	0.50000	ON03	ON03	-0.04550	False
IA00383 (EL623959265)	2	0.63000	0.64000	ON03	ON03	0.01068	False
IA00699 (EL632808123)	2	0.76000	0.77000	ON03	ON03	0.01724	False

Table 2.7.7
Beta Analysis
Mathematics Grade 3

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00834 (MA293457)	2	0.80000	0.74000	ON09	ON09	0.00476	False
IA00838 (MA293524)	2	0.79000	0.78000	ON14	ON14	0.01841	False
IA00850 (MA297405)	2	0.71000	0.68000	ON17	ON06	0.01955	False
IA00852 (MA297438)	2	0.65000	0.60000	ON13	ON13	-0.00981	False
IA00924 (MA306310)	2	0.49000	0.50000	ON05	ON05	0.01717	False
IA00930 (MA306359)	2	0.68000	0.64000	ON03	ON03	-0.01184	False
IA00932 (MA306375)	2	0.51000	0.47000	ON12	ON12	-0.01668	False
IA01019 (MA311277)	2	0.81000	0.77000	ON14	ON14	-0.02712	False
IA01081 (MA623656013)	4	0.96000	0.87000	ON02	ON02	0.01107	False
IA02323 (MA301611A)	2	0.87000	0.80000	ON12	ON12	-0.01610	False
IA04659 (MA307303)	2	0.43000	0.44000	ON03	ON05	0.02724	False
IA04760 (MA713752330)	2	0.80000	0.79000	ON07	ON07	0.03403	False
IA04813 (MA735572247)	2	0.72000	0.69000	ON10	ON10	0.00576	False
IA04828 (MA735653938)	2	0.59000	0.50000	ON06	ON06	-0.04038	False
IA04844 (MA735735757)	2	0.62000	0.59000	ON15	ON15	-0.00853	False
IA07855 (MA900579464)	2	0.47000	0.47000	ON04	ON04	0.04494	False
IA10020 (MA001639117)	4	1.76000	1.41000	ON11	ON11	-0.03055	False
IA12584 (MA900373094)	2	0.60000	0.56000	ON10	ON10	-0.00554	False
IA14263 (MA231151939)	2	0.52000	0.50000	ON05	ON08	-0.01097	False
IA14291 (MA231376908)	2	0.61000	0.62000	ON15	ON08	0.04092	False

Table 2.7.8

Beta Analysis

Mathematics Grade 4

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00828 (MA287237)	2	0.81000	0.75000	ON11	ON06	0.01953	False
IA00841 (MA293718)	2	0.74000	0.67000	ON09	ON09	0.01688	False
IA00861 (MA297629)	2	0.87000	0.86000	ON04	ON09	0.04299	False
IA00906 (MA301811)	2	0.74000	0.71000	ON08	ON08	0.01130	False
IA00958 (MA307055)	2	0.50000	0.48000	ON05	ON05	0.02085	False
IA00961 (MA307081)	2	0.56000	0.54000	ON09	ON08	0.03716	False
IA00963 (MA307085)	2	0.69000	0.63000	ON12	ON12	0.02720	False
IA01049 (MA311537)	2	0.68000	0.63000	ON13	ON13	0.00650	False
IA01055 (MA311572)	2	0.61000	0.61000	ON14	ON14	0.07800	True
IA01057 (MA311581)	5	2.25000	1.81000	ON02	ON02	-0.03925	False
IA01093 (MA623879088)	3	1.49000	1.14000	ON03	ON03	-0.05288	True
IA02175 (MA286769)	2	0.78000	0.72000	ON06	ON06	0.03092	False
IA02819 (MA713583365)	2	0.68000	0.49000	ON06	ON11	-0.14572	True
IA02841 (MA713774890)	2	0.55000	0.47000	ON17	ON13	-0.02310	False
IA02902 (MA714251321)	2	0.45000	0.42000	ON16	ON12	0.03098	False
IA04945 (MA800744502)	2	0.63000	0.60000	ON04	ON04	0.04173	False
IA04965 (MA800867144)	2	0.63000	0.58000	ON15	ON15	0.05701	True
IA12273 (MA200267672)	2	0.56000	0.57000	ON05	ON05	0.02528	False
IA12278 (MA200334941)	2	0.52000	0.49000	ON07	ON07	0.04297	False
IA12525 (MA713956124)	5	1.97000	1.47000	ON10	ON10	-0.02098	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	0.78000	0.74000	ON05	ON05	0.01838	False
IA00776 (MA221207)	2	0.67000	0.63000	ON05	ON05	0.02124	False
IA00803 (MA262207)	2	0.73000	0.65000	ON06	ON06	-0.02608	False
IA00806 (MA272292)	2	0.48000	0.48000	ON09	ON09	0.02845	False
IA00880 (MA298106)	2	0.24000	0.23000	ON14	ON14	-0.00817	False
IA00885 (MA299556)	2	0.75000	0.66000	ON04	ON04	-0.03956	False
IA00943 (MA306466)	2	0.68000	0.66000	ON06	ON06	0.02394	False
IA00989 (MA307638)	2	0.85000	0.83000	ON07	ON07	0.01947	False
IA01020 (MA311280)	2	0.45000	0.41000	ON17	ON14	-0.02538	False
IA01029 (MA311337)	2	0.85000	0.82000	ON07	ON07	0.01338	False
IA01149 (MA624347774)	2	0.43000	0.45000	ON08	ON08	0.04331	False
IA01155 (MA624357395)	3	0.75000	0.72000	ON03	ON03	0.00966	False
IA01159 (MA624377498)	5	2.30000	1.79000	ON02	ON02	-0.02757	False
IA02552 (MA311324)	2	0.43000	0.45000	ON13	ON13	0.07321	True
IA02725 (MA704359215)	2	0.57000	0.52000	ON12	ON12	0.00484	False
IA02736 (MA704359678)	5	2.10000	1.68000	ON10	ON10	-0.04230	False
IA02917 (MA715102107)	2	0.85000	0.83000	ON15	ON15	0.00643	False
IA04970 (MA800974344)	2	0.74000	0.70000	ON08	ON04	0.00451	False
IA05002 (MA801652356)	2	0.56000	0.53000	ON11	ON11	0.07395	True
IA05015 (MA801763240)	2	0.26000	0.24000	ON16	ON13	0.03280	False

Table 2.7.10

Beta Analysis

Mathematics Grade 6

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00778 (MA221669)	2	0.81000	0.75000	ON11	ON11	0.00347	False
IA00804 (MA264305)	2	0.77000	0.72000	ON17	ON13	-0.01227	False
IA00817 (MA280989)	2	0.49000	0.48000	ON14	ON14	0.00271	False
IA00818 (MA282268)	2	0.41000	0.37000	ON05	ON05	-0.02352	False
IA00819 (MA282277)	2	0.45000	0.39000	ON11	ON11	-0.01643	False
IA00827 (MA287186)	2	0.58000	0.53000	ON09	ON09	-0.01033	False
IA00845 (MA296349)	2	0.57000	0.54000	ON16	ON09	-0.01833	False
IA00972 (MA307339)	5	2.86000	2.76000	ON02	ON02	0.00747	False
IA01058 (MA311658)	2	0.44000	0.44000	ON13	ON13	0.02044	False
IA02037 (MA217493)	2	0.68000	0.63000	ON14	ON14	-0.03023	False
IA02078 (MA251350)	5	1.45000	1.36000	ON03	ON03	0.00208	False
IA02597 (MA311693)	2	0.80000	0.71000	ON04	ON04	-0.00340	False
IA02698 (MA703179529)	2	0.23000	0.22000	ON15	ON15	0.00481	False
IA02906 (MA714281467)	2	0.64000	0.56000	ON10	ON10	-0.03986	False
IA04745 (MA703231515)	3	1.07000	1.08000	ON08	ON08	0.02766	False
IA04884 (MA736365836)	2	0.67000	0.61000	ON12	ON12	0.00006	False
IA04899 (MA736511626)	2	0.60000	0.55000	ON10	ON10	-0.00649	False
IA05126 (MA805103779)	2	0.57000	0.47000	ON07	ON07	-0.00854	False
IA05134 (MA805167086)	2	0.31000	0.33000	ON06	ON06	0.03579	False
IA05135 (MA805171807)	2	0.59000	0.57000	ON08	ON08	0.00547	False

Table 2.7.11

Beta Analysis

Mathematics Grade 7

Item Id	NumScoreCats	Old Mean	New Mean	Old Form	Form	Beta	Flag Beta
IA00831 (MA288414)	2	0.72000	0.73000	ON07	ON07	0.02060	False
IA00847 (MA296358)	2	0.53000	0.48000	ON09	ON09	-0.00883	False
IA00909 (MA301846)	2	0.88000	0.82000	ON13	ON12	-0.01391	False
IA00948 (MA306600)	2	0.84000	0.81000	ON11	ON11	-0.01140	False
IA00949 (MA306605)	2	0.51000	0.44000	ON05	ON05	-0.02918	False
IA01006 (MA311093)	2	0.86000	0.82000	ON05	ON05	-0.00753	False
IA01011 (MA311109)	2	0.47000	0.45000	ON12	ON12	0.02320	False
IA01018 (MA311140)	2	0.38000	0.37000	ON10	ON10	0.00691	False
IA01069 (MA316886)	5	2.08000	1.89000	ON03	ON03	-0.00286	False
IA01097 (MA623950280)	2	0.35000	0.31000	ON06	ON06	-0.01010	False
IA02722 (MA703943185)	5	1.87000	1.85000	ON02	ON02	-0.01191	False
IA04486 (MA227988)	2	0.69000	0.65000	ON09	ON09	-0.00297	False
IA04538 (MA282218)	2	0.62000	0.62000	ON08	ON08	0.03888	False
IA04591 (MA298183)	2	0.52000	0.49000	ON03	ON03	0.00082	False
IA04593 (MA298192)	2	0.19000	0.18000	ON10	ON10	-0.01431	False
IA04629 (MA303730)	2	0.50000	0.50000	ON07	ON07	-0.00126	False
IA07614 (MA306636)	2	0.47000	0.42000	ON08	ON08	-0.02985	False
IA08208 (MA904265644)	3	0.85000	0.82000	ON05	ON05	0.00628	False
IA10259 (MA005177715)	2	0.40000	0.37000	ON04	ON04	-0.01907	False
IA10268 (MA005205470)	2	0.36000	0.46000	ON14	ON02	0.10510	True

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	0.64000	0.54000	ON06	ON06	-0.04246	False
IA00858 (MA297513)	2	0.78000	0.75000	ON04	ON04	0.01167	False
IA00864 (MA297652)	5	2.81000	2.63000	ON03	ON03	0.00412	False
IA00865 (MA297656)	2	0.56000	0.54000	ON13	ON09	0.00960	False
IA00905 (MA301702)	2	0.52000	0.46000	ON07	ON07	-0.04014	False
IA00979 (MA307472)	2	0.65000	0.64000	ON10	ON10	0.03342	False
IA00985 (MA307570)	2	0.54000	0.55000	ON05	ON05	0.02812	False
IA01033 (MA311384)	2	0.59000	0.57000	ON08	ON08	0.00808	False
IA01037 (MA311414)	2	0.41000	0.33000	ON10	ON10	-0.01412	False
IA01042 (MA311448)	2	0.51000	0.50000	ON13	ON02	0.03387	False
IA01125 (MA624247061)	3	0.85000	0.84000	ON14	ON05	0.01039	False
IA02495 (MA309741)	2	0.45000	0.37000	ON08	ON08	-0.04428	False
IA04665 (MA307399)	2	0.57000	0.48000	ON06	ON06	-0.07028	True
IA04678 (MA309738)	2	0.39000	0.39000	ON04	ON04	0.01294	False
IA04957 (MA800770988)	2	0.39000	0.40000	ON11	ON11	0.05090	True
IA05059 (MA803856627)	2	0.67000	0.68000	ON09	ON09	0.02792	False
IA05070 (MA804042487)	2	0.33000	0.34000	ON07	ON07	0.01931	False
IA10142 (MA002243189)	2	0.84000	0.74000	ON12	ON12	-0.03395	False
IA12109 (MA002234564)	5	1.70000	1.56000	ON02	ON02	0.00102	False
IA14628 (MA311461)	2	0.64000	0.60000	ON10	ON03	-0.01942	False

Section 2.8

Final Item Parameters

Table 2.8.1
IRT Parameters for Dichotomous Items
English Language Arts Grade 3

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA13367 (EL305848619)	0.98654	0.01964	-1.15910	0.03238	0.26140	0.01420
IA13368 (EL305848846)	0.65420	0.01689	0.20530	0.02812	0.18570	0.00890
IA13369 (EL305849356)	0.38413	0.02444	1.79290	0.04886	0.22360	0.01300
IA13373 (EL305855968)	0.40065	0.00582	-1.92260	0.03929	0.00250	0.01110
IA13374 (EL305856828)	0.49489	0.01709	0.00510	0.05935	0.20010	0.01630
IA13375 (EL305857179)	0.57078	0.01148	-1.40280	0.05935	0.03130	0.02430
IA13376 (EL305857557)	0.62922	0.01811	0.39020	0.02962	0.22460	0.00890
IA13380 (EL305867369)	0.93327	0.01592	-0.41980	0.01832	0.12080	0.00730
IA13385 (EL305871854)	0.76872	0.01398	-0.75470	0.02869	0.08360	0.01190
IA13432 (EL306536295)	1.05985	0.01709	-1.22080	0.02339	0.08750	0.01250
IA13433 (EL306538230)	1.14056	0.02173	-0.77250	0.02132	0.26640	0.00890
IA13439 (EL306542014)	0.84838	0.01740	0.34860	0.01590	0.15100	0.00530
IA13441 (EL306542408)	0.70752	0.01423	-0.76810	0.03653	0.14390	0.01390
IA13443 (EL306542695)	0.61252	0.01194	0.12060	0.02143	0.01980	0.00740
IA13453 (EL306551827)	0.70035	0.01479	-1.22260	0.04886	0.11740	0.02070
IA13745 (EL311051383)	0.56837	0.01464	-1.13990	0.07121	0.14040	0.02530
IA13763 (EL311056839)	0.81146	0.01648	-0.58170	0.02846	0.18960	0.01070
IA13767 (EL311058588)	0.38907	0.00883	-0.34090	0.05462	0.00690	0.01520
IA13769 (EL311059365)	0.83798	0.01877	-0.76600	0.03365	0.24120	0.01270
IA13777 (EL311132236)	0.79500	0.01418	-1.73120	0.04598	0.04440	0.02550
IA13779 (EL311137156)	0.92698	0.01719	-0.47360	0.02097	0.16850	0.00830
IA13782 (EL311155968)	0.51434	0.00990	-1.36180	0.05935	0.01760	0.02250
IA13784 (EL311158115)	0.67084	0.01816	0.26200	0.02869	0.24620	0.00860
IA13787 (EL311160399)	0.65632	0.01449	-0.59330	0.03826	0.12730	0.01380

