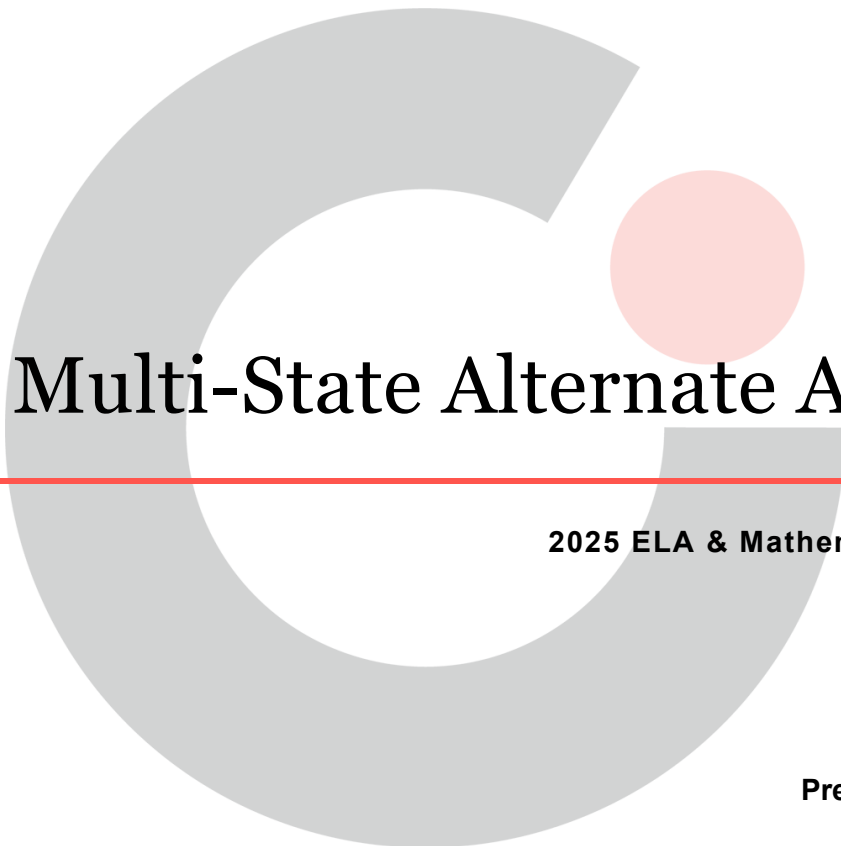




Multi-State Alternate Assessment

2025 ELA & Mathematics Technical Report

2/9/26

Prepared by Cognia for MSAA

TABLE OF CONTENTS

CHAPTER 1. OVERVIEW OF MSAA AND 2025 UPDATES	6
1.1 PURPOSES AND USES OF THE MSAA	6
1.2 INTENDED MSAA SCORE INTERPRETATIONS AND USES	6
1.3 VALIDITY EVIDENCE AND ASSUMPTIONS FOR THE MSAA	7
1.4 UPDATES FOR THE 2025 PROGRAM	8
1.4.1 TRACKING MSAA LONGITUDINAL STUDENT PERFORMANCE	8
CHAPTER 2. OVERVIEW OF THE MSAA	11
2.1 HISTORY OF THE MSAA	11
2.1.1 DEVELOPMENTS IN 2016-17	12
2.1.2 DEVELOPMENTS IN 2017-18	12
2.1.3 DEVELOPMENTS IN 2018-19	13
2.1.4 DEVELOPMENTS IN 2020-21	13
2.1.5 DEVELOPMENTS IN 2022-23	14
2.1.6 DEVELOPMENTS IN 2023-24	14
2.2 STAKEHOLDERS	14
2.3 MSAA PARTICIPATION	16
2.4 MSAA PARTICIPATION RATE	18
CHAPTER 3. TEST DEVELOPMENT-CONTENT AND ADMINISTRATION	19
3.1 HISTORY OF ALTERNATE ACADEMIC ACHIEVEMENT STANDARDS AND CORE CONTENT CONNECTORS	19
3.1.1 THE LEARNING PROGRESSION FRAMEWORKS	19
3.1.2 CORE CONTENT CONNECTORS	19
3.2 ALIGNMENT AND LINKAGES	21
3.3 2025 MSAA ASSESSMENT DESIGN	22
3.3.1 OPERATIONAL DESIGN	22
3.3.2 OPERATIONAL ITEMS AND EMBEDDED ITEMS IN FIELD TEST POSITIONS	24
3.3.3 ITEM DESIGN AND ADMINISTRATION	25
3.3.4 ITEM COMPONENTS	27
3.3.4.1 SELECTED-RESPONSE: READING, WRITING (STAND-ALONE ITEMS AND MULTIPLE-PART SELECTED-RESPONSE WRITING PROMPT), MATHEMATICS	27
3.3.4.2 CONSTRUCTED-RESPONSE: MATHEMATICS, ELA	28
3.3.4.3 OPEN-RESPONSE: WRITING PROMPT	28
3.4 CONTENT AND BLUEPRINTS	28
3.4.1 ENGLISH LANGUAGE ARTS	28
3.4.2 MATHEMATICS	29
CHAPTER 4. TEST DEVELOPMENT-STAKEHOLDER INVOLVEMENT	31
4.1 GENERAL PHILOSOPHY AND ROLE OF THE ITEM DEVELOPMENT AND PSYCHOMETRIC SUBCOMMITTEES AND OTHER STAKEHOLDERS IN TEST DEVELOPMENT	31
4.2 SAMPLE ITEM TEACHER GUIDE	33

CHAPTER 5. TRAINING AND ADMINISTRATION	35
5.1 TEST ADMINISTRATOR AND TEST COORDINATOR TRAINING	35
5.2 TEST ADMINISTRATOR TRAINING MODULES	35
5.3 TEST COORDINATOR TRAINING MODULES	36
5.4 BEST PRACTICE VIDEOS	36
5.5 TEST ADMINISTRATION MANUAL	37
5.6 DIRECTIONS FOR TEST ADMINISTRATION (DTA)	37
5.7 TEST COORDINATOR AND TEST ADMINISTRATOR USER GUIDES	38
5.8 OPERATIONAL ADMINISTRATION	38
5.8.1 MSAA SERVICE CENTER	38
5.8.2 ADDITIONAL SUPPORTS	39
5.8.3 MONITORING AND QUALITY CONTROL	39
5.8.4 END OF TEST SURVEY RESULTS	40
CHAPTER 6. SCORING	43
6.1 SELECTED-RESPONSE AND CONSTRUCTED-RESPONSE ITEM SCORING PROCESSES	43
6.1.1 OVERVIEW OF SCORING PROCESS WITHIN THE SYSTEM AND TEST ADMINISTRATOR/SCORER TRAINING	43
6.2 OPEN-RESPONSE WRITING PROMPTS SCORING PROCESSES	44
6.2.1 OVERVIEW OF OPEN-RESPONSE WRITING ENTRY PROCESS WITHIN THE ASSESSMENT SYSTEM AND TEST ADMINISTRATOR TRAINING	44
6.2.2 BENCHMARKING AND IDENTIFICATION OF SCORING MATERIALS	44
6.2.3 SCORER RECRUITMENT AND QUALIFICATIONS	45
6.2.4 COGNIA STAFF AND SCORING LEADERSHIP	45
6.2.5 TRAINING	45
6.2.6 QUALIFICATION	46
6.2.7 METHOD FOR SCORING OPERATIONAL OPEN-RESPONSE WRITING PROMPTS	47
6.2.8 MONITORING OF SCORING QUALITY CONTROL	49
6.2.8.1 CALIBRATION SETS	50
6.2.8.2 READ-BEHIND SCORING	50
6.2.8.3 DOUBLE-BLIND SCORING	50
6.2.8.4 FINAL SCORE RESOLUTION	51
6.2.9 QUALITY AND PRODUCTION MANAGEMENT REPORTS	51
6.2.10 INTERRATER AGREEMENT	52
CHAPTER 7. REPORTING	53
7.1 DEVELOPMENT AND APPROVAL OF REPORT SPECIFIC DOCUMENTS	53
7.2 SPECIFIC PRIMARY REPORTS GENERATED FOR SCHOOLS, DISTRICTS, AND STATES	54
7.2.1 STUDENT REPORT	55
7.2.2 SCHOOL ROSTER REPORT	55
7.2.3 SUMMARY REPORTS	56
7.2.4 QUALITY ASSURANCE	56
CHAPTER 8. PRELIMINARY STATISTICAL ANALYSES	58
8.1 DIFFERENTIAL ITEM FUNCTIONING (DIF)	58

8.2 DIMENSIONALITY ANALYSIS	60
CHAPTER 9. ITEM RESPONSE THEORY SCALING AND EQUATING	64
9.1 ITEM RESPONSE THEORY	64
9.2 CALIBRATION PROCEDURE	66
9.3 ITEM RESPONSE THEORY RESULTS	66
9.4 EQUATING	70
9.5 REPORTED SCALE SCORES	71
9.6 MSAA PERFORMANCE LEVELS, CUT SCORES, AND STANDARDS VALIDATION	73
CHAPTER 10. RELIABILITY	76
10.1 IRT MARGINAL RELIABILITY	76
10.2 SUBGROUP RELIABILITY	77
10.3 RELIABILITY OF PERFORMANCE-LEVEL CATEGORIZATION: ACCURACY AND CONSISTENCY	78
CHAPTER 11. VALIDITY OF THE MSAA PRIMARY AND SECONDARY CLAIMS	82
11.1 THE MSAA VALIDATION FRAMEWORK, EVIDENCE SYNTHESIS, AND SUFFICIENCY	82
11.2 EVALUATING THE PRIMARY SCORE CLAIM	85
11.2.1 CATEGORY A: CONSTRUCT CONCEPTUALIZATION AND REPRESENTATION.....	85
11.2.2 CATEGORY B: FAIRNESS	88
11.2.3 CATEGORY C: ADMINISTRATION AND SCORING	89
11.2.4 CATEGORY D: PRECISION AND GENERALIZATION	92
11.2.5 CATEGORY E: SCORE REPORTING	95
REFERENCES.....	98
APPENDICES.....	101

APPENDIX A ACCOMMODATION FREQUENCIES

APPENDIX B TEST PARTICIPATION

APPENDIX C TEST DESIGN BLUEPRINTS

APPENDIX D PANELISTS AND COMMITTEE MEMBERS

APPENDIX E TEACHER GUIDE SAMPLE ITEMS

APPENDIX F PROCESSING AND REPORTING BUSINESS REQUIREMENTS

APPENDIX G MSAA 2025 GUIDE FOR SCORE REPORT INTERPRETATION

APPENDIX H DIFFERENTIAL ITEM FUNCTIONING RESULTS

APPENDIX I ITEM RESPONSE THEORY PARAMETERS

APPENDIX J TEST CHARACTERISTIC CURVES & TEST INFORMATION FUNCTIONS

APPENDIX K RAW TO SCALED SCORE LOOKUP TABLES

APPENDIX L SCORE DISTRIBUTIONS

APPENDIX M IRT SUBGROUP RELIABILITY

APPENDIX N

DECISION ACCURACY AND CONSISTENCY RESULTS

APPENDIX O

LIST OF ACRONYMS & COMMONLY USED TERMS

APPENDIX P

MSAA VALIDITY SURVEY

Appendix References by Chapter

Chapter	1	2	3	4	5	6	7	8	9	10	11
Appendix	N/A	A, B	C	D, E		F, G	G	H	I, J, K, L, M	M, N	A, C, E, G, H, J, K, L, M, N, P

Chapter 1. Overview of MSAA and 2025 Updates

1.1 Purposes and Uses of the MSAA

The Multi-State Alternate Assessment (MSAA) is a comprehensive, two-stage adaptive summative assessment system developed to support higher academic achievement for students with the most significant cognitive disabilities, with the goal of preparing them for a wider range of post-secondary opportunities. The MSAA assesses grade-level academic content that is aligned with and derived from the content standards of MSAA Partner States. The assessment includes a variety of embedded supports, enabling students to use familiar materials and independently demonstrate their knowledge and skills to the extent possible. The MSAA is administered in English language arts (ELA) and mathematics for students in grades 3–8 and high school.

The long-term goal of the MSAA Partners is to support students with the most significant cognitive disabilities in achieving progressively higher academic outcomes and graduating from high school prepared to pursue meaningful post-secondary opportunities. A high-quality summative assessment, while important, is not sufficient on its own to accomplish this goal. The MSAA functions as one part of a broader system that includes curriculum, instruction, and professional development, all designed to ensure students can access and engage with grade-level academic content aligned to state standards.

The MSAA is aligned with Alternate Academic Achievement Standards (AA-AAAS) as outlined in the Elementary and Secondary Education Act (ESEA), which requires that all students participate in assessments that measure achievement of grade-level content standards. The MSAA was specifically developed to ensure that students with the most significant cognitive disabilities have access to a summative assessment that accurately reflects what they know and can do in relation to these standards. To demonstrate that MSAA English Language Arts and Mathematics scores validly reflect proficiency with the AA-AAAS, this technical report presents supporting evidence drawn from test design documentation, administration procedures, and psychometric analyses.

1.2 Intended MSAA Score Interpretations and Uses

MSAA is designed, developed, and implemented to support four primary intended score interpretations and uses, described in the following sections.

Primary Intended MSAA Score Interpretation (Primary Claim)

MSAA scores provide trustworthy information about the grade-level knowledge and skills of students with significant cognitive disabilities in the subjects of Mathematics and English Language Arts.

Primary Intended MSAA Score Uses (Secondary Claims)

1. Schools and districts use the MSAA and its results to monitor trends in student performance.
2. Teachers use MSAA test results and supporting resources to inform instructional planning.
3. A parent or guardian uses MSAA test results to
 - (a) better understand what a student knows and can do, and

(b) as a source of evidence about the progress a student is making on annual learning goals

The primary and secondary claims stated here align with the original statements of intended score interpretations and uses in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* and were further refined in workshops with TAC members, MSAA partners, and the Center for Assessment in 2018 and again in 2025 (See the “claim” and “uses” statements on page 8.)

A synthesis of the evidence and assumptions underlying the primary and secondary claims of the MSAA are presented in Chapter 11, which includes a holistic appraisal of the evidence supporting each assumption of the primary claim. Additionally, chapters 3-10 will each conclude with a summary of the relevant validity of evidence presented, explicitly linking the chapter to the primary claim assumptions it supports.

1.3 Validity Evidence and Assumptions for the MSAA

The 2025 technical report describes several procedural and psychometric processes of the MSAA program. These processes contribute to the accumulation of validity evidence to support MSAA score interpretations and uses. This report presents documentation to help substantiate the intended interpretations and uses of MSAA test scores (AERA et al., 2014). Each section in this report contributes important information about the MSAA tests: test design and development, test alignment, test administration, scoring, reliability, performance levels, and reporting. We use the word assumption in the following sense: what are we assuming to be true when we make a claim? A key idea here is that any assumption is one that is subject to corroboration (or falsification) through logical or empirical analysis. When an assumption is supported by evidence, it provides a “warrant” or “backing” for the claim.

The phrase “intended score interpretations for uses” appears several times in the *Standards for Educational and Psychological Testing* (AERA et al., 2014) and is an important representation of the field’s views on validity. For the MSAA and other assessment programs, the phrase refers broadly to **test scores** (e.g., test scale scores, aggregations of test scores, the percentage of students at or above Level 3), and other test performance information elements (e.g., the definition of Level 3 in the performance-level descriptors). The *Standards for Educational and Psychological Testing* provides a framework for describing areas of evidence that should be considered when constructing a validity argument: test content, response processes, internal structure, relationship to other variables, and consequences of testing. These sources address different aspects of evidence supporting the reasonableness of assumptions that are critical to the valid interpretation and use of test scores; they are not distinct types of validity.

Supporting those assumptions contributes to a body of evidence regarding the overall validity of score interpretations and uses (sometimes referred to as a validity argument). Moreover, these sources represent only a partial list of sources of evidence from the MSAA design, development, test administration, analysis, and reporting processes that are relevant to the primary and secondary claims of the MSAA. While technical advisors, psychometricians and peer reviewers look to their professional experience and industry norms to judge the sufficiency of evidence supporting the interpretation and use of test scores, readers should understand that validity evidence is never complete; it always can be added to, refreshed, replicated, and improved upon. The evaluation and sufficiency (and suggestions for further study) of the MSAA featured in Chapter 11, therefore, are best seen through that lens of continuous improvement.

1.4 Updates for the 2025 Program

The MSAA Partners for 2025 include American Samoa, Arizona, the Bureau of Indian Education, the Department of Defense Education Activity (DoDEA), Guam, Maine, Montana, the Commonwealth of the Northern Mariana Islands (CNMI), South Dakota, the U.S. Virgin Islands, Vermont, and Washington, D.C.

In 2025 the forms administered were a reuse of the forms from 2023. The program continued the practice of each grade having its own Directions for Test Administration (DTA) that corresponds with the items in the online system to emulate and standardize the student testing experience.

The 2025 administration test documentation was updated to reflect changes in the *Test Administration Manual (TAM)*, *MSAA Online Assessment System User Guide for Test Administrators*, *MSAA Online Assessment System User Guide for Test Coordinators*, *Directions for Test Administration (DTA)*, and the *MSAA 2025 Guide for Score Report Interpretation*. These documents and the online training modules were revised to streamline information and provide more clarity to Test Administrators (TAs) and Test Coordinators (TCs). Additional detailed information is available in Chapter 5.

The 2025 End-of-Test Survey (EOTS) was updated to include a required question asking respondents to identify the state or entity for which they administered the MSAA. This addition provided value to MSAA Partners by allowing them to view feedback specific to their state or entity.

1.4.1 Tracking MSAA Longitudinal Student Performance

Each year, MSAA students receive a scale score and performance level for each assessment. To track performance trends, we monitor the percentage of students in each performance level over time, aiming for a shift toward higher levels as teaching improves and assessments are equated with consistent cut scores. This analysis offers insights into overall improvements and teaching effectiveness across MSAA states and entities. Figures 1-1 and 1-2 show student performance trends in ELA and Mathematics from 2021 to 2025, with each line representing the percentage of students at each performance level. For more information about each performance level and how these levels are determined, see Chapter 9 (9.2).

Figures 1-1 and 1-2 depict the performance level trends for students across various grades (3-8, and High School) in ELA and Mathematics over the years 2021 to 2025. Each grade-level chart shows the percentage of students in four performance levels (1-4), represented by different colored lines:

- Red Line with circles (Level 1): Represents students performing at Level 1.0
- Green Line with triangles (Level 2): Represents students performing at Level 2.
- Teal Line with squares (Level 3): Represents students performing at Level 3, which indicates meeting expectations.
- Purple Line with stars (Level 4): Represents students performing at Level 4, which also indicates meeting expectations.

Overall, the trends suggest slightly decreasing performance across the years (the percentage of students in level 1 increasing and in level 4 decreasing), with slight variations in each performance level for different grade levels.

Figure 1-1 ELA Student Performance Over Time

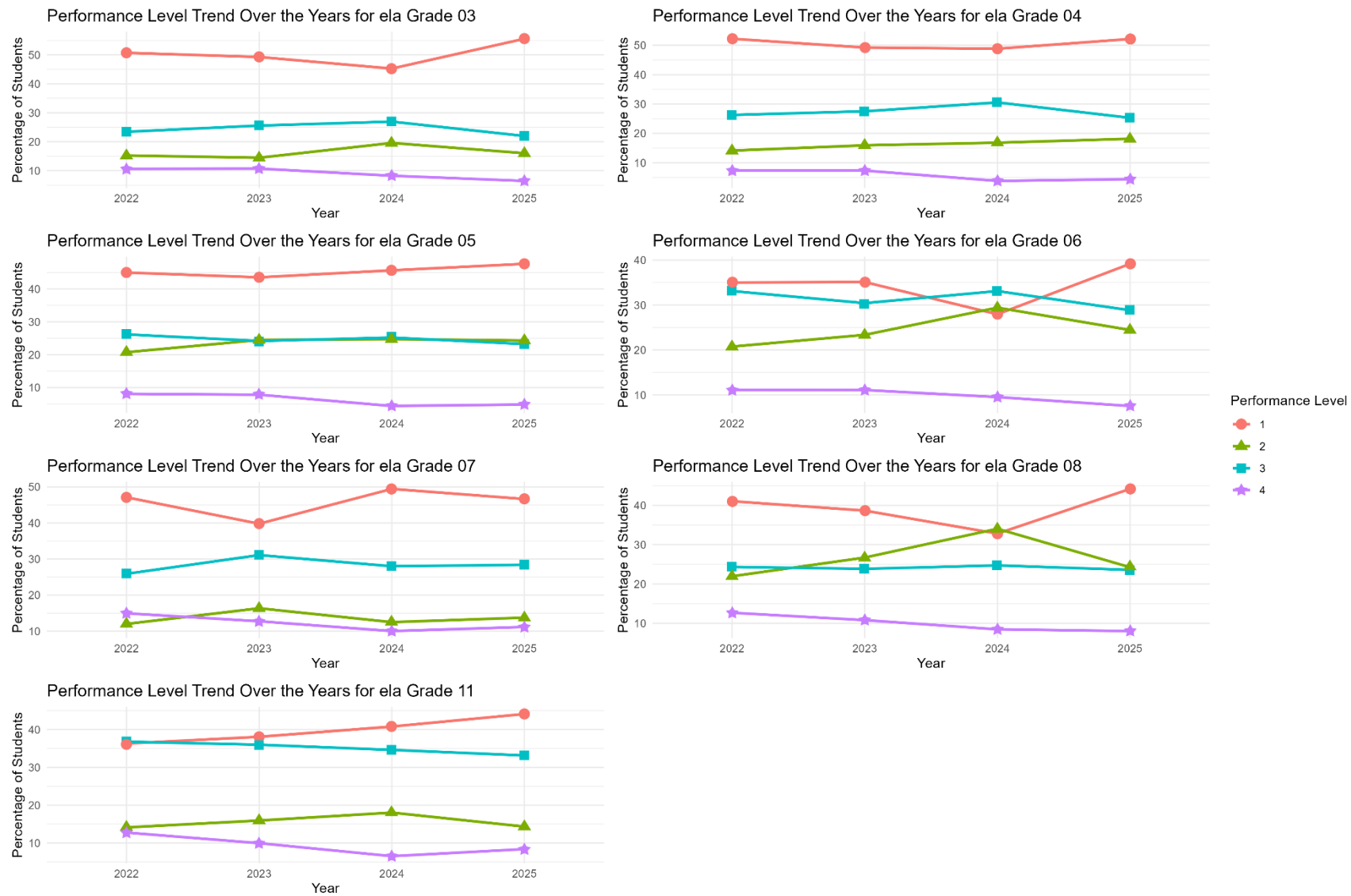
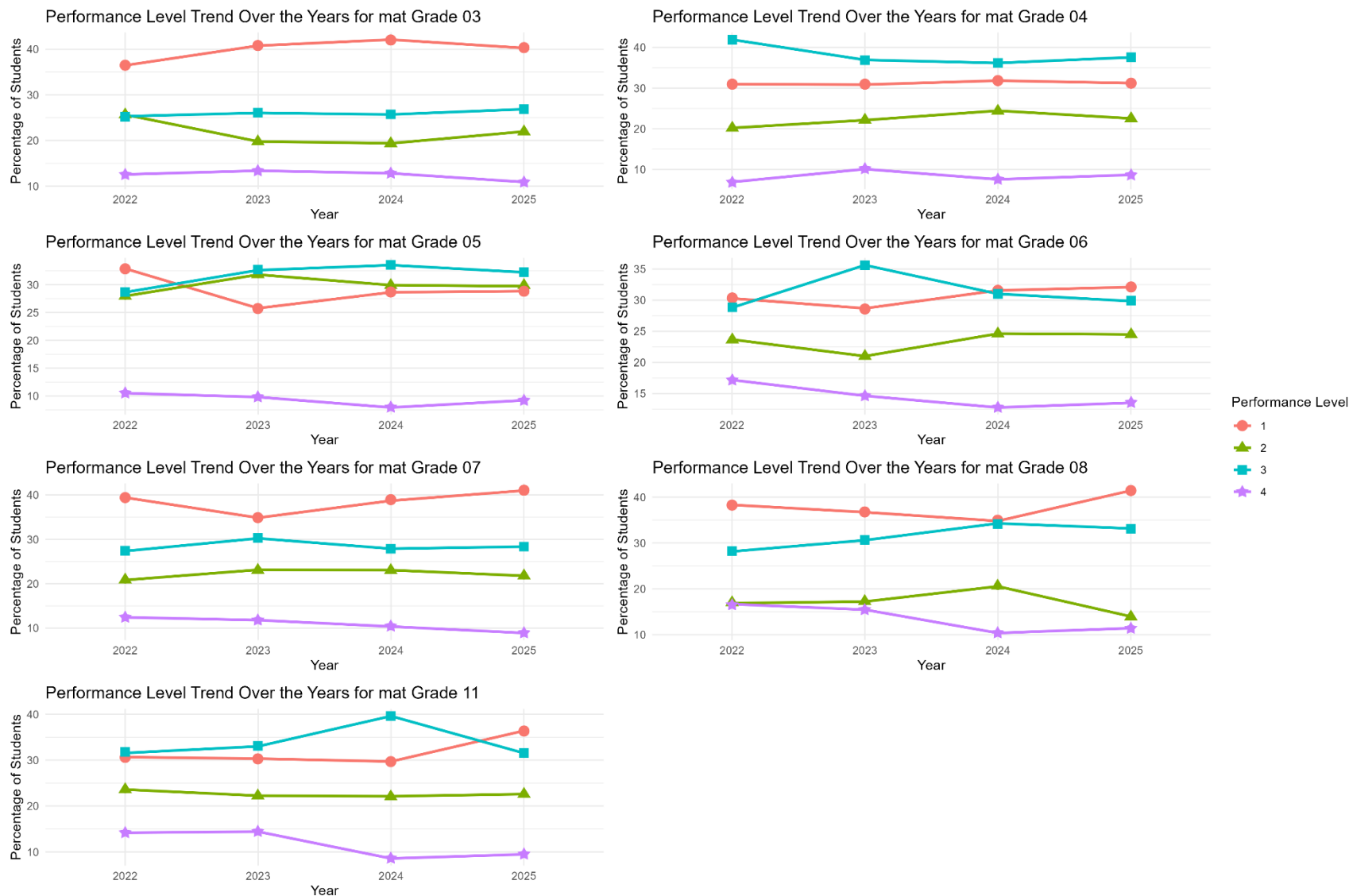


Figure 1-2 Mathematics Student Performance Over Time



Chapter 2. Overview of the MSAA

The MSAA assesses ELA and mathematics at grades 3–8 and HS and is aligned with the state content standards and the MSAA Core Content Connectors (CCCs). The MSAA is a computer-based, on-demand, two-stage adaptive assessment consisting primarily of selected-response items, along with some constructed-response items and open-response writing prompts. These item types are written at varying levels of complexity to reflect different stages of student skill and knowledge development.