Table 2.8.2

IRT Parameters for Polytomous Items

English Language Arts Grade 3

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA13383 (EL305868518)	0.81093	0.00740	0.09185	0.00641	0.32575	0.01067	-0.32575	0.01151	0.00000	0.00000
IA13384 (EL305871312)	0.41481	0.00505	-1.21335	0.01311	0.76245	0.02622	-0.76245	0.01856	0.00000	0.00000
IA13389 (EL305885122)	0.75344	0.00643	0.51093	0.00983	2.15533	0.01559	-0.03667	0.01438	-2.11867	0.02666
IA13436 (EL306539868)	0.65467	0.00607	0.24170	0.00763	0.49740	0.01215	-0.49740	0.01421	0.00000	0.00000
IA13765 (EL311057786)	0.49489	0.00561	-1.29090	0.01274	0.78620	0.02573	-0.78620	0.01765	0.00000	0.00000
IA13771A (EL311063921#SCORE_TRAIT_CONV)	0.97848	0.01035	0.86343	0.00915	1.50473	0.01319	-0.05717	0.01462	-1.44757	0.02485
IA13771D (EL311063921#SCORE_TRAIT_IDEADEV)	1.00223	0.01070	1.52340	0.01377	2.03640	0.01651	0.60800	0.01782	-0.57410	0.02478
IA13774 (EL311078713)	0.66790	0.00612	-0.31010	0.00835	1.01540	0.01461	-1.01540	0.01433	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)	Parameters and Measures of Standard Error d5	Parameters and Measures of Standard Error SE(d5)
IA13389 (EL305885122)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13771A (EL311063921#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13771D (EL311063921#SCORE_TRAIT_IDEADEV)	-2.07030	0.05090	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA06357 (EL291009)	1.24650	0.02399	-1.07040	0.02094	0.24610	0.00990
IA06359 (EL291015)	0.36138	0.01093	-0.99650	0.11013	0.03360	0.02980
IA06908 (EL909549145)	0.90882	0.01865	0.23130	0.01550	0.16850	0.00530
IA06909 (EL909550937)	1.16226	0.02094	0.29610	0.01122	0.14760	0.00370
IA07261 (EL912955945)	0.83128	0.02104	0.54600	0.01805	0.23850	0.00550
IA07361 (EL916346163)	1.00106	0.01855	-1.60070	0.03471	0.15760	0.01910
IA11771 (EL208629059)	1.33110	0.02246	-0.61860	0.01400	0.20570	0.00610
IA11772 (EL208629200)	0.88560	0.01591	-0.67740	0.02267	0.14470	0.00930
IA11775 (EL208629667)	0.86073	0.01779	-0.33620	0.02337	0.24170	0.00820
IA11776 (EL208629781)	0.59600	0.00976	-1.27480	0.04037	0.01140	0.01650
IA11778 (EL208630729)	0.62252	0.01311	-0.96110	0.04465	0.07350	0.01720
IA11780 (EL208631197)	0.38777	0.00772	-1.32280	0.06918	0.01050	0.02050
IA11784 (EL208631698)	0.86567	0.01768	-1.47150	0.04118	0.16460	0.01990
IA11786 (EL208632061)	1.34644	0.02505	-1.68410	0.02522	0.13530	0.01660
IA11787 (EL208632192)	1.17596	0.01982	-1.39950	0.02186	0.07800	0.01260
IA13548 (EL307283777)	0.71899	0.01682	0.04770	0.02453	0.18020	0.00820
IA13551 (EL307365872)	1.16549	0.02002	-1.27920	0.02210	0.14400	0.01180
IA13553 (EL307370948)	0.50764	0.01194	-0.95580	0.06154	0.05060	0.02120
IA13556 (EL307385093)	1.09189	0.02043	-0.47680	0.01758	0.23610	0.00690
IA13557 (EL307400674)	1.14621	0.02048	-0.78960	0.01874	0.22620	0.00810
IA13558 (EL307470135)	1.07948	0.02145	-0.09530	0.01596	0.26060	0.00560
IA13563 (EL307500095)	0.56449	0.01301	-0.83140	0.05160	0.09630	0.01800
IA13574 (EL307581198)	1.08571	0.01865	-1.31740	0.02476	0.14540	0.01300
IA13578 (EL307583130)	1.00500	0.03283	0.63840	0.01932	0.42960	0.00480

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA06362 (EL291024)	0.72657	0.00681	0.18255	0.00694	0.37365	0.01129	-0.37365	0.01271	0.00000	0.00000
IA11779 (EL208630869)	0.95926	0.00910	-1.91175	0.01069	0.88095	0.02218	-0.88095	0.01391	0.00000	0.00000
IA11782A (EL208631509#SCORE_TRAIT_CONV)	0.89700	0.00929	0.19080	0.00823	1.65790	0.01522	-0.10700	0.01280	-1.55090	0.02043
IA11782D (EL208631509#SCORE_TRAIT_IDEADEV)	0.88618	0.00917	0.78530	0.00997	2.44750	0.01735	0.39420	0.01419	-0.85500	0.02031
IA11785 (EL208631896)	0.59430	0.00650	-2.12660	0.01632	1.06080	0.03452	-1.06080	0.02014	0.00000	0.00000
IA13554 (EL307375298)	0.66455	0.00574	0.31653	0.00994	2.23363	0.01736	-0.12627	0.01473	-2.10737	0.02582
IA13562 (EL307485646)	0.90388	0.00757	-0.55390	0.00635	0.61810	0.01142	-0.61810	0.01057	0.00000	0.00000
IA13579 (EL307584265)	0.51775	0.00584	-1.69415	0.01256	0.50955	0.02462	-0.50955	0.01847	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)	Parameters and Measures of Standard Error d5	Parameters and Measures of Standard Error SE(d5)
IA11782A (EL208631509#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11782D (EL208631509#SCORE_TRAIT_IDEADEV)	-1.98670	0.03275	0.00000	0.00000	n/a	n/a
IA13554 (EL307375298)	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA03728 (EL736467737)	0.48072	0.01082	-2.12400	0.09759	0.04810	0.03750
IA03730 (EL736469533)	0.84068	0.02083	0.52410	0.01788	0.22650	0.00560
IA03731 (EL736469872)	1.27043	0.02264	-0.94630	0.01811	0.23470	0.00850
IA03733 (EL736470482)	0.54780	0.01348	0.03380	0.03436	0.09480	0.01110
IA03734 (EL736471910)	0.59206	0.01469	-1.08390	0.06042	0.17830	0.02080
IA03735 (EL736472217)	0.40312	0.00840	-1.29290	0.07024	0.01260	0.02150
IA03737 (EL736473790)	0.69024	0.01413	-0.03750	0.02302	0.10880	0.00810
IA03738 (EL736474369)	0.39871	0.01263	-1.00050	0.10857	0.07380	0.03070
IA04458 (EL827627427)	0.60024	0.01876	1.15080	0.02326	0.17520	0.00630
IA11592 (EL207164820)	0.68901	0.01293	-0.32880	0.02548	0.11120	0.00920
IA11595 (EL207274050)	0.79830	0.01534	-0.00830	0.01858	0.12360	0.00680
IA11599 (EL207278877)	1.12669	0.01987	-2.20750	0.03565	0.03630	0.02600
IA11600 (EL207341702)	0.30094	0.00775	-1.27290	0.10869	0.01590	0.02580
IA11602 (EL207342693)	0.32816	0.00760	-1.19270	0.08812	0.01230	0.02250
IA11603 (EL207360230)	0.93439	0.01690	-1.22220	0.02828	0.13900	0.01350
IA13509 (EL307061151)	0.80382	0.01504	-0.87810	0.02840	0.12060	0.01190
IA13510 (EL307063923)	1.18195	0.02314	-0.33740	0.01589	0.25360	0.00630
IA13511 (EL307077265)	0.71828	0.01675	-0.26490	0.03085	0.24000	0.01010
IA13512 (EL307137436)	0.57437	0.02249	1.64520	0.02863	0.17820	0.00600
IA13513 (EL307138100)	0.97255	0.01590	-2.39590	0.04126	0.02880	0.02940
IA13526 (EL307153223)	1.14233	0.02274	-0.68240	0.02080	0.32990	0.00810
IA13531 (EL307175118)	1.30547	0.02223	-1.61210	0.02209	0.11660	0.01340
IA13540 (EL307249807)	0.99318	0.01751	-1.32270	0.02723	0.14170	0.01360
IA13543 (EL307253066)	1.28924	0.02259	-1.50810	0.02244	0.17060	0.01290

Table 2.8.6

IRT Parameters for Polytomous Items

English Language Arts Grade 5

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA03740 (EL736475157)	0.34115	0.00458	0.87625	0.01843	1.37505	0.02411	-1.37505	0.03817	0.00000	0.00000
IA03742 (EL736477190)	0.57543	0.00584	0.38045	0.00866	0.57815	0.01336	-0.57815	0.01648	0.00000	0.00000
IA03744A (EL736478825#SCORE_TRAIT_CONV)	1.00106	0.01035	0.01157	0.00759	1.63077	0.01462	-0.04643	0.01200	-1.58433	0.01825
IA03744D (EL736478825#SCORE_TRAIT_IDEADEV)	1.09971	0.01111	0.55545	0.00746	1.89385	0.01329	0.61565	0.01116	-0.72135	0.01533
IA11597 (EL207277420)	0.47378	0.00558	0.26210	0.00913	0.32450	0.01483	-0.32450	0.01673	0.00000	0.00000
IA13529 (EL307156922)	0.70935	0.00659	-1.62985	0.01051	1.01665	0.02155	-1.01665	0.01406	0.00000	0.00000
IA13532 (EL307176181)	0.77648	0.00755	-0.49040	0.00636	0.12090	0.01102	-0.12090	0.01102	0.00000	0.00000
IA13544A (EL307267042#SCORE_TRAIT_CONV)	1.09824	0.01105	-0.50800	0.00633	1.28490	0.01491	-0.05460	0.01052	-1.23030	0.01218
IA13544D (EL307267042#SCORE_TRAIT_IDEADEV)	1.06490	0.01052	0.09178	0.00741	2.23478	0.01727	0.53418	0.01128	-0.66333	0.01285

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)	Parameters and Measures of Standard Error d5	Parameters and Measures of Standard Error SE(d5)
IA03744A (EL736478825#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA03744D (EL736478825#SCORE_TRAIT_IDEADEV)	-1.78815	0.02399	0.00000	0.00000	n/a	n/a
IA13544A (EL307267042#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13544D (EL307267042#SCORE_TRAIT_IDEADEV)	-2.10563	0.02255	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA13606 (EL308558989)	0.38771	0.00719	-1.26270	0.06123	0.01080	0.01660
IA13607 (EL308559785)	0.41634	0.00554	-1.77530	0.03548	0.00200	0.00950
IA13610 (EL308576679)	0.69259	0.01366	-0.15360	0.02455	0.20340	0.00690
IA13613 (EL308630437)	0.38107	0.01437	1.09430	0.04926	0.21090	0.01070
IA13616 (EL308672079)	0.76132	0.01198	-1.66140	0.03099	0.05180	0.01260
IA13858 (EL312142482)	0.47360	0.01052	-0.79090	0.05255	0.10130	0.01450
IA13875 (EL312978480)	0.93145	0.01798	-0.58880	0.02156	0.27050	0.00670
IA13876 (EL312980576)	0.58789	0.01205	-0.23730	0.03039	0.15380	0.00850
IA13895 (EL314063698)	0.70852	0.01433	-0.05500	0.02351	0.21530	0.00650
IA13898 (EL314148707)	0.42634	0.01131	-1.06540	0.08085	0.15440	0.02040
IA13899 (EL314158725)	0.20417	0.01115	-0.16660	0.26530	0.08530	0.03940
IA13900 (EL314164847)	0.67078	0.01280	-0.76650	0.03144	0.17510	0.00960
IA13913 (EL314429542)	0.49871	0.01366	0.02230	0.04551	0.26910	0.01030
IA13914 (EL314478387)	0.62863	0.01272	-1.11240	0.04252	0.21650	0.01280
IA13918 (EL314482379)	0.40764	0.01209	0.97550	0.03638	0.11810	0.00880
IA13928 (EL315274742)	0.51393	0.01633	0.88080	0.03354	0.27330	0.00750
IA13930 (EL315278077)	0.77102	0.01716	-1.32150	0.04087	0.32590	0.01240
IA13931 (EL315281344)	0.26937	0.00883	-2.57090	0.24209	0.09340	0.05090
IA13933 (EL315301385)	0.58219	0.01096	-0.75420	0.03473	0.10420	0.01060
IA13938 (EL315381843)	0.89730	0.01626	-1.17970	0.02605	0.21380	0.00920
IA13940 (EL315462998)	0.60517	0.01123	-0.69210	0.03309	0.14750	0.00980
IA13942 (EL315482249)	0.55191	0.01323	-0.32590	0.04252	0.27720	0.01020
IA13944 (EL315486389)	0.93616	0.01567	-0.24270	0.01557	0.14560	0.00480
IA13957 (EL319246845)	0.62069	0.01355	0.74250	0.02036	0.15050	0.00510

Table 2.8.8

IRT Parameters for Polytomous Items

English Language Arts Grade 6

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA13601 (EL308553441)	0.75338	0.00691	-1.98520	0.01013	0.61010	0.01959	-0.61010	0.01524	0.00000	0.00000
IA13897 (EL314140609)	0.59412	0.00526	-1.32930	0.00982	0.69590	0.01879	-0.69590	0.01503	0.00000	0.00000
IA13907 (EL314379145)	0.40617	0.00408	-0.22540	0.01382	1.62910	0.02375	-1.62910	0.02411	0.00000	0.00000
IA13921A (EL314484549#SCORE_TRAIT_CONV)	0.91352	0.00899	-0.68270	0.00749	1.49650	0.01758	-0.07780	0.01282	-1.41870	0.01415
IA13921D (EL314484549#SCORE_TRAIT_IDEADEV)	0.90700	0.00841	0.72282	0.00984	2.89092	0.01878	1.33592	0.01453	0.01462	0.01521
IA13929 (EL315276310)	0.84051	0.00844	-2.14750	0.01008	0.22070	0.01830	-0.22070	0.01659	0.00000	0.00000
IA13932 (EL315300123)	0.53028	0.00479	-1.76185	0.01216	1.16555	0.02449	-1.16555	0.01697	0.00000	0.00000
IA13946A (EL315605273#SCORE_TRAIT_CONV)	0.94591	0.00929	-0.84203	0.00741	1.41007	0.01800	-0.08173	0.01277	-1.32833	0.01310
IA13946D (EL315605273#SCORE_TRAIT_IDEADEV)	0.89212	0.00841	0.52648	0.00930	2.83768	0.01946	1.27768	0.01433	0.06578	0.01433

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)	Parameters and Measures of Standard Error d5	Parameters and Measures of Standard Error SE(d5)
IA13921A (EL314484549#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13921D (EL314484549#SCORE_TRAIT_IDEADEV)	-1.22608	0.02025	-3.01538	0.04119	0.00000	0.00000
IA13946A (EL315605273#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13946D (EL315605273#SCORE_TRAIT_IDEADEV)	-1.13232	0.01868	-3.04882	0.03815	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA07202 (EL912338456)	0.59735	0.01868	1.03960	0.02475	0.24390	0.00600
IA07203 (EL912344338)	0.53233	0.01323	-1.16680	0.06474	0.17810	0.01940
IA07230 (EL912500351)	0.85485	0.01885	-0.31640	0.02434	0.33010	0.00700
IA07232 (EL912501540)	1.14691	0.02158	-1.18470	0.02220	0.26320	0.00890
IA07241 (EL912582905)	0.99706	0.01956	-1.25320	0.02782	0.27520	0.01060
IA07246 (EL912620913)	0.41046	0.01310	-0.13420	0.06648	0.13610	0.01650
IA07247 (EL912621697)	0.94286	0.01639	-0.94330	0.02234	0.19270	0.00840
IA07248 (EL912622150)	0.93563	0.01903	-0.98770	0.02796	0.32140	0.00940
IA07249 (EL912623426)	0.74750	0.01951	0.68740	0.01993	0.25710	0.00520
IA13233 (EL206078126)	0.86825	0.01644	-0.35790	0.02047	0.22370	0.00660
IA13234 (EL206078710)	1.06326	0.02426	0.45090	0.01471	0.27010	0.00410
IA13235 (EL206079400)	0.91964	0.01670	-0.19880	0.01685	0.18080	0.00550
IA13269 (EL213939721)	0.63980	0.01156	-0.85380	0.03210	0.09070	0.01090
IA13347 (EL305230669)	0.77795	0.01600	-0.51230	0.02675	0.25020	0.00830
IA13348 (EL305244758)	0.89224	0.01613	-0.33120	0.01899	0.21110	0.00620
IA13398 (EL306347661)	0.79694	0.01705	-1.27820	0.03866	0.26370	0.01350
IA13400 (EL306348894)	0.70035	0.01798	0.42150	0.02341	0.27300	0.00620
IA13417 (EL306443148)	1.12651	0.01978	-1.38990	0.02287	0.19620	0.01000
IA13464 (EL306569606)	0.55420	0.01305	-1.43610	0.06661	0.16460	0.02150
IA13465 (EL306571635)	0.55432	0.01006	-1.94600	0.06086	0.03090	0.02410
IA13472 (EL306633578)	0.54068	0.00782	-1.97300	0.04615	0.00950	0.01790
IA13477 (EL306636364)	0.64397	0.01042	-1.67780	0.04120	0.02320	0.01700
IA13478 (EL306637041)	0.95985	0.01692	-0.80160	0.02060	0.18710	0.00760
IA13479 (EL306637866)	0.99318	0.01767	-0.52020	0.01779	0.19810	0.00620

Table 2.8.10

IRT Parameters for Polytomous Items

English Language Arts Grade 7

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA07205A (EL912346409#SCORE_TRAIT_CONV)	0.89735	0.00905	-1.00630	0.00817	1.63910	0.02050	-0.09190	0.01354	-1.54720	0.01402
IA07205D (EL912346409#SCORE_TRAIT_IDEADEV)	0.93815	0.00899	0.66504	0.01952	3.49014	0.02766	1.53844	0.02212	0.37334	0.02207
IA07242 (EL912583971)	0.32863	0.00426	-1.03270	0.02507	2.77230	0.04686	-2.77230	0.03969	0.00000	0.00000
IA07398 (EL917655389)	0.48536	0.00532	-0.51125	0.00856	0.12935	0.01477	-0.12935	0.01488	0.00000	0.00000
IA13350 (EL305250618)	0.63392	0.00554	-0.53220	0.01004	1.49270	0.01792	-1.49270	0.01683	0.00000	0.00000
IA13411A (EL306400886#SCORE_TRAIT_CONV)	0.98718	0.00982	-1.02027	0.00748	1.38703	0.01885	-0.07607	0.01281	-1.31097	0.01233
IA13411D (EL306400886#SCORE_TRAIT_IDEADEV)	0.99124	0.00941	0.99454	0.02186	3.34994	0.02745	1.72394	0.02404	0.33624	0.02429
IA13428 (EL306463000)	0.67784	0.00607	-1.16820	0.00811	0.56190	0.01519	-0.56190	0.01279	0.00000	0.00000
IA13463 (EL306568260)	0.40323	0.00444	0.51140	0.01595	1.50830	0.02179	-1.50830	0.03242	0.00000	0.00000

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)	Parameters and Measures of Standard Error d5	Parameters and Measures of Standard Error SE(d5)
IA07205A (EL912346409#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA07205D (EL912346409#SCORE_TRAIT_IDEADEV)	-1.24006	0.02703	-4.16186	0.09463	0.00000	0.00000
IA13411A (EL306400886#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13411D (EL306400886#SCORE_TRAIT_IDEADEV)	-1.53336	0.03246	-3.87676	0.10657	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA11517 (EL206801492)	0.75350	0.01324	-1.85670	0.04265	0.10910	0.01870
IA11519 (EL206802113)	0.55632	0.00822	-2.33280	0.05356	0.01580	0.02250
IA11523 (EL206804483)	0.84785	0.01456	-1.78530	0.03542	0.11990	0.01610
IA11524 (EL206804808)	0.39641	0.00880	-1.59970	0.08856	0.03000	0.02590
IA11525 (EL206805580)	0.71634	0.01365	-0.66810	0.03032	0.22030	0.00950
IA11526 (EL206806025)	0.42916	0.00722	-2.33680	0.07425	0.01740	0.02530
IA11530 (EL206808101)	0.49906	0.01124	-0.33480	0.04477	0.15280	0.01200
IA11731 (EL208447958)	0.51358	0.01448	0.43360	0.03840	0.25730	0.00910
IA11738 (EL208451039)	0.84450	0.01141	-1.93310	0.02834	0.01260	0.01370
IA11922 (EL209839730)	1.00100	0.02431	0.60320	0.01672	0.32220	0.00430
IA11925 (EL209877427)	0.84392	0.01792	0.26860	0.01984	0.28570	0.00530
IA11927 (EL209878098)	0.99212	0.01896	-0.69230	0.02253	0.30590	0.00740
IA11928 (EL209878664)	0.99976	0.01705	-1.08340	0.02267	0.21050	0.00880
IA11992 (EL213257312)	1.16714	0.01987	-0.12750	0.01360	0.19540	0.00440
IA12014 (EL215725121)	0.69900	0.01195	-1.50170	0.03797	0.08640	0.01510
IA13429 (EL306531694)	0.28219	0.00838	-1.48620	0.15600	0.03960	0.03380
IA13430 (EL306532662)	0.38419	0.00805	-1.61670	0.08629	0.02490	0.02470
IA13438 (EL306541779)	0.62869	0.01581	1.26020	0.01998	0.15470	0.00450
IA13440 (EL306542147)	0.67078	0.01282	-1.09270	0.03840	0.16450	0.01300
IA13445 (EL306543569)	0.92299	0.01714	-0.06530	0.01800	0.23410	0.00550
IA13450 (EL306550178)	0.71493	0.01767	0.94680	0.01927	0.21580	0.00480
IA13456 (EL306557022)	0.41693	0.00929	-1.50080	0.08459	0.04880	0.02500
IA13466 (EL306619088)	0.98583	0.01523	-0.23030	0.01474	0.14770	0.00480
IA13467 (EL306620254)	0.93998	0.01597	-0.00090	0.01516	0.15510	0.00470

Table 2.8.12
IRT Parameters for Polytomous Items
English Language Arts Grade 8

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA11532 (EL206809440)	0.69636	0.00660	-0.90405	0.00777	0.18455	0.01362	-0.18455	0.01328	0.00000	0.00000
IA11535A (EL206813674#SCORE_TRAIT_CONV)	0.89494	0.00876	-0.97013	0.00784	1.39427	0.01954	-0.14023	0.01359	-1.25403	0.01311
IA11535D (EL206813674#SCORE_TRAIT_IDEADEV)	0.87584	0.00817	0.44266	0.01034	2.83866	0.02110	1.30896	0.01539	0.22826	0.01466
IA11743 (EL208451735)	0.61752	0.00548	-1.72980	0.01119	1.13480	0.02265	-1.13480	0.01543	0.00000	0.00000
IA11930 (EL209978403)	0.46902	0.00465	-0.70705	0.00986	0.58945	0.01809	-0.58945	0.01601	0.00000	0.00000
IA13473 (EL306633886)	0.58624	0.00577	-1.96385	0.01165	0.58715	0.02240	-0.58715	0.01770	0.00000	0.00000
IA13475 (EL306634908)	0.55209	0.00510	-1.12895	0.01073	0.96715	0.02107	-0.96715	0.01571	0.00000	0.00000
IA13480A (EL306640647#SCORE_TRAIT_CONV)	0.92081	0.00923	-0.84670	0.00736	1.23450	0.01743	0.06420	0.01340	-1.29870	0.01290
IA13480D (EL306640647#SCORE_TRAIT_IDEADEV)	0.95932	0.00905	1.02860	0.03066	3.02390	0.03413	1.86550	0.03244	0.57880	0.03231

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)	Parameters and Measures of Standard Error d5	Parameters and Measures of Standard Error SE(d5)
IA11535A (EL206813674#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA11535D (EL206813674#SCORE_TRAIT_IDEADEV)	-0.91964	0.01756	-3.45624	0.04471	0.00000	0.00000
IA13480A (EL306640647#SCORE_TRAIT_CONV)	0.00000	0.00000	n/a	n/a	n/a	n/a
IA13480D (EL306640647#SCORE_TRAIT_IDEADEV)	-0.77320	0.03458	-4.69500	0.15408	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA00993 (MA310834)	0.95792	0.01478	-0.23370	0.01606	0.15540	0.00600
IA02228 (MA297412)	0.78630	0.01580	-0.37314	0.02884	0.28990	0.00920
IA02509 (MA310842)	0.92159	0.01469	-0.37059	0.01837	0.15980	0.00700
IA02544 (MA311274)	0.69267	0.01145	-1.59269	0.04392	0.03260	0.02080
IA04837 (MA735665944)	0.98063	0.01503	-0.13600	0.01533	0.17670	0.00560
IA04840 (MA735732140)	0.81094	0.01532	-1.53100	0.04259	0.12740	0.02060
IA04850 (MA735755736)	1.11553	0.02198	-0.09280	0.01716	0.32630	0.00580
IA08090 (MA902374248)	0.77041	0.01343	-0.03769	0.02117	0.18580	0.00700
IA09963 (MA001336857)	0.87781	0.01797	-0.45223	0.02701	0.33530	0.00880
IA10457 (MA902654429)	0.91830	0.02295	1.08209	0.01618	0.25730	0.00410
IA12053 (MA001034206)	0.56081	0.01073	-1.68832	0.06473	0.03520	0.02700
IA12163 (MA134676677)	0.77311	0.01305	0.66328	0.01375	0.06950	0.00390
IA12173 (MA134830909)	1.45554	0.02324	0.94666	0.00973	0.07340	0.00190
IA12541 (MA735734501)	0.65503	0.01416	-0.84488	0.04599	0.22740	0.01540
IA14197 (MA134867959)	1.02460	0.01860	0.08022	0.01582	0.24640	0.00540
IA14264 (MA231152219)	0.72866	0.01430	0.19277	0.02166	0.17580	0.00700
IA14267 (MA231157499)	0.63362	0.01324	0.32065	0.02409	0.13880	0.00760
IA14557 (MA293459)	0.83327	0.01309	-0.18613	0.01740	0.09760	0.00640
IA02233 (MA297504)	0.53530	0.00589	-0.89355	0.01253	0.00000	0.00000
IA02516 (MA310884)	0.94874	0.00846	-0.36317	0.00827	0.00000	0.00000
IA04863 (MA736028646)	0.97561	0.00894	-0.66784	0.00864	0.00000	0.00000
IA07798 (MA900371363)	0.44635	0.00531	0.43015	0.01375	0.00000	0.00000
IA09960 (MA001330803)	0.97493	0.00908	-0.25122	0.00827	0.00000	0.00000
IA10016 (MA001631751)	0.83878	0.00807	-1.26636	0.01119	0.00000	0.00000
IA10180 (MA002751340)	0.27835	0.00474	1.34916	0.02932	0.00000	0.00000
IA10391 (MA703073138)	0.92430	0.00817	0.35703	0.00888	0.00000	0.00000
IA10408 (MA735754028)	1.20385	0.01126	-1.18350	0.00913	0.00000	0.00000
IA12189 (MA134940265)	0.62125	0.00686	-1.70560	0.01618	0.00000	0.00000
IA12190 (MA134942387)	0.89028	0.00821	-0.64350	0.00900	0.00000	0.00000
IA12651 (MA935137496)	0.75147	0.00715	0.04408	0.00925	0.00000	0.00000
IA14145 (MA001137079)	0.90178	0.00846	-0.59593	0.00876	0.00000	0.00000
IA14198 (MA134872363)	0.88941	0.00826	-0.31304	0.00852	0.00000	0.00000
IA14275 (MA231230636)	0.67870	0.00672	-0.60311	0.00998	0.00000	0.00000
IA14679 (MA735848960)	0.77190	0.00705	0.46288	0.00998	0.00000	0.00000
IA14715 (MA900371312)	0.77693	0.00749	-0.85692	0.01034	0.00000	0.00000
IA14719 (MA900573685)	0.90265	0.00821	-0.59082	0.00888	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA04859 (MA735951978)	1.01262	0.00773	0.35094	0.00545	0.96135	0.00991	0.01046	0.00991	-0.97182	0.01267
IA07527 (MA261857A)	1.07079	0.00763	-0.03598	0.00515	1.09677	0.01048	-0.01764	0.00934	-1.07913	0.01101
IA12411 (MA260584A)	0.91584	0.00672	-0.44960	0.00611	1.40457	0.01383	0.09742	0.01048	-1.50199	0.01212
IA14199 (MA134930940)	1.04789	0.00759	-0.01607	0.00551	1.20393	0.01097	0.17627	0.00954	-1.38020	0.01237

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)
IA04859 (MA735951978)	0.00000	0.00000	n/a	n/a
IA07527 (MA261857A)	0.00000	0.00000	n/a	n/a
IA12411 (MA260584A)	0.00000	0.00000	n/a	n/a
IA14199 (MA134930940)	0.00000	0.00000	n/a	n/a

Table 2.8.15

IRT Parameters for Dichotomous Items

Mathematics Grade 4

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA00960 (MA307075)	0.94798	0.01810	-2.09506	0.04923	0.10950	0.03220
IA01052 (MA311554)	1.36985	0.02723	0.08005	0.01383	0.35150	0.00480
IA04952 (MA800763627)	1.14792	0.01448	-0.02428	0.00991	0.05440	0.00340
IA07659 (MA311543)	1.01009	0.01800	-0.21370	0.01810	0.24060	0.00690
IA07920 (MA900754381)	1.08377	0.01525	0.21851	0.01084	0.08470	0.00370
IA12240 (MA136447421)	1.12997	0.01922	-0.39770	0.01614	0.18890	0.00690
IA12335 (MA202029218)	0.74726	0.02030	0.32561	0.02963	0.39510	0.00760
IA12439 (MA297984)	1.42523	0.02045	-0.16240	0.01038	0.15080	0.00420
IA14320 (MA232239522)	0.87286	0.02208	0.49669	0.02087	0.34740	0.00600
IA14350 (MA233049856)	1.22395	0.02157	0.31443	0.01234	0.23510	0.00420
IA14351 (MA233051799)	0.54676	0.01158	-1.95960	0.08520	0.06310	0.03650
IA14607 (MA307314)	0.80958	0.01402	-0.71496	0.02790	0.18060	0.01120
IA14728 (MA900750340)	0.80203	0.01474	-0.66182	0.02859	0.17730	0.01130
IA02086 (MA258228)	0.71774	0.00739	-1.30742	0.01268	0.00000	0.00000
IA04618 (MA303317)	0.77082	0.00755	0.01676	0.00899	0.00000	0.00000
IA04934 (MA800652951)	0.78184	0.00750	0.84117	0.01095	0.00000	0.00000
IA07901 (MA900740880)	0.82110	0.00770	0.10576	0.00876	0.00000	0.00000
IA07902 (MA900741771)	0.87964	0.00831	-0.32668	0.00842	0.00000	0.00000
IA07915 (MA900751271)	0.55711	0.00612	-0.76396	0.01211	0.00000	0.00000
IA07945 (MA900776517)	0.61218	0.00648	0.22139	0.01049	0.00000	0.00000
IA09849 (MA000732007)	0.75226	0.00719	-0.31838	0.00911	0.00000	0.00000
IA10120 (MA002162929)	0.55155	0.00602	0.26324	0.01141	0.00000	0.00000
IA10289 (MA010534486)	0.69607	0.00668	0.04270	0.00945	0.00000	0.00000
IA12241 (MA136448521)	0.93395	0.00867	-0.09253	0.00807	0.00000	0.00000
IA12281 (MA200344609)	0.76562	0.00785	-1.00606	0.01049	0.00000	0.00000
IA12605 (MA900750085)	0.91498	0.00852	-0.15836	0.00819	0.00000	0.00000
IA14302 (MA231836735)	1.09010	0.01015	-0.35562	0.00772	0.00000	0.00000
IA14307 (MA231843918)	0.96333	0.00887	-0.45073	0.00819	0.00000	0.00000
IA14325 (MA232254177)	0.73554	0.00719	-0.36818	0.00922	0.00000	0.00000
IA14326 (MA232261850)	0.67929	0.00780	-1.60463	0.01476	0.00000	0.00000
IA14338 (MA232725235)	0.95889	0.00913	-0.94231	0.00934	0.00000	0.00000
IA14340 (MA232909389)	0.81917	0.00770	0.04408	0.00865	0.00000	0.00000
IA14349 (MA232979874)	0.90580	0.00989	-1.75577	0.01383	0.00000	0.00000
IA15968 (MA250533)	0.68673	0.00694	-0.29359	0.00957	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA04621 (MA303335)	0.92396	0.00658	-0.03734	0.00535	1.78622	0.01334	0.37350	0.00968	-0.55479	0.01027
IA04757 (MA713677363)	0.71672	0.00637	-1.28949	0.00934	0.86113	0.01875	-0.86113	0.01313	0.00000	0.00000
IA04948 (MA800750938)	1.18107	0.00918	-0.12602	0.00546	0.53303	0.00946	-0.53303	0.00946	0.00000	0.00000
IA14311 (MA231875780)	1.06745	0.00745	-0.43364	0.00511	1.71952	0.01410	0.52516	0.00995	-0.53248	0.00917
IA14694 (MA800780887)	1.24522	0.00862	-0.25376	0.00449	1.53498	0.01184	0.37774	0.00873	-0.48484	0.00864
IA14707 (MA803957909)	0.97245	0.00688	0.35285	0.00583	2.08037	0.01333	0.41552	0.00977	-0.67855	0.01140

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)
IA04621 (MA303335)	-1.60493	0.01398	0.00000	0.00000
IA14311 (MA231875780)	-1.71220	0.01188	0.00000	0.00000
IA14694 (MA800780887)	-1.42788	0.01057	0.00000	0.00000
IA14707 (MA803957909)	-1.81734	0.01662	0.00000	0.00000