Students with the most significant cognitive disabilities often require materials and instructional strategies that are substantially adapted and scaffolded, providing built-in supports (e.g., embedded graphic support to aid in comprehension) to meet their individual needs. When students begin to learn a new skill or acquire new knowledge, they need more support and scaffolding. As students learn and develop mastery of that skill or knowledge, they need less support.

The MSAA levels of complexity are designed to follow instructional practices. The MSAA mirrors these instructional practices by embedding supports and scaffolds directly into test items. These supports are designed to help students demonstrate what they know and can do as independently as possible. Additional individualized supports—such as varied item presentation methods—may also be provided by Test Administrators (TAs), based on student needs.

The MSAA is a one-on-one administration and is delivered primarily online, with paper-pencil or hybrid formats available as accommodations when appropriate. The assessment platform includes built-in accessibility features, and students may use additional accommodations such as assistive technology, scribes, braille, or sign language. Appendix A summarizes the frequency of accommodation use during the 2025 administration. TAs have flexibility in managing the testing schedule and may pause or resume testing sessions based on the student's engagement and stamina.

The mathematics assessment includes 35 operational items, primarily selected-response, along with 4 to 7 constructed-response items. The ELA assessment consists of 39 to 42 operational items, also primarily selected-response items and includes 1 to 4 constructed-response items, a multiple-part selected-response writing prompt, and an open-response writing prompt at each grade level. Both content areas use a two-stage adaptive format, with students completing Session 1 followed by Session 2. The MSAA is administered over a seven-week window, giving Test Administrators the flexibility to determine how many items a student completes in each sitting. Additionally, Session 1 includes embedded field-test items for every grade and content area.

2.1 History of the MSAA

Development efforts leading to the MSAA began in late 2010 with the launch of the National Center and State Collaborative (NCSC), which initiated work on the NCSC Alternate Assessment. This assessment was designed to meet the requirements of the Elementary and Secondary Education Act (ESEA) and the Individuals with Disabilities Education Act (IDEA) and was based on alternate academic achievement standards for students with the most significant cognitive disabilities. These efforts culminated in the first operational administration of the NCSC assessment in spring 2015. Following that administration, the work of NCSC concluded. The first MSAA administration took place in spring 2016, and it is worth noting that the MSAA was not administered in 2020 due to widespread school closures resulting from the global impact of the COVID-19 pandemic.

Building on the foundation established by NCSC, the MSAA Partners carried forward the work, maintaining many of the same principles, goals, and core values. The principle beliefs that underlie the MSAA began with NCSC and were laid out in the prior planning and development of the AA-AAAS. As recorded in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual*, states and organizational partners implementing the NCSC development plan found they needed to come to a consensus on topics that were a mix of practice and theory in the comprehensive context of teaching and learning. A blend of policy, educational, and technical solutions was required. Through policy discussions and in iterative research and design steps, the partners arrived at a shared philosophy and guiding principles that were reflected in the overall project resources. These project resources included a comprehensive system of curriculum, instruction, classroom assessment, and professional development, as well as the operational assessment design. For more information about the NCSC assessment, please refer to the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* (see “References” for the URL) or contact the MSAA Partners at MSAA@azed.gov.

The MSAA Partners share the belief, carried forward from their NCSC predecessors, that accessibility is fundamental to the validity of the assessment. They hold that meaningful access to content aligned with college- and career-ready academic standards starts with a rigorous curriculum, quality instructional resources, and effective teacher training. The original design of NCSC curriculum and instruction resources was informed by extant research and iterative small studies to ensure inclusive accessibility and appropriately high expectations for learning. Then, the NCSC assessments were based on the same model of learning as reflected in classroom resources. Finally, the NCSC project provided resources for intervention in communicative competence to ensure that all students have a way first to learn the concepts, and then to show what they know and what they can do on the assessment. The NCSC Theory of Action and Validity Approach, available at [the MSAA states’ website](#), was developed to explain the basis for these resources, how they were intended to relate to one another and to college- and career-ready academic standards, and, ultimately, how they relate to the goal of having all students with the most significant cognitive disabilities leave high school ready to participate in college, careers, and their communities.

Practice-focused summaries of the foundational components reflected in the design of the NCSC assessment, known as the NCSC Brief series, are available to orient readers to the larger context of the comprehensive NCSC system of curriculum, instruction, assessment, and professional development. The NCSC Brief series can be found in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* (see “References” for URL) located [on the Arizona Department of Education’s website](#), or by contacting the MSAA Partners at MSAA@azed.gov.

2.1.1 Developments in 2016-17

The MSAA was initially administered as fixed forms in 2016. In 2017 the MSAA became an adaptive test employing multistage testing methodology. The 2016 MSAA English Language Arts (ELA) tests for grades 3 and 4 included a distinctive set of five foundational items that were dichotomously scored as a single unit. This particular approach was discontinued in 2017, and assessment of foundational ELA knowledge and skills was changed to using more traditional multiple-choice items.

2.1.2 Developments in 2017-18

Soon after the initial development and implementation of the first MSAA in spring 2016, the assessment program partners committed to adding an essay writing component to the English Language Arts (ELA)

tests for every grade level. After several years of field testing and subsequent analysis and discussion, this writing-prompt component was introduced operationally in 2018. See section 3.3.3 for more information about the knowledge, skills, and testing formats associated with the essay writing prompts.

Because of the introduction of the essay writing components, a standards validation workshop was held in the summer of 2018 to evaluate the degree to which the introduction of the essays may have affected the performance-level cut scores for each grade. The standards validation was also expanded to include the mathematics tests as the cut scores for both the ELA and mathematics tests were evaluated for alignment across the different grade levels. The standards validation resulted in changes for some performance-level cut scores.

2.1.3 Developments in 2018-19

The technical report for the 2019 test underwent a transition to a validity-argument technical report, including the introduction of the validity argument model. This resulted in major changes to Chapter 11 to more clearly delineate the assumptions underlying the MSAA and the supporting evidence for those assumptions.

2.1.4 Developments in 2020-21

A change in the item development and complexity categorization process was initiated for the construction of the 2021 test. Originally, the items were developed and categorized with respect to four ordinal tiers, where tier 4 represented the highest level of complexity and tier 1 the lowest. MSAA TAC and Partners monitored item performance across 4 tiers during the first two years of MSAA development and operational administrations and found items at Tiers 3 and 4 did not significantly differentiate between performance levels. Beginning in 2021, newly developed field-test items were written to three target levels. In particular, Tier 4 adhered to existing guidelines; Tier 1 remained mostly the same with minor adjustments, and Tiers 2 and 3 merged with changes to option presentation and modeling demonstrations aligned with focal knowledge, skills, and abilities (FKSAs, see Chapter 3 for more detail on FKSAs). This shift still aligned with the original complexity intent. Level 3 (formerly Tier 4) became the most complex; Level 2 (Formerly Tiers 2 and 3) less complex, and Level 1 (formerly Tier 1) is the least complex. The three writing-prompt levels remained unchanged. Updated guidelines provided a structure for item design. The mapping of tiers to levels is provided in Table 2-1. As a result of this change, the operational items included in operational test forms from 2022 onward reflect a mix: some items were developed using the original four-tier structure, while others were developed following the updated three-level design.

Table 2-1. Mapping Tiers and Complexity Levels

Tier 4	Item complexity Level 3
Tier 3	Item complexity
Tier 2	Level 2
Tier 1	Item complexity Level 1

The *ELA Level Guidelines*, *Mathematics Level Guidelines*, and item specifications for each grade and content were updated to reflect the three item complexity levels. A primary distinction among items written at item complexity levels 1–3 is (a) their connection to content standards, and (b) the scaffolded supports provided at each level.

- Complexity Level 3 items target the Core Content Connectors, with minimal support provided during item administration.
- Complexity Level 2 items target Core Content Connectors, with content supports (e.g., graphics, examples, definitions) provided during item administration.
- Complexity Level 1 items target Essential Understandings, with content supports (e.g., graphics, simplified language) and item supports (e.g., two response options provided during item administration).

2.1.5 Developments in 2022-23

In 2022 MSAA partners collaborated closely with Cognia and the MSAA Technical Advisory Committee (TAC) to create and execute the *MSAA Validity Survey*. This survey aimed to assess the utilization of MSAA test scores in the field, as well as the professional development opportunities provided to equip educators and school and district leaders with the skills to interpret data effectively. The validity survey was developed in two phases: a pilot phase (completed June 2023), which tested the survey instrument and informed the sampling plan, and an operational phase (completed fall 2023), which focused on collecting data. The survey resulted in the gathering of large amounts of evidence that are now cited in appropriate sections of the technical report, with particular emphasis on Chapter 11 which provides an overall summary of all the validity evidence for the MSAA.

2.1.6 Developments in 2023-24

In 2023, as new partners joined MSAA and existing partners showed interest in expanding MSAA to include additional high school grades (e.g., 9th grade), Cognia conducted a study to evaluate the suitability of administering the Grade 11 assessment to 9th-grade students. The analysis focused on content alignment and student performance data.

Beginning in 2024, some MSAA students enrolled in grade 9 participated in the Grade 11 ELA and mathematics assessments. Overall, for high school assessments, Grade 9 students showed lower performance compared to their peers who took ELA and Mathematics in 11th grade elsewhere in the MSAA. Despite this, the percentage of Grade 9 students scoring in Levels 3 and 4 indicates strong exposure to content standards and mastery.

2.2 Stakeholders

Many stakeholders are involved in the development of the MSAA. MSAA Partners are key representatives who together compose the decision-making body for MSAA. The partners have developed subcommittees to focus on specific aspects of the program. Each subcommittee holds decision-making authority on behalf of the MSAA Partners within its respective domain. Table 2-2 shows the 2025 state representation for each subcommittee, followed by a description of their respective responsibilities.

Table 2-2 Subcommittee Representation

Subcommittee	State/ Entity Representation
Item Development	Maine, BIE, Arizona, Vermont, South Dakota, and CNMI
Manuals, User Guides, and Training	Arizona, Guam, Maine, South Dakota, Vermont, DC, and BIE
Psychometric and Test Construction	Arizona, Maine, BIE, Vermont, South Dakota, CNMI and American Samoa
Platform	Arizona, BIE, South Dakota, Arizona, and CNMI
Scoring	Arizona, South Dakota, and CNMI
Reports	Guam, Arizona, and South Dakota
Interim Assessment (NEW)	Guam, Maine, BIE, Vermont, South Dakota, DC, Arizona, and CNMI

Although item development did not happen in 2025, the MSAA Item Development Subcommittee provides overall input and direction related to new item development of field-test items. This may include reviewing item development tasks; participating in development planning, and item and passage reviews; reviewing alternative text; participating in Accessible Portable Item Protocol (APIP) reviews, plus the computer-based and paper-based materials reviews; and providing direction on updates to the graphics and editorial style guides, teacher directions, and front matter for the *Directions for Test Administration* (DTA). In addition to the Item Development Subcommittee members, stakeholders from schools and districts across the MSAA Partners participate in the process during new field-test item development through passage content and bias meetings and the item content and bias meetings. Additional detailed information is available in Chapter 4.

The Manuals, User Guides, and Training Subcommittee oversees development of the *Test Administration Manual* (TAM), *MSAA Online Assessment System User Guide for Test Administrators*, *MSAA Online Assessment System User Guide for Test Coordinators*, online training modules, and the final quiz that is required for TAs.

The Psychometric and Test Construction Subcommittee provides oversight and planning for the Technical Advisory Committee (TAC) meetings and contributes to key psychometric and test design decisions. These responsibilities typically include reviewing item performance data, approving test blueprints and designs, and making decisions related to the construction of operational and field-test item sets following Data Review. However, due to the reuse of forms, the subcommittee's work related to test construction for the 2025 assessment was completed in preparation for the 2023 administration. The subcommittee continues to support the program through its role in TAC meeting oversight and is also responsible for determining the content of the End-of-Test Survey, reviewing survey results after administration, advising on relevant policy decisions, and guiding the structure of the technical report.

The Platform Subcommittee determines development priorities for the online assessment platform each year and on an as-needed basis. This group also reviews recommendations and developments pertaining to the security of the online platform and ultimately approves all changes made to the platform.

The Scoring Subcommittee reviews and approves the scoring specifications and scorer training materials, observes scoring processes, reviews daily scorer quality control and production management reports, and participates in the virtual Scoring Observation meeting (open to all Partners).

The Reports Subcommittee is responsible for decisions pertaining to report revisions and design. This group also approves all changes made to the overall layout of the student results files and the final processing and reporting business requirements implemented for MSAA reporting. This subcommittee also reviews and approves the *MSAA Parent Guides* in English and Spanish and the *Guide for Score Report Interpretation*.

The Interim Assessment Subcommittee was established in 2025 in response to interest from several MSAA Partners in exploring the development of an interim assessment for this student population. The subcommittee convened to begin identifying the requirements for a potential beginning-of-year interim assessment, with implementation targeted for fall 2026 and beyond. Members engaged in discussions around the intended purpose and use of the assessment, as well as considerations related to test design and reporting. These conversations are ongoing, as the solution remains in the early stages of development.

2.3 MSAA Participation

The criteria for student participation in the 2025 MSAA reflect the pervasive nature of a significant cognitive disability. All content areas are considered when determining who should participate in this assessment. Table 2-3 below shows the participation criteria, and the descriptors used to determine eligibility for each student.

Table 2-3 Participation Criteria

Participation Criteria	Participation Criteria Descriptors
The student has a significant cognitive disability.	Review of student records indicates a disability or multiple disabilities that significantly affect intellectual functioning and adaptive behavior. *
The student is learning content linked to grade-level content standards.	Goals and instruction listed in the IEP for this student are linked to the enrolled grade-level content standards and address knowledge and skills that are appropriate and challenging for this student.
The student requires extensive, direct, individualized instruction and substantial supports to achieve measurable gains in a grade- and age-appropriate curriculum.	The student (a) requires extensive, repeated, individualized instruction and support that is not of a temporary or transient nature, and (b) uses substantially adapted materials and individualized methods of accessing information in alternative ways to acquire, maintain, generalize, demonstrate, and transfer skills across multiple settings.

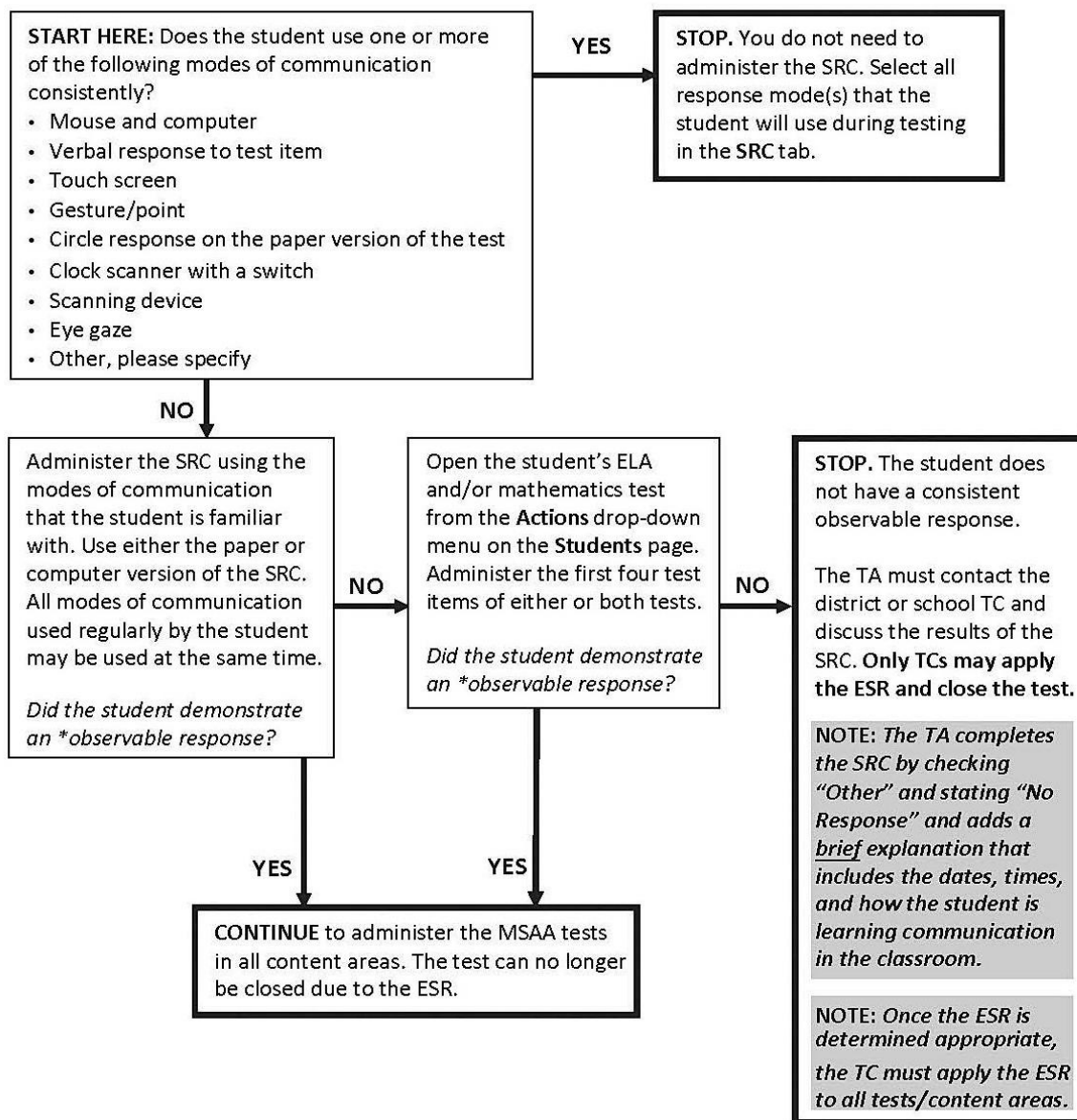
**Adaptive behavior is defined as essential for someone to live independently and to function safely in daily life.*

Appendix B shows the 2025 summary of students who participated in the MSAA for both mathematics and ELA by demographic category.

Assessments for students with the most significant cognitive disabilities rely on a foundation of communicative competence. Without adequate receptive and expressive communication skills, students are unlikely to effectively demonstrate their knowledge and abilities during an assessment. The assessment process includes identifying students who do not yet have an appropriate mode of communication. To participate meaningfully in the MSAA, students must be able to communicate using an observable response mode—a consistent and intentional behavior or movement that can be interpreted by a communication partner as purposeful communication. The Student Response Check (SRC) is a task during which a student is asked to demonstrate their preferred mode(s) of communication. In these cases, the SRC aids in gathering information that is needed to determine whether there are communication barriers to meaningful participation in the MSAA assessment. If a student's responses to test items are not clearly observable, or understood by the TA or scribe, the testing experience may need to be ended early. This process is called the Early Stopping Rule (ESR). To end the test for a student, the ESR procedures must be followed.

Figure 2-1 shows the procedure for determining if the SRC is appropriate to administer and, if so, how to proceed in determining if the student has an observable, interpretable mode of communication that can be used throughout testing. If clear, intentional communication is not shown, the ESR may be applied. This graphic from the *2025 MSAA TA User Guide* shows the process of implementing the ESR.

Figure 2-1. Student Response Check (SRC) Flowchart: When to Apply the ESR



**An observable response is defined as a predictable and consistent behavior or movement that is able to be understood by a communication partner as intentional communication.*

Following the assessment, teachers may use the Communication Tool Kit developed by NCSC to support these students in developing an appropriate mode of communication. The Tool Kit is available on the MSAA Website.

2.4 MSAA Participation Rate

Under the “Every Student Succeeds Act” (ESSA), participation in a state’s alternate assessment based on alternate academic achievement standards (AA-AAAS) is capped at one percent of the total student population. For this reason, MSAA rosters are managed locally—typically by the state, district, or school test coordinator, depending on the Partner’s organizational structure. During the administration window, student rosters are regularly updated, including additions, transfers, and deletions as needed. The data below are drawn from the overall test status summary report produced by the MSAA System and provide statistics by Partner, grade level, and content area. Estimated enrollment is determined by subtracting the number of canceled assessments from the total number of students on the roster. The number of students who attempted the assessment is based on the count of submitted tests. Table 2-4 displays participation rates by content area, broken down by grade level.

Table 2-4 MSAA Participation Rate by Subject and Grade

Subject	Grade	The Total Number of Students Enrolled in the MSAA Platform	Number of students attempted the test	Attempt Rate
ELA	3	1,689	1,622	96%
ELA	4	1,661	1,556	94%
ELA	5	1,626	1,528	94%
ELA	6	1,584	1,490	94%
ELA	7	1,490	1,396	94%
ELA	8	1,395	1,298	93%
ELA	HS	1,389	1,278	92%
Mathematics	3	1,688	1,615	96%
Mathematics	4	1,662	1,559	94%
Mathematics	5	1,626	1,528	94%
Mathematics	6	1,584	1,490	94%
Mathematics	7	1,490	1,395	94%
Mathematics	8	1,395	1,297	93%
Mathematics	HS	1,389	1,270	91%

Chapter 3. Test Development-Content and Administration

3.1 History of Alternate Academic Achievement Standards and Core Content Connectors

As noted in Chapter 2, MSAA has evolved from the work of NCSC. As such, MSAA's history is firmly planted in the foundation of the NCSC AA-AAAS and follows the original Theory of Action ([PDF linked here](#).) Designed specifically for students with the most significant cognitive disabilities, the NCSC AA-AAAS was a performance-based test that was aligned with grade-level state content standards for ELA and mathematics and tested student performance based on alternate academic achievement standards.

The NCSC state and center partners, comprised of content and special education experts, focused on defining the constructs of reading, writing, and mathematics to reflect an appropriate expectation of instruction and learning throughout a student's educational experience. Furthermore, the experts sought to make those constructs adaptable to the way in which students with the most significant cognitive disabilities demonstrate acquired knowledge and skills. NCSC established overarching content definitions by examining: (a) existing content definitions in general education; (b) the content, concepts, terminology, and tools of each domain; (c) a body of extant research; and (d) the Common Core State Standards (CCSS). These content definitions became central to the development of assessment items.

NCSC developers revised and refined the NCSC AA-AAAS design using cycles of continuous feedback from state and center partners. Developers evaluated proposed designs through iterative item and test development steps, special studies, and pilot testing, all of which were central to the final NCSC assessment model implemented through the first administration of the operational test in spring 2015.

Prior to the start of item development, the Core Content Connectors (CCCs) connecting the Learning Progression Frameworks (LPFs) to the CCSS were developed.

3.1.1 The Learning Progression Frameworks

The LPFs present a broad description of the essential content and general sequencing for student learning and skill development (Hess, 2010). The LPFs provide the educational logic to help move students with the most significant cognitive disabilities along with their peers, based on researched teaching and learning, toward mastering skills for college and career readiness. Experts at NCSC looked at these learning targets together with grade-level content expectations from the CCSS to identify and clarify the most salient grade-level core academic content to guide instruction and assessment from kindergarten through high school for students with the most significant cognitive disabilities. This core academic content is referred to as the CCCs. The LPFs can be found on the [MSAA website](#).

3.1.2 Core Content Connectors

The CCCs were defined by NCSC as the academic content designed to frame the instruction and assessment of students with the most significant cognitive disabilities. This identified core content serves as a connection or stage between the LPFs (designed for typically developing students) and the CCSS (which define grade-level content and achievement). The CCCs are intentionally dually aligned with both.

The CCCs are designed to contribute to a fully aligned system of content, instruction, and assessment that focuses on the core content, knowledge, and skills needed at each grade to ensure success at the next grade level.