Table 2.8.17
IRT Parameters for Dichotomous Items
Mathematics Grade 5

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA02390 (MA306430)	1.37487	0.01926	0.00208	0.00982	0.13810	0.00360
IA02545 (MA311279)	0.97949	0.01602	-0.55170	0.01815	0.14150	0.00790
IA04981 (MA801230425)	1.29344	0.02611	0.88548	0.01153	0.19570	0.00290
IA05004 (MA801654509)	0.90774	0.01499	-0.11685	0.01564	0.11040	0.00610
IA07994 (MA901137741)	1.17553	0.02292	0.41947	0.01381	0.30300	0.00420
IA08212 (MA904353319)	0.69954	0.01267	-2.04651	0.05889	0.04400	0.03200
IA09867 (MA000859040)	0.81147	0.02148	-0.25358	0.03344	0.42360	0.00920
IA10336 (MA298013)	1.13881	0.01674	-1.40565	0.02111	0.06510	0.01240
IA12299 (MA201552031)	0.57139	0.01818	-0.02086	0.04976	0.30310	0.01330
IA12333 (MA202000453)	0.89950	0.01674	-1.55847	0.03858	0.13070	0.02070
IA12423 (MA282124)	1.55526	0.04409	0.80250	0.01358	0.43520	0.00330
IA12613 (MA900980248)	0.78927	0.01736	-0.25392	0.02694	0.23070	0.00950
IA14141 (MA000952312)	0.49423	0.01102	-0.39648	0.04497	0.01830	0.01510
IA14143 (MA000963503)	0.79777	0.02066	-0.01664	0.02979	0.40170	0.00820
IA14362 (MA233233270)	0.73560	0.01849	0.50062	0.02260	0.24610	0.00690
IA14644 (MA704359445A)	0.61862	0.01597	0.41878	0.02773	0.17800	0.00870
IA14720 (MA900660071)	1.12882	0.01736	0.24347	0.01130	0.13610	0.00380
IA14722 (MA900662425)	0.76542	0.01731	-0.33576	0.03036	0.27020	0.01020
IA16002 (MA298014)	1.23920	0.02096	-1.23912	0.02214	0.23410	0.01170
IA02734 (MA704359624)	0.42423	0.00567	-1.86675	0.02271	0.00000	0.00000
IA02921 (MA715102228)	1.02245	0.00994	0.09407	0.00788	0.00000	0.00000
IA04704 (MA311283)	0.81312	0.00804	-1.08630	0.01039	0.00000	0.00000
IA04938 (MA800662477)	0.66354	0.00737	0.77488	0.01221	0.00000	0.00000
IA05010 (MA801673695)	0.65432	0.00670	1.00212	0.01370	0.00000	0.00000
IA05022 (MA802302084)	0.78721	0.00731	-0.02896	0.00879	0.00000	0.00000
IA08025 (MA901364051)	0.66112	0.00701	0.62639	0.01176	0.00000	0.00000
IA08213 (MA904453014)	0.74410	0.00742	-0.65910	0.00936	0.00000	0.00000
IA09863 (MA000846693)	0.91758	0.00855	-0.19069	0.00810	0.00000	0.00000
IA09914 (MA001060832)	0.73209	0.00726	-0.31259	0.00902	0.00000	0.00000
IA09928 (MA001145679)	0.70464	0.00762	0.75856	0.01176	0.00000	0.00000
IA10481 (MA935150419)	0.91459	0.00865	-0.35139	0.00822	0.00000	0.00000
IA14371 (MA233369406)	0.99031	0.00896	-0.35641	0.00788	0.00000	0.00000
IA14382 (MA233458877)	0.83516	0.00814	0.35042	0.00913	0.00000	0.00000
IA14732 (MA900835182)	0.50206	0.00603	-0.00808	0.01164	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA02929 (MA715102462)	1.13551	0.00778	-0.23968	0.00435	1.26047	0.01053	0.54838	0.00880	-0.58040	0.00890
IA05024 (MA802306160)	0.72107	0.00644	-1.01234	0.01017	1.36824	0.02050	-1.36824	0.01413	0.00000	0.00000
IA07728 (MA802371654)	1.09024	0.00752	-0.30779	0.00462	1.42062	0.01168	0.56188	0.00913	-0.52045	0.00893
IA08179 (MA903746975)	1.03471	0.00742	-0.44464	0.00476	1.47838	0.01300	0.35758	0.00940	-0.46042	0.00901
IA09886 (MA000965213)	0.64994	0.00587	0.09584	0.00905	1.17723	0.01448	-1.17723	0.01680	0.00000	0.00000
IA14336 (MA232642946)	1.36395	0.00968	0.18549	0.00386	0.95279	0.00826	0.19574	0.00786	-0.29332	0.00846

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)
IA02929 (MA715102462)	-1.22845	0.01053	0.00000	0.00000
IA07728 (MA802371654)	-1.46205	0.01126	0.00000	0.00000
IA08179 (MA903746975)	-1.37554	0.01070	0.00000	0.00000
IA14336 (MA232642946)	-0.85520	0.00981	0.00000	0.00000

Table 2.8.19
IRT Parameters for Dichotomous Items
Mathematics Grade 6

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA02145 (MA282127)	1.05465	0.01658	0.06719	0.01345	0.17610	0.00480
IA02224 (MA296377)	0.78762	0.01719	-0.79089	0.03792	0.31570	0.01310
IA02470 (MA307363)	0.49473	0.01120	-2.09910	0.10181	0.06040	0.04090
IA04627 (MA303705)	0.63222	0.01501	0.45483	0.02435	0.11750	0.00820
IA04657 (MA307250)	1.12989	0.01891	0.55200	0.01125	0.14020	0.00330
IA04722 (MA311652)	1.29664	0.01947	0.26860	0.01044	0.16990	0.00340
IA04894 (MA736452061)	0.96907	0.01491	0.44114	0.01171	0.08820	0.00360
IA04895 (MA736453104)	0.86985	0.01982	0.18813	0.02261	0.36020	0.00650
IA04896 (MA736454897)	0.78275	0.02312	1.00851	0.02087	0.32650	0.00520
IA07740 (MA805104566)	1.30186	0.02292	0.40207	0.01160	0.23320	0.00370
IA07742 (MA805111429)	0.84187	0.01967	0.28379	0.02215	0.33030	0.00640
IA07769 (MA900281418)	0.71081	0.01095	-1.02175	0.02992	0.03120	0.01310
IA08238 (MA908142878)	1.08801	0.02459	1.03054	0.01391	0.25630	0.00330
IA12193 (MA135360248)	0.71223	0.01242	-2.12136	0.05902	0.05070	0.03310
IA12436 (MA296382)	1.16731	0.02616	0.66969	0.01449	0.34690	0.00390
IA15971 (MA272290)	0.87553	0.01607	-0.11220	0.02018	0.23410	0.00700
IA00974 (MA307362)	1.16066	0.01085	-0.96459	0.00835	0.00000	0.00000
IA02691 (MA703149512)	0.62903	0.00644	-0.85165	0.01090	0.00000	0.00000
IA02826 (MA713677108)	0.82833	0.00755	0.23289	0.00904	0.00000	0.00000
IA04893 (MA736449649)	1.17369	0.01095	0.55571	0.00835	0.00000	0.00000
IA05128 (MA805109765)	0.53199	0.00654	-0.17470	0.01276	0.00000	0.00000
IA05140 (MA805280133)	0.55394	0.00598	0.77034	0.01357	0.00000	0.00000
IA05141 (MA805280170)	0.78412	0.00745	-0.74091	0.00939	0.00000	0.00000
IA08050 (MA901737396)	0.75654	0.00700	0.04052	0.00893	0.00000	0.00000
IA10001 (MA001554177)	0.63020	0.00654	-0.47618	0.00997	0.00000	0.00000
IA10004 (MA001556929)	1.06246	0.01055	-1.38898	0.01009	0.00000	0.00000
IA10011 (MA001577731)	0.83766	0.00791	-0.52616	0.00858	0.00000	0.00000
IA10013 (MA001603294)	1.18992	0.01191	-1.34759	0.00951	0.00000	0.00000
IA10155 (MA002526793)	0.44393	0.00548	-0.85037	0.01403	0.00000	0.00000
IA10210 (MA003477341)	1.02971	0.00928	-0.43235	0.00788	0.00000	0.00000
IA12626 (MA902348152)	0.68211	0.00679	-0.13098	0.00951	0.00000	0.00000
IA14406 (MA233849848)	1.08188	0.00897	-0.00169	0.00777	0.00000	0.00000
IA14717 (MA900432715)	0.81205	0.00826	0.62621	0.01009	0.00000	0.00000
IA14740 (MA902761648)	0.44398	0.00598	0.54272	0.01797	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA04560 (MA287883)	1.18150	0.00852	0.20798	0.00403	0.81496	0.00844	0.28621	0.00824	-0.19594	0.00875
IA04912 (MA800203270)	1.25963	0.00948	0.26263	0.00552	0.52395	0.00897	-0.52395	0.01010	0.00000	0.00000
IA05143 (MA805301630)	0.47668	0.00527	-0.62507	0.00931	0.48632	0.01713	-0.48632	0.01505	0.00000	0.00000
IA07810 (MA900378821)	1.17897	0.00801	-0.16928	0.00479	1.91973	0.01308	0.12567	0.00864	-0.61274	0.00913
IA10412 (MA800259417)	1.17004	0.00781	-0.49772	0.00442	1.22023	0.01144	0.50270	0.00924	-0.26922	0.00844
IA14401 (MA233752432)	1.15565	0.00796	-0.12402	0.00436	1.30589	0.01046	0.43878	0.00861	-0.35320	0.00861

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)
IA04560 (MA287883)	-0.90524	0.01043	0.00000	0.00000
IA07810 (MA900378821)	-1.43266	0.01138	0.00000	0.00000
IA10412 (MA800259417)	-1.45371	0.01017	0.00000	0.00000
IA14401 (MA233752432)	-1.39147	0.01110	0.00000	0.00000

Table 2.8.21

IRT Parameters for Dichotomous Items

Mathematics Grade 7

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA00883 (MA298180)	0.56061	0.01081	-1.85596	0.07039	0.04200	0.03090
IA00950 (MA306624)	1.62247	0.06101	1.89050	0.01678	0.22050	0.00200
IA02410 (MA306595)	1.12049	0.01923	-0.99634	0.02149	0.23380	0.01000
IA02526 (MA311074)	1.01687	0.02108	0.80439	0.01485	0.26450	0.00370
IA04513 (MA272764)	0.97820	0.02517	0.75899	0.01896	0.41280	0.00430
IA04529 (MA281422)	1.34491	0.01948	0.19662	0.00990	0.12140	0.00310
IA07841 (MA900556478)	1.68743	0.03379	0.93840	0.01038	0.19890	0.00230
IA08110 (MA903153837)	1.45545	0.03730	1.15173	0.01292	0.29500	0.00270
IA08197 (MA904159889)	0.72889	0.01154	-0.66843	0.02511	0.06850	0.01010
IA09888 (MA000971342)	1.04827	0.01768	0.15871	0.01352	0.16000	0.00450
IA10031 (MA001684984)	1.43758	0.02625	0.95530	0.01087	0.17780	0.00240
IA12152 (MA005204778)	1.49645	0.02449	0.17211	0.01050	0.21170	0.00350
IA12259 (MA200144334)	0.75801	0.01514	-0.87416	0.03682	0.23900	0.01370
IA12270 (MA200256305)	1.06624	0.01850	-0.65841	0.01908	0.21080	0.00810
IA12593 (MA900552355)	1.21037	0.02274	0.11078	0.01449	0.30820	0.00460
IA12594 (MA900566128)	0.73113	0.01529	1.22309	0.01618	0.09170	0.00360
IA14186 (MA005200422)	0.78995	0.02323	0.94854	0.02294	0.38720	0.00500
IA14229 (MA201106705)	1.01823	0.01875	0.41044	0.01473	0.24350	0.00430
IA14426 (MA234550897)	1.22771	0.02220	0.19058	0.01244	0.21750	0.00420
IA14723 (MA900665080)	1.74673	0.03384	0.79280	0.01014	0.18730	0.00240
IA16014 (MA301838)	2.20957	0.04358	0.60530	0.00918	0.23500	0.00250
IA07848 (MA900567252)	0.89465	0.00803	0.10945	0.00857	0.00000	0.00000
IA07897 (MA900739359)	0.97022	0.00881	-0.29078	0.00821	0.00000	0.00000
IA09940 (MA001178265)	0.87098	0.00813	0.65625	0.01014	0.00000	0.00000
IA10027 (MA001678587)	1.12190	0.00993	0.14277	0.00785	0.00000	0.00000
IA10416 (MA801105441)	0.70333	0.00701	0.75223	0.01171	0.00000	0.00000
IA12147 (MA005077768)	0.92849	0.00838	-0.42262	0.00833	0.00000	0.00000
IA12294 (MA201203484)	1.42575	0.01334	0.29357	0.00749	0.00000	0.00000
IA14427 (MA234551702)	0.65167	0.00614	0.62739	0.01195	0.00000	0.00000
IA14429 (MA234553959)	1.19645	0.01154	-1.49943	0.00990	0.00000	0.00000
IA14432 (MA234555649)	0.88973	0.00842	-0.65732	0.00869	0.00000	0.00000
IA14437 (MA234561626)	0.99179	0.00852	0.46513	0.00930	0.00000	0.00000
IA14497 (MA236233939)	0.98366	0.01013	-1.31073	0.01002	0.00000	0.00000
IA14748 (MA904001752)	0.30706	0.00482	1.33839	0.02801	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA02711 (MA703876323)	1.50468	0.01363	1.00987	0.00636	0.31451	0.01010	-0.31451	0.01187	0.00000	0.00000
IA07679 (MA717248260)	1.12808	0.00750	0.28994	0.00520	1.82125	0.01118	0.40204	0.00912	-0.48341	0.01034
IA07730 (MA802907874)	1.27562	0.00881	0.48490	0.00445	1.16878	0.00871	0.44970	0.00850	-0.52449	0.01009
IA10445 (MA900765087)	1.00849	0.00789	-0.73393	0.00754	1.22864	0.01444	-1.22864	0.01151	0.00000	0.00000
IA14465 (MA235900389)	1.00849	0.00701	0.21503	0.00552	1.89580	0.01219	0.54976	0.00959	-0.64562	0.01102
IA14466 (MA235905181)	1.43349	0.00984	-0.08568	0.00445	1.86988	0.01163	0.13290	0.00830	-0.64076	0.00892

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)
IA07679 (MA717248260)	-1.73988	0.01505	0.00000	0.00000
IA07730 (MA802907874)	-1.09399	0.01208	0.00000	0.00000
IA14465 (MA235900389)	-1.79994	0.01573	0.00000	0.00000
IA14466 (MA235905181)	-1.36201	0.01064	0.00000	0.00000

Table 2.8.23

IRT Parameters for Dichotomous Items

Mathematics Grade 8

Item ID	Parameters and Measures of Standard Error a	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error c	Parameters and Measures of Standard Error SE(c)
IA00988 (MA307600)	0.48614	0.01696	0.53206	0.04766	0.25440	0.01190
IA02065 (MA228311)	1.18078	0.02073	-1.23191	0.02246	0.19940	0.01190
IA08003 (MA901143033)	1.31333	0.02784	0.30344	0.01311	0.33180	0.00390
IA10039 (MA001736920)	0.82095	0.01645	-0.12529	0.02246	0.24010	0.00750
IA10118 (MA002159196)	1.10535	0.01634	0.31884	0.01072	0.10320	0.00320
IA10134 (MA002181298)	1.05023	0.01923	-0.48550	0.01870	0.24370	0.00730
IA10311 (MA264730)	0.46644	0.00634	-1.50397	0.02771	0.00150	0.00860
IA12106 (MA002178533)	1.54998	0.02511	0.18372	0.01038	0.24770	0.00330
IA12376 (MA203023828)	0.99940	0.01908	-1.54434	0.03740	0.28860	0.01820
IA12480 (MA307586)	0.55476	0.00629	-0.67045	0.01391	0.00020	0.00330
IA14460 (MA235755845)	1.56720	0.03026	0.63047	0.01095	0.27110	0.00290
IA14523 (MA236476077)	1.27874	0.02031	-1.63875	0.02303	0.09660	0.01540
IA14599 (MA306554)	1.26311	0.02124	-0.44365	0.01437	0.24470	0.00580
IA14624 (MA311391)	1.75745	0.05383	0.82910	0.01414	0.52760	0.00280
IA14692 (MA800750287)	1.67315	0.02392	-0.37695	0.00924	0.13100	0.00370
IA16011 (MA301682)	0.71160	0.01660	-0.08983	0.02965	0.26100	0.00910
IA16012 (MA301698)	1.13747	0.01866	-0.44160	0.01574	0.22390	0.00620
IA00897 (MA301481)	1.39768	0.01232	0.37517	0.00775	0.00000	0.00000
IA02753 (MA704839576)	1.27477	0.01170	0.85316	0.00912	0.00000	0.00000
IA02935 (MA715919560)	1.02002	0.01062	0.75909	0.00958	0.00000	0.00000
IA02938 (MA715919661)	0.39426	0.00572	-1.63750	0.02041	0.00000	0.00000
IA02943 (MA715919788)	0.96454	0.00887	0.07836	0.00855	0.00000	0.00000
IA07707 (MA800659905)	0.59147	0.00593	0.71599	0.01311	0.00000	0.00000
IA07733 (MA804043870)	1.04606	0.00964	0.03902	0.00798	0.00000	0.00000
IA08010 (MA901253257)	0.49387	0.00541	0.19569	0.01277	0.00000	0.00000
IA08068 (MA902265010)	0.73305	0.00675	0.39877	0.01015	0.00000	0.00000
IA10041 (MA001737758)	0.81245	0.00758	-0.46361	0.00867	0.00000	0.00000
IA10122 (MA002170014)	0.99568	0.00918	-0.70260	0.00810	0.00000	0.00000
IA10201 (MA003128642)	0.65061	0.00639	-0.24593	0.00969	0.00000	0.00000
IA12348 (MA202726454)	0.77125	0.00737	-0.54456	0.00889	0.00000	0.00000
IA12355 (MA202828823)	1.22800	0.01098	-0.36805	0.00764	0.00000	0.00000
IA14459 (MA235753974)	0.86416	0.00789	0.16650	0.00878	0.00000	0.00000
IA14521 (MA236473688)	0.58028	0.00598	-0.04103	0.01060	0.00000	0.00000
IA14693 (MA800756470)	0.85277	0.00784	0.30515	0.00912	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	Parameters and Measures of Standard Error SE(a)	Parameters and Measures of Standard Error b	Parameters and Measures of Standard Error SE(b)	Parameters and Measures of Standard Error d0	Parameters and Measures of Standard Error SE(d0)	Parameters and Measures of Standard Error d1	Parameters and Measures of Standard Error SE(d1)	Parameters and Measures of Standard Error d2	Parameters and Measures of Standard Error SE(d2)
IA04717 (MA311433)	1.23027	0.00820	0.15778	0.00457	1.52138	0.01010	0.36425	0.00852	-0.49253	0.00950
IA04967 (MA800880893)	0.68386	0.00619	-0.73818	0.00761	0.54835	0.01400	-0.54835	0.01233	0.00000	0.00000
IA12383 (MA203085405)	0.82271	0.00598	-1.44206	0.00704	1.86648	0.02213	0.63626	0.01488	-0.48005	0.01152
IA14245 (MA203016132)	0.71475	0.00608	-0.93738	0.00843	0.85131	0.01617	-0.85131	0.01285	0.00000	0.00000
IA14533 (MA236558384)	1.45388	0.00995	-0.18119	0.00380	1.13224	0.00915	0.11924	0.00793	-0.33880	0.00803
IA14684 (MA800473031)	1.25394	0.00876	0.43651	0.00451	1.06728	0.00868	0.48689	0.00848	-0.32702	0.00967