Each CCC represents a teachable and assessable part of the content. Related CCCs are addressed during instruction to create deeper understanding of grade-specific academic content. The CCCs are specifically intended to promote success as students advance alongside peers without disabilities from grade level to grade level. They are the starting point for instruction; they do not represent everything an individual student can and should learn.

The CCCs preserve the sequence of learning outlined in the LPFs, to the extent possible, while deconstructing the progress indicators (which describe concepts and skills along the learning continuum for each grade span in the learning progression) into smaller segments of content. The CCCs and corresponding Curriculum Resource Guides were developed to help explain and promote how students can engage in the CCSS while following the LPFs. To demonstrate the content sequence maintained by the CCCs, Table 3-1 shows a series of CCCs developed for multiple grades by NCSC for one big idea within the mathematics strand of geometry.

Table 3-1. CCCs Developed for Geometry Big Idea: Shapes and Figures—Their Attributes, Properties, and Corresponding Parts

Grades	Geometry Core Content Connectors	Geometry Core Content Connectors	Geometry Core Content Connectors	Geometry Core Content Connectors
K–2	K.G.M1a1 Recognize two-dimensional shapes (e.g., circle, square, triangle, rectangle) regardless of orientation or size	K.GM.1a2 Recognize two-dimensional shapes in environment regardless of orientation or size	K.GM.1a3 Use spatial language (e.g., above, below) to describe two-dimensional shapes	2.GM.1a4 Identify two-dimensional shapes such as rhombus, pentagon, hexagon, oval, and equilateral, isosceles, and scalene triangles
3–4	3.GM.1h1 Identify shared attributes of shapes	4.GM.1h2 Classify two-dimensional shapes based on attributes (number of angles)	N/A	N/A
5–6	5.GM.1a1 Recognize properties of simple plane figures	5.GM.1b1 Distinguish plane figures by their properties	N/A	N/A
7–8	7.GM.1e Construct or draw plane figures using properties	8.GM.1g1 Recognize congruent and similar figures	N/A	N/A
HS	H.GM.1e Make formal geometric constructions with a variety of tools and methods	H.GM.1b Use definitions to determine congruency and similarity of figures	N/A	N/A

The CCCs reference the *Learning Progressions Frameworks Designed for Use with the Common Core State Standards in Mathematics K–12* (Hess, 2010). The letter/number in each box provides a cross-reference to the letter/number in the original learning progressions. For example, for 3.GM.1h1, the 3 means third grade, the GM means geometry, the 1h relates to the specific progress indicator in the original learning progression, and the 1 means that it is the first in a series of connectors.

Table 3-1 shows how learner understanding builds across years. For example, in the second column, the student recognizes shapes, then compares shapes based on attributes, then distinguishes plane figures by properties, then recognizes congruent/similar figures, and finally by high school can use definitions to determine congruency/similarity of figures. These skills all promote the big idea about shapes—their attributes, properties, and corresponding parts (Wakeman, Lee, & Browder, 2012).

The MSAA Partners adopted the CCCs as the standards that the students are taught and assessed against as a participating consortia member. These CCCs are the measured academic content on the MSAA and are aligned to and derived from each participating state’s content standards.

3.2 Alignment and Linkages

Evidence that test content reflects the concepts that were meant to be measured is one of the critical sources of information necessary to support valid interpretations of test scores (AERA et al., 2014). *Alignment* refers to coherent connections within and across a system (Forte, 2013a, 2013b). Traditional alignment procedures describe the degree of intersection, overlap, or relationship among academic content embedded in state content standards, assessment, and instruction (Webb, 2005).

As part of the assessment development process, NCSC conducted a series of studies to answer several key questions related to the alignment of the assessment. These efforts were meant to ensure that students’ scores can be interpreted as reflecting the knowledge and skills defined in the standards and claims (developed by NCSC, see *National Center and State Collaborative 2015 Operational Assessment Technical Manual*). The alignment questions were:

1. What is the degree of alignment between the CCCs and the grade-level CCSS?
2. What is the degree of alignment between instructional student learning expectations and measurement targets (expectations for assessed knowledge and skills)?
3. To what degree do the assessment tasks and items align to the grade-level CCSS?
4. To what degree do the assessment tasks and items align to the performance-level descriptors (PLDs)?
5. How well do the claims align with grade-level content and provide useful information for tracking student progress toward achieving the knowledge and skills in the grade-level standards?

To address the five alignment questions, various studies were conducted between 2012 and 2015 at different points in the development process to ensure system coherence. Table 3-2 lists the studies, when each was conducted, and the alignment question being addressed. Each MSAA partner may have conducted their own reviews in addition to those listed below.

Table 3-2. Studies Related to Evidence of System Coherence

Study	Conducted	Claim for Which Evidence Is Provided
Relationship Studies	Mathematics: Summer 2012; Reading: Winter 2013; Writing: Summer 2013	Evidence for Alignment Question #1. The content and skills in the CCCs represent an adequate and appropriate sample of the grade-level CCSS.
UMASS Study of Coherence	Fall 2013	Evidence for Alignment Question #2. The targets for measurement provide information useful for tracking student progress in the CCSS and to teachers for providing instruction focused on academic expectations.
Task/Item Alignment Study	Summer 2015	Evidence for Alignment Question #3. The content and skills assessed by the NCSC AA-AAAS represent an adequate and appropriate sample of the grade-level CCSS.
Item Mapping Study	Summer 2015	Evidence for Alignment Question #4. The score reports are accurate and support appropriate inferences about student knowledge and skills.
Vertical Coherence Study	Summer 2015	Evidence for Alignment Question #5. The targets for measurement provide information useful for tracking student progress in the CCSS and for providing instruction focused on academic expectations.

MSAA has carefully and gradually evolved from NCSC, ensuring the alignment has been maintained while still allowing for adjustments. For example, the prioritized CCCs and reporting categories have remained the same for mathematics and for ELA (apart from one adjustment to the Reading Foundational CCC at grades 3 and 4) from those originally identified by NCSC. Section 3.4 provides details on the contents and blueprints. The various alignment studies noted in Table 3-2 are applicable for the MSAA, as the NCSC Theory of Action serves as the foundation for the MSAA program. MSAA has implemented test design adjustments that were outlined by NCSC during the original test design planning phase. One example of this is the implementation of the stage-adaptive test design. Section 3.3 below provides detailed information about the assessment design.

3.3 2025 MSAA Assessment Design

3.3.1 Operational Design

The operational MSAA is designed to produce valid and reliable mathematics and ELA scores for the intended uses. The mathematics and reading portions of the test are composed primarily of selected-response items. In mathematics, all grade levels also include constructed-response items that require students to work through a process to solve a problem. Writing is composed of selected-response stand-alone items, a multiple-part selected-response writing prompt, and an open-response writing prompt.

The operational items vary in complexity following the *Mathematics Level Guidelines*, *ELA Level Guidelines*, and *Passage and Item Sloping Guidelines*, where each level at a given standard addresses both the content complexity and the degree of scaffolding and support provided with the items. The items assess grade-level academic concepts defined by either the focal knowledge, skill, and ability (FKSA; level 2–3) or Essential Understanding (EU; level 1). The FKSAs allow for items of graduated complexity where the same FKSA is addressed but with increased levels of support and/or decreased levels of complexity so that students with various levels of cognitive ability can access the content. The EUs define entry-level skills based on a grade-specific CCC that builds increasing understanding of the grade-level content. The levels provide three decreasingly complex versions (items) of the task referred to as Level 3

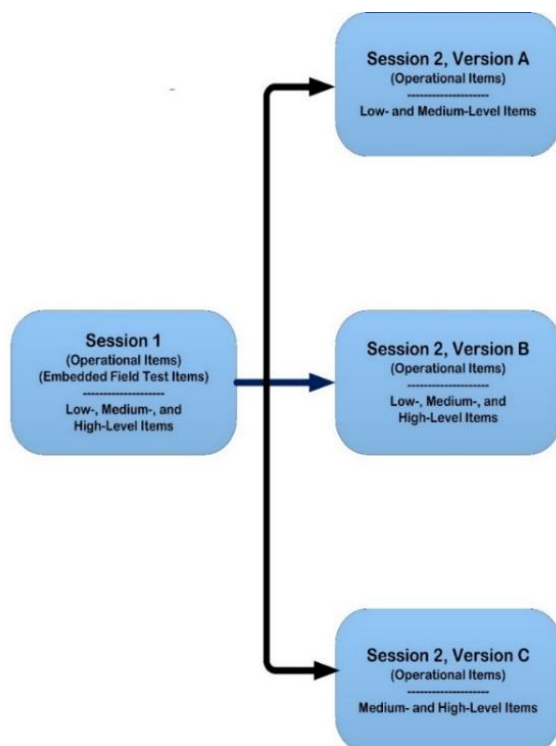
(most complex), Level 2 (less complex), and Level 1 (least complex). The writing prompts also use three levels of items. Level 1 is a multiple-part selected-response item cluster, where all items build on each other toward the creation of a final product. Level 2 and Level 3 are open-response writing prompts that vary in complexity based on the amount of support provided at each level.

For the 2025 assessment, the test forms used were a reuse of the multistage adaptive forms originally developed for the 2023 administration. These included three two-stage adaptive forms for both ELA and mathematics. The forms were developed to align with guidelines informed by the respective content-area test blueprints (discussed in Section 3.4). Operational items appear in both Session 1 and Session 2. All students receive the same operational items in Session 1, which serves as Stage 1 of the adaptive test. Based on student performance in Stage 1, they are assigned one of three versions of Session 2 (Stage 2), which vary in difficulty. Version C is the most challenging, followed by Version B, then Version A. Each version of Session 2 includes a Level 1 writing prompt; additionally, Session 2A includes a Level 2 writing prompt, while Sessions 2B and 2C include a Level 3 writing prompt. As a result, students can follow one of three possible testing paths through the multistage design. Although there is moderate item overlap across the versions of Session 2, there is sufficient variation to maintain the intended separation in test information functions (TIFs) across the paths.

Students are assigned to a testing path based on predefined routing cut scores. Additional details about TIFs can be found in Chapter 9, and information about the measurement reliability associated with these TIF values is provided in Chapter 10.

Figure 3-1 illustrates the two-stage adaptive design, showing the placement of field-test items in Session 1 as well as the item levels used across sessions. Each of the three field-test forms includes the same three possible operational test paths: Session 1 followed by Session 2A, Session 2B, or Session 2C.

Figure 3.1 Two-Stage Adaptive Design



3.3.2 Operational Items and Embedded Items in Field Test Positions

As discussed earlier, there are three versions of Session 2. Several items are the same across Sessions 2A, 2B, and 2C.

As shown in Table 3-3, the ELA tests administer 39–42 operational items including two writing prompts, across the testing sessions per grade. Additionally, each form has 11–13 field-test items for a total of 33–39 items across two field-test forms. As discussed earlier, there are three versions of Session 2, which consist of 3 passage sets and 5–7 writing items. In Session 2 a total of 7–21 items overlap across Sessions 2A, 2B, and 2C, depending on the grade. The items that overlap are not always the same ones across all 3 versions (e.g., one passage set may overlap across 2A and 2B, and a different passage set may overlap across 2B and 2C).

Because of the reuse of the 2023 forms in 2025, items in field test positions were not scored. The field test data from these items was analyzed in 2023.

Table 3-3. ELA Operational and Items in Field-Test Positions

Grade	Total Operational Items Administered to Each Student	Writing Prompt Operational Items	Items in Field Test Positions Total Across Two Forms
3	42	2	20
4	42	2	22
5	40	2	20
6	39	2	19
7	39	2	22
8	39	2	20
HS	39	2	22

As shown in Table 3-4, the mathematics tests consist of 35 operational items across the testing sessions per grade. Additionally, each field-test form has 10 different items in field test positions for a total of 20 items across the two field-test forms. As discussed earlier, Sessions 2A, 2B, and 2C each have 20 items. A subset of the items in Session 2A are common with items in Session 2B. There can be up to 10 items that are common between Sessions 2A and 2B. A subset of the items in Session 2C is common with items in Session 2B. There can be up to 10 items that are common between Sessions 2C and 2B. There are no common items between Session 2A and Session 2C.

Table 3-4. Mathematics Operational and Embedded Field-Test Items

Grade	Total Operational Items Administered to Each Student	Items in Field Test Positions Total Across Two Forms
3	35	20
4	35	20
5	35	20
6	35	20
7	35	20
8	35	20
HS	35	20

The 2025 items in field-test positions were selected in 2023 according to the following criteria:

- mathematics and ELA items represent a variety of item complexity levels (including the writing stand-alone component and a Level 1 writing prompt);
- ELA passage or writing topics are unique to the form and provide a variety of genres; and
- the passage and items are engaging, accurate, and free of regional bias.

The items on each of the forms are reviewed by psychometricians for any statistical concerns. The *AY23 Test Construction Process* provides the procedures to follow in constructing the test including the psychometric parameters that form the criteria each constructed test should meet. This document is used as the guiding resource to replicate MSAA test construction processes across administration years. The test construction process occurs following data review of the field-test items from the previous administration. The content specialists create the test forms based on the test blueprints and criteria provided by the psychometricians. The forms are then evaluated by the psychometricians and revision loops occur as needed. Once the psychometricians provide approval of a constructed test it is then also reviewed by the MSAA Psychometric and Test Construction Subcommittee. All constructed tests, as well as the field-test items, are posted on a secure FTP site for the MSAA Psychometric and Test Construction Subcommittee review and approval. A webinar is held with the MSAA subcommittee to explain the test construction process and to review the Test Construction Design document, which provides information specific to each content area about the items selected. The MSAA subcommittee then has an opportunity to provide input and final approval.

The 2025 assessment reused the test forms originally developed for the 2023 administration, which included field-test items in both mathematics and ELA (reading and writing) across varying levels of complexity. To address overlap in item difficulty observed across Tiers 2, 3, and 4, MSAA Partners revised the item design by consolidating these tiers into Levels 2 and 3, as previously described and shown in Section 2.1.3. Beginning in 2021, newly developed field-test items were written to three target levels. As a result, the operational items included in the 2025 (reused 2023) test forms reflect a mix: some items were developed using the original four-tier structure, while others follow the updated three-level design. Writing prompts, which were already aligned to three tiers, were simply renamed to reflect the level-based structure.

3.3.3 Item Design and Administration

The MSAA item design and administration is intended to capture student performance at different levels of skill and knowledge acquisition. The assessment items incorporate important aspects of item design related to both varying levels of content complexity and the degree and type of scaffolds and supports. The MSAA Partners follow NCSC's intentional assessment development process to address the targeted grade-level academic content linked to evidence-based curricular and instructional materials.

The MSAA content development processes address levels of cognitive and language complexity, specifically addressing the state content standards, and the heterogeneous characteristics of the target student population. The assessment items vary systematically in complexity yet remain aligned with the FKSA or the EU behind the CCCs. The items are designed to capture student performance by varying two characteristics: (1) levels of content complexity and (2) degrees and types of scaffolds and supports. The scaffolds and supports (e.g., reminders, examples, and models) are provided to focus the student on the task and elicit a response without guiding the student's response.

Overall Item Structure

A range of item levels is developed for each CCC, as described in Section 3.3.1. Each level provides variable features and supports that offer multiple entry points for a variety of students to demonstrate their knowledge and skill. All items assess grade-level academic concepts defined by either the FKSA or the

Essential Understandings (EUs). Items follow the level guidelines and item specifications. As outlined in the level guidelines, items of graduated complexity address the same FKSA but provide increased levels of support and/or decreased levels of complexity, and at the lowest level address the EU that has the most decreased level of complexity, and, also, as part of the item, provide the greatest level of support. Additionally, the MSAA item specifications are consistent with design patterns and task template guidelines that were originally developed by NCSC. The item types, as outlined in the MSAA item specifications, are selected-response, multiple-part selected-response, constructed-response, and open-response. Regardless of level or item type, all items include scripted teacher directives.

Mathematics and ELA (reading, language, and stand-alone writing) selected-response items are multiple-choice items where a student selects a response from three options (two options at Level 1); the answer is worth 0 or 1 point. ELA multiple-part selected-response item clusters are multiple-choice items that are clustered together and connected to a single CCC. For each item, the student selects a response from three options (two options at Level 1); the answer is worth 0 or 1 point. The overall cluster could, then, be worth more than 1 point. There are two- and three-part item clusters. A typical example of a multiple-part selected-response item cluster would be an initial item in the cluster that asks the student to identify the main idea and then a second item that asks for a supporting detail. In contrast, the mathematics MSAA item specifications and level guidelines require CCCs with multiple components to be addressed with unique items. For example, a CCC asking a student to identify and solve an equation might be evaluated using one item that requires the student to identify the correct equation for a word problem and a separate item that requires solving an equation.

Mathematics and ELA constructed-response items require the student to interact in some way with a set of materials to provide a response. These items are scored as correct or incorrect by the TA following the directions provided in the *Directions for Test Administration* (DTA). For example, students might construct a graph, solve a problem, or complete a table, chart, or graphic organizer. These items are worth 0 or 1 point because the items ask the student to show whether a single concept is understood.

ELA reading foundational items focus on comprehension skills and are administered as selected-response items worth 0 or 1 point. Students are asked to read five words and select the most appropriate word to complete the sentence provided.

ELA writing prompt items require students to compose a permanent product about a specific topic, following the writing process. The Level 1 writing prompt is a multiple-part selected-response item cluster where the items build on each other toward the creation of a final product. For each item in the cluster, the student selects a response from two options, with the answer worth 0 or 1 point. Writing prompt item clusters may have four to six parts, depending on the grade.

Unlike Level 1, the Level 2 and Level 3 writing prompts are open-response writing prompts that vary in complexity with the amount of support provided at each level. The Level 2 writing prompt provides a graphic organizer and a template with sentence starters that a student utilizes to create a product based on information he or she included in the graphic organizer. The Level 3 writing prompt provides a graphic organizer and a template that does not have sentence starters; the student completes his or her product within the template based on information he or she included in the graphic organizer. For Levels 2 and 3, the student response is evaluated against a grade- and level-specific rubric. Open-response writing prompt items were developed for Levels 2 and 3 only. As outlined in Chapter 1, the writing prompt items are operational in each grade for the 2025 MSAA. For reference, the specific writing rubrics are included as an appendix in the *MSAA 2025 Guide for Score Report Interpretation*.

Administration

For every grade level, the ELA and mathematics tests contain two test Sessions. Test Session 1 contains operational and field-test items. Students are then routed to Session 2A, 2B, or 2C in accordance with their performance on Session 1. TAs begin with Session 1 of either the ELA test or the mathematics test. Descriptions of the test sessions are shown in Tables 3-6 and 3-7.

Table 3-6. ELA Test Sessions

Session 1: ELA	Session 2: ELA, Includes Writing Prompts
Literary and informational reading passages and associated selected-response and constructed-response reading items	Literary and informational reading passages and associated selected-response and constructed-response reading items
Selected-response and constructed response writing stand-alone items	Stand-alone writing items, including constructed response items
Reading Foundational items (grades 3 and 4 only)	One open-response writing prompt
One multiple-part selected-response writing prompt	N/A
Items in field-test positions	N/A

Table 3-7. Mathematics Test Sessions

Session 1: Mathematics	Session 2: Mathematics
Selected-response mathematics items	Selected-response mathematics items
Constructed-response mathematics items*	Constructed-response mathematics items*
Items in field-test positions	N/A

**Constructed-response mathematics items are dichotomously scored.*

3.3.4 Item Components

3.3.4.1 Selected-Response: Reading, Writing (Stand-Alone Items and Multiple-Part Selected-Response Writing Prompt), Mathematics

All directions and materials needed for administering selected-response items are provided in the secure grade-, content-, and form-specific DTA. Selected-response items are presented to students in a standardized and consistent format. Every item is presented in the following order:

- item stimulus (which may include a passage, passage part, picture, graphic, or other illustration);
- item question; and
- response options presented in vertical or horizontal formation depending on the size of the response options.

Students select a response from the options in a variety of ways (e.g., using the computer mouse, verbalizing, gesturing, using eye gaze or communication devices, using assistive technology). Students' responses are entered into the MSAA System. If a student has the scribe accommodation, the scribe enters the student-selected response on behalf of the student.

3.3.4.2 Constructed-Response: Mathematics, ELA

The secure grade-, content-, and form-specific DTA contains the directions as well as the materials and manipulatives needed by the TA to assess the student on the constructed-response items. The TA prints out the materials and manipulatives with which the student will interact. Each item is presented to the student in a standardized, scripted sequence of steps, culminating in the TA scoring the student's performance using the required scoring rubrics. The scoring rubrics provide scoring standards that must be used in evaluating student responses. The constructed-response item is scored by the TA as correct or incorrect based on the scoring rubric for that item. The TA enters the student constructed-response score into the MSAA System.

3.3.4.3 Open-Response: Writing Prompt

All open-response writing prompt directions and stimulus materials, including the response template, are included in the secure grade-, content-, and form-specific DTA. TAs print or prepare any writing stimulus materials that they would need to use for the test. The open-response writing prompt is presented to the student by the TA in a standardized, scripted sequence of steps.

The student, or a scribe, records the response to the writing prompt either on the response template in the online MSAA System or on the paper response template included in the DTA. If the student uses a paper version of the response template, the TA

- uploads the response template, including any annotations, into the MSAA System, or
- transcribes or types (exactly) the student's writing response, including any annotations, into the MSAA System.

If the student's writing response includes inventive spelling, hard-to-read penmanship, or use of symbols, TAs are directed to annotate the response so that an external scorer can understand it. For more information about scoring, see Chapter 6.

3.4 Content and Blueprints

The test blueprints followed by MSAA are consistent with the original NCSC Theory of Action, the evidence-centered design undertaken to develop the summative assessment, and with best practices in educational measurement. Tables 3-8 and 3-9 show the broad targets developed to guide the item development process and to inform test construction. The tables provide general guidance for identifying areas of emphasis in the development of the mathematics and ELA tests. The test blueprints in Appendix C incorporate the overall content distributions used for the development of the operational tests. Each grade level/content area is represented by a table that first describes the domain (e.g., operations and algebraic thinking) or text type (e.g., reading informational text), weights by domain and ELA strands and text types, CCC, item types, and number of items. To continuously improve the assessment following each administration, the items' statistics for each test in each grade and content area are revisited to balance both the content requirements of the blueprints and the psychometric characteristics of the items for the subsequent administration. The core set of operational items on each two-stage adaptive test is established from this balanced approach.

3.4.1 English Language Arts

For the 2025 MSAA, the ELA items in reading and writing are aligned with prioritized CCCs, which are in turn connected to the CCSS and state content standards, as well as to the LPFs. The distribution of ELA items related to various text types (e.g., literary, informational, and argument) aligns to the text type emphasis in reading and writing outlined in the CCSS and state content standards.

Item alignment to prioritized CCCs is maintained in several ways. Item specifications are revised by Cognia and reviewed by MSAA partners annually to reflect instructional shifts that could affect item alignment. One example of this is described in Section 3.3.3; item specifications were revised in response to a review and recommendations by Diane Browder. In addition, the alignment of new development is reviewed annually by general and special education teachers, administrators, and other education specialists as part of the item review process.

For the 2025 MSAA, reading comprehension assessment items are presented as a single selected-response or multiple-part selected-response item as described in Section 3.3.3.

In grades 5–8 and HS, some prioritized content standards require evaluation of content across more than one passage. These skills are measured using paired passage sets. All paired passages are written in the informational text type. Tables in the test blueprints identify which CCCs require paired passages.

In grades 3 and 4, the reading vocabulary and foundational content category addresses the standard of fluency. In 2025, the reading foundational items are used operationally.

The three CCCs prioritized for writing at each grade level consist of one CCC operationally assessed by a multiple-part selected-response writing prompt and an open-response writing prompt, and two CCCs operationally assessed by selected-response writing stand-alone items. The selected-response writing stand-alone items are designed to assess discrete basic writing skills. The multiple-part selected-response writing prompt and the open-response writing prompt are designed to measure a student's ability to generate a permanent product to represent organized ideas specific to a writing mode, supported with details or facts to develop those ideas or clarify meaning, and measure the use of standard English conventions (for the open-response writing prompt only).

Table 3-8. Guidelines for Distribution of ELA Content by Grade Level

ELA Content Category	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	HS
Reading Literary	24–32%	24–32%	25–33%	21–30%	17–26%	17–26%	17–26%
Reading Informational	18–26%	18–26%	25–33%	26–34%	32–36%	32–36%	32–36%
Reading Vocabulary and Foundational (G3 and G4)	12–16%	12–16%	6–10%	9–11%	6–9%	6–9%	6–9%
Writing	36–38%	32–38%	31–40%	36–40%	36–40%	36–40%	36–40%

3.4.2 Mathematics

Mathematics items are aligned with prioritized CCCs, which are in turn connected to the CCSS and state content standards, as well as to the LPFs. Mathematical knowledge across the CCCs is assessed through selected-response items and constructed-response items. The need for constructed-response items is determined by the FKSA associated with a given CCC.

Table 3-9. Guidelines for Distribution of Mathematics Content by Grade Level

Mathematics Content Category	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	HS
Operations and Algebraic Thinking	28–32%	28–32%	9–11%	N/A	N/A	N/A	N/A
Number and Operations Base Ten	17–23%	9–11%	34–40%	N/A	N/A	N/A	N/A
Number and Operations Fractions	17–23%	28–32%	17–23%	N/A	N/A	N/A	N/A
Measurement and Data	17–23%	17–23%	17–23%	N/A	N/A	N/A	N/A
Geometry	9–11%	9–11%	9–11%	9–11%	17–23%	28–32%	9–11%
Ratio and Proportions	N/A	N/A	N/A	28–32%	34–40%	N/A	N/A
Expressions and Equations	N/A	N/A	N/A	17–23%	9–11%	17–23%	N/A
The Number System	N/A	N/A	N/A	28–32%	17–23%	9–11%	N/A
Statistics and Probability	N/A	N/A	N/A	9–11%	9–11%	17–23%	17–23%
Functions	N/A	N/A	N/A	N/A	N/A	17–23%	N/A
Algebra and Functions	N/A	N/A	N/A	N/A	N/A	N/A	47–52%
Number and Quantity	N/A	N/A	N/A	N/A	N/A	N/A	17–23%

In some cases, the selected FKSAs are best addressed by separating the skill into two parts, creating two unique items to fully address a single content standard. Tables in Appendix C identify which CCCs require two items.

In addition, there are items identified as not allowing the use of calculators. These items are related to computation, where the construct being assessed would be masked using a calculator.

Chapter 4. Test Development-Stakeholder Involvement

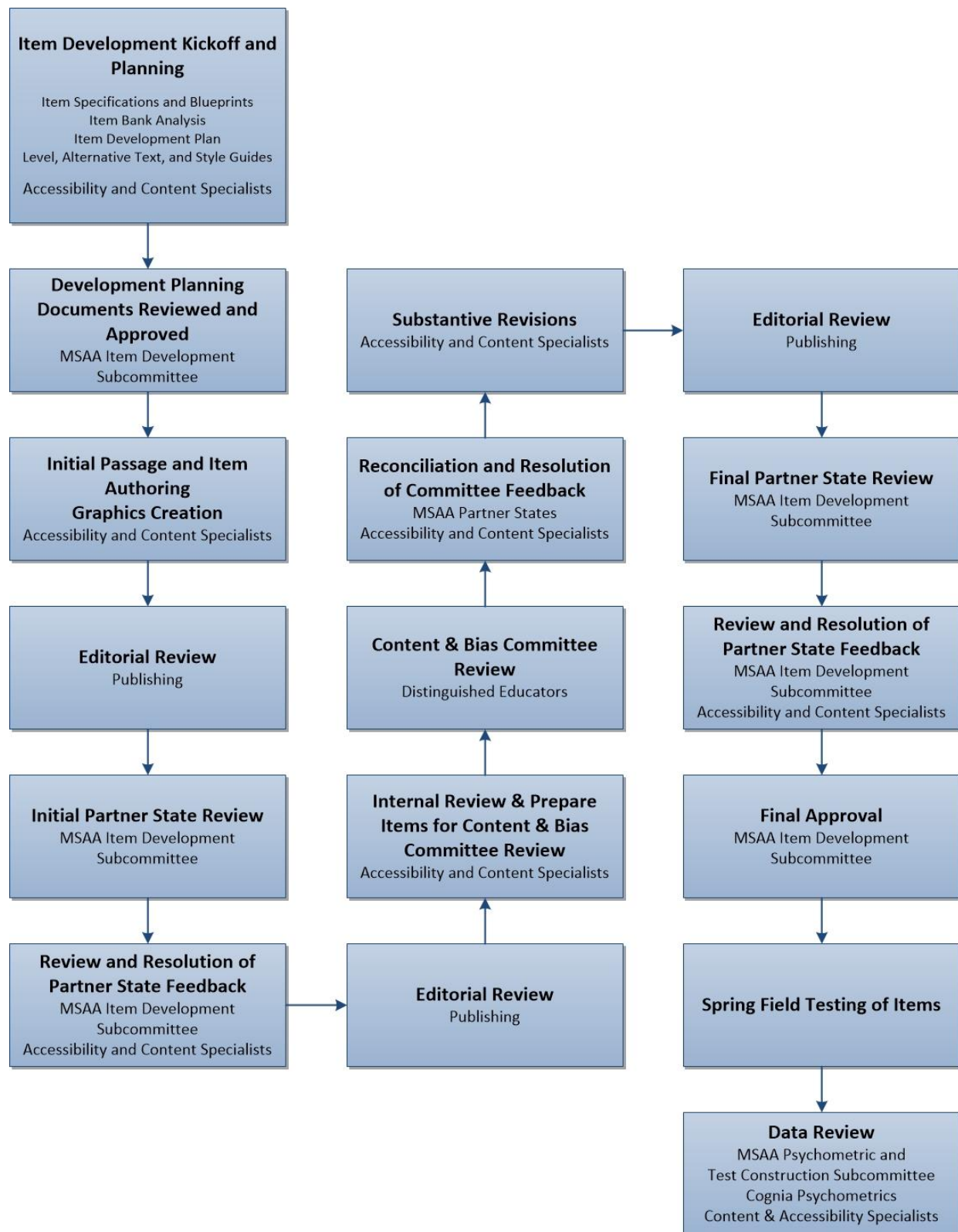
4.1 General Philosophy and Role of the Item Development and Psychometric Subcommittees and Other Stakeholders in Test Development

As discussed previously, the MSAA is a comprehensive assessment system designed to promote increasingly higher academic outcomes for students with the most significant cognitive disabilities in preparation for a broader array of post-secondary outcomes. Given the wide diversity of the student population, great emphasis is placed on ensuring that the MSAA is appropriate and accessible to all eligible students.

The MSAA operational items are from the previous NCSC 2015 administration, as well as the 2016, 2017, 2018, 2019, 2021, 2022 and 2023 MSAA administrations. As described in Chapter 3, MSAA has followed an iterative item development process that allows for multiple rounds of review by a range of stakeholders, including MSAA Partners, content experts, partner-selected representative reviewers, and external reviewers for content and bias. Items are field-tested during the spring administration following their development and initial review. After field testing, items undergo data analysis and a formal data review process with MSAA Partners.

While the 2025 assessment reused content from the 2023 administration, that content was originally developed through this rigorous item development and review process. Figure 4-1 outlines the full item development cycle that was followed in preparing the items used in both administrations.

Figure 4.1 Item Development Process



Before item development begins for the ELA assessment, general and special education teachers, administrators, and education specialists are selected to review reading passages for content accuracy and potential bias or sensitivity concerns. In addition, an item content and bias review committee (Appendix D) convenes each summer to evaluate newly developed ELA and mathematics items. Content review groups examine items for alignment to the Foundational Knowledge, Skills, and Abilities (FKSA) or Essential Understandings (EU), depth of knowledge, clarity, and consistency in teacher directions. Separate bias and sensitivity committees review items for potential bias, sensitivity, and accessibility issues.

Based on the committees' recommendations, 100% of ELA and mathematics items were accepted or accepted with revisions. Likewise, all reviewed passages were accepted or accepted with revisions. Final passage and item approval was conducted by MSAA Partners. **Note:** *these reviews did not take place for the 2025 assessment because the test content was reused from the 2023 administration. All content included in the 2025 test had already undergone the necessary reviews during its original development for the 2023 test.*

Item Development Subcommittee Involvement

The MSAA Item Development Subcommittee, made up of MSAA Partners, provides overall direction and guidance regarding field-test item development. This multistage development and review process by the subcommittee and educator review committees provides ample opportunity to evaluate items for their accessibility, appropriateness, and adherence to the principles of Universal Design for Learning. Documentation that guides these reviews has been developed and updated collaboratively with the Subcommittees throughout the life of MSAA. This documentation includes:

- item specifications
- level guidelines
- alternative text guidelines
- item review checklists
- bias checklists

Through these reviews, ensuring items are accessible is a primary area of consideration throughout the item development process. This focus on accessibility is critical in developing an assessment that allows for the widest range of student participation, as educators seek to provide access to the general education curriculum and foster higher expectations for students with the most significant cognitive disabilities.

4.2 Sample Item Teacher Guide

A resource developed in 2021, called the *Sample Item Teacher Guides*, was created to help teachers use the sample items as an additional assessment tool. This allowed teachers to further understand what students may know and be able to do based on their performance on these sample items. They responded to this information by applying instructional strategies and scaffolding suggestions outlined in the Teacher Guides. An example from the *Sample Item Teacher Guides* is included in Appendix E.

The Teacher Guides have a blueprint table at each grade that outlines the items in each sample test. The ELA blueprint table/overview can be used to help select the sample item(s) that will provide the best evidence of student learning. The learning targets differentiate between the types of evidence each item will provide. The item type describes how the student will engage with the item: through multiple choice,

constructed response, or open response. For ELA, items that address reading standards are grouped by passage set; each passage set primarily addresses standards in genre-specific content categories. The passages for items that assess reading standards are accessed in the *Directions for Test Administration* (DTA) and computer-based testing platform.

To obtain evidence of understanding for each grade-level standard, teachers can use the Teacher Guides to do the following:

- Access the sample items for the students' grade levels.
- Use items individually as the learning targets are covered in class.
- Use the items in small groups to address a series of learning targets that focus on one standard.
- Use the entire sample item set to measure students' understanding of learning targets before, during, or after instruction.
- Review sample item sets from lower grades to build understanding of prerequisite skills for a given standard.
- Review sample item sets from higher grades to know how standard and item information build from the target grade.
- Use the sample items as models to create additional items to assess the standards.

Chapter 5. Training and Administration

5.1 Test Administrator and Test Coordinator Training

The MSAA Partners adhere to the premise from the testing standards (AERA et al., 2014) that a key consideration in developing test administration procedures and manuals is that test administration should be fair to all examinees. When all Test Administrators (TAs) utilize the same well-defined administration procedures and the provided training, manuals, and supporting documents, administration is prescribed, standardized, and poised to be fair to all examinees. Test Coordinators (TCs) are directly responsible for supporting TAs in understanding and following the administration procedures. Comprehensive TC training and materials targeted to their role and responsibility ensure that they are appropriately prepared to support the TAs.

As the MSAA is a computer-administered test, the administration procedures are consistent with the hardware and software requirements of the test specifications. MSAA requires completion of training by all TCs and TAs to support standardized-test processes and procedures. MSAA provides ancillary testing materials each year outlining specific practices and policies including (a) *the Test Administration Manual (TAM)*; (b) MSAA Online Test Administration Training; (c) *the MSAA Online Assessment System User Guide for Test Administrators*; (d) *the MSAA Online Assessment System User Guide for Test Coordinators*; and (e) grade-, content-, and form-specific *Directions for Test Administration (DTA)*. The online training and supporting documents are both comprehensive and prescriptive, while also clearly outlining where Test Administrators (TAs) have flexibility in administering the MSAA. TCs and TAs receive these resources to support consistent implementation and uphold the validity of assessment results. In addition, standardized training and documentation assist MSAA Partners in maintaining testing integrity by helping to prevent, identify, and address irregularities in academic testing, particularly in technology-based environments.

5.2 Test Administrator Training Modules

Traditionally, the online training modules for Test Administrators (TAs) have been available starting two weeks before the testing window and remained accessible throughout the testing period. For the AY25 administration, MSAA Partners extended the pre-administration training window by one additional week, allowing TAs to begin accessing the training modules up to three weeks before the start of testing. The training modules are customized to address the specific responsibilities of the TA and to provide important information from the three documents TAs are required to use: the (1) *TAM*, (2) *DTA*, and (3) *MSAA Online Assessment System User Guide for Test Administrators*. These training modules were updated for the 2025 administration in conjunction with the updates to the required documents. There are six modules (see Table 5-1). Each module requires approximately 17–34 minutes to complete.

Table 5-1. Training Modules for Test Administrators

Module 1:	MSAA Overview
Module 2:	Navigating the MSAA Online Assessment System
Module 3:	Test Administrator and Test Coordinator Responsibilities
Module 4:	The Writing Prompt
Module 5:	Accessibility Features and Accommodations
Module 6:	Student Response Check and Early Stopping Rule

TAs are required to view the training modules (accessed through the MSAA System) in sequence and to successfully complete a final quiz after viewing all modules. Each module must be viewed before the link for the subsequent module becomes accessible.

Questions pertaining to information in the module follow each online training module for TAs. These questions are included as a review of the content to prepare TAs for the final quiz. TAs must obtain a score of 80% or higher on the final quiz to be certified to access the secure test administration materials. The TAs are notified within the MSAA System whether they pass the final quiz. They are allowed multiple attempts to obtain a score of 80% or higher on the final quiz. TAs are allowed access to the secure test materials only after fulfilling this certification requirement.

In addition to the module training, TAs are instructed to become familiar with the online system by accessing sample items. In addition to the sample items, which were developed by content and measurement experts for teachers, administrators, and policymakers for the NCSC assessment, MSAA added sample items for the 2025 administration that are representative of current MSAA item development. The sample items do not address all assessed content at each grade level and are not representative of every type of item. Rather, the sample items provide a preview of the array of items and illustrate multiple item features that allow students with a wide range of learner characteristics to interact with the assessment process.

5.3 Test Coordinator Training Modules

Similar to the TA training modules, online modules tailored for Test Coordinators (TCs) are available both prior to and during the testing window. These modules highlight key information from the primary resources TCs are expected to use: the (1) *Test Administration Manual (TAM)*, (2) *MSAA Online Assessment System User Guide for Test Coordinators*, and (3) *MSAA Online Assessment System User Guide for Test Coordinators*. The TC training modules are updated annually to reflect revisions identified through reviews conducted by Cognia and MSAA Partners. There are six modules, each ranging from 17 to 34 minutes in length (see Table 5-2)

Table 5-2. Training Modules for Test Coordinators

Module 1:	MSAA Overview
Module 2:	Navigating the MSAA Online Assessment System
Module 3:	Test Administrator and Test Coordinator Responsibilities
Module 4:	The Writing Prompt
Module 5:	Creating and Managing Users and Classrooms
Module 6:	Student Response Check and Early Stopping Rule

TCs are required to view the online training modules (accessed through the MSAA System) in sequence. Each module must be viewed before the link to the subsequent module becomes accessible. There are questions at the end of each module as a review of the content of that module. TCs are required to complete the online training but not required to take a final quiz.

5.4 Best Practice Videos

The best practice videos are accessed through the MSAA System and provide TAs with targeted information about the MSAA. Video 1 focuses on (1) reviewing assessment features that are available within the MSAA online system, (2) how to go to full screen mode and zoom within the browser, and (3) procedures to follow when using the hybrid approach to administration (i.e., both online and paper-pencil formats). Video 2 focuses on the purpose and steps of conducting the student response check (SRC) and

on how to implement the early stopping rule (ESR). Videos 3 and 4 focus on administration of the open-response writing prompts. In each of these videos a mock student-TA interaction is used to provide TAs with a true picture of these administration processes (see Table 5-3).

Table 5-3. Best Practice Videos

Video 1:	How to Administer an Item
Video 2:	How to Administer the SRC and Implement the ESR
Video 3:	How to Administer a Level 2 Writing Prompt
Video 4:	How to Administer a Level 3 Writing Prompt

5.5 Test Administration Manual

The *Test Administration Manual* (TAM) provides an overview of, and the guidelines for, planning and managing the MSAA administration for district and school personnel. Additionally, the TAM defines the roles and responsibilities of the TA, TC, and State MSAA Coordinator, who are involved in and oversee the administration of the MSAA. It is organized according to the following tasks:

- providing an overview of the MSAA and the required documents (i.e., TAM, DTA, *MSAA Online Assessment System User Guide for Test Administrators*, *MSAA Online Assessment System User Guide for Test Coordinators*);
- defining the roles and responsibilities of the TA and TC, as well as training requirements;
- describing the accessibility features for both online and paper administration as well as the allowable accommodations (i.e., assistive technology, paper version, scribe, sign language); and
- providing detailed information about how to maintain test security and what constitutes a test irregularity.

The TAM also contains appendices for scribe accommodation and sign language accommodation protocols, the procedures for annotations, and guidelines regarding the use of augmentative and alternative communication by students taking the MSAA. The TAM is accessible to TAs and TCs through the MSAA System and is made available approximately 2 weeks prior to the beginning of the testing window, as well as throughout the testing window.

5.6 Directions for Test Administration (DTA)

The secure grade-, content-, and form-specific DTAs are required to be used by TAs when administering the MSAA. Each DTA is accessible through the MSAA System once a TA has been certified. The DTAs must be used by the TA for MSAA administration. The following elements are provided as part of each DTA (as applicable for a content area):

- standardized directions and scripts to be followed exactly as written for each item, including alternative text as appropriate;
- details about manipulatives required to administer a test item, such as calculators and counters;
- reference sheets that contain important graphics;
- materials required for administration of the mathematics and ELA constructed-response items;
- scoring rubrics for mathematics and ELA constructed-response items;
- writing prompt scripts, graphic organizers, student response templates, and stimulus materials for all writing prompts in each grade-level ELA DTA; and
- specific directions on how to administer the braille versions of ELA foundational reading items in grades 3 and 4.

While the TA has some flexibility in presentation and response mode to ensure the MSAA is accessible to a student, the DTAs are designed to provide standardization to ensure a TA is not changing what is being measured.

5.7 Test Coordinator and Test Administrator User Guides

The *MSAA Online Assessment System User Guide for Test Coordinators* and *MSAA Online Assessment System User Guide for Test Administrators* provide technical information and troubleshooting tips, plus step-by-step instructions to navigate the MSAA System. Each user guide contains specific information relevant to the role of the TA and the TC. The user guides provide many efficient screenshots that demonstrate the functionality of the MSAA System. The user guides also contain appendices that describe accessibility features, assistive technology compatibility, and the MSAA System technology requirements.

As with the TAM, the user guides are accessible to TAs and TCs through the MSAA System and are available approximately 2 weeks prior to the beginning of the testing window, as well as throughout the testing window.

5.8 Operational Administration

The administration window for the MSAA was March 10–April 25, 2025. Both the ELA and mathematics assessments were completed within the same administration window. Regardless of administration format (i.e., online or paper), the student assessments were submitted electronically by the TA on or before April 25, 2025. The MSAA is not a timed test. Testing time varies for each student, with testing paused and resumed based on a student's needs. If a student becomes sick or exhibits frustration, lack of engagement, or refusal to participate during the administration of the MSAA, TAs are directed to pause the testing and take a break, which can last for a few minutes or a few days, depending on the student's needs. The MSAA protocols allow the TA to pause and resume the administration of the test as often as necessary during the testing window, based on a student's needs.

Throughout the administration window, monitoring and quality control processes are ongoing, as part of the MSAA. Support is provided to TCs and TAs through the MSAA Service Center, additional support built into the MSAA System functionality, and the MSAA Partner States. TA feedback is gathered through an End-of-Test Survey; additional information regarding this survey can be found in Chapter 5.8.4. Review of the service center logs, and analysis of the test survey results inform MSAA Partners about areas where clarification and further support may be needed.

5.8.1 MSAA Service Center

To provide support to schools before, during, and after testing, Cognia operates and provides tiered technical support through the MSAA Service Center. The MSAA Service Center is available year-round from 8:00 a.m. to 5:00 p.m. EST, Monday through Friday, to accommodate the multiple time zones in which the test is administered. One week prior to the administration window and ending one week after the close of the administration window, the MSAA Service Center has extended hours from 6:00 a.m. to 10:00 p.m. EST, Monday through Friday,

The TAM directs TAs and TCs to contact the MSAA Service Center with questions pertaining to the MSAA System and test administration procedures. The MSAA Service Center's toll-free support number, e-mail address, and chat link are disseminated to the field through the MSAA System and related communications.

Functionally, support is provided in a tiered manner, where Tier 1 support involves direct support to the caller by MSAA Service Center representatives, Tier 2 support includes support by the program management team for items such as policy questions, and Tier 3 support applies to technical requests, which are escalated to the technology vendor for attention.

All activity is tracked in the MSAA Service Center ticketing system, ServiceNow, and is included in weekly status reports that are provided to MSAA Partners. These reports summarize ticket activity, call analysis data (e.g., call duration, hold time), and per-grade/content and per-state test status summaries throughout the administration window.

5.8.2 Additional Supports

In addition to the MSAA Service Center, the Cognia program management team periodically provides direct phone and e-mail support where logistical or procedural support is needed by the field. Cases with policy or consortium-wide implications are directed to MSAA Partners and related policy documentation.

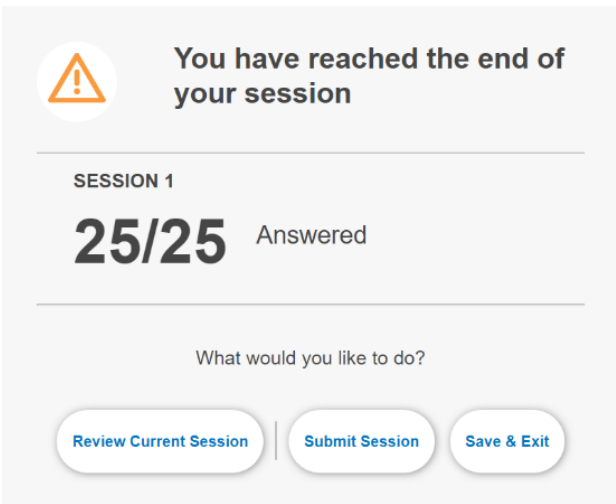
Furthermore, a banner messaging system in the MSAA System is implemented, as needed, to notify users of important information during the administration window. When the messaging system is activated, a banner message appears at the top of the screen upon login to notify users of system information and upcoming system activities, such as known issues and scheduled system maintenance, as well as upcoming test administration deadlines.

5.8.3 Monitoring and Quality Control

To ensure that proper testing procedures and appropriate test practices are maintained throughout administration, numerous measures are taken both to communicate participants' responsibilities and to monitor the appropriateness, accuracy, and completion of key procedures and tasks. The TAM outlines the procedure for reporting any violation or suspected violation of test security or confidentiality by notifying the school or district TC. TCs are then instructed to follow state procedures regarding reporting the issue or suspected issue; however, district TCs are informed that they must report to the State MSAA Coordinator any incidents involving alleged or suspected violations that are considered serious irregularities. The TAM further explains that the consequences for inappropriate test practices are determined by the individual state's professional codes of ethics and state law.

The online MSAA System contains built-in measures to ensure proper testing procedures, as seen in the session-based test design. When the TA clicks the Next button on the last question of a session, a prompt appears notifying the TA that he or she has reached the end of the session, displaying the number of answered items, and presenting options for the TA to proceed to the next phase of the test (either Session 2 or final submission of the completed test, as appropriate), return to the current session, or save and exit the test.

Figure 5.1 End of Session Prompt



If the TA clicks the *Save & Exit* button, the test will resume the next time on the last item answered. If the TA clicks the *Submit Session* button, the session is submitted and cannot be re-opened, and the TA is permitted to continue to the next phase of the test. This prompt reduces the risk of users accidentally submitting a session without properly understanding the implications.

Throughout the administration window, Cognia monitors activity and provides weekly updates to MSAA Partners on test status trends identified in support calls. These updates provide a mechanism for concerns to be identified early and the appropriate measures to be taken, such as creation of assessment-wide or state-level materials and communications. This high level of communication and collaboration throughout the assessment process contributes to a proper, and valid administration of the MSAA.

5.8.4 End of Test Survey Results

The End-of-Test Survey (EOTS) is designed to capture feedback from Test Administrators (TAs) to help MSAA Partners better understand the test administration experience. Each TA is instructed to complete one EOTS after administering the assessment to all assigned students. At the December 2023 TAC meeting, the committee emphasized the need to shorten the survey to reduce the TAs burden. In response, the Psychometric subcommittee worked collaboratively with Cognia to revise the 2024 EOTS to a five-question format. Following the December 2024 TAC meeting, the survey was further refined for the 2025 EOTS, expanding the survey to include a total of eight questions. Several key questions were retained in the 2025 version. These questions are outlined below along with the data from responses collected.

- *I understood and was able to perform my role as a test administrator (Strongly Agree, Agree, Disagree, Strongly Disagree).*
- **Answered: 3,418 / Skipped: 65**

STRONGLY AGREE	AGREE	DISAGREE	STRONGLY DISAGREE	TOTAL
65.36% 2234	33.06% 1130	0.12% 4	1.46% 50	3,418

- *Select your level of agreement with the statement that best reflects your experience with each of these training materials (Strongly Agree, Agree, Disagree, Strongly Disagree).*

- a. The **User Guide** was helpful for understanding and performing my role as a TA.
- b. The **Test Administration Manual** was helpful for understanding and performing my role as a TA.
- c. The **Training Modules** were helpful for understanding and performing my role as a TA.