Item ID	Parameters and Measures of Standard Error d3	Parameters and Measures of Standard Error SE(d3)	Parameters and Measures of Standard Error d4	Parameters and Measures of Standard Error SE(d4)
IA04717 (MA311433)	-1.39311	0.01239	0.00000	0.00000
IA12383 (MA203085405)	-2.02270	0.01217	0.00000	0.00000
IA14533 (MA236558384)	-0.91269	0.00884	0.00000	0.00000
IA14684 (MA800473031)	-1.22714	0.01290	0.00000	0.00000

Section 2.9

Decision Accuracy and Consistency (DAC)

Table 2.9.1
DAC Results
English Language Arts Grade 3

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
63649	0.90	0.59	Overall	0.80	0.72	0.10	0.09
N/A	N/A	N/A	Cut 1	0.94	0.92	0.02	0.03
N/A	N/A	N/A	Cut 2	0.90	0.87	0.05	0.05
N/A	N/A	N/A	Cut 3	0.96	0.94	0.03	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.85	0.75	N/A	N/A
N/A	N/A	N/A	Perf 2	0.79	0.72	N/A	N/A
N/A	N/A	N/A	Perf 3	0.81	0.75	N/A	N/A
N/A	N/A	N/A	Perf 4	0.70	0.50	N/A	N/A

Table 2.9.2
DAC Results
English Language Arts Grade 4

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
62216	0.91	0.61	Overall	0.82	0.74	0.09	0.09
N/A	N/A	N/A	Cut 1	0.94	0.92	0.02	0.03
N/A	N/A	N/A	Cut 2	0.91	0.87	0.05	0.04
N/A	N/A	N/A	Cut 3	0.97	0.95	0.02	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.85	0.76	N/A	N/A
N/A	N/A	N/A	Perf 2	0.81	0.75	N/A	N/A
N/A	N/A	N/A	Perf 3	0.83	0.77	N/A	N/A
N/A	N/A	N/A	Perf 4	0.67	0.48	N/A	N/A

Table 2.9.3
DAC Results
English Language Arts Grade 5

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
63664	0.92	0.63	Overall	0.83	0.76	0.08	0.08
N/A	N/A	N/A	Cut 1	0.95	0.93	0.02	0.03
N/A	N/A	N/A	Cut 2	0.91	0.88	0.04	0.05
N/A	N/A	N/A	Cut 3	0.97	0.96	0.02	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.83	0.72	N/A	N/A
N/A	N/A	N/A	Perf 2	0.86	0.81	N/A	N/A
N/A	N/A	N/A	Perf 3	0.81	0.75	N/A	N/A
N/A	N/A	N/A	Perf 4	0.78	0.61	N/A	N/A

Table 2.9.4
DAC Results
English Language Arts Grade 6

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
64470	0.92	0.59	Overall	0.80	0.71	0.10	0.10
N/A	N/A	N/A	Cut 1	0.94	0.91	0.03	0.03
N/A	N/A	N/A	Cut 2	0.91	0.88	0.04	0.05
N/A	N/A	N/A	Cut 3	0.95	0.92	0.03	0.03
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.85	0.76	N/A	N/A
N/A	N/A	N/A	Perf 2	0.80	0.72	N/A	N/A
N/A	N/A	N/A	Perf 3	0.78	0.70	N/A	N/A
N/A	N/A	N/A	Perf 4	0.73	0.59	N/A	N/A

Table 2.9.5

DAC Results

English Language Arts Grade 7

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
64737	0.92	0.61	Overall	0.81	0.73	0.10	0.10
N/A	N/A	N/A	Cut 1	0.94	0.92	0.03	0.03
N/A	N/A	N/A	Cut 2	0.92	0.88	0.04	0.04
N/A	N/A	N/A	Cut 3	0.95	0.92	0.03	0.02
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.87	0.79	N/A	N/A
N/A	N/A	N/A	Perf 2	0.81	0.74	N/A	N/A
N/A	N/A	N/A	Perf 3	0.79	0.72	N/A	N/A
N/A	N/A	N/A	Perf 4	0.66	0.49	N/A	N/A

Table 2.9.6

DAC Results

English Language Arts Grade 8

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
65482	0.93	0.61	Overall	0.80	0.72	0.10	0.10
N/A	N/A	N/A	Cut 1	0.94	0.92	0.03	0.03
N/A	N/A	N/A	Cut 2	0.91	0.88	0.04	0.04
N/A	N/A	N/A	Cut 3	0.95	0.92	0.03	0.03
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.86	0.78	N/A	N/A
N/A	N/A	N/A	Perf 2	0.79	0.71	N/A	N/A
N/A	N/A	N/A	Perf 3	0.79	0.72	N/A	N/A
N/A	N/A	N/A	Perf 4	0.76	0.64	N/A	N/A

Table 2.9.7

DAC Results

Mathematics Grade 3

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
63570	0.94	0.65	Overall	0.83	0.75	0.09	0.09
N/A	N/A	N/A	Cut 1	0.95	0.92	0.02	0.03
N/A	N/A	N/A	Cut 2	0.92	0.89	0.04	0.04
N/A	N/A	N/A	Cut 3	0.96	0.94	0.02	0.02
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.87	0.79	N/A	N/A
N/A	N/A	N/A	Perf 2	0.81	0.74	N/A	N/A
N/A	N/A	N/A	Perf 3	0.83	0.76	N/A	N/A
N/A	N/A	N/A	Perf 4	0.80	0.69	N/A	N/A

Table 2.9.8

DAC Results

Mathematics Grade 4

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
61946	0.94	0.67	Overall	0.84	0.78	0.08	0.08
N/A	N/A	N/A	Cut 1	0.95	0.93	0.02	0.02
N/A	N/A	N/A	Cut 2	0.92	0.89	0.04	0.04
N/A	N/A	N/A	Cut 3	0.96	0.95	0.02	0.02
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.86	0.77	N/A	N/A
N/A	N/A	N/A	Perf 2	0.85	0.79	N/A	N/A
N/A	N/A	N/A	Perf 3	0.84	0.79	N/A	N/A
N/A	N/A	N/A	Perf 4	0.79	0.67	N/A	N/A

Table 2.9.9
DAC Results
Mathematics Grade 5

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
63371	0.94	0.67	Overall	0.85	0.78	0.08	0.08
N/A	N/A	N/A	Cut 1	0.95	0.93	0.02	0.03
N/A	N/A	N/A	Cut 2	0.92	0.89	0.04	0.04
N/A	N/A	N/A	Cut 3	0.97	0.96	0.02	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.85	0.76	N/A	N/A
N/A	N/A	N/A	Perf 2	0.85	0.80	N/A	N/A
N/A	N/A	N/A	Perf 3	0.84	0.78	N/A	N/A
N/A	N/A	N/A	Perf 4	0.82	0.69	N/A	N/A

Table 2.9.10
DAC Results
Mathematics Grade 6

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
64578	0.94	0.67	Overall	0.85	0.78	0.08	0.08
N/A	N/A	N/A	Cut 1	0.95	0.93	0.02	0.03
N/A	N/A	N/A	Cut 2	0.92	0.89	0.04	0.04
N/A	N/A	N/A	Cut 3	0.97	0.96	0.02	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.85	0.77	N/A	N/A
N/A	N/A	N/A	Perf 2	0.85	0.79	N/A	N/A
N/A	N/A	N/A	Perf 3	0.85	0.79	N/A	N/A
N/A	N/A	N/A	Perf 4	0.81	0.69	N/A	N/A

Table 2.9.11
DAC Results
Mathematics Grade 7

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
64726	0.94	0.68	Overall	0.84	0.78	0.08	0.08
N/A	N/A	N/A	Cut 1	0.95	0.92	0.03	0.03
N/A	N/A	N/A	Cut 2	0.93	0.90	0.04	0.04
N/A	N/A	N/A	Cut 3	0.97	0.96	0.02	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.86	0.79	N/A	N/A
N/A	N/A	N/A	Perf 2	0.85	0.79	N/A	N/A
N/A	N/A	N/A	Perf 3	0.83	0.76	N/A	N/A
N/A	N/A	N/A	Perf 4	0.85	0.74	N/A	N/A

Table 2.9.12
DAC Results
Mathematics Grade 8

N	Reliability	Kappa	Cuts/Perf Lvl	Accuracy	Consistency	F Pos	F Neg
65585	0.94	0.68	Overall	0.84	0.78	0.08	0.08
N/A	N/A	N/A	Cut 1	0.94	0.92	0.03	0.03
N/A	N/A	N/A	Cut 2	0.93	0.90	0.04	0.04
N/A	N/A	N/A	Cut 3	0.97	0.96	0.02	0.01
N/A	N/A	N/A	Cut 4	1.00	1.00	0.00	0.00
N/A	N/A	N/A	Perf 1	0.87	0.79	N/A	N/A
N/A	N/A	N/A	Perf 2	0.84	0.78	N/A	N/A
N/A	N/A	N/A	Perf 3	0.84	0.77	N/A	N/A
N/A	N/A	N/A	Perf 4	0.83	0.72	N/A	N/A

Section 2.10

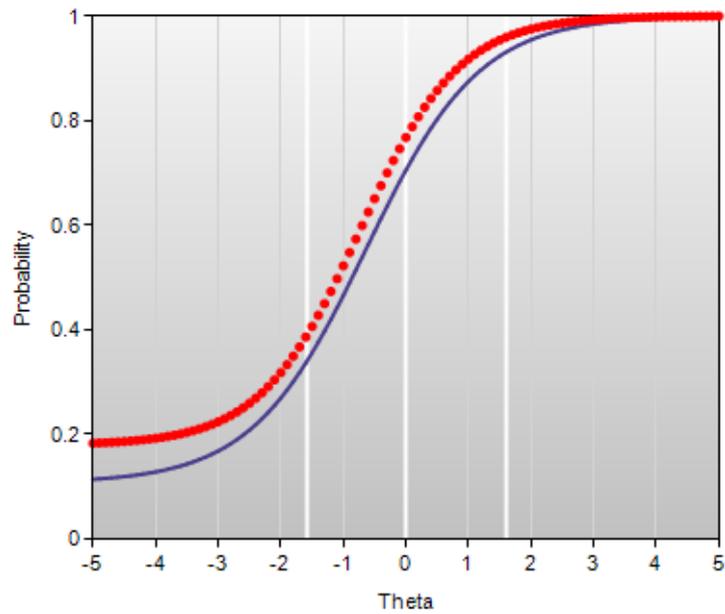Fit Plots and Beta Plots of Watchlist Items

Initial Calibration

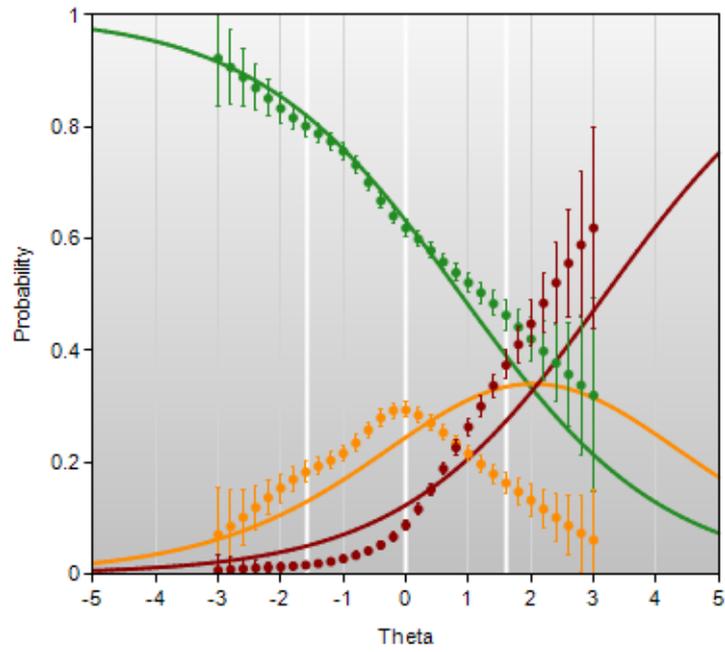English Language Arts Grade 3: IA00280
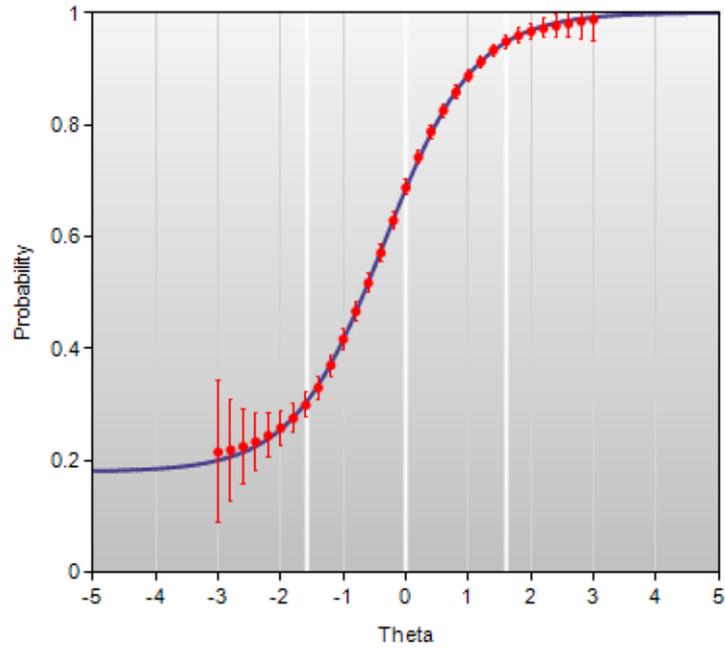
Beta Chart

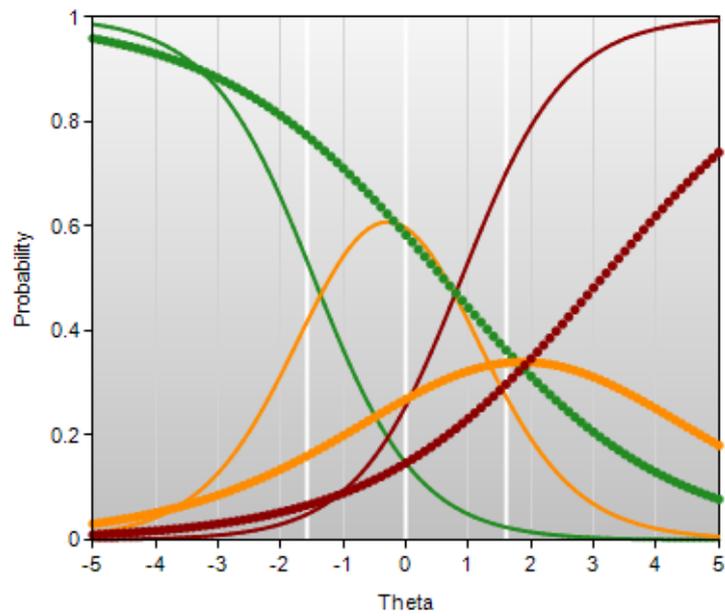English Language Arts Grade 3: IA00280
(EL308824)