<i>Answered: 3373 / Skipped 110</i>	STRONGLY AGREE	AGREE	DISAGREE	STRONGLY DISAGREE	TOTAL
The User Guide was helpful for understanding and performing my role as a TA.	49.33% 1,660	48.26% 1,624	1.49% 50	0.92% 31	3,365
The Test Administration Manual was helpful for understanding and performing my role as a TA.	51.41% 1,731	46.09% 1,552	1.57% 53	0.92% 31	3,367
The Training Modules were helpful for understanding and performing my role as a TA.	51.70% 1,737	45.24% 1,520	2.17% 73	0.89% 30	3,360

The 2024 survey included the following question: *"Please share what aspects of these resources you found most useful to prepare for test administration or provide suggestions for improving these resources."*

For the 2025 version, this open response question was revised and split into two separate prompts to allow for more targeted feedback:

- *What aspects of these resources did you find most useful in preparing for test administration? (Open Response)*
- *What specific suggestions do you have for improving these resources? (Open Response)*

Key updates to the 2025 EOTS included the addition of a required first question that stated:

- *Please select the state or entity for which you administered the 2025 Multi-State Alternate Assessment.*

ANSWER CHOICES	RESPONSES	QUANTITY OF RESPONSES
Arizona	50.70%	1,766
American Samoa	0.72%	25
Bureau of Indian Education	1.78%	62
Commonwealth of the Northern Mariana Islands	0.72%	25
District of Columbia	7.38%	257
Department of Defense Education Activity	4.08%	142
Guam	1.55%	54
Maine	8.50%	296
Montana	8.27%	288
South Dakota	9.10%	317
United States Virgin Islands	0.34%	12
Vermont	6.86%	239
TOTAL	--	3,483

Respondents were required to answer the state/entity identifier question to access the remaining (optional) questions. MSAA Partners agreed that collecting state/entity-specific responses was essential for highlighting areas of success and identifying opportunities for improvement. In addition, Partners collaborated to include two new questions focused on the use of MSAA score reports. These questions aim to gather information on how TAs engage with families using the MSAA score reports after the assessment is complete. The questions and the data collected (from the selected response question) are provided below:

- *Do you plan to discuss your student's results with their parents or guardians?*
- **Answered: 3,349 / Skipped 134**

ANSWER CHOICES	RESPONSES	QUANTITY OF RESPONSES
Yes	67.06%	2,246
No	32.94%	1,103
TOTAL	--	3,349

- *What specific aspects of the student's results do you plan to highlight during those discussions?*
(Open Response)

Chapter 6. Scoring

6.1 Selected-Response and Constructed-Response Item Scoring Processes

6.1.1 Overview of Scoring Process Within the System and Test Administrator/Scorer Training

Overview of Scoring Process Within the Assessment System

The MSAA System provides automated machine scoring for all item types, aside from the open-response writing prompt and mathematics constructed-response items, which require human scoring. The selected-response and constructed-response item types were described in detail in Chapter 3. The student may provide their responses to the items within the MSAA System. The system also allows teachers to enter responses for a student from the paper-based test delivery. The selected-response items are scored according to the answer keys provided in each test package. The constructed-response items are scored as a correct or incorrect student response based on established scoring rubrics, which is then entered by the Test Administrator (TA). At the completion of the operational test, all test data are extracted from the system and are then compiled to generate full result sets for each student's tests.

All item responses are exported from the system and are provided to the Cognia Information Technology Reporting (IT-Reporting) Department. The exported items go through a key verification check, conducted by a data analyst, to confirm that the selected-response and constructed-response item keys were entered correctly. Any items that may be flagged are provided to the content specialists to conduct a blind key check. The content specialists review the actual item and mark the key in the flagged file. Any mismatches are researched by the content specialist, and updates are made following a problem-item notice process to update and correct the key. In cases where no mismatches are found, the content specialist notifies the data analyst, and the file is released for final processing.

Items are scored in the MSAA testing system as correct or incorrect, with each item contributing a score of 1 or 0 to the content-area raw score. Non-responses (blank responses) to any item are scored as 0 points. Detailed score assignments and comprehensive data analysis requirements are provided in the MSAA Assessments Reporting Services Deliverables Decision Rules document, which can be reviewed in Appendix F.

Test Administrator/Scorer Training and Support

All TAs must participate in training modules and pass a final quiz to be certified to administer the MSAA, as described in detail in Chapter 5. During the test administration, TAs use the grade, content, and form-specific DTAs to administer each item. When TA scoring is required, such as in the case of the mathematics constructed-response items, the DTA includes the teacher scripting and directions related to any item setup and administration specifics, any templates required by the items, and the rubrics used to score the items. Once the item is administered, the TA enters the response as correct or incorrect in the MSAA System.

The *MSAA Online Assessment System User Guide for Test Administrators* provides further directions to TAs on entering item responses in the MSAA. The guide outlines the use of the system, including how to enter student responses and submit each content-area test.

For support related to the administration, scoring, entry of student responses, and submission of student responses during the administration window, TAs can call or e-mail the MSAA Service Center with any questions.

6.2 Open-Response Writing Prompts Scoring Processes

6.2.1 Overview of Open-Response Writing Entry Process Within the Assessment System and Test Administrator Training

Open-Response Writing Entry Process

As described in Chapter 1, the open-response writing prompts in grades 3–8 and HS are being operationally administered in the 2025 MSAA. The open-response writing prompts are described in detail in Chapter 3. The student, or a qualified scribe, records the response on either the response template in the MSAA System or the paper response template included in the writing DTA. TAs can upload the student's final writing response template directly in the system, retype the student response within the item response field of the item, or upload the template and retype it within the item response field. The item responses (no matter how they are entered) are then extracted from the online system and provided to Cognia for human scoring.

Test Administrator Training and Support

All TAs are required to participate in administration training modules and pass a final quiz to be certified to administer the MSAA assessment, as described in Chapter 5. The TA training includes review of the parameters for the administration of the open-response writing prompt, as well as entry of the student responses into the MSAA System. In addition, the best practice videos provide a student-TA representation that gives TAs a true picture of the processes involved in conducting the open-response writing prompt. During the test administration, TAs use the grade-, content-, and form-specific DTAs to administer each open-response writing prompt. The DTAs include the teacher scripting and directions related to any item setup, administration specifics, and the materials for the open-response writing prompt.

The *MSAA System User Guide for Test Administrators* provides further direction on entry of student responses to the open-response writing prompt. Additionally, the MSAA Service Center provides support for TAs.

6.2.2 Benchmarking and Identification of Scoring Materials

The open-response writing prompts were benchmarked during the 2015, 2016, and 2017 field tests. During the benchmarking activity, Cognia scoring experts (Scoring Supervisors and Scoring Team Leaders [STLs], defined below) worked collaboratively with NCSC representatives in 2015 and with MSAA representatives from the Scoring Subcommittee in 2016 and 2017 to review student responses, assign a score based on the MSAA grade- and level-specific rubrics for each trait (i.e., organization, idea development, conventions), and identify item-specific writing anchors and practice sets.

The final scores for the anchor and practice sets were recorded, and representatives from NCSC (2015) and the MSAA Scoring Subcommittee (2016 and 2017) acknowledged their consensus on the sign-off

document for each prompt. Also, development of a scoring decisions document began in 2017. It was reviewed by the MSAA Scoring Subcommittee, which provided rationale and decision points to be used during scoring by the Scoring Supervisors and STLs.

Following the identification of the anchor sets, two qualification sets were identified for each prompt. Each qualification set consisted of 10 responses; scores were based on anchor responses and scoring decisions made during the benchmarking meetings. The MSAA Scoring Subcommittee reviewed and approved the scores and responses used for qualification sets.

6.2.3 Scorer Recruitment and Qualifications

The MSAA scorers are a diverse group of individuals with a broad range of backgrounds, including teachers, business professionals, graduate students, and retired educators. All selected scorers hold the required minimum of a four-year college degree that includes ELA or writing coursework. Preference is given to those who have some experience with an alternate-assessment student population. In 2025, 76% of the scoring team had previous experience with scoring the MSAA assessment. Within the scoring team, 27% hold a master's degree and one person holds a Ph.D. All scorers sign a nondisclosure/confidentiality agreement.

6.2.4 Cognia Staff and Scoring Leadership

The MSAA operational open-response writing prompts were scored between May 7 and May 20, 2025. All Level 2 and Level 3 scoring was conducted in a remote setting via zoom during scheduled office hours.

The following staff members participated:

- ELA Group Manager for Scoring: Responsible for managing scoring-related activities and monitoring reports, as well as leadership and training of scorers to ensure overall consistency of scoring.
- Accessibility Assessment Specialist: Responsible for overseeing scoring activities and acting as the accessibility lead in coordination with the Cognia scoring staff.
- *iScore* Operations Manager: Responsible for setup and maintenance of *iScore* scoring system and for coordinating technical communication.
- Scoring Content Specialist: Responsible for selecting calibration responses, training STLs and scorers, resolving arbitrations, and supervising the consistency of scoring for items in assigned grades.
- Scoring Team Leader (STL): Responsible for performing quality-control measures, resolving arbitrations, and monitoring the accuracy of a small group of scorers, usually consisting of not more than six.

6.2.5 Training

Scoring Content Specialists thoroughly review the decisions and materials that result from the benchmarking meetings in preparation for training STLs and scorers. One Scoring Content Specialist is assigned to all Level 2 writing prompts and one Scoring Content Specialist is assigned to the Level 3 writing prompts.

The Scoring Content Specialists review prerecorded training modules for use in training. Leadership training for 2025 took place from April 30th through May 6. STLs are required to meet or exceed the accuracy standard of 80% exact agreement on all items and at least 90% exact/adjacent¹ agreement on

¹ "Adjacent agreement" means that the two scores differed by only one score point.

each trait. This requirement applies to each of the three writing traits² individually across qualification sets 1 and 2. The STLs are also present during scorer training, which further reinforces their understanding of the rubrics and training materials.

Scoring Content Specialists conduct training on each open-response writing prompt before scorers are allowed access to student responses. Scorers are divided into two groups with one group focusing on Level 2 items and the other on Level 3 items. The ELA Group Manager oversees the training sessions for scorers. Training sessions are facilitated by the Scoring Content Specialists and are conducted in the following manner:

- Training commences with an introduction to scoring and an overview to explain the purpose and goal of the testing program and any unique features of the test and/or testing population.
- A general discussion addresses the security, confidentiality, and proprietary nature of testing, scoring materials, and procedures.
- Initial item training consists of a pre-recorded module that focuses on the following:
 - the three traits of the MSAA analytic rubrics for writing and how the scoring for each trait is applied to student work (See “Writing Scoring Rubrics,” an appendix to the *MSAA 2025 Guide for Score Report Interpretation*, provided in Appendix G of this report)
 - pertinent information on the testing instructions and item stimulus
 - actual responses with an item-specific anchor set, averaging 10 responses representing a range of scores across traits
 - anchor exemplars (presented in a predetermined order) that consist of responses that are typical, rather than unusual or uncommon; solid, rather than controversial or borderline; and true
 - the anchor response score and the scoring rationale, allowing scorers to internalize typical characteristics of each score point
- Scorers are instructed to refer to the anchor set frequently during scoring.
- After completing the module, training continues with the Scoring Content Specialist presenting the supplementary training materials. These consist of practice responses representing all score points across traits, when possible, and often containing responses that are more unusual and/or less solid (e.g., are shorter than normal, employ atypical approaches, or contain both very low and very high attributes). None of the practice papers contain responses that would require identification as nonscorable responses.
- During the review of practice responses, the trainer(s) often focuses on the distinction between adjacent score points or clarification of other scoring issues that are traditionally difficult for scorers to internalize.
- After scorers independently read and score each practice response, the trainer(s) discusses the actual score and explains the rationale.
- A question-and-answer segment addresses any remaining questions from scorers and provides clarification prior to the qualification process.

6.2.6 Qualification

Following the training for each prompt, scorers are required to complete a qualification set to determine eligibility to score student work. There are two qualification sets in each grade and level consisting of 10 responses each. The responses, which represent a range of score points, are randomly distributed to scorers through *iScore*.

Scorers have two opportunities to qualify. If scorers attain a score match of at least 70% exact and 90% exact/adjacent agreement on all three traits for the first qualification set, they are considered a “qualified

² The three writing traits are organization, idea development, and conventions. See rubrics embedded in Appendix G.

scorer” and permitted to score live student responses. If they do not attain the required percentages, the Content Specialist conducts retraining. Following this retraining, scorers are assigned qualification set 2. Since scorers qualify at the trait level, a scorer who qualifies on the first and third trait in qualification set 1, for example, receives the retraining referenced above. However, this scorer would only be required to qualify on trait 2 in qualification set 2. If the data indicates a qualified scorer has shown weakness in a particular trait, they receive additional training prior to the start of scoring.

Scorers who fail to achieve the minimum levels of agreement are not allowed to score. When scorers demonstrate a level of understanding and the ability to apply feedback during the training and qualification process on a certain writing prompt, Scoring Leadership may choose to include the scorer in future trainings on a different writing prompt.

When the queue for the first open-response writing prompt is cleared, the training process is repeated for the next grade and level. This process continues until all 14 open-response writing prompts are scored. (See Section 6.2.1.) Scorer qualification success rates tend to improve as they train additional items/grades. They increase their understanding of the rubric and its use over time. In most instances, the initial team of scorers who begin on the project remain on the project. However, some attrition occurs over scoring, and additional scorers are added to maintain the scoring schedule. Often, the Level 2 scorers will join the Level 3 scorers near the project’s end and only after they have completed scoring of Level 2. The qualification results are listed in Table 6-1. This table includes the total number of scorers and leadership who attempted to qualify on the item and the total number who successfully did so. It also includes the percentage of people who passed all three traits on either the first or second qualification set combined.

Table 6-1. Scorer Qualification Rates

Grade	Level 2# Attempted	Level 2# Passed	Level 2% Passed	Level 3# Attempted	Level 3# Passed	Level 3% Passed
3	5	5	100%	12	10	83%
4	7	5	71%	14	11	78%
5	9	9	100%	17	16	94%
6	9	9	100%	14	13	92%
7	8	8	100%	11	7	63%
8	10	10	100%	10	10	100%
HS	4	4	100%	16	14	87%

6.2.7 Method for Scoring Operational Open-Response Writing Prompts

Student responses to the open-response writing prompts and any uploaded material are exported from the platform and imported to the Cognia *iScore* scoring system. Through *iScore*, qualified scorers read and evaluate student responses, submitting scores electronically. The processes by which images are logged in, scanned, and uploaded into *iScore* provide anonymity to individual students and ensure random distribution of all responses during scoring.

All student responses are scored from uploaded evidence and/or computer-generated text, defined as student work directly entered in the MSAA System. For Level 2 prompts, when both uploaded and computer-generated text is available, the uploaded evidence is scored first, and the computer-generated text is used for clarification and confirmation of the uploaded student writing evidence. When there is only uploaded writing evidence but no computer-generated text to provide clarification and confirmation, the uploaded writing evidence is scored. When there is only computer-generated text but no uploaded writing

evidence, the computer-generated text is scored. For Level 3 prompts, the computer-generated text and the uploaded evidence serve to provide a holistic demonstration of student ability and are considered together when both are available. When only one portion is available, the prompt is scored like a Level 2 prompt.

The following processes are in place during the scoring of the MSAA operational open-response writing prompts:

- The *iScore* system forces scorers to review all available pages before allowing a score to be submitted.
- All scoring is “blind.” Only booklet numbers within *iScore* are linked to student responses; no student names are visible to scorers unless a name appears on material uploaded by the TA.
- Cognia maintains security during scoring by using a highly secure server-to-server interface to ensure that access to all student response images is limited to those who are scoring or working for Cognia in a scoring management capacity.
- During scoring, *iScore* enables constant measuring and monitoring of scorers for scoring accuracy and consistency. Each scorer’s reading rate and total number of responses scored are also monitored.
- Scorers are required to maintain an acceptable scoring accuracy rate (80% exact averaged across the three traits with no trait less than 70%). This is measured through read-behinds, and daily calibration sets. (These measures are described below.)
- Scorers who repeatedly fall below standard are retrained or dismissed from scoring that item as detailed in section 6.2.8.
- Scoring rules are in place to determine the final score of record, or when a final score is to be provided by Scoring Leadership. (For examples of scoring resolutions, see Section 6.2.8.4.)

Table 6-2 represents the total number of student responses scored by writing prompt in each grade.

Table 6-2. Student Responses per Grade

Grade	Number of Student Responses for WRCC002	Number of Student Responses for WRCC003	Number of Student Responses Total
3	785	697	1,482
4	713	722	1,435
5	708	713	1,421
6	509	875	1,384
7	401	888	1,289
8	546	662	1,208
HS	468	718	1,186

Note: For identification purposes in iScore, Level 2 prompts were designated as WRCC002 across all grades, and Level 3 prompts were designated WRCC003.

Scoring Rules

All open-response writing prompts are scored against a three-trait rubric (see rubrics in Appendix G). The scoring scale options of 0, 1, 2, and 3 are applied to each trait. (Note: for determining a student’s total raw score to be transformed to a scaled score, the score categories of 1 and 2 were combined to be converted to a 1, and score category 3 was converted to a 2. These converted trait scores were the scores used in the psychometric analyses.) When a response does not conform to score point parameters, scorers can designate the response as a Blank, Unreadable, or No Score.

Table 6-3 displays the resolution process for each of the responses mentioned above.

Table 6-3. Scoring Resolution Process

Designation	Resolution Process
Blank	Responses scored Blank are sent to another scorer for a second read. Responses scored Blank twice are converted to zeros ("0") for reporting purposes. Any discrepancies are resolved by Scoring Leadership.
Unreadable	Responses judged unreadable are forwarded to a special queue within <i>iScore</i> to be reviewed by a Scoring Supervisor, who resolves the student score. (If the response remains unreadable after review, the Scoring Supervisor assigns a score of "0.")
No Score	Responses where the uploaded evidence is a mismatch to the typed response are escalated to Scoring Leadership for response appraisal and scoring. Responses that legitimately respond to another item are escalated for review by Scoring Leadership. Any student response indicating administrative inconsistencies, potential cheating, and/or security lapses before, during, or after the test administration is scored based on its merits and then forwarded for review. If further attention is warranted, the State Services team notifies the appropriate MSAA Partner State. Responses determined to be nonscorable are resolved by the Cognia leadership team in consultation with the MSAA Scoring Subcommittee, if necessary.

Scorers also have the option of flagging a response as an "Alert," requiring immediate review and possible immediate action by Scoring Leadership and an MSAA Partner State. "Alert" responses can include, but are not limited to, those that suggest one or more of the following problems:

- thoughts of suicide
- criminal activity
- alcohol or drug use
- extreme depression
- violence
- rape, sexual or physical abuse
- self-harm or intent to harm others
- neglect

In 2025, there were no "Alert" responses flagged during scoring.

6.2.8 Monitoring of Scoring Quality Control

Scorers are continuously monitored to ensure that scoring is accurate and consistent. Throughout the scoring process, read-behind scoring, double-blind scoring, and calibration sets are used as quality-control measures. MSAA Scoring Subcommittee representatives, along with the Cognia Accessibility and Scoring teams, may monitor reports daily. Read-behind and double-blind statistics are reviewed daily. Calibration sets are administered and reviewed repeatedly during the project. Scoring Leadership and Content Specialists from the Scoring Services and Content Development/Accessibility departments at Cognia pay close attention to the disaggregated read-behind, double-blind, and calibration statistics.

Scorers needing more clarification on applying scores to specific traits are coached by Scoring Leadership. This continuous training allows Scoring Leadership an opportunity to resolve issues, reiterate scoring guidelines, and establish parameters for atypical student responses. Scorers who demonstrate

inaccurate or inconsistent scoring are retrained and allowed to resume scoring under increased supervision. Scoring Leadership removes scorers who continue to fall below accuracy standards. On any day that a scorer falls below accuracy standards, the work is voided and rescored by other qualified scorers. During 2025 MSAA scoring the void process occurred 45 times. In addition, there were 6 occasions during the scoring of L3, where scorers were released from one item but were successful in scoring at a different grade level. There were no scorers released from an L2 grade that moved on to score other grades.

6.2.8.1 Calibration Sets

To determine whether scorers are still calibrating to the scoring standard, they are required to complete a trio of online calibration sets at the start of each day, beginning with the second day of scoring. Scoring Leadership selects the responses for the sets, with each calibration set consisting of five responses representing a range of scores. Scorers who assign at least 12 out of 15 scores exactly can then begin scoring for the day. Scorers who fail to meet that standard are retrained by discussing the calibration responses in terms of the rubric and the anchor set. Scoring Leadership determines if these retrained scorers should be allowed to begin scoring; though if they are, these scorers continue to be closely monitored. Over the course of scoring, only three scorers (across all seven grades and 14 items), required retraining after completing a calibration set. Scorers who received retraining successfully were permitted to resume scoring.

6.2.8.2 Read-Behind Scoring

Read-behinds provide a crucial tool in verifying scorer accuracy. STLs complete read-behinds on individual scorers on a daily basis. The STL's evaluation of each response is performed with no knowledge of the scores assigned across traits. The scores are only available to the STLs after they have also scored the response. If there is a difference in scores, either adjacent (one score point difference) or discrepant (more than one score point difference), the STL score is the score of record. If the scores are discrepant, or if there are a significant number of adjacent scores between the scorer and the STL, the STL discusses the rationale with the scorer.

The average number of read-behinds for each scorer is 5–10 reads a day, but this number varies depending on the accuracy of each scorer and the total number of responses. In this administration, 5.3% of all responses had a read-behind performed on the score. The read-behinds provide an immediate means of identifying scorers in need of further clarification on how to effectively apply the scoring rubrics to student responses. If scorers fall consistently below the 70% exact and 90% exact/adjacent (combined) threshold Scoring Leadership voids their scores for the day and may release them from scoring that item. Scoring Leadership monitors scoring accuracy and consistency by reviewing the read-behinds performed by the STLs and completing read-behinds on the STLs when possible. On average, for L2 grades the read-behind exact agreement rate per trait was 92% for Organization, 92% for Idea Development and 93% for Conventions. On average for L3 grades, the read-behind exact agreement rate per trait was 88% for Organization, 83% for Idea Development and 86% for Conventions.

6.2.8.3 Double-Blind Scoring

While read-behinds measure scorer accuracy in relationship to STL scores, double-blind scoring provides statistics on scorer-to-scorer agreement. Double-blind scoring is the practice of having two scorers independently score a response, without knowing either the identity of the other scorer or the score the other scorer assigned. In double-blind scoring, neither scorer knows which response will be (or already has been) scored by another randomly selected scorer. All responses for MSAA are 100% double-blind scored.

In addition to monitoring interrater agreement rates, double-blind scoring allows Scoring Leadership to resolve arbitrations when two scorers' double-blind scores do not agree across any of the three traits. If there is not exact agreement, iScore automatically places the response into an arbitration queue. Scoring Leadership, with no prior knowledge of the scores assigned, evaluates the response, with the leadership score becoming the score of record. The double-blind statistics provide an overview of agreement rate among the entire pool of scorers and assists in identifying any need for retraining.

6.2.8.4 Final Score Resolution

If scorers are adjacent in their scoring of a response, the two scores are averaged and rounded up for the score of record. If the scorers are discrepant in their scoring, the response will be sent to an STL for arbitration. The STL will review the response, provide the final score of record, and counsel scorers as needed. During the arbitration by scoring leadership, all three traits are evaluated and the final score of record for each trait is supplied by scoring leadership.

In read-behind cases, the Scoring Supervisor/STL score is the final score of record. For adjacent and discrepant scorer scores, the read-behind score is the final score of record. If a response gets more than one read-behind and the two scores supplied by the STLs do not agree, a resolution score is needed. In the unlikely event that a resolution is required, the Scoring Supervisor provides a final score for all three traits during the post-scoring edit process.

6.2.9 Quality and Production Management Reports

Reports generated through iScore are essential during the scoring of the MSAA. Reports provide real-time statistics for review by the Cognia Scoring team and the MSAA Scoring Subcommittee to closely monitor scoring, thereby ensuring that

- scorer data (individual level) is monitored in real time to allow early scorer intervention when necessary;
- overall accuracy, consistency, and reliability of scoring (group level) is maintained;
- individual traits in need of further clarification are identified; and
- scoring schedules are upheld.

The reports listed in Table 6-5 provide the comprehensive tools and statistical information needed to execute quality control and manage production.

Table 6-4. Scoring Quality Control and Production Management

Report	Description
Read-Behind Disaggregated Summary	The Read-Behind Disaggregated Summary report shows the total number of read-behind responses read by both the scorer and the STL, and notes the number and percentage of exact, adjacent, and discrepant scores across each trait.
Double-Blind Disaggregated Summary	The Double-Blind Disaggregated Summary report shows the total number of double-blind responses read by a scorer and notes the number and percentage of exact, adjacent, and discrepant scores across each trait.
Compilation Report	The Compilation Report shows, for each scorer, the total number of responses scored, the number of calibration responses scored, and the percentage of exact, adjacent, and discrepant read-behind scores across each trait.
Summary Report	The Summary Report lists the total number of student responses loaded into iScore. This report includes the number of reads completed to date and the number of reads that remain.

6.2.10 Interrater Agreement

Kappa statistics (kappa coefficients) measure the agreement among two or more raters. The calculation is based on the difference between the level of agreement observed and the level of agreement that would be expected by chance alone. Kappa is a measure of this difference standardized to lie on a -1 to 1 scale, where 1 is perfect agreement, 0 is exactly what would be expected by chance, and negative values indicate disagreement. The kappa information in Table 6-6 shows agreement between raters at Substantial Agreement or Almost Perfect Agreement ranges for most of the open-response writing prompts across grades. In seven cases, the kappa agreement rate is in the Moderate Agreement range.