Initial Calibration
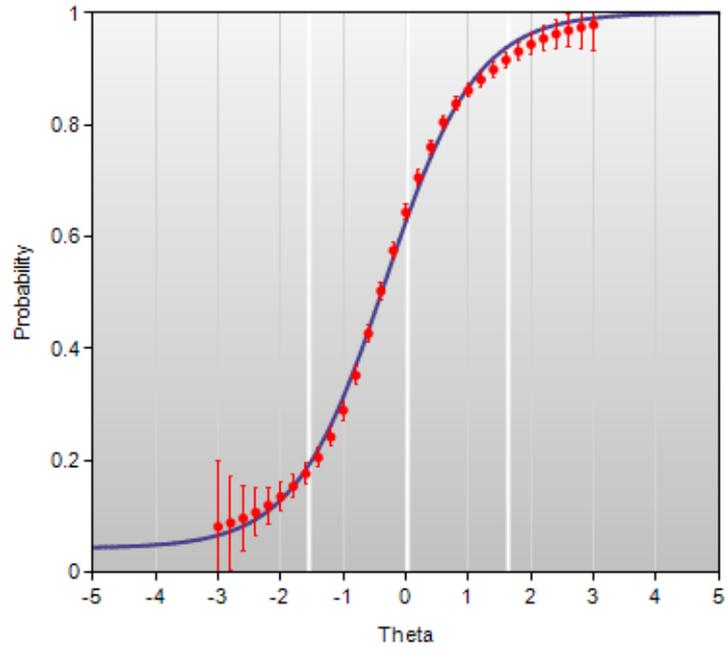
English Language Arts Grade 3: IA00445

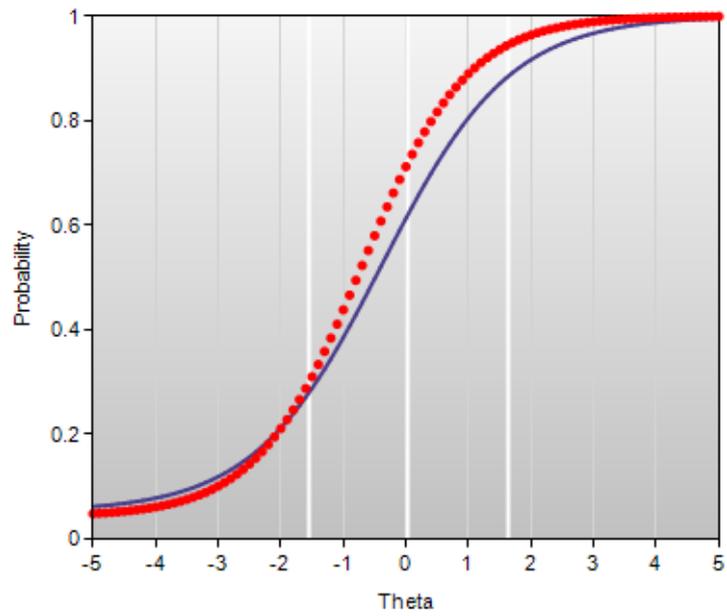

Beta Chart

English Language Arts Grade 3: IA00445
(EL626043435)
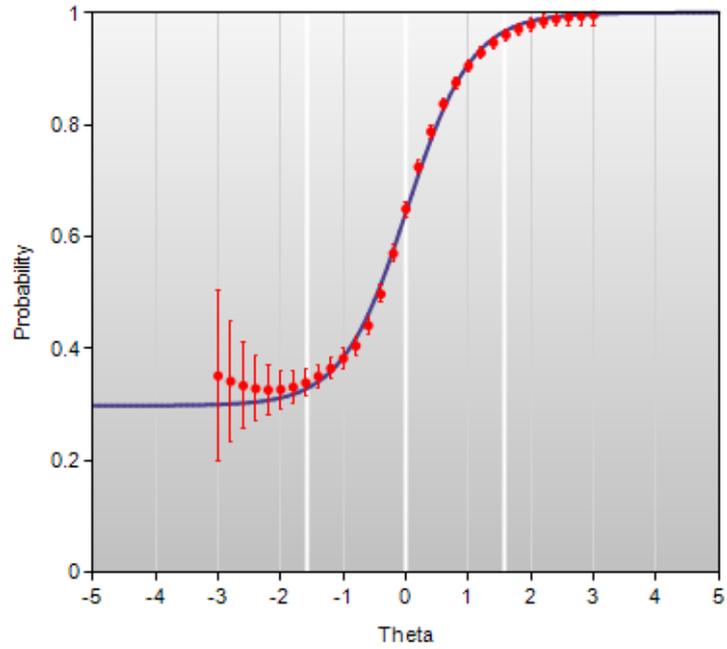
Initial Calibration

English Language Arts Grade 4: IA00408

Beta Chart
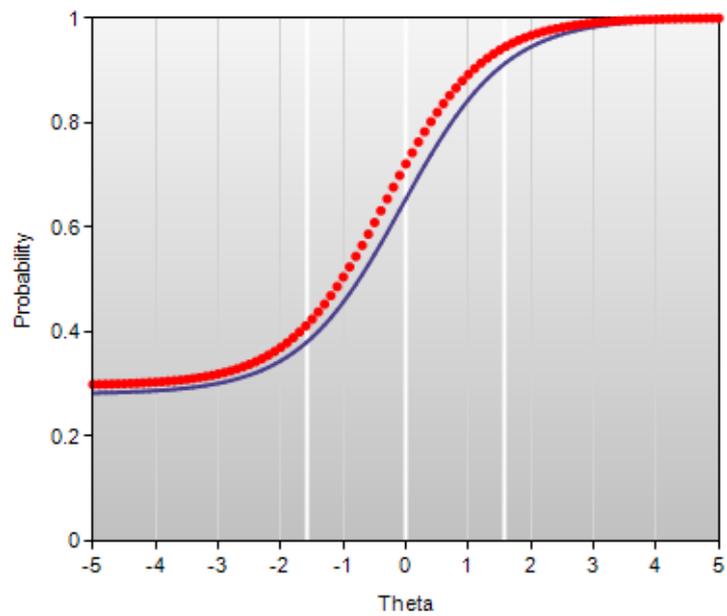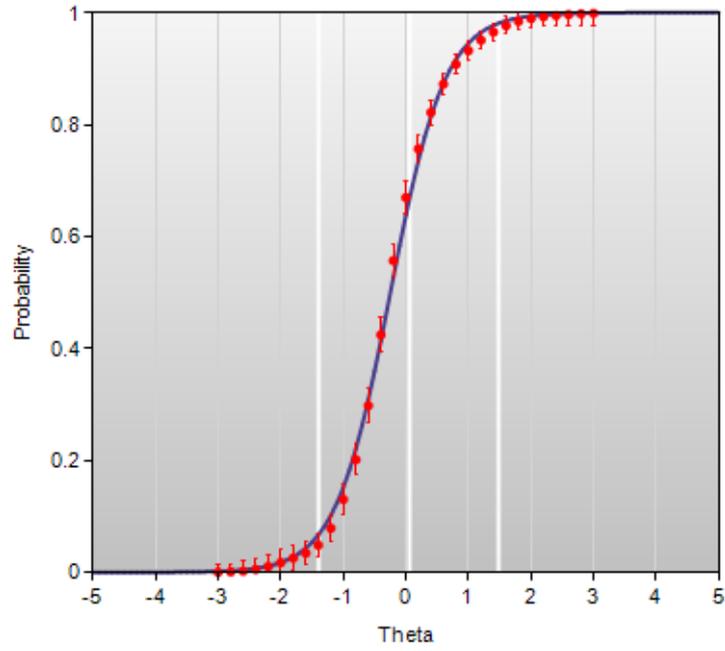
English Language Arts Grade 4: IA00408 (EL624647580)

Initial Calibration

English Language Arts Grade 6: IA00175

Beta Chart

English Language Arts Grade 6: IA00175 (EL303504)
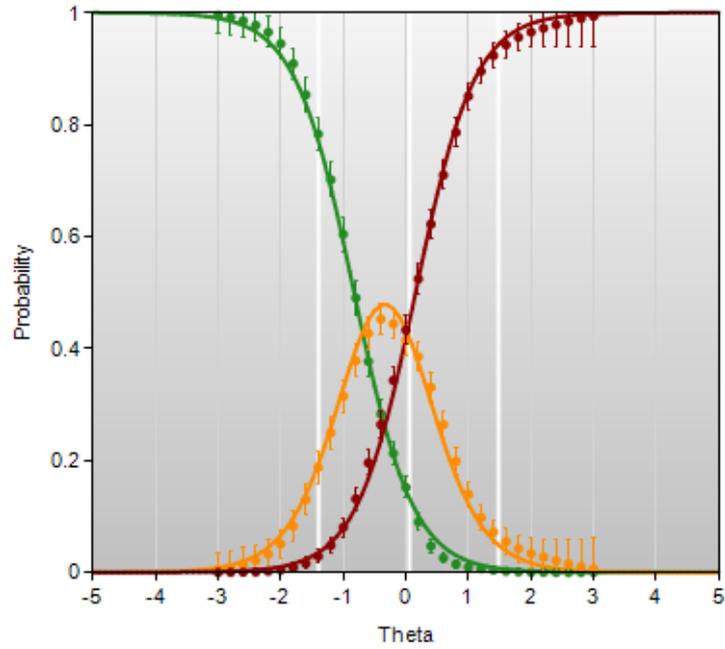
Initial Calibration

Mathematics Grade 4: IA01055
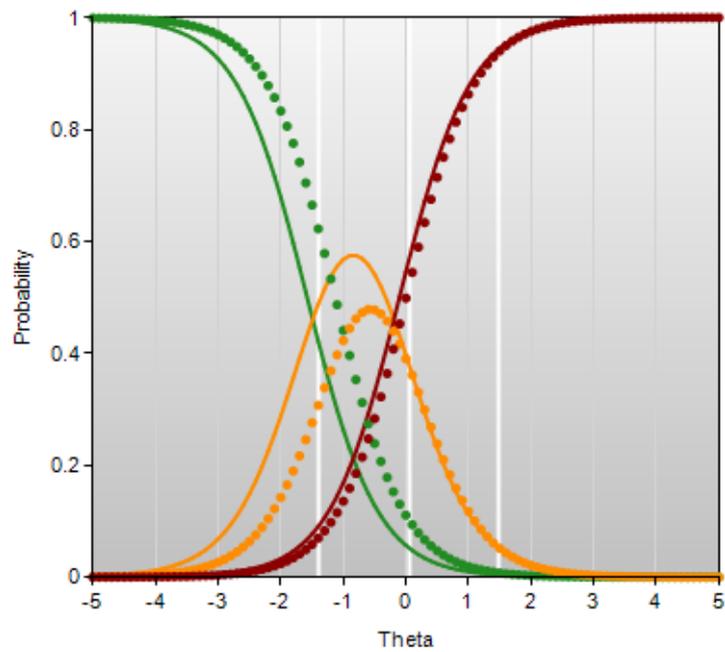

Beta Chart

Mathematics Grade 4: IA01055 (MA311572)
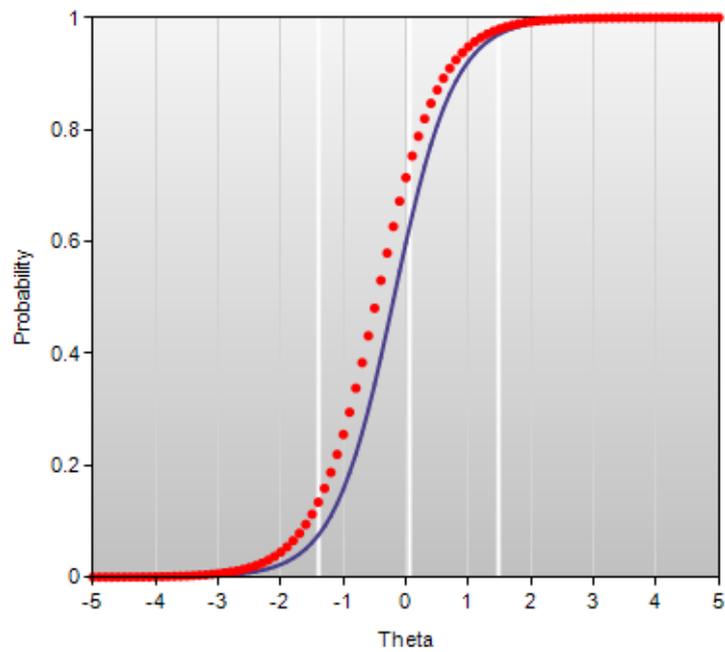
Initial Calibration

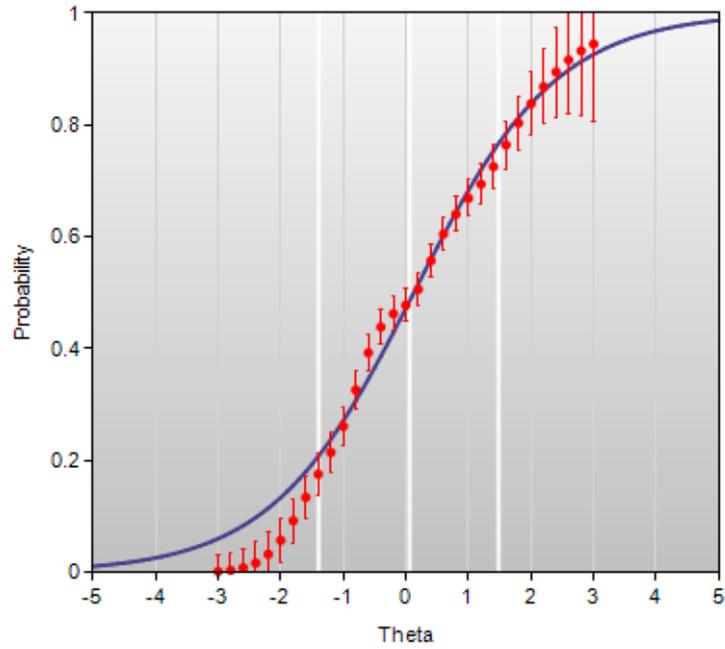## Mathematics Grade 4: IA01093

Beta Chart

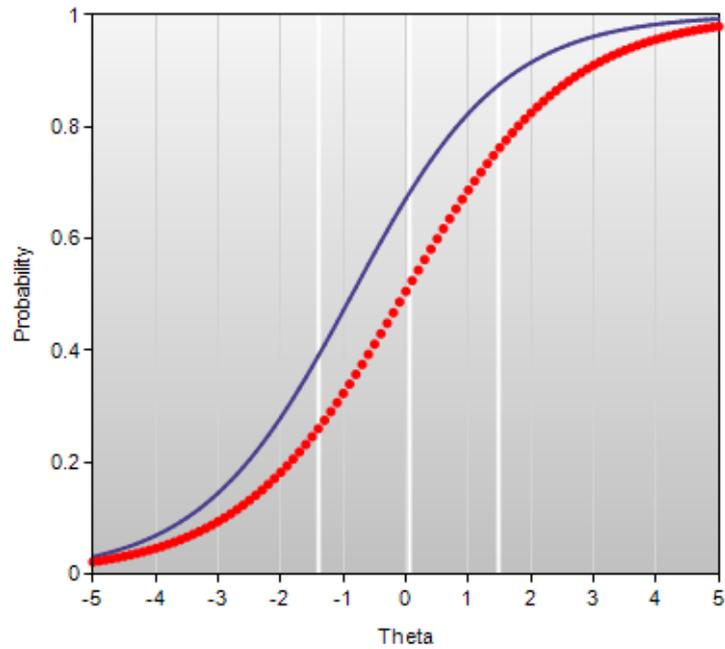## Mathematics Grade 4: IA01093 (MA623879088)