Table 6-5. Kappa Agreement—Operational Open-Response Writing

Grade	Item	Organization Trait 1	Idea Development Trait 2	Conventions Trait 3
3	WRCC002	0.71	0.63	0.80
3	WRCC003	0.69	0.66	0.80
4	WRCC002	0.68	0.67	0.79
4	WRCC003	0.73	0.71	0.82
5	WRCC002	0.55	0.54	0.78
5	WRCC003	0.76	0.80	0.78
6	WRCC002	0.65	0.71	0.78
6	WRCC003	0.52	0.51	0.77
7	WRCC002	0.62	0.64	0.81
7	WRCC003	0.58	0.57	0.79
8	WRCC002	0.75	0.70	0.80
8	WRCC003	0.62	0.60	0.79
HS	WRCC002	0.73	0.56	0.74
HS	WRCC003	0.71	0.71	0.74

Note: For identification purposes in iScore, Level 2 prompts are designated as WRCC002 across all grades and Level 3 prompts are designated as WRCC003.

Agreement Ranges:

< 0 Disagreement
 0 = Chance Agreement
 0.01–0.20 Slight Agreement
 0.21–0.40 Fair Agreement
 0.41–0.60 Moderate Agreement
 0.61–0.80 Substantial Agreement
 0.81–0.99 Almost Perfect Agreement

Chapter 7. Reporting

7.1 Development and Approval of Report Specific Documents

The *Reporting Services Deliverables Decision Rules* document ensures that reported results for MSAA are accurate relative to collected data. The *Reporting Services Deliverables Decision Rules* document delineating processing rules is prepared, edited in collaboration with the MSAA Reports Subcommittee, and then approved by all participating MSAA Partners prior to processing of the results. The processing and reporting business requirements and participation status structure provide the framework for the reporting requirements, which are defined for each unique report and similarly edited in collaboration with the MSAA Reports Subcommittee. The *Reporting Services Deliverables Decision Rules* are then approved by the MSAA Reports Subcommittee prior to reporting.

The *Reporting Services Deliverables Decision Rules* document contains the hierarchy by which the participation statuses are assigned for each individual test, incorporating data elements collected by the test platform and directly from the MSAA Partners. The reporting requirements and corresponding report design templates were developed by Cognia with the guidance of the MSAA Reports Subcommittee. Both documents underwent iterative review processes that included draft reviews by the appropriate subcommittee, incorporation of edits, draft reviews by all participating MSAA Partner States, and subcommittee review and integration of feedback, until final revisions were approved by all participating MSAA Partner States.

Creating the Report Design Templates

To develop the report design templates, Cognia worked with the MSAA Report Subcommittee to identify modifications to the templates used last year that would ensure that the data elements, layout, and report text were meaningful for reporting the spring 2025 MSAA results. Once finalized, the results of this collaborative process were presented to participating MSAA Partners for final approval. There were no report design changes for spring 2025 MSAA static reports.

MSAA 2025 Guide for Score Report Interpretation

Cognia uses an iterative process to annually update the *Guide for Score Report Interpretation* with the MSAA Reports Subcommittee. Updates are made to ensure that the guide provides the most helpful information to the district and school staff as they review reports for their own knowledge and as they discuss the reports with parents or guardians. The guide includes an overview of the MSAA, student participation criteria, score reporting overview, and samples of the various types of reports available to schools and districts. Guidelines inform the interpretation and utilization of MSAA scores. The guide also includes explanations for all special reporting codes and messages, as well as performance-level scale score ranges. States are permitted to remove codes not used in their state. Appendices included in this guide contain the Performance-Level Descriptors (PLDs) for ELA and mathematics, a sample individual student report, and the writing prompt scoring rubrics. The final, approved *MSAA 2025 Guide for Score Report Interpretation* is delivered electronically to the MSAA Partners for state-specific revisions and distribution.

7.2 Specific Primary Reports Generated for Schools, Districts, and States

Cognia, in collaboration with the MSAA Reports Subcommittee, annually reviews and updates the following primary reports:

- Student reports
- School and district roster reports
- School, district, and state summary reports

Reports are generated for each school, district, or state that has results, as defined by the MSAA processing and reporting business requirements and reporting requirements. These reports, along with student results data files, are posted online via the MSAA Online Assessment System's secure data and reporting portal. As determined by the MSAA Partners, only Test Coordinators (TCs) are granted access to the online reports. Access is controlled by user-permissioned accounts, as illustrated in Table 7-1.

Table 7-1. Report/File Availability by Role

Reports and Files	State Test Coordinator	District Test Coordinator	School Test Coordinator
Student Report	Yes	Yes	Yes
School Roster Report	Yes	Yes	Yes
District Roster Report	Yes	Yes	No
School Summary Report	Yes	Yes	Yes
District Summary Report	Yes	Yes	No
State Summary Report	Yes	No	No
School Data Files	Yes	Yes	Yes
District Data Files	Yes	Yes	No
State Data Files	Yes	No	No

For the purposes of the assessment system, MSAA Partners are regarded as State TCs. As such, they can add new district and school TCs to the online system and block access to any users no longer in the TC role. For 2025, these reports were provided in July to schools, districts, and to parents as soon as possible at the beginning of the school year.

The primary results reported are the students' scale score and performance-level classification for mathematics and ELA. The performance-level classifications, with cut scores determined through the original standard setting and subsequent standards validation processes (see Chapter 9 for more information), are reported under the generic labels, Level 1, Level 2, Level 3, and Level 4. Level 4 is the highest attainable performance level.

The average scale score and the percentage of students in each performance level are summarized by school, district, and state on both the roster and summary reports. These summaries allow for comparing individual student performance to overall state performance and comparing school and district results with the overall state results.

BIE and VT both have state-specific options for additional, dynamic reporting services. BIE is contracted with eMetric and uses Lighthouse's Reporting/Data Interaction, which is eMetric's K-12 assessment reporting and analytics platform. The disaggregate statistics displayed in the platform were provided to eMetric by Cognia. Vermont is contracted with FocalPoint for a similar service, using FocalPoint's reporting platform, LENS. Disaggregate statistics displayed by FocalPoint were provided to FocalPoint by

Cognia. The Partner states all have access to the same disaggregate statistics in their state student results file, but only BIE and VT have additional platforms that allow them dynamic reporting functionality.

7.2.1 Student Report

The student report is a two-sided, single-page document generated for each student eligible to receive a performance level in at least one content area, as defined by the student report requirements. The report contains results for both ELA and mathematics content areas and was developed for parents and guardians of students who participated in MSAA. Reports are organized by school and posted via the secure-access portal for authorized users to download, print, and disseminate to parents and guardians. Each report contains the student name, test grade, and school on the front of the report. The back page contains the student name, state student ID, school, and test grade. Sample student reports are included in the *MSAA 2025 Guide for Score Report Interpretation*.

Page 1 of the report contains the scale score, performance level, and associated performance-level descriptor for the level obtained by the student for each content area. A sentence below the graphical display explains the standard error of measurement (SEM) in layperson's terms by providing the expected range of scores the student would likely earn if tested again. For example, "If your child were to be tested again, it is likely that they would receive a score between 1228 and 1236."

Page 2 contains a brief overview of MSAA, including examples of some of the built-in supports available during testing, and highlights the compatibility of the assessment with various modes of communication. Parents and guardians are encouraged to discuss with their child's teacher the supports their child used on the MSAA.

Tests for students unable to show an observable mode of communication are closed using the Early Stopping Rule, and the lowest scale score is assigned and displayed along with the Level 1 performance level. This is annotated, and in place of the Level 1 performance-level descriptor, the following text is displayed: Your child did not show a consistent observable mode of communication during the test, and the test was closed by the teacher. Since your child did not complete the test, the results may not be an accurate representation of your child's skills. If you have additional questions, please contact your child's teacher.

If a student receives a student report but does not receive results for one of the two content areas, results for the missing content area are replaced with text encouraging parents or guardians to contact the child's teacher or school for more information.

7.2.2 School Roster Report

The school roster report is organized at the school level and provides a by-grade list of all students enrolled in MSAA, with a snapshot of their participation/test status and results for both content areas. The number of tested students, the average scale score, and the percentage of students by performance level are summarized for the school, district, and state at the top of the roster. The processing and reporting business requirements and roster report requirements identify which participation status codes are included on the roster and which of the participation test status codes are included in each calculation.

The summary information at the top of the school roster report supports interpretation of results by users, typically those at the school and district levels. Given that many schools have a relatively small number of students in this population, MSAA Partners do not suppress information when the number of students participating is small. Additionally, these rosters are confidential to authorized school and district

personnel only. This practice places an added responsibility on users to understand the data in the context of small numbers and to use all the provided information to understand the results, as explained in the *MSAA 2025 Guide for Score Report Interpretation*.

Student results are listed below the summary section and identified by name and state student identification number. It is intended that these data points be used in conjunction with the *MSAA 2025 Guide for Score Report Interpretation*. For each content area, the following student-level elements are reported:

- Participation/Test Status
- State Compare (comparison to state average)
- Scale Score
- Performance Level

7.2.3 Summary Reports

Summary reports are organized at the school, district, and state levels for each entity with at least one student included in summary report calculations. Inclusion in these calculations is defined by the processing and reporting business requirements and summary report requirements. The following information is summarized by grade and content area and displayed for the school, district, and state based on the level of the report:

- Enrolled (number of students enrolled)
- Tested (number of valid student tests)
- Did Not Test (number of enrolled students who did not test)
- Average Scale Score
- Performance Level (number and percentage at each performance level by grade in the state, district, and school)

This summary provides a comparative snapshot of results and participation information at a high level and includes both participation and performance summary information, allowing users to evaluate both aspects of their assessment results as guided by the *MSAA 2025 Guide for Score Report Interpretation*. Sample Student, School Roster, and Summary Reports appear in this document in Appendix G.

7.2.4 Quality Assurance

Proprietary quality-assurance measures at Cognia are embedded throughout the entire process of data capture, analysis, and reporting. The data processors and data analysts who work on the project implement quality-control checks of their respective computer programs. Moreover, when the data are handed off to different teams within the IT-Reporting Department, the sending team verifies that the data are accurate prior to handoff. Additionally, when a team receives a data set, the first step is to verify the data for accuracy.

A second level of quality-assurance measurement is parallel processing. One data analyst is responsible for writing all programs required to populate the student and aggregate reporting tables for the administration. Each reporting table is assigned to another data analyst on staff who uses the processing and reporting business requirements to independently program the reporting table. The production and quality-assurance tables are compared, and only after 100% agreement are the tables released for report generation.

The third aspect of quality control at Cognia involves the Software Quality Assurance (SQA) team, which works together with the data processing and data analysis teams to ensure quality data is captured and delivered accurately. Quality control checks are being performed by the data processors and data analysts as the data is handed off via multiple internal software tools. These quality checks initialize the accuracy of the data being ingested into the database and subsequent tables/columns. SQA develops a test plan that includes previously agreed upon report designs and decision rule documents. Test cases housed in an internal test cases repository are then executed in a process including but not limited to the following steps:

- testing data counts of data imported
- testing data quality of individual fields for valid values, such as Gender, Ethnicity, etc.
- validation scripts are developed by the software developers to ensure that they match business requirements and technical specifications

In this testing effort to ensure the quality of the data, the SQA team uses a sample of schools and districts selected based on multiple criteria, such as:

- unique student testing records
- students completed testing
- students partially completed testing
- invalidated students

Working with the data processing and data analysis teams allows for timely and precise turnaround if any data anomalies are found. To allow full transparency and cohesive teamwork in data validation, test cases are tied to tickets outlining required work.

Finally, the SQA team executes test cases validating student printed reports in comparison to the previously agreed upon report design specifications. Once all the test cases have passed, the SQA team notifies the Cognia Portfolio Management team for final sign-off and communication.

Additionally, Breakthrough Technologies (BT), our partner vendor, has a designated QA team that assists with ensuring testing and reporting data are accurate. It starts with the BT team performing QA validations on the CBT extracts that are handed off and used for reporting. If needed, there is back and forth between Cognia DP and BT to investigate and resolve any anomalies seen in the data. Once the Cognia Reporting team has completed the reporting cycle and produced all report deliverables, they are handed off to BT via Cognia SFTP site. BT completes a roll up of files creating school, district, and state level zip files. These zip files are posted and available for download in the MSAA System for active Test Coordinator users. BT's QA team does validations on the zip files as well as testing of the platform to ensure user permissions and org hierarchies that are assigned to users are functioning as expected prior to the online reporting window going live in the platform. Partner TCs are given time prior to online reporting window opening, during which they can access and review their users and take the appropriate action to ensure access is granted to the appropriate people at the appropriate level.

Chapter 8. Preliminary Statistical Analyses

A complete evaluation of a test's quality must include an evaluation of each test item. Both *Standards for Educational and Psychological Testing* (AERA et al., 2014) and *Code of Fair Testing Practices in Education* (Joint Committee on Testing Practices, 2004) include standards for identifying quality items. Items should only elicit the knowledge or skills that are identified as part of the domain being tested. To this end, items should be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students, particularly racial, ethnic, or gender groups.

Both qualitative and quantitative analyses have been conducted to ensure that the 2025 MSAA ELA (reading and writing) and mathematics items meet these standards. Qualitative analyses are described in chapters 3–6 of this report; this chapter focuses on quantitative evaluations. Statistical evaluations are presented in two parts: (1) differential item functioning (DIF) statistics and (2) dimensionality analysis of inter-item correlations. The item analyses presented here are based on the administration of the MSAA in spring 2025.

Note that classical difficulty (p-value) and discrimination (point-biserial) indices are often used in testing programs to compare the quality of items. However, such indices are not appropriate for a multistage adaptive test. The inappropriateness of these statistics stems from the fact that when two items are on two different stages or on different levels of a stage, the students taking one of the items will tend to have a higher overall ability distribution (as measured by scaled score) compared to the students taking the other item. As an example of the inappropriateness, consider a case where the two items have similar p-values, but one item is from Stage 2A, and the other is from 2C. This similarity would lead to the misleading inference that the two items are comparable in difficulty when, in fact, the 2C item is likely to be much harder than the 2A item. Thus, the classical difficulty and discrimination statistics are not included in the evaluation of item quality presented in this chapter. Chapter 9 provides IRT (population invariant) item parameter details more appropriate for evaluating item difficulty.

8.1 Differential Item Functioning (DIF)

The *Code of Fair Testing Practices in Education* (Joint Committee on Testing Practices, 2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit and that actions should be taken to ensure that differences in performance are due to construct-relevant, rather than irrelevant, factors. Chapter 3 of *Standards for Educational and Psychological Testing* (AERA et al., 2014) includes similar guidelines. As part of the effort to identify such problems, MSAA items were evaluated in terms of DIF statistics.

For the 2025 administration, the standardization DIF procedure (Dorans & Kulick, 1986) was employed to evaluate subgroup differences. The standardization DIF procedure is designed to identify items for which subgroups of interest perform differently, 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 conditional on scale score. Then an overall average is calculated, weighting by the pooled scale score distribution so that it is the same for the two groups.

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 item bias, e.g., cause by construct irrelevant factors. On the other hand, 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 should be reconsidered.

For the 2025 MSAA, four subgroup comparisons were evaluated for DIF:

- Male compared with Female
- White compared with Black
- White compared with Hispanic
- Not economically disadvantaged status compared with economically disadvantaged

The DIF statistics were calculated based only on the members of the subgroup in question in the computations; values were calculated only for subgroups with 100 or more students. The tables in Appendix H present the number of items classified as either “low” or “high” DIF, overall and by group favored. Computed DIF indices have a theoretical range from -1.0 to 1.0 for selected-response items. Dorans and Holland (1993) suggested that index values between -0.05 and 0.05 should be considered negligible. The preponderance of MSAA items fell within this range (see Appendix H). Dorans and Holland further state that items with values between -0.10 and -0.05 and those with values 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; thus, content experts conducted a review of items flagged for DIF.

The number of items with a “high” DIF index for each level (the cognitive complexity of the item; refer to information in Chapter 3 for further detail regarding the levels) is shown in Tables 8-1 and 8-2. Since an item can exhibit DIF for multiple comparisons, an item was counted once if any of the comparisons showed “high” DIF. Tables 8-1 and 8-2 show that only a few items were classified as “high” DIF for each grade and each level. These operational items continue to be monitored for content bias as well as evaluated for potential construct irrelevant factors and any issues to be addressed. These results indicate that the content bias reviews for ELA and Mathematics were conducted thoroughly. Please note that items listed in Tables 8-1 and 8-2 are operational items, which means they have passed the initial content and bias review during the field test. These high DIF operational items were monitored closely by both Cognia content and psychometric teams once identified. For instance, if an item exhibits significant Differential Item Functioning (DIF) in year 1 but doesn't display such a pattern in year 2, that item will be retained in the item bank. Nevertheless, these items were not removed from the test form until further substantiating evidence was obtained to justify their removal.

Table 8-1. Number of Items with “High” DIF by Level—ELA

Grade	N_Item	Level 1	Level 2	Level 3
3	4	1	1	2
4	5	0	3	2
5	2	1	1	0
7	4	1	1	2
6	4	0	3	1
8	3	0	2	1
11	5	0	4	1

Table 8-2. Number of Items with “High” DIF by Level—Mathematics

Grade	N_Item	Level 1	Level 2	Level 3
3	2	0	1	1
4	5	0	3	2
5	2	0	2	0
6	1	0	1	0
7	4	2	2	2
8	4	1	2	2
11	2	0	2	0

8.2 Dimensionality Analysis

Because tests are constructed with multiple content-area subcategories, and their associated knowledge and skills, the potential exists for a large number of dimensions being invoked beyond the common primary dimension. Generally, the subcategories are highly correlated with each other; therefore, the primary dimension they share typically explains an overwhelming majority of variance in test scores. In fact, the presence of just such a dominant primary dimension is the psychometric assumption that provides the foundation for the unidimensional item response theory (IRT) models that have been used for calibrating, linking, scaling, and equating the 2025 MSAA operational tests.

The purpose of dimensionality analysis is to study test item responses for evidence of violations of test unidimensionality and, if such evidence is found, to understand what it is telling us about possible multidimensionality. In practice, the most common approach is to look for statistically significant violations of local independence (LI), also known as local item dependence (LID). Because LID (i.e., violations of LI) can occur for reasons other than multidimensionality, if evidence of LID is found, the next step is to study the LID to determine its source (or sources), including the possibility of multidimensionality. Hence, we first conducted hypothesis tests to detect statistically significant LID and if it was found: (a) estimated the size of the LID and (b) studied the nature of the LID with particular emphasis on possible multidimensionality. Our findings are reported below. (Note: Only operational items were analyzed since they are used for score reporting.)

The dimensionality analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout et al., 2001) and DETECT (Zhang & Stout, 1999). Both methods use as their basic statistical building block the estimated average conditional covariances for item pairs. A conditional covariance is the covariance between two items conditioned on expected total score for the rest of the test, and the average conditional covariance is obtained by averaging across every possible conditioning score. When a test is strictly unidimensional, all conditional covariances are expected to take on values of zero, indicating statistically independent item responses for examinees with equal expected total test scores. Nonzero conditional covariances are essentially evidence of LID, which often implies multidimensionality. Thus, nonrandom patterns of positive and negative conditional covariances are indicative of LID, which may imply multidimensionality.

DIMTEST is a hypothesis-testing procedure for detecting LID. The data are first 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 display the greatest evidence of LID. The cross-validation sample is then used to test whether the conditional covariances of the selected cluster of items display LID, conditioned on total score on the non-clustered items. The DIMTEST statistic follows a standard normal distribution under the null hypothesis of unidimensionality.

The DETECT statistic is an effect-size measure for the size of the LID (the size of the violation of LI). As with DIMTEST, the data are first divided into training and cross-validation samples. The training sample is used to find a set of mutually exclusive and collectively exhaustive clusters of items that best fit a pattern of positive conditional covariances for pairs of items from the same cluster and negative conditional covariances from different clusters. Next, the clusters from the training sample are used with the cross-validation sample 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 size of the LID for an item pair. DETECT values less than 0.2 indicate very weak LID (near unidimensionality); values of 0.2 to 0.4, weak to moderate LID; values of 0.4 to 1.0, moderate to strong LID; and values greater than 1.0, very strong LID (Roussos & Ozbek, 2006).

Note that the goal of the dimensionality analysis is to evaluate the assumption of unidimensionality in the IRT model used for the calibration. In 2018 an additional data cleaning procedure was introduced to determine which data would be used to calibrate the operational items. The new procedure was introduced in response to the repeated finding in dimensionality analyses from previous years that a small (but nontrivial) percentage of the students, referred to as “R9-stringers,” were exhibiting response behavior incompatible with the assumptions of the psychometric model. R9-stringers are students who respond to nine (or more) consecutive multiple-choice items with the exact same option. Because the calibration data had the R9-stringers removed, the same data were also chosen for the dimensionality analyses. Table 8-3 summarizes the dimensionality analysis sample sizes both prior to and after removing the R9-stringers. Table 8-3 also includes the percentages of stringers from the previous year as a point of reference.

Table 8-3. Summary of 2025 Testing Population

Subject	Grade	Total Before Removing Stringers	Total After Removing Stringers	Number of Stringers	Percent Stringers	Percent Stringers in 2024
ELA	3	1,721	1,474	247	14	15
ELA	4	1,677	1,428	249	15	15
ELA	5	1,679	1,415	264	16	19
ELA	6	1,596	1,380	216	14	16
ELA	7	1,448	1,282	166	11	18
ELA	8	1,409	1,199	210	15	16
ELA	HS	1,364	1,180	184	13	16
Mathematics	3	1,702	1,460	242	14	13
Mathematics	4	1,654	1,427	227	14	17
Mathematics	5	1,601	1,411	190	12	15
Mathematics	6	1,536	1,376	160	10	11
Mathematics	7	1,408	1,279	129	9	10
Mathematics	8	1,359	1,200	159	12	9
Mathematics	HS	1,310	1,171	139	11	14

DIMTEST and DETECT were separately applied to the three operational paths of each grade on the 2025 MSAA ELA and mathematics tests. The three paths resulted in three datasets to be analyzed for each ELA and mathematics grade-level test, a total of 42 analyses. First, each dataset was split into a training sample and a cross-validation sample. The sample sizes across the 42 analyses varied from a low of 131 (grade 4, ELA, Path C) to a high of 290 (grade 3, ELA, Path A). A rough tabulation of the sample size distribution is given in Table 8-4, including a comparison to the most recent past administration in 2024. The sample sizes in 2025 were smaller than those in 2024. The overall average sample size for 2024 was about 394, whereas the overall average sample size for 2025 was about 190.

Table 8-4. Dataset Sample Sizes Used for Dimensionality Analyses by Year*

Sample Size	Number of Datasets for Path A	Number of Datasets for Path A	Number of Datasets for Path B	Number of Datasets for Path B	Number of Datasets for Path C	Number of Datasets for Path C
Year	2025	2024	2025	2024	2025	2024
< 600	14	14	14	14	14	14
600 to 800	0	0	0	0	0	0
800 to 1000	0	0	0	0	0	0
> 1000	0	0	0	0	0	0

*Stringers not included in these analyses.

DIMTEST was then applied to every dataset. Even though the sample sizes were not large for the MSAA test paths, the DIMTEST null hypothesis was rejected at a significance level of 0.05 for every dataset. Next, DETECT was used to estimate the effect size for the violations of local independence for all the tests. Table 8-5 displays the effect size estimates from DETECT.

Table 8-5. Average Multidimensional Effect Sizes by Content Area and Grade Across Years*

Path	Content Area	Grade	Multidimensionality 2025	Effect Size 2024
A	ELA	3	0.39	0.33
A	ELA	4	0.39	0.46
A	ELA	5	0.40	0.38
A	ELA	6	0.44	0.73
A	ELA	7	0.23	0.41
A	ELA	8	0.45	0.44
A	ELA	HS	0.35	1.06
A	ELA	Average	0.38	0.54
A	Mathematics	3	0.60	0.42
A	Mathematics	4	0.50	0.38
A	Mathematics	5	0.59	0.66
A	Mathematics	6	0.75	0.70
A	Mathematics	7	0.63	0.87
A	Mathematics	8	0.48	0.30
A	Mathematics	HS	0.54	0.86
A	Mathematics	Average	0.58	0.60
B	ELA	3	0.35	0.53
B	ELA	4	0.24	0.33
B	ELA	5	0.64	0.39
B	ELA	6	0.39	0.63
B	ELA	7	0.72	0.51
B	ELA	8	0.37	0.83
B	ELA	HS	0.56	0.45
B	ELA	Average	0.47	0.52
B	Mathematics	3	0.43	0.54
B	Mathematics	4	0.56	0.76
B	Mathematics	5	0.86	0.42
B	Mathematics	6	0.50	0.39
B	Mathematics	7	0.26	0.60
B	Mathematics	8	0.48	0.58
B	Mathematics	HS	0.49	0.83
B	Mathematics	Average	0.51	0.59
C	ELA	3	0.23	0.27
C	ELA	4	0.14	0.19

Path	Content Area	Grade	Multidimensionality 2025	Effect Size 2024
C	ELA	5	0.32	0.19
C	ELA	6	0.37	0.30
C	ELA	7	0.18	0.35
C	ELA	8	0.26	0.15
C	ELA	HS	0.20	0.18
C	ELA	Average	0.24	0.23
C	Mathematics	3	0.29	0.38
C	Mathematics	4	0.51	0.27
C	Mathematics	5	0.33	0.32
C	Mathematics	6	0.21	0.23
C	Mathematics	7	0.32	0.40
C	Mathematics	8	0.25	0.35
C	Mathematics	HS	0.50	0.58
C	Mathematics	Average	0.34	0.36

* DNR = Did not reject DIMTEST null hypothesis test

The results for 2025 are displayed in Table 8-5 along with results from 2024 as reference. The data shows the following trends: Overall Increase: The DETECT indices are relatively stable in both ELA and Mathematics due to the reduction in student count, with the overall trend being for averages have decreased slightly. ELA tests generally have lower DETECT indices than Mathematics tests, as in prior administrations. In 2025, no tests had DETECT values above 1.

- Path A:
 - ELA: Mostly moderate to moderately strong DETECT values; 7 grades had effect sizes between .2 and .6; only 3 grades had values above .8
 - Mathematics: five grades with moderate and seven grades with moderately strong DETECT values; 5 grades had effects above .8
- Path B:
 - ELA: 5 grades had effect sizes between .2 and .6; only 1 grade with DETECT values above .8.
 - Mathematics: Only one grade had an effect size lower than .4; 6 grades had DETECT values between .4 and .8 and 6 grades had DETECT values above .8
- Path C:
 - ELA: Consistently lower DETECT indices; the lowest among all paths with some non-rejection of the DIMTEST statistic.
 - Mathematics: Lower DETECT indices compared to Paths A and B, but moderate overall.

Additional analysis was conducted to identify sources of LID and explain the DIMTEST and DETECT results. This involved examining how DETECT divided tests into clusters to identify patterns related to substantive item characteristics. Two main factors contributed to increased DETECT values:

1. Key-Option Clustering: Removing the R9-stringers significantly reduced, but did not eliminate, key-option clustering, a pattern that persisted when compared with previous administrations.
2. Secondary Dimension from CR Items: Constructed response items introduced a weak secondary dimension to the mathematics tests.

Chapter 9. Item Response Theory Scaling and Equating

This chapter describes the procedures used to calibrate, equate, and scale the 2025 MSAA. Through these psychometric analyses, several quality-control procedures and checks on the processes were implemented. These procedures included evaluation of item parameters and their standard errors for reasonableness, examination of test characteristic curves (TCCs) and test information functions (TIFs) for reasonableness, evaluation of model fit, and evaluation of the scaling results (e.g., parallel processing by the Data and Reporting Services and the Psychometrics and Research Departments, and comparison of lookup tables to the previous year's lookup tables).

9.1 Item Response Theory

All MSAA items were calibrated using item response theory (IRT). IRT uses a mathematical model to define a relationship between an unobserved measure of student performance, usually referred to as theta (θ), and the probability ($P(\theta)$) of obtaining a particular score on an item. This mathematical relationship is referred to as the item characteristic curve (ICC). In IRT, all items are assumed to be unique 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 $P(\theta)$ (Hambleton & Swaminathan, 1985; Hambleton & van der Linden, 1997). The process of estimating the specific 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 relationship between θ and $P(\theta)$. Once the item parameters are known, an estimate of θ for each student can be calculated based on the student's observed responses to the items. This estimate, $\hat{\theta}$, is an estimate of the student's location on a latent ability continuum.

For the 2025 MSAA tests, the two-parameter logistic (2PL) model was used to estimate the ICC for dichotomous items, and the graded-response model (GRM) was used for polytomous items (Nering & Ostini, 2010). The 2PL model for dichotomous items can be defined as:

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

Where:

- U represents the scored response on an item,
- i indexes the items,
- j indexes students,
- a represents item discrimination,
- b represents item difficulty,
- θ is the student proficiency, and
- D is a normalizing constant equal to 1.701.

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 two-parameter logistic form:

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

Where:

U represents the scored response on an item,
 i indexes the items,
 j indexes students,
 k indexes threshold,
 θ is the student ability,
 α represents item discrimination,
 b represents item difficulty,
 d represents an item-category 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 responding to a particular category given θ , are derived by subtracting adjacent ICTCs:

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

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:

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

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 is expressed as:

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

Where:

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

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

9.2 Calibration Procedure

Because the 2025 MSAA was a pre-equated assessment program, the item parameters for the 2025 operational administration (re-used from the 2023 administration) came from calibrations conducted in years previous to 2023. Items previously used only as field-test items were calibrated in the corresponding field-test calibration that occurred after the calibration of the operational items. No new calibrations were run for the 2025 MSAA prior to the reporting of scores. The procedures used to conduct the calibrations discussed above are described in this section.

As described in Section 8-2, in preparation for previous operational and field-test calibrations, the R9 stringers were removed from the data. In calibrating the operational items, first, an off-scale calibration was conducted on all the operational items using PARSCALE (Muraki & Bock, 2003). At this point, each item was carefully examined for model fit. A visual inspection of the item fit plots was conducted. The empirical proportions of correct responses at a given level of ability must follow the shape of the model-based curve. In addition, the item parameter estimates were inspected. The discrimination parameters should not be extreme in either direction (neither greater than 3 nor less than 0.35); the difficulty parameters should also not be extreme (generally between -3 and 3, and definitely between -4 and 4); and the standard error of the difficulty parameters should generally be less than 0.2. Items that violate any of these conditions are automatically marked as “Do Not Use,”

Field-test items were first calibrated and then evaluated for model-fit and item parameter estimates, following the same procedures as for operational items. Items with model-fit issues or parameter estimates outside acceptable criteria were classified as Do-Not-Use (DNU). Eligible items not classified as DNU were added to the item bank.

9.3 Item Response Theory Results

The tables in Appendix I give the IRT item parameters for all the operational items on the 2025 MSAA tests by grade and content area based on their pre-equated calibrations. The statistics for the operational items are summarized in Tables 9-1 through 9-4. The mean item parameter estimates shown in the tables below are within the generally acceptable and expected ranges. For easy reference, Table 9-1 displays the means and standard deviations averaged across all dichotomously scored operational items for each grade and content area.

Table 9-1. IRT Summary Statistics for Dichotomously Scored Items

Content Area	Grade	n_items	a_mean	a_sd	b_mean	b_sd
ELA	3	64	0.94	0.28	-0.31	0.76
ELA	4	63	0.98	0.33	-0.26	0.68
ELA	5	68	0.99	0.41	-0.34	0.74
ELA	6	66	1.10	0.36	-0.18	0.51
ELA	7	69	0.95	0.34	-0.29	0.55
ELA	8	65	1.11	0.44	-0.26	0.60
ELA	HS	73	1.21	0.46	-0.38	0.58
Mathematics	3	67	0.81	0.30	-0.14	0.68
Mathematics	4	70	0.95	0.31	0.09	0.58
Mathematics	5	71	0.84	0.24	0.08	0.74
Mathematics	6	69	1.01	0.30	-0.17	0.52
Mathematics	7	69	0.90	0.34	-0.20	0.75
Mathematics	8	69	0.84	0.25	-0.09	0.67
Mathematics	HS	69	1.07	0.40	-0.10	0.47

Because the items were developed to correspond to different levels (Levels 1, 2, and 3), the item statistics have also been summarized by item level for ELA (Table 9-2 for the dichotomous items and Table 9-3 for the writing-prompt traits) and for mathematics (Table 9-4).

Table 9-2. IRT Summary Statistics by Grade and Level—ELA Dichotomous Items

Grade	Level	n_items	a_mean	a_sd	b_mean	b_sd
3	Level 1	23	1.05	0.27	-1.01	0.21
3	Level 2	14	0.90	0.29	-0.29	0.47
3	Level 3	27	0.87	0.25	0.28	0.66
3	#N/A	23	1.21	0.25	-0.94	0.24
4	Level 1	23	0.83	0.29	-0.09	0.42
4	Level 2	17	0.86	0.29	0.44	0.48
4	Level 3	23	1.29	0.43	-1.07	0.19
4	#N/A	23	0.91	0.35	-0.14	0.40
5	Level 1	22	0.77	0.21	0.23	0.75
5	Level 2	22	1.39	0.27	-0.62	0.43
5	Level 3	24	0.99	0.36	-0.08	0.37
5	#N/A	20	0.91	0.20	0.18	0.38
6	Level 1	17	1.15	0.33	-0.90	0.26
6	Level 2	33	0.97	0.33	-0.28	0.37
6	Level 3	19	0.74	0.22	0.24	0.44
6	#N/A	21	1.47	0.37	-0.92	0.25
7	Level 1	19	1.05	0.45	-0.23	0.37
7	Level 2	25	0.85	0.25	0.27	0.37
7	Level 3	23	1.64	0.47	-1.01	0.18
7	#N/A	25	1.02	0.35	-0.29	0.34
8	Level 1	25	0.99	0.20	0.10	0.50
8	Level 2	23	1.05	0.27	-1.01	0.21
8	Level 3	14	0.90	0.29	-0.29	0.47
8	#N/A	27	0.87	0.25	0.28	0.66
HS	Level 1	23	1.21	0.25	-0.94	0.24
HS	Level 2	23	0.83	0.29	-0.09	0.42
HS	Level 3	17	0.86	0.29	0.44	0.48
HS	#N/A	23	1.29	0.43	-1.07	0.19

Table 9-3. IRT Summary Statistics by Trait and Level—ELA Writing Prompt Items

WP	Item	Number of	a	a	b	b	d0	d0	d1	d1	b-d0	b-d0	b-d1	b-d1
Trait	Level	Items	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
C	2	7	0.80	0.08	-0.15	0.19	0.68	0.11	-0.68	0.11	-0.83	0.16	0.53	0.27
C	3	7	0.80	0.08	0.30	0.22	0.92	0.11	-0.92	0.11	-0.62	0.21	1.22	0.27
I	2	7	0.81	0.10	0.23	0.25	0.69	0.19	-0.69	0.19	-0.46	0.17	0.92	0.41
I	3	7	0.84	0.18	1.41	0.46	0.96	0.24	-0.96	0.24	0.45	0.59	2.36	0.43
O	2	7	0.76	0.08	0.35	0.31	1.19	0.14	-1.19	0.14	-0.84	0.24	1.54	0.41
O	3	7	0.91	0.19	1.36	0.45	1.38	0.31	-1.38	0.31	-0.02	0.54	2.74	0.56

Note. C = Conventions, I = Idea Development, O = Organization.

Table 9-4. IRT Summary Statistics by Grade and Level—Mathematics

Grade	Level	n_items	a_mean	a_sd	b_mean	b_sd
3	1	22	0.65	0.14	-0.70	0.69
3	2	35	0.89	0.35	0.05	0.47
3	3	8	0.88	0.29	0.44	0.41
3	#N/A	22	0.90	0.16	-0.48	0.33
4	1	36	0.95	0.36	0.39	0.44
4	2	12	1.01	0.33	0.22	0.61
4	3	26	0.82	0.26	-0.71	0.35
4	#N/A	40	0.87	0.24	0.53	0.46
5	1	5	0.79	0.18	0.54	0.55
5	2	21	1.05	0.25	-0.73	0.31
5	3	37	0.99	0.35	-0.01	0.34
5	#N/A	11	0.99	0.17	0.36	0.44
6	1	21	0.90	0.23	-1.10	0.25
6	2	36	0.92	0.41	0.05	0.44
6	3	12	0.87	0.33	0.61	0.55
6	#N/A	23	0.86	0.30	-0.78	0.59
7	1	36	0.83	0.21	0.18	0.37
7	2	10	0.82	0.26	0.54	0.36
7	3	22	1.04	0.21	-0.65	0.25
7	#N/A	40	1.12	0.47	0.12	0.30
8	1	7	0.89	0.40	0.38	0.21
8	2	22	0.65	0.14	-0.70	0.69
8	3	35	0.89	0.35	0.05	0.47
8	#N/A	8	0.88	0.29	0.44	0.41
HS	1	22	0.90	0.16	-0.48	0.33
HS	2	36	0.95	0.36	0.39	0.44
HS	3	12	1.01	0.33	0.22	0.61
HS	#N/A	26	0.82	0.26	-0.71	0.35

Results for the dichotomously scored items are as follows. As seen in Tables 9-2 and 9-4, item difficulty tends to have a positive relationship with level: as item level increases, the items tend to be more difficult (as intended). In nearly all cases, the average difficulty increased from Level 1 to Level 2 and from Level 2 to Level 3. The largest differences were clearly the Level 1 to Level 2 differences for all grade levels for both ELA and mathematics. To investigate these tendencies more rigorously, a one-way analysis of variance (ANOVA) was conducted on item difficulty with level as the factor. Separate ANOVAs were run for ELA and mathematics. ANOVA results revealed statistically significant differences by item level for both ELA and mathematics, with R-squared values of 57% and 52%, respectively. Tukey paired-comparison tests confirmed significant differences between Level 1 and all other levels in both subjects. Additionally, Level 2 versus Level 3 comparisons were significant across all grades. For grade 4

mathematics, Level 3 items were generally less challenging than Level 2 items, with a larger standard error indicating a broader difficulty range within Level 3.

Regarding polytomously scored writing prompt traits (Table 9-3), Level 3 traits were more difficult than Level 2 traits. The differences were more pronounced for Idea Development and Organization traits compared to the Conventions trait.

The IRT statistics were also summarized by different paths (Tables 9-5 and 9-6).

Table 9-5. IRT Summary Statistics by Grade by Path—ELA Dichotomous Items

Grade	Path	Number of Items	a_Mean	a_SD	b_Mean	b_SD
3	A	41	0.94	0.26	-0.63	0.63
3	B	41	0.89	0.28	-0.28	0.69
3	C	41	0.87	0.26	-0.07	0.77
4	A	41	0.99	0.34	-0.58	0.55
4	B	41	0.88	0.29	-0.06	0.63
4	C	41	0.86	0.30	0.00	0.63
5	A	39	1.08	0.44	-0.72	0.54
5	B	39	0.88	0.31	-0.31	0.49
5	C	39	0.83	0.28	-0.04	0.76
6	A	38	1.00	0.36	0.01	0.35
6	B	38	1.20	0.37	-0.40	0.50
6	C	38	0.94	0.29	0.08	0.40
7	A	38	1.03	0.36	-0.63	0.40
7	B	38	0.79	0.23	-0.08	0.52
7	C	38	0.91	0.31	-0.28	0.35
8	A	38	1.00	0.38	-0.14	0.39
8	B	38	1.22	0.45	-0.54	0.52
8	C	38	0.93	0.30	0.05	0.49
HS	A	38	1.34	0.55	-0.76	0.39
HS	B	38	1.05	0.36	-0.41	0.34
HS	C	38	0.99	0.31	-0.13	0.59

Table 9-6. IRT Summary Statistics by Grade by Path—Mathematics

Grade	Path	Number of Items	a_Mean	a_SD	b_Mean	b_SD
3	A	35	0.75	0.23	-0.47	0.57
3	B	35	0.79	0.28	0.25	0.61
3	C	35	0.85	0.36	-0.16	0.59
4	A	35	0.91	0.21	-0.28	0.42
4	B	35	0.88	0.27	0.01	0.43
4	C	35	0.95	0.37	0.33	0.59
5	A	35	0.85	0.26	-0.31	0.73
5	B	35	0.82	0.22	0.16	0.64
5	C	35	0.84	0.22	0.27	0.69
6	A	35	1.00	0.30	-0.48	0.43
6	B	35	0.98	0.24	0.08	0.50
6	C	35	0.98	0.28	-0.18	0.38
7	A	35	0.87	0.35	-0.61	0.58
7	B	35	0.90	0.34	-0.28	0.63
7	C	35	0.89	0.28	0.09	0.73
8	A	35	0.85	0.28	-0.42	0.68
8	B	35	0.83	0.23	-0.08	0.52
8	C	35	0.81	0.20	0.15	0.61
HS	A	35	1.00	0.22	-0.39	0.39
HS	B	35	1.08	0.50	0.11	0.42
HS	C	35	1.00	0.28	-0.10	0.36

The average item difficulty increased significantly from Path A to Path B in 6 out of 7 ELA cases and for most mathematics tests, as intended. Difficulty increased from Path B to Path C in 4 out of 7 mathematics cases and for all ELA tests, although the differences were generally small.

The TCCs provide a more complete picture of the various paths. TCCs display the expected (average) raw score associated with each θ_j value between -2.0 and 2.0. Mathematically, the TCC is computed by summing the expected score on all the ICCs of all items that contribute to the raw score. Using the notation introduced in the previous section, the expected raw score at a given value of θ_j is

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

Where:

X indexes total raw test score,

U_i indexes the scored response on an item,

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

j indexes students (here, θ_j runs from -2 to 2), and

$E(X|\theta_j)$ is the expected raw score on the test 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”—flatter at the ends of the distribution and steeper in the middle.

The TIF, $I(\theta)$ (see Lord, 1980, for theoretical definitions and examples of equations), displays the amount of statistical information 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). 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:

$$SEM(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}$$

Appendix J shows graphs of the TCCs and TIFs for each grade and content area, while Figures 3-2 and 3-3 display test information for a grade 6 ELA test for Path C and a Grade 5 mathematics test for Path B, respectively.

9.4 Equating

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year, as well as to equate one year’s forms to those given in the previous year. Equating ensures that students are not

advantaged or disadvantaged because the test form they took is easier or harder than those taken by other students.

All 2025 MSAA tests used Fixed Common Item Parameter (FCIP) pre-equating methodology as described in Kolen and Brennan (2014). Item pre-equating allows the raw-to-scaled score conversion to be produced before a form is administered, which in turn allows for faster reporting and turnaround times. In item pre-equating, new forms are built from a pool of pre-existing IRT-calibrated items. In addition to these operational items, non-operational items were also included on the forms. These items were a repeat of the items that were field tested in 2023, because of the reuse of that form for the 2025 administration. The operational items were used as a set of common items for transforming the item parameters of the non-operational items so that they would be on the same theta scale as the IRT-calibrated item pool. This allows for the item pool to be expanded continually.

However, with pre-equating, several cautions need to be taken into consideration. For example, Kolen and Brennan (2014) suggest that, to ensure that items behave the same on each administration, the items should appear in the same contexts and positions operationally as they did non-operationally. Thus, care must be taken to avoid significant shifts in position and context. Any drift must be carefully monitored and controlled to ensure comparability between forms of the test. Psychometricians continue to monitor item drift in the pre-equated model. Items showing drift over repeated years are examined, referred to content experts to consider potential construct-irrelevant rationales for differences in item behavior over time.

No items were identified during drift analyses warranting further consideration, thus the scoring tables were identical in 2024 and 2025. Raw score to scaled score lookups based on the pre-equated model for the items used in the 2025 administration are displayed in Appendix K.

Post-equating procedures are conducted to evaluate drift after every operational administration. For any equating design, it is critical that rigorous procedures are implemented to monitor the quality of the equating and to check that the assumptions underlying the equating are not violated. The equating data are analyzed in detail for scale drift through traditional b-b analyses.

9.5 Reported Scale Scores

Because the θ scale used in IRT calibrations is not readily understood by most stakeholders, reporting scales were developed for MSAA. The reporting scales are simple linear transformations of the underlying θ scale. The reporting scales range from 1200 to 1290 for all grade/content-area combinations. The Level 3 cut score was originally fixed at the August 2015 standard setting to be 1240 for each grade level, but some of the scale score cuts, including some of the second cuts, were adjusted during the July 2018 standards validation, as evidenced in Table 9-8. The Highest Obtainable Scale Score was fixed at 1290.

By providing more specific information about the position of a student's results, scale scores supplement performance-level scores. Students' raw scores (i.e., total number of points) on the 2021 MSAA tests were translated to scale scores using a data analysis process called scaling, which simply converts from one scale to another.

It is important to note that converting from raw scores to scale scores does not change students' performance-level classifications. Scale scores make for more consistent reporting of results. Raw scores are not comparable from year to year (nor across Paths A, B, and C) because they are affected by

differences in group ability and/or difficulty of the items that appear on each test form. Equating is a statistical procedure that is used to adjust for differences in form difficulty so that scores on alternate forms can be used interchangeably (Kolen & Brennan, 2014). Since the scale is used for equating, scale scores are comparable from one year to the next.

The scale 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 scale score metric. Students' ability estimates are based on their raw scores and are found by mapping through the TCC. Scale scores are calculated using the following linear equation:

$$SS = m\hat{\theta} + b$$

Where:

m is the slope

b is the intercept

For MSAA, the base-form operational scale was set so that the theta corresponding to the proficient cut from the August 2015 standard setting was transformed to a scale score of 1240, and so that the standard deviation of the scale scores in the base-year was 15. The lowest obtainable scale score (LOSS) was set at 1200, and the highest obtainable scale score (HOSS) was set at 1290. A separate linear transformation is used for each grade and content-area combination. Because only one point within the θ scale score space and the standard deviation of the scale was fixed, the scale score cutpoints between Level 1 and Level 2 and between Level 3 and Level 4 were free to vary across the grade and content-area combinations. When the standards validation was conducted in July 2018, the transformation constants established in the base year were not modified, but some of the theta cuts were modified, including some of the Level 2/Level 3 cuts (i.e., the proficient cuts). Thus, scale score value for the proficient cut is no longer equal to 1240 for some tests (as seen in Table 9-8).

Table 9-7 shows the slope and intercept values used to calculate the scale scores for each content area and grade. Note that the values in Table 9-7 will not change unless the standards are reset.

Table 9-7. Scale Score Slope and Intercept by Content Area and Grade

Content Area	Grade	Slope	Intercept
ELA	3	11.7202	1242.0537
ELA	4	12.0593	1240.0910
ELA	5	12.4236	1241.6149
ELA	6	12.3522	1237.8126
ELA	7	12.2964	1242.4332
ELA	8	12.6082	1239.4570
ELA	HS	11.4922	1244.2240
Mathematics	3	13.0552	1243.6651
Mathematics	4	13.1002	1239.8674
Mathematics	5	13.0769	1241.4102
Mathematics	6	12.8203	1241.2532
Mathematics	7	12.9093	1243.2438
Mathematics	8	13.0213	1242.3583
Mathematics	HS	12.9897	1242.4799

Appendix K contains raw score to scale score lookup tables for the 2025 MSAA tests. These are the actual tables used to determine student scale scores, error bands, and performance levels. Graphs of the scale score cumulative frequency distributions, for the 2025 MSAA tests and for the most recent past test are presented in Appendix L.

9.6 MSAA Performance Levels, Cut Scores, and Standards Validation

Cut scores for MSAA in ELA and mathematics were originally set in a standard setting process that took place in August 2015. Details of the standard setting procedures can be found in the *Standard Setting Report (Measured Progress, 2015)*. In July 2018, Cognia and the MSAA Psychometric Subcommittee conducted a standards validation. Standards validation does not change the scale; its purpose is only to determine whether adjustments to the cut scores are needed.

The standards validation process for the 2018 MSAA was necessary to ensure that cut scores, set in 2015 for the assessments, continue to provide valid interpretation of ELA and mathematics performance using the Performance-Level Descriptors (PLDs). The standards for both ELA and mathematics were vertically articulated, using 2017 performance data, to update the performance standards and provide a coherent basis for interpreting 2018 scores and performance, and in preparation for validating the ELA standards further. No additional steps were necessary to validate the mathematics performance standards. The validation process for the ELA performance standards was necessitated by the addition of the open-response writing prompt scores to the existing ELA score scale in 2018.

A complete description of the standards articulation and validation processes appears in the 2018 *MSAA Standards Validation Report*. (See Appendix M of the 2018 *MSAA Technical Report* located online here: cms.azed.gov/home/GetDocumentFile?id=5cb0b3b61dcb2511e88cfef7).

Final cut scores, after mathematics and ELA vertical articulation and ELA standards validation for the 2017-18 MSAA, appear in Table 9-8.

Table 9-8. Cut Scores on the Theta Metric and Reporting Scale

Content Area	Grade	Theta Cut1	Theta Cut2	Theta Cut3	Theta Minimum	Scale Score Cut1	Scale Score Cut2	Scale Score Cut3	Scale Score Maximum
ELA	3	-0.70318	-0.21788	0.97664	1200	1234	1240	1254	1290
ELA	4	-0.53007	-0.00755	1.52654	1200	1234	1240	1259	1290
ELA	5	-0.83676	-0.12999	1.15500	1200	1232	1240	1256	1290
ELA	6	-0.63000	-0.10626	1.02714	1200	1231	1237	1251	1290
ELA	7	-0.59215	-0.19788	0.94792	1200	1236	1240	1255	1290
ELA	8	-0.75241	-0.15521	0.78177	1200	1230	1238	1250	1290
ELA	HS	-0.76610	-0.41106	0.89860	1200	1236	1240	1255	1290
Mathematics	3	-0.70202	-0.16584	0.76660	1200	1235	1242	1254	1290
Mathematics	4	-0.63872	-0.10438	0.81776	1200	1232	1239	1251	1290
Mathematics	5	-0.75784	-0.10784	0.84805	1200	1232	1240	1253	1290
Mathematics	6	-0.68276	-0.21475	0.72127	1200	1233	1239	1251	1290
Mathematics	7	-0.75478	-0.25128	0.76727	1200	1234	1240	1254	1290
Mathematics	8	-0.65755	-0.21950	0.62527	1200	1234	1240	1251	1290
Mathematics	HS	-0.61432	-0.22940	0.54044	1200	1235	1240	1250	1290

Table 9-9 shows the percentage of students by performance-level categories along with the average and standard deviation of the scale scores for each grade/content-area combination. Also, the percentages of

Levels 3 and 4 (levels corresponding to Proficient and above, which are the levels of critical interest for federal accountability purposes) within each grade and content area are provided in the table.