Initial Calibration

Mathematics Grade 4: IA02819
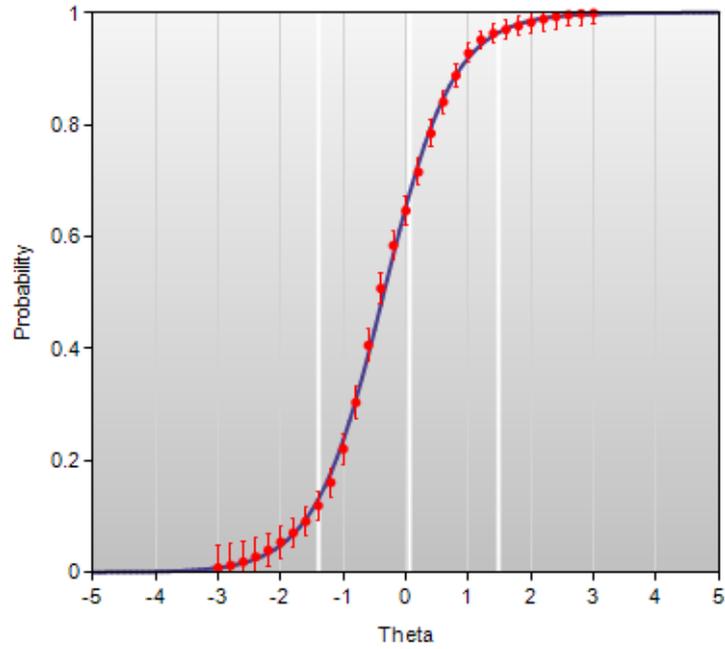

Beta Chart

Mathematics Grade 4: IA02819 (MA713583365)

Initial Calibration

Mathematics Grade 4: IA04965

Beta Chart

Mathematics Grade 4: IA04965 (MA800867144)

Initial Calibration

Mathematics Grade 5: IA02552
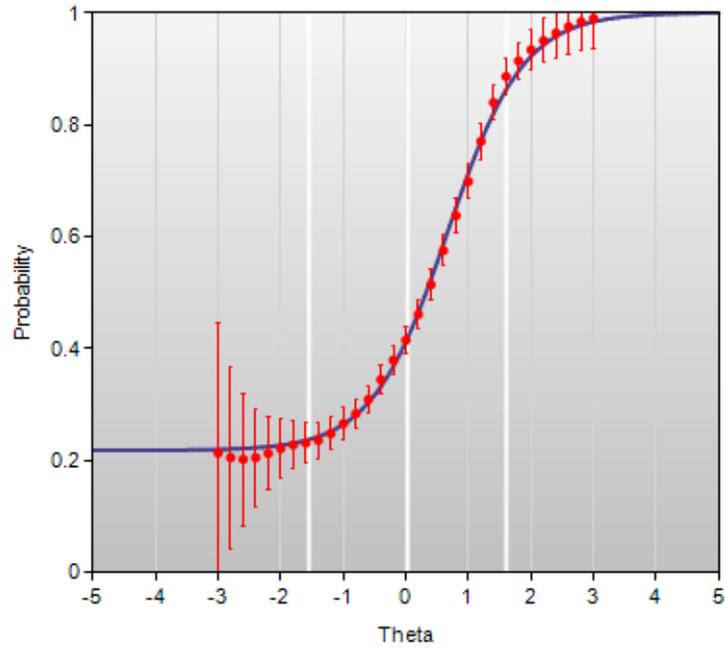
Beta Chart

Mathematics Grade 5: IA02552 (MA311324)
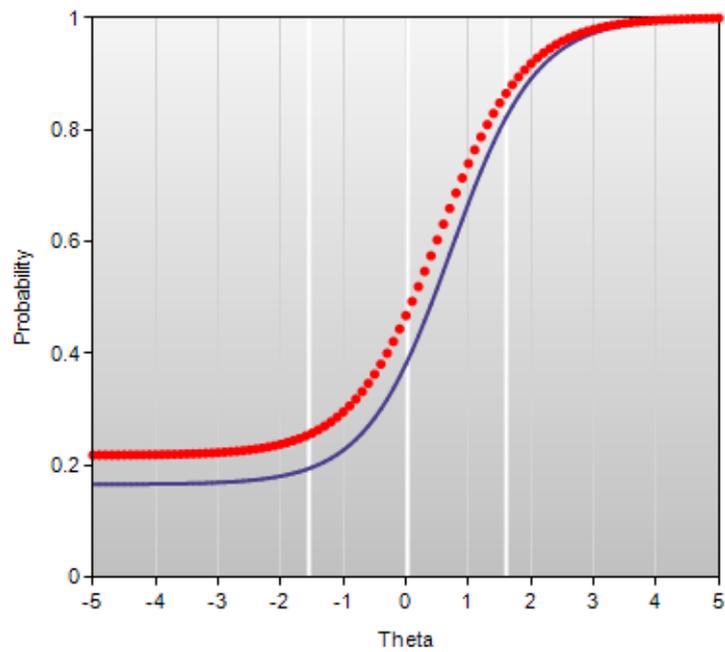
Initial Calibration

Mathematics Grade 5: IA05002

Beta Chart

Mathematics Grade 5: IA05002 (MA801652356)
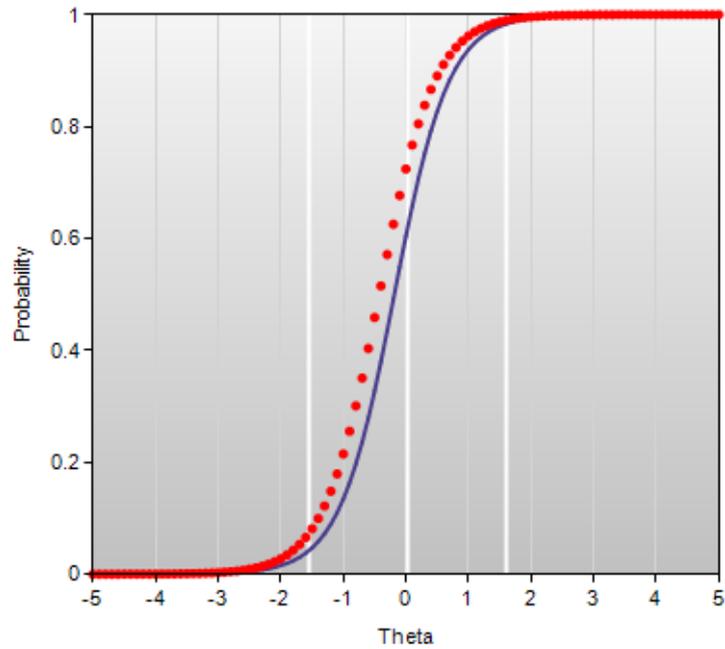
Initial Calibration

Mathematics Grade 7: IA10268
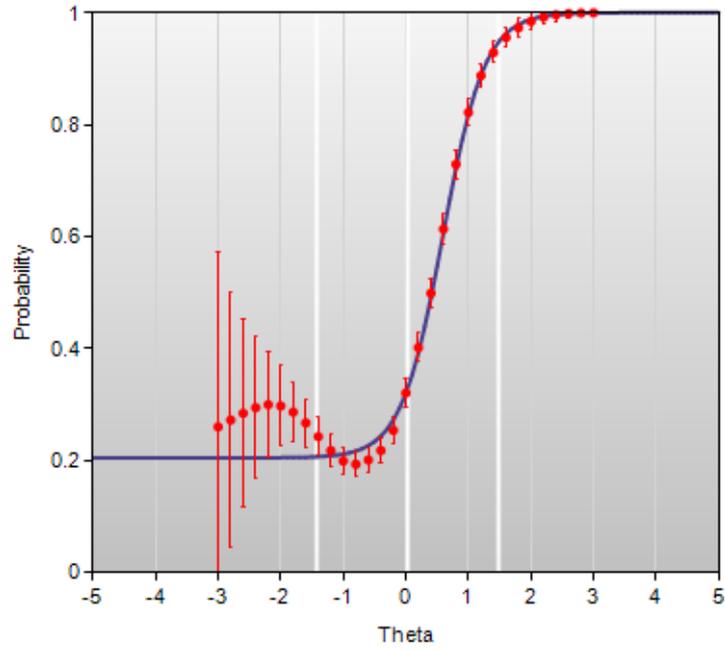
Beta Chart

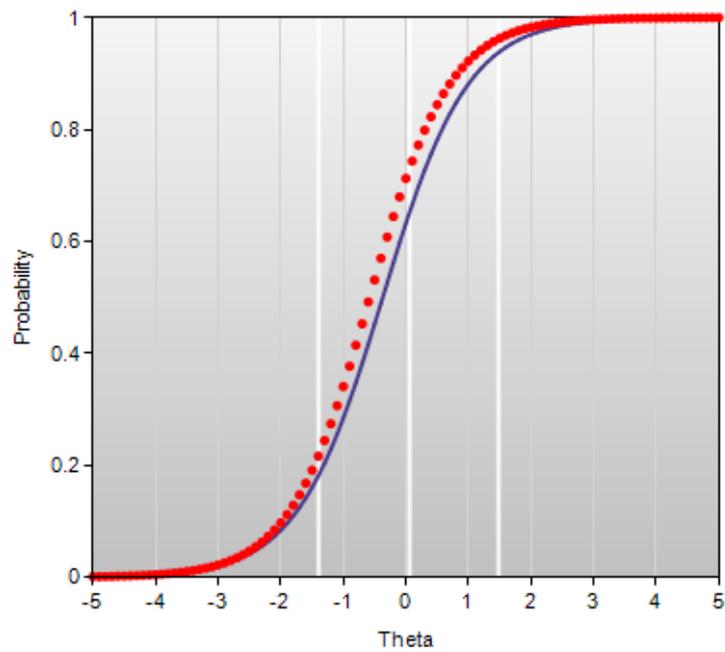## Mathematics Grade 7: IA10268 (MA005205470)

Initial Calibration

Mathematics Grade 8: IA04665

Beta Chart

Mathematics Grade 8: IA04665 (MA307399)

Initial Calibration

Mathematics Grade 8: IA04957

Beta Chart

Mathematics Grade 8: IA04957 (MA800770988)
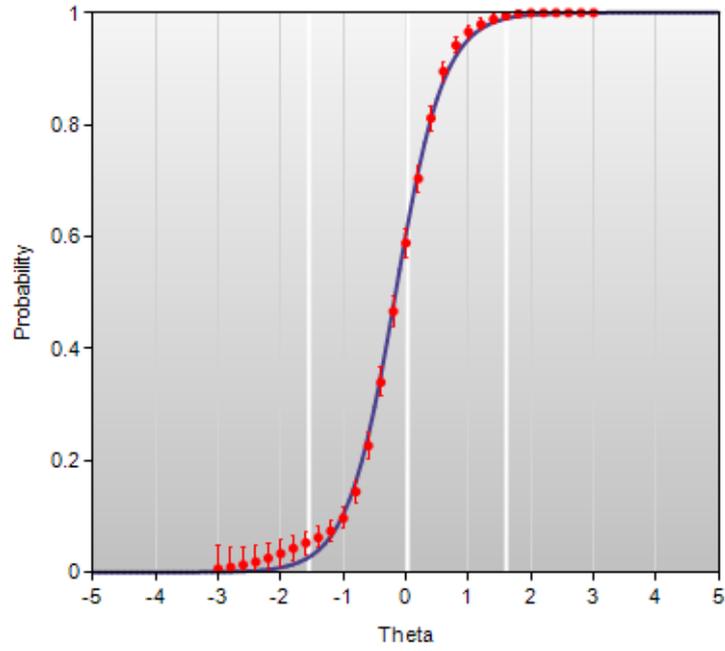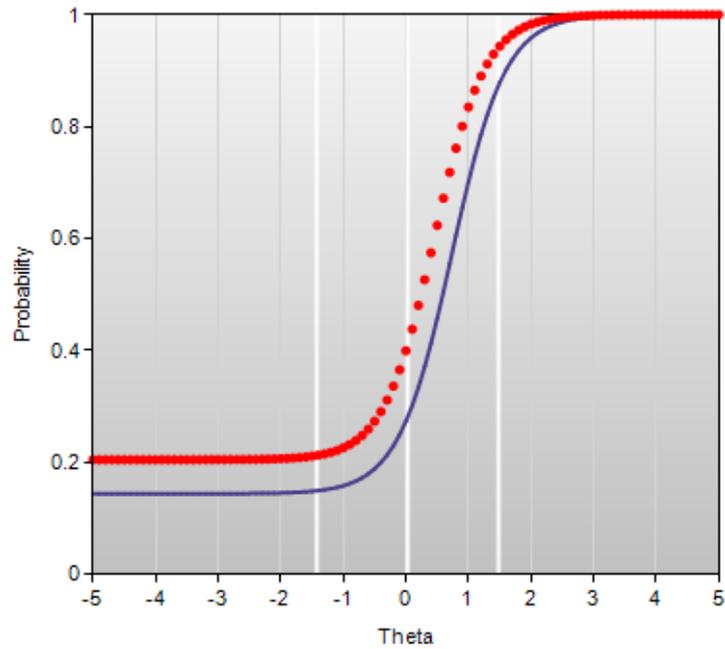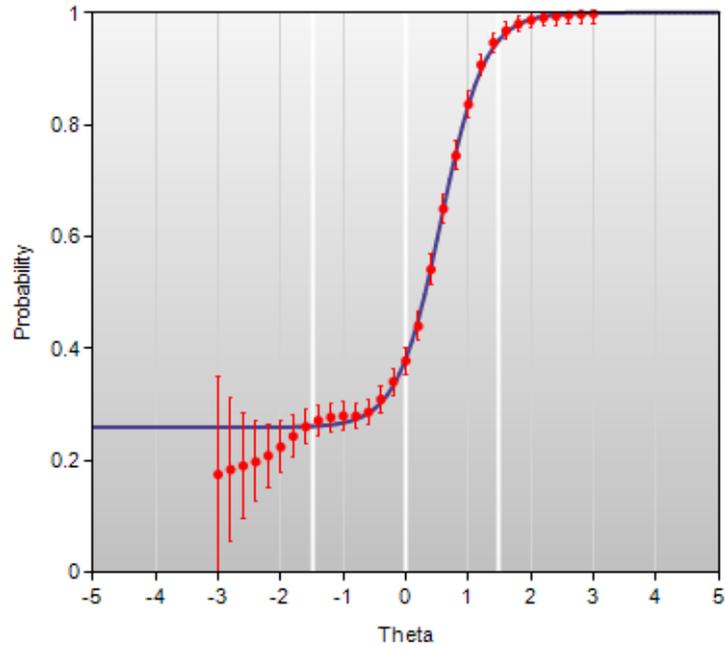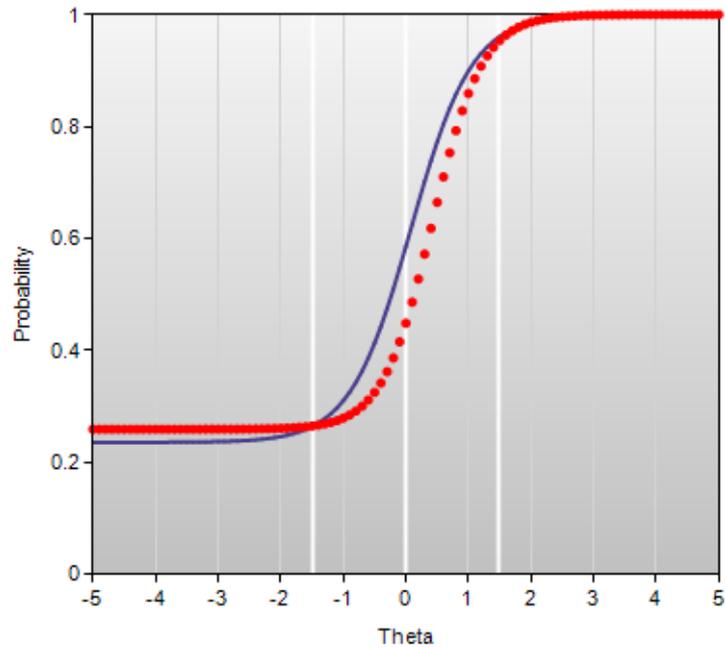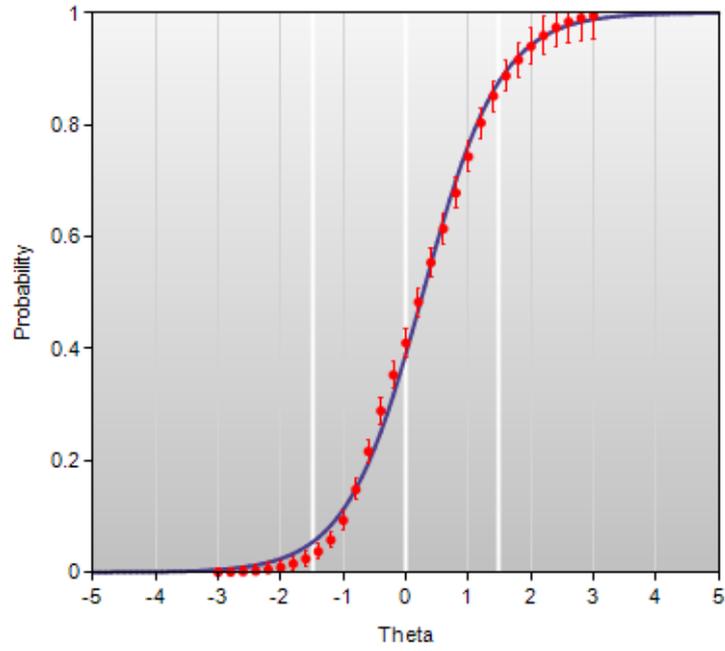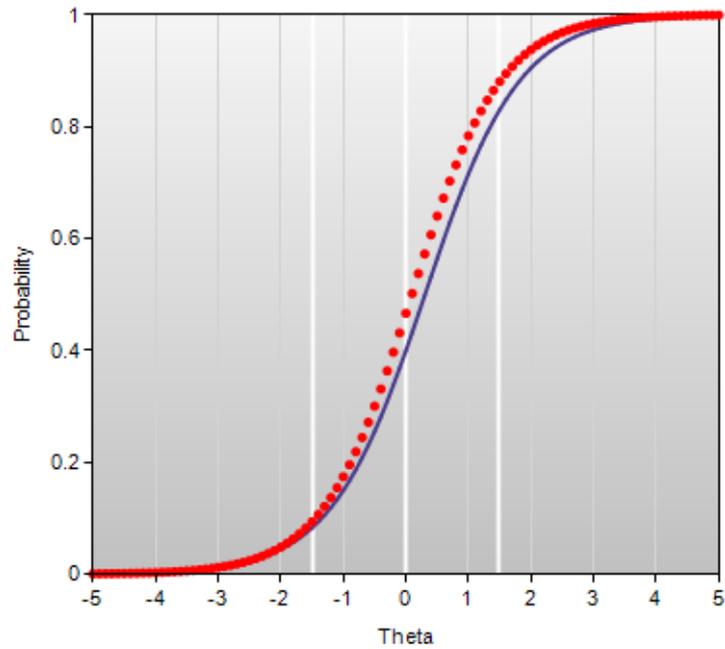
APPENDIX I

RELIABILITY

Table I-1: Subgroup Reliabilities—ELA

Grade	Subgroup	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	SEM
3	All Students	9,849	44	21.49	9.22	0.90	2.98
3	ELL	1,687	44	15.54	7.25	0.84	2.92
3	Economically Disadvantaged	5,716	44	18.50	8.43	0.88	2.96
3	African American	875	44	18.43	8.58	0.88	2.94
3	Asian	288	44	25.53	8.77	0.89	2.96
3	Hispanic	3,098	44	18.26	8.34	0.87	2.96
3	Native American/Alaska Native	63	44	14.52	7.82	0.87	2.81
3	White	4,921	44	23.90	9.02	0.89	2.97
3	Pacific Islander/Hawaiian	18	44	19.61	8.19	0.87	2.95
3	Multiracial	586	44	21.68	9.46	0.90	2.96
3	Male	4,962	44	20.94	9.15	0.90	2.96
3	Female	4,886	44	22.04	9.26	0.90	2.99
3	Special Education	1,969	44	14.50	7.55	0.86	2.84
4	All Students	9,495	44	24.46	9.25	0.91	2.84
4	ELL	1,478	44	17.96	7.76	0.86	2.92
4	Economically Disadvantaged	5,470	44	21.74	8.87	0.89	2.88
4	African American	833	44	21.32	8.95	0.90	2.87
4	Asian	325	44	27.40	9.38	0.91	2.75
4	Hispanic	2,930	44	21.49	8.81	0.89	2.90
4	Native American/Alaska Native	73	44	19.79	8.10	0.88	2.84
4	White	4,769	44	26.65	8.88	0.90	2.80
4	Pacific Islander/Hawaiian	16	44	17.63	9.08	0.90	2.82
4	Multiracial	549	44	25.06	9.02	0.90	2.80
4	Male	4,895	44	24.05	9.37	0.91	2.83
4	Female	4,600	44	24.90	9.10	0.90	2.84
4	Special Education	1,842	44	16.35	7.85	0.87	2.86
5	All Students	9,749	48	24.66	9.75	0.90	3.13
5	ELL	1,370	48	16.09	7.18	0.82	3.01
5	Economically Disadvantaged	5,589	48	21.48	9.08	0.88	3.11
5	African American	865	48	21.60	8.99	0.88	3.11
5	Asian	297	48	28.51	10.03	0.90	3.10
5	Hispanic	3,056	48	21.37	9.21	0.89	3.11
5	Native American/Alaska Native	79	48	17.96	8.09	0.86	3.07
5	White	4,876	48	27.20	9.35	0.89	3.12
5	Pacific Islander/Hawaiian	24	48	20.33	9.29	0.88	3.17
5	Multiracial	552	48	24.30	9.87	0.90	3.14
5	Male	5,022	48	24.21	9.91	0.90	3.13
5	Female	4,727	48	25.14	9.56	0.89	3.11
5	Special Education	1,834	48	15.85	7.97	0.86	2.95
6	All Students	9,583	50	26.06	11.11	0.91	3.27
6	ELL	1,147	50	14.47	7.67	0.84	3.10
6	Economically Disadvantaged	5,504	50	22.47	10.59	0.90	3.28
6	African American	843	50	23.09	10.49	0.90	3.27
6	Asian	333	50	30.80	10.92	0.92	3.18
6	Hispanic	3,013	50	22.21	10.87	0.91	3.29
6	Native American/Alaska Native	73	50	19.19	10.19	0.90	3.23
6	White	4,778	50	28.79	10.42	0.91	3.21
6	Pacific Islander/Hawaiian	13	50	26.92	12.14	0.94	3.08
6	Multiracial	530	50	26.04	11.22	0.92	3.24
6	Male	4,786	50	24.30	11.02	0.91	3.27
6	Female	4,795	50	27.81	10.92	0.91	3.22
6	Special Education	1,671	50	15.15	8.17	0.86	3.05
7	All Students	9,680	50	25.92	10.80	0.91	3.15
7	ELL	1,241	50	15.80	7.43	0.83	3.08
7	Economically Disadvantaged	5,571	50	22.18	10.05	0.90	3.17
7	African American	894	50	22.55	9.83	0.90	3.15

Grade	Subgroup	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	SEM
7	Asian	313	50	30.34	10.83	0.92	3.06
7	Hispanic	2,993	50	21.73	10.07	0.90	3.16
7	Native American/Alaska Native	79	50	18.87	9.23	0.88	3.18
7	White	4,896	50	29.08	10.25	0.91	3.10
7	Pacific Islander/Hawaiian	12	50	22.92	13.59	0.95	3.17
7	Multiracial	493	50	24.57	10.65	0.91	3.15
7	Male	4,971	50	24.45	10.80	0.91	3.17
7	Female	4,707	50	27.48	10.57	0.91	3.11
7	Special Education	1,652	50	15.79	7.92	0.85	3.03
8	All Students	9,800	50	27.39	11.26	0.91	3.30
8	ELL	1,309	50	15.97	7.86	0.84	3.18
8	Economically Disadvantaged	5,570	50	23.34	10.63	0.90	3.31
8	African American	1,017	50	22.98	10.66	0.90	3.31
8	Asian	313	50	31.98	10.49	0.91	3.17
8	Hispanic	2,876	50	22.84	10.76	0.90	3.32
8	Native American/Alaska Native	81	50	21.05	9.54	0.89	3.20
8	White	5,024	50	30.74	10.43	0.91	3.21
8	Pacific Islander/Hawaiian	22	50	27.41	10.95	0.91	3.29
8	Multiracial	467	50	27.00	11.11	0.91	3.29
8	Male	5,043	50	25.91	11.13	0.91	3.32
8	Female	4,754	50	28.95	11.18	0.92	3.24
8	Special Education	1,670	50	17.24	8.54	0.86	3.15

Table I-2. Subgroup Reliabilities—Mathematics

Grade	Subgroup	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	SEM
3	All Students	9,797	48	24.95	11.94	0.93	3.05
3	ELL	1,645	48	18.80	10.29	0.91	3.00
3	Economically Disadvantaged	5,650	48	21.11	11.11	0.93	3.04
3	African American	891	48	20.38	11.06	0.93	3.01
3	Asian	302	48	30.67	11.58	0.94	2.94
3	Hispanic	3,022	48	21.16	11.03	0.92	3.04
3	Native American/Alaska Native	58	48	17.12	11.08	0.93	2.87
3	White	4,923	48	27.89	11.64	0.93	3.02
3	Pacific Islander/Hawaiian	17	48	19.71	9.80	0.90	3.09
3	Multiracial	584	48	24.72	12.13	0.94	3.04
3	Male	4,953	48	25.69	12.06	0.94	3.04
3	Female	4,843	48	24.19	11.77	0.93	3.04
3	Special Education	1,951	48	16.21	10.91	0.93	2.91
4	All Students	9,403	54	27.52	13.03	0.94	3.32
4	ELL	1,396	54	20.36	10.98	0.91	3.21
4	Economically Disadvantaged	5,368	54	23.54	12.19	0.93	3.29
4	African American	843	54	22.58	12.15	0.93	3.27
4	Asian	330	54	32.98	13.25	0.94	3.22
4	Hispanic	2,823	54	23.45	12.03	0.93	3.29
4	Native American/Alaska Native	66	54	21.06	10.59	0.91	3.15
4	White	4,774	54	30.55	12.72	0.93	3.31
4	Pacific Islander/Hawaiian	17	54	16.65	10.05	0.90	3.21
4	Multiracial	550	54	27.48	13.38	0.94	3.29
4	Male	4,851	54	28.63	13.28	0.94	3.31
4	Female	4,552	54	26.34	12.66	0.93	3.31
4	Special Education	1,829	54	17.08	10.91	0.92	3.10
5	All Students	9,716	54	25.38	12.34	0.92	3.50
5	ELL	1,342	54	16.65	8.54	0.86	3.16
5	Economically Disadvantaged	5,546	54	21.19	10.86	0.90	3.40

Grade	Subgroup	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	SEM
5	African American	889	54	20.79	10.78	0.90	3.36
5	Asian	309	54	31.80	13.17	0.93	3.49
5	Hispanic	2,999	54	21.11	10.78	0.90	3.39
5	Native American/Alaska Native	75	54	18.17	9.68	0.89	3.28
5	White	4,869	54	28.69	12.28	0.92	3.53
5	Pacific Islander/Hawaiian	24	54	20.00	9.92	0.89	3.28
5	Multiracial	551	54	24.41	12.50	0.92	3.49
5	Male	5,014	54	26.35	12.90	0.93	3.52
5	Female	4,702	54	24.34	11.63	0.91	3.47
5	Special Education	1,831	54	15.50	9.34	0.89	3.14
6	All Students	9,515	54	23.98	12.54	0.92	3.54
6	ELL	1,082	54	14.01	8.08	0.86	3.06
6	Economically Disadvantaged	5,411	54	19.53	10.69	0.90	3.40
6	African American	863	54	18.99	10.70	0.90	3.38
6	Asian	342	54	32.07	13.53	0.93	3.54
6	Hispanic	2,909	54	19.44	10.75	0.90	3.40
6	Native American/Alaska Native	74	54	17.20	10.54	0.90	3.25
6	White	4,785	54	27.18	12.49	0.92	3.57
6	Pacific Islander/Hawaiian	13	54	27.69	13.99	0.93	3.63
6	Multiracial	529	54	23.82	12.73	0.92	3.55
6	Male	4,754	54	24.18	12.86	0.92	3.54
6	Female	4,759	54	23.79	12.21	0.92	3.52
6	Special Education	1,667	54	13.29	8.19	0.86	3.03
7	All Students	9,560	54	20.20	11.69	0.93	3.19
7	ELL	1,163	54	11.89	6.22	0.80	2.79
7	Economically Disadvantaged	5,436	54	16.05	9.18	0.89	3.03
7	African American	909	54	15.57	8.97	0.89	2.99
7	Asian	320	54	27.39	13.95	0.94	3.32
7	Hispanic	2,874	54	15.57	8.97	0.89	3.01
7	Native American/Alaska Native	76	54	12.99	8.07	0.88	2.75
7	White	4,883	54	23.57	12.09	0.93	3.25
7	Pacific Islander/Hawaiian	13	54	18.00	13.91	0.95	3.02
7	Multiracial	485	54	18.85	11.17	0.92	3.12
7	Male	4,906	54	20.83	12.03	0.93	3.19
7	Female	4,652	54	19.54	11.29	0.92	3.19
7	Special Education	1,616	54	11.67	6.65	0.83	2.77
8	All Students	9,664	54	23.26	12.64	0.93	3.41
8	ELL	1,212	54	13.27	7.00	0.82	2.98
8	Economically Disadvantaged	5,427	54	18.53	10.54	0.90	3.28
8	African American	1,036	54	17.61	10.69	0.91	3.26
8	Asian	325	54	29.95	14.44	0.94	3.41
8	Hispanic	2,736	54	18.34	10.36	0.90	3.29
8	Native American/Alaska Native	76	54	16.92	10.15	0.91	3.11
8	White	5,002	54	26.84	12.64	0.93	3.43
8	Pacific Islander/Hawaiian	22	54	23.73	12.90	0.93	3.39
8	Multiracial	467	54	22.54	12.42	0.93	3.38
8	Male	4,976	54	23.04	12.85	0.93	3.39
8	Female	4,685	54	23.49	12.40	0.92	3.42
8	Special Education	1,659	54	13.20	7.50	0.84	2.96

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

Grade	Item Reporting Category	Label	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	SEM
3	1	Reading	22	28	14.67	6.12	0.84	2.44
3	2	Language	9	12	6.05	2.89	0.71	1.54
3	3	Writing	1	4	0.76	0.89	NaN	NaN
4	1	Reading	24	31	18.87	6.85	0.88	2.42
4	2	Language	7	9	4.47	2.18	0.64	1.30
4	3	Writing	1	4	1.11	0.89	NaN	NaN
5	1	Reading	23	28	15.16	5.91	0.82	2.50
5	2	Language	8	12	6.73	2.82	0.73	1.45
5	3	Writing	2	8	2.77	1.89	0.85	0.72
6	1	Reading	23	28	16.69	6.43	0.86	2.41
6	2	Language	8	12	6.17	3.22	0.73	1.67
6	3	Writing	2	10	3.20	2.34	0.90	0.75
7	1	Reading	23	28	16.28	6.44	0.86	2.37
7	2	Language	8	12	6.36	3.07	0.71	1.66
7	3	Writing	2	10	3.28	2.09	0.90	0.67
8	1	Reading	24	29	17.49	6.66	0.86	2.49
8	2	Language	7	11	6.45	3.05	0.73	1.57
8	3	Writing	2	10	3.45	2.41	0.90	0.76

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

Grade	Item Reporting Category	Label	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	SEM
3	1	Operations and Algebraic Thinking	13	15	9.14	3.87	0.82	1.65
3	2	Number and Operations in Base Ten	6	8	3.82	2.61	0.76	1.27
3	3	Number and Operations-Fractions	7	9	4.36	2.42	0.72	1.29
3	4	Measurement and Data	10	12	6.00	3.28	0.78	1.56
3	5	Geometry	4	4	1.63	1.22	0.49	0.87
4	1	Operations and Algebraic Thinking	8	11	5.18	3.06	0.74	1.55
4	2	Number and Operations in Base Ten	9	10	5.78	2.70	0.75	1.34
4	3	Number and Operations-Fractions	12	16	8.29	4.21	0.82	1.79
4	4	Measurement and Data	8	11	5.69	3.00	0.73	1.54
4	5	Geometry	3	6	2.58	1.59	0.52	1.10
5	1	Operations and Algebraic Thinking	5	8	3.70	2.25	0.59	1.44
5	2	Number and Operations in Base Ten	12	16	7.69	4.24	0.82	1.78
5	3	Number and Operations-Fractions	11	14	6.72	3.39	0.70	1.85
5	4	Measurement and Data	6	10	4.20	2.63	0.62	1.62
5	5	Geometry	6	6	3.07	1.51	0.56	1.00
6	1	Ratios and Proportional Relationships	8	11	5.05	2.88	0.66	1.69
6	2	The Number System	8	11	6.10	2.90	0.67	1.66
6	3	Expressions and Equations	12	16	6.46	4.00	0.80	1.79
6	4	Geometry	5	8	3.14	2.40	0.67	1.37
6	5	Statistics and Probability	7	8	3.23	2.11	0.59	1.34
7	1	Ratios and Proportional Relationships	10	11	5.12	2.70	0.74	1.38
7	2	The Number System	8	11	4.06	2.65	0.69	1.48
7	3	Expressions and Equations	9	13	4.06	3.18	0.78	1.49
7	4	Geometry	5	8	2.47	2.13	0.60	1.34
7	5	Statistics and Probability	8	11	4.49	2.49	0.65	1.48
8	1	The Number System and Expressions and Equations	17	21	8.10	5.06	0.83	2.08
8	2	Functions	8	11	5.28	2.96	0.66	1.73
8	3	Geometry	12	16	6.44	4.13	0.82	1.74
8	4	Statistics and Probability	3	6	3.44	1.73	0.57	1.14