Table 9-9. Percentage of Students by Performance-Level Categories and Scale Score Summary

Content Area	Grade	Number of Students	Level 1	Level 2	Level 3	Level 4	Average Scale Score	SD of Scale Score
ELA	3	1,599	47%	19%	26%	4%	1234	14
ELA	4	1,539	44%	25%	24%	5%	1235	15
ELA	5	1,509	35%	26%	30%	8%	1234	14
ELA	6	1,478	40%	13%	28%	11%	1234	13
ELA	7	1,378	37%	24%	24%	8%	1237	15
ELA	8	1,282	38%	15%	35%	9%	1233	13
ELA	HS	1,260	46%	30%	36%	15%	1238	13
Mathematics	3	1,592	26%	23%	39%	9%	1237	16
Mathematics	4	1,539	23%	32%	34%	9%	1236	14
Mathematics	5	1,509	27%	26%	31%	14%	1238	14
Mathematics	6	1,477	34%	22%	28%	9%	1237	14
Mathematics	7	1,372	34%	14%	33%	11%	1237	14
Mathematics	8	1,283	29%	24%	34%	10%	1237	14
Mathematics	HS	1,253	30%	25%	35%	10%	1237	13

Tables 9-10 (ELA) and 9-11 (mathematics) show the percentage of students in each performance-level category by path, along with the average and standard deviation of the scale scores for each grade/content-area combination. Note that the percentage of examinees being classified as Level 3 and Level 4 (levels of Proficient or above) increased from Path A to Path C. This trend was expected due to the stage adaptive nature of the MSAA.

Table 9-10. Performance-Level Distributions by Path—ELA

Grade	Path	Number of Students	Level 1	Level 2	Level 3	Level 4	Average Scale Score	SD of Scale Score
3	A	545	87%	12%	1%	0%	1,229	6
3	B	411	5%	54%	42%	0%	1,239	4
3	C	355	0%	0%	66%	34%	1,253	7
4	A	688	84%	15%	0%	0%	1,229	6
4	B	452	2%	37%	62%	0%	1,241	4
4	C	262	0%	0%	79%	21%	1,254	7
5	A	672	77%	23%	0%	0%	1,228	6
5	B	434	0%	51%	48%	0%	1,240	5
5	C	237	0%	0%	72%	28%	1,253	6
6	A	263	81%	18%	1%	0%	1,227	5
6	B	658	4%	59%	37%	0%	1,236	4
6	C	413	0%	1%	65%	33%	1,249	7
7	A	591	85%	12%	2%	0%	1,231	5
7	B	299	4%	35%	61%	0%	1,241	4
7	C	346	0%	0%	59%	41%	1,254	7
8	A	359	77%	23%	0%	0%	1,228	5
8	B	587	2%	70%	29%	0%	1,236	4
8	C	312	0%	1%	61%	38%	1,248	6
HS	A	440	80%	17%	3%	0%	2,464	4
HS	B	397	7%	39%	54%	0%	2,480	4
HS	C	320	0%	1%	73%	26%	2,509	6

Table 9-11. Performance-Level Distributions by Path—Mathematics

Grade	Path	Number of Students	Level 1	Level 2	Level 3	Level 4	Average Scale Score	SD of Scale Score
3	A	505	80%	20%	1%	0%	1230	8
3	B	393	8%	48%	45%	0%	1240	4
3	C	405	0%	2%	51%	46%	1254	9
4	A	395	72%	26%	2%	0%	1227	8
4	B	582	4%	49%	47%	0%	1238	4
4	C	421	0%	1%	72%	27%	1249	7
5	A	363	66%	33%	1%	0%	1229	7
5	B	544	3%	60%	38%	0%	1239	4
5	C	438	0%	4%	69%	27%	1251	9
6	A	475	58%	39%	4%	0%	1231	6
6	B	379	5%	51%	45%	0%	1238	4
6	C	480	0%	1%	59%	40%	1252	10
7	A	503	71%	26%	2%	0%	1230	6
7	B	382	2%	50%	47%	0%	1240	4
7	C	356	0%	1%	58%	40%	1253	9
8	A	346	79%	18%	2%	0%	1230	6
8	B	508	7%	44%	48%	0%	1239	4
8	C	402	0%	2%	61%	37%	1251	9
HS	A	307	68%	32%	1%	0%	1231	6
HS	B	404	5%	46%	49%	0%	1239	3
HS	C	450	0%	2%	72%	25%	1248	8

Chapter 10. Reliability

Although the psychometric characteristics of individual items' performance are an important focus for evaluation, a complete evaluation of an assessment must also address the way items function together. Tests that function well provide a dependable assessment of the student's level of ability. Unfortunately, no test can do this perfectly. A variety of factors can contribute to a given student's score being either higher or lower than his or her true ability. For example, a student may misread an item or mistakenly fill in the wrong bubble when he or she knew the right answer. Collectively, extraneous factors that affect a student's score are referred to as "measurement error." Any assessment includes some amount of measurement error. This is true of all academic assessments—some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. Assessments that have less measurement error (i.e., errors made are small on average and student scores on such a test will consistently represent their ability) are described as "reliable."

Test reliability is often estimated using Cronbach's alpha, but for the 2025 MSAA with three test paths (A, B, and C), this method would underestimate reliability due to differing ability distributions. Instead, an item response theory (IRT) estimate was used to provide a single reliability value per test.

10.1 IRT Marginal Reliability

IRT marginal reliability estimation is based on applying the standard classical test theory (CTT) formula, relating variances of true score, observed score, and measurement error, in the IRT setting. In CTT, the relationship between these variances is given by:

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

where σ_X^2 is the observed-score variance, σ_T^2 is the true-score variance, and σ_E^2 is the error variance.

Starting from this basic equation, it can be shown that the formula for CTT reliability can be expressed by:

$$CTT\ Reliability = 1 - \frac{\sigma_E^2}{\sigma_X^2}$$

IRT marginal reliability is based on extending the CTT model to an IRT framework (Samejima, 1994) and provides an IRT-based estimate of the overall test reliability. Error variance is estimated as the mean squared conditional standard error of measurement (CSEM) of the theta estimates across students within a grade. Observed score variance is estimated as the variance of the theta estimates across students within a grade. Equivalently, the mean squared CSEM of the scale scores and the variance of the scale scores can be used in place of the CSEM of the theta estimates and the variance of the theta estimates, respectively. IRT marginal reliability is then given by the following formula:

$$IRT \text{ Marginal Reliability} = 1 - \frac{\overline{CSEM(\theta)^2}}{Var(\hat{\theta})} = 1 - \frac{\overline{CSEM(SS)^2}}{Var(SS)}$$

Where:

$\overline{CSEM(\theta)^2}$ is the mean squared CSEM,
 $\overline{CSEM(SS)^2}$ is the mean squared scale CSEM,
 $Var(\hat{\theta})$ is the variance of theta estimates, and
 $Var(SS)$ is the scale score variance.

Using this formula, IRT marginal reliability estimates were calculated for each multistage test in ELA and mathematics, using the scale scores (and their standard errors) for all the students across all three paths. The reliability of a test can also be evaluated by simply examining directly the CSEMs themselves. CSEMs facilitate the interpretation of individual scale scores. With any given scale score estimate for a student, the reasonable limits of the true scale score for the student can be calculated by using the CSEM for the scale score.

Tables 10-1 and 10-2 present descriptive scale score statistics, IRT-based reliability, and mean scale score CSEMs for ELA and mathematics by grade. (Statistics are based on operational items, which counted toward students' reported scores only.) As shown in the tables, all the values are .92 or higher for ELA and 0.85 or higher for mathematics which indicates adequate reliability.

Table 10-1. IRT Marginal Reliability by Grade—ELA

Grade	Min	Max	Mean	SD	IRT Marginal Reliability	Mean Scaled CSEM
3	1200	1290	1236.64	10.83	0.93	2.84
4	1200	1290	1237.21	11.60	0.93	3.00
5	1200	1290	1236.07	10.64	0.92	2.83
6	1203	1284	1236.25	9.62	0.92	2.62
7	1201	1282	1240.28	11.01	0.92	3.04
8	1210	1271	1235.52	9.49	0.92	2.57
HS	1218	1277	1240.51	9.06	0.93	2.31

Table 10-2. IRT Marginal Reliability by Grade—Mathematics

Grade	Min	Max	Mean	SD	IRT Marginal Reliability	Mean Scaled CSEM
3	1200	1290	1240.12	11.18	0.86	3.92
4	1203	1283	1238.31	9.49	0.87	3.33
5	1200	1290	1240.20	9.79	0.85	3.76
6	1205	1290	1240.23	10.60	0.89	3.17
7	1200	1290	1239.39	10.70	0.87	3.60
8	1200	1290	1239.89	10.82	0.88	3.54
HS	1200	1290	1239.98	9.24	0.88	2.95

10.2 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on all students who took a particular 2025 MSAA test. As an alternate assessment program, it is likely that there are reliability differences across subgroups. For this reason, reliability coefficients for different subgroups were

calculated, including gender, ethnicity, LEP status, socioeconomic status, migrant status, and various disability groups. Appendix M presents reliabilities for various subgroups of interest. Subgroup reliabilities were calculated using the IRT-based formula (defined above) based only on the members of the subgroup in question in the computations; values were calculated only for subgroups with 100 or more students.

Table 10-3. Subgroup Reliability Summary by Subject and Grade

Subject	Grade	Min	Max
ELA	3	0.85	0.93
ELA	4	0.85	0.93
ELA	5	0.81	0.94
ELA	6	0.88	0.93
ELA	7	0.85	0.92
ELA	8	0.88	0.93
ELA	HS	0.88	0.94
Mathematics	3	0.81	0.86
Mathematics	4	0.79	0.88
Mathematics	5	0.71	0.87
Mathematics	6	0.86	0.89
Mathematics	7	0.81	0.88
Mathematics	8	0.84	0.89
Mathematics	HS	0.75	0.89

For several reasons, the results relating to subgroup reliability should be interpreted with caution. First, reliability is dependent not only on the measurement properties of a test but on the statistical distribution of the studied subgroup. For example, it can readily be seen in Appendix M that subgroup sample sizes varied considerably, which results in a natural variation in reliability coefficients. Alternatively, reliability, which is a type of correlation coefficient, may be artificially depressed when there is a restriction of range (Draper & Smith, 1998) as occurs for subgroups with little variability. Second, there is no industry standard to interpret the strength of a reliability coefficient, especially when the population of interest is a single subgroup. Comparison of subgroup minimum maximum values in Table 10-3 to Tables 10-1 and 10-2 are even more confounding then, due to the dramatic difference in both sample size and variability between the analyses.

10.3 Reliability of Performance-Level Categorization: Accuracy and Consistency

While related to reliability, the accuracy and consistency of student classification into performance categories are even more important statistics in a standards-based reporting framework (Livingston & Lewis, 1995). After the performance levels were specified and students' performances were classified into those levels, analyses were conducted to determine the statistical accuracy and consistency of the classifications. For the MSAA, students are classified into one of four performance levels: Level 1, Level 2, Level 3, or Level 4. This section of the report explains the methodologies used to assess the reliability of classification decisions, and results are provided.

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Consistency measures the extent to which classification decisions based on test scores match the decisions 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 given to the same group of students. In operational test programs, however, such a design is usually impractical.

However, techniques have been developed to estimate both the accuracy and the consistency of classification decisions based on a single administration of a test. The Rudner (2001, 2005) technique was used for the 2025 MSAA because it can be easily applied to data that are scored in the IRT theta metric or any linear transformation of this metric, such as the MSAA scale scores. The applicability of the Rudner technique to IRT-based metrics distinguishes this method from methods based on observed scores, such as the Lewis and Livingston (1995) method. Thus, the Rudner method can be used to provide a single index for a multistage test, whereas an observed score method would need to be separately applied to each path of a multistage test.

For details of the Rudner method, refer to Rudner (2001, 2005); The Rudner method focuses on estimating the accuracy of classifying examinees based on their test scores. It involves creating a classification table of true and expected scores for polytomously scored items under item response theory (IRT). The method calculates the expected proportion of examinees correctly classified into different score categories by comparing observed scores to true scores across the theta continuum. Rudner extends the procedure for dichotomous items to polytomous ones and provides a simplified approach for estimating accuracy using standard errors at cut scores.

For the classification accuracy tables, cell $[i, j]$ represents the estimated proportion of students whose true scale score fell into classification i (where $i = 1$ to 4, for the four achievement levels) and whose observed scale 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.

For the classification consistency tables, cell $[i, j]$ of this table represents the estimated proportion of students whose observed scale score on the first of the two hypothetical parallel multistage tests would fall into classification i (where $i = 1$ to 4) and whose observed scale score on the second hypothetical parallel multistage test 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.

Another way to measure consistency is to use 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:

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

Where;

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

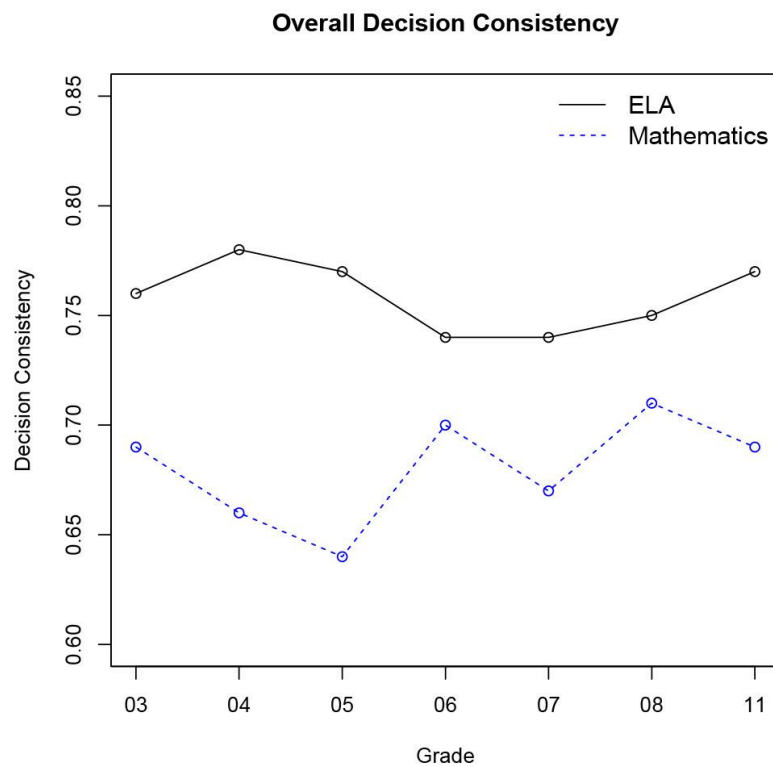
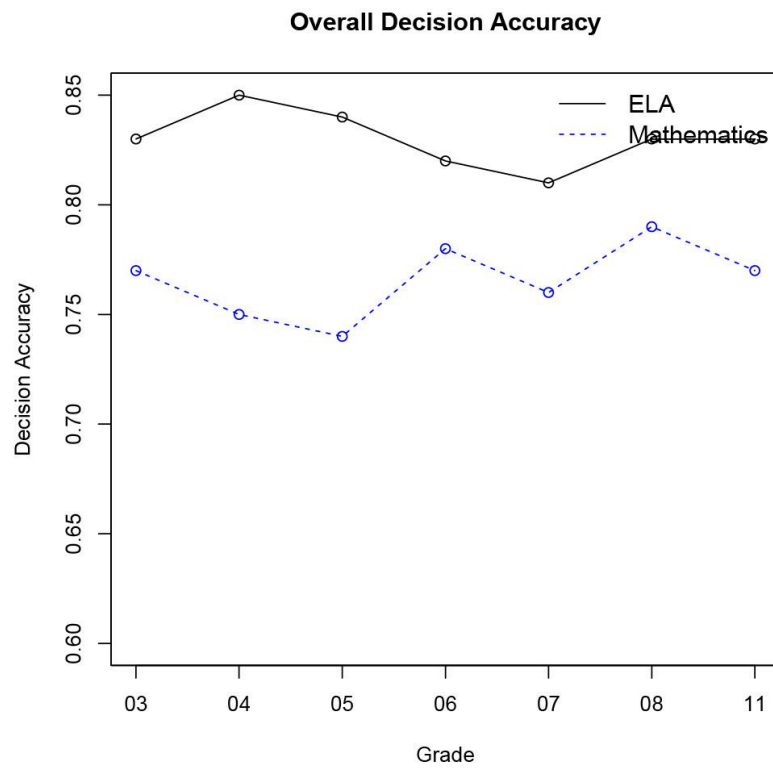
Because κ is corrected for chance, its values are lower than other consistency estimates.

Figure 10-1 shows the overall decision accuracy for ELA and mathematics by grade level. ELA overall has higher decision accuracy (above 0.8) than mathematics (above 0.7). Across all grades, more than 80% and 70% of ELA and Mathematics students would be expected to be at the same performance level again when categorized according to their observed scale score, respectively. More details on decision accuracy and consistency (DAC) are provided in Appendix N. Table N-1 in Appendix N includes overall accuracy and consistency indices, along with kappa. Accuracy and consistency values conditional on performance level are also provided in Table N-1. For these calculations, the denominator is the proportion of students associated with a given performance level. Following is an example from Table N1, looking at Level 1 for grade 3 ELA.

- The conditional *accuracy* value was 0.83. This indicates that among the students whose *true scale scores* placed them in Level 1, 83% would be expected to be in this same level again when categorized according to their observed scale scores.
- The *consistency* value was 0.76. This indicates that among the students whose *observed scale scores* placed them in Level 1, 76% would be expected to be in this same level again if a second parallel test form were used.

For some testing situations, the greatest concern may be decisions regarding level thresholds. For example, in testing done for Every Student Succeeds Act accountability purposes, the primary concern is distinguishing between students who are proficient and those who are not yet proficient. For the 2025 MSAA, Table N-2 in Appendix N provides accuracy, and consistency estimates at each cutpoint, as well as false positive and false negative decision rates. A false positive rate is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative rate is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.

Figure 10-1. Overall Decision Accuracy and Consistency by Content Area by Grade



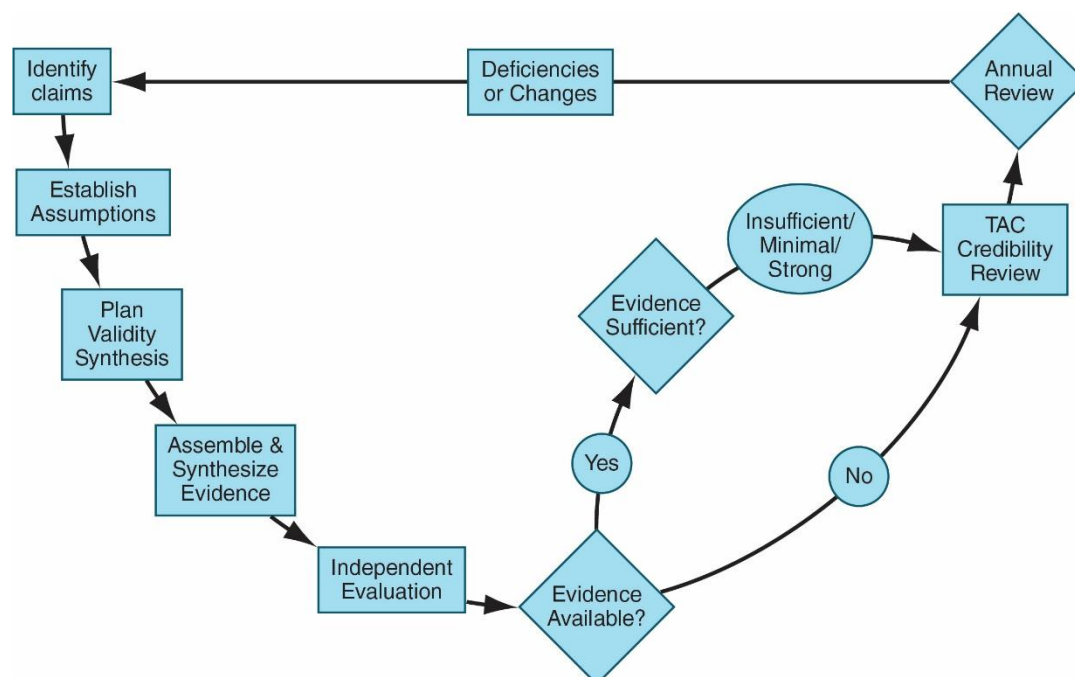
Chapter 11. Validity of the MSAA Primary and Secondary Claims

Chapter 11 provides an overview of the claims of the MSAA assessment, including a synthesis and evaluation of the assumptions and evidence supporting them. This chapter establishes an explicit link between the assumptions and sources of evidence, using a holistic approach to synthesize and evaluate the sufficiency of the evidence supporting each assumption.

11.1 The MSAA Validation Framework, Evidence Synthesis, and Sufficiency

In 2025, a workshop consisting of TAC members was held at the Center for Assessment to refine and re-establish the validation framework for the primary intended score interpretation (claim) of MSAA test scores. The recommended MSAA validation framework that emerged out of this workshop, depicted in Figure 11-1, connects the primary and secondary claims to their associated assumptions, synthesizes evidence supporting the reasonableness of those assumptions, evaluates the sufficiency of the evidence and then provides feedback to inform continuous improvement of the validity argument. An explicit connection of each assumption (organized by category) to its related evidence can be found in Table 11-1: Evidence Assumption Crosswalk, which is a useful resource to help readers connect assumptions to supporting evidence within the report. As added clarity, we use the word “assumption” in the following sense: what are we assuming to be true when we make a claim? By definition, it is subject to corroboration (or falsification) through logical or empirical analysis. When an assumption is supported by evidence, it provides a “warrant” or “backing” for the claim.

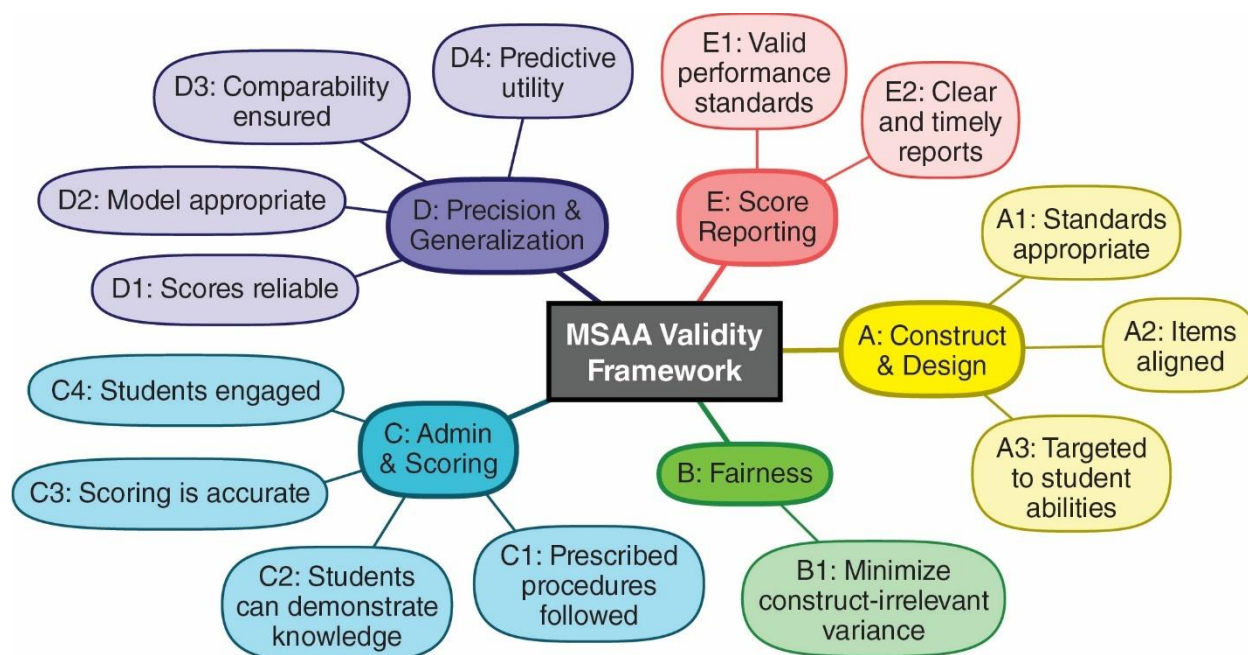
Figure 11-1. MSAA Validation Framework



Workshop members further elaborated the above framework using these descriptions:

- I. **Primary Claim:** *MSAA scores provide trustworthy information about the grade-level knowledge and skills of students with significant cognitive disabilities in the subjects of Mathematics and English Language Arts*
- II. **Assumptions:** straightforward statements about the desired assertions and conditions necessary to address the primary claim within each of the categories in part III
- III. **Categories of Assumptions (Figure 11-2):** groups of assumptions that are logically related
 - a. Construct Conceptualization and Test Design
 - b. Fairness
 - c. Administration and Scoring
 - d. Precision and Generalization
 - e. Score Reporting

Figure 11-2. MSAA Assumptions by Category



- IV. **Support:** specific information describing each assumption, accompanied by illustrative or required evidence as well as a holistic judgment of evidentiary sufficiency AND an overall judgment as follows
 1. Insufficient [little to no credible evidence]
 2. Limited [some relevant evidence]
 3. Partial [evidence covers most major assumption components]
 4. Substantial [generally acceptable body of evidence]
 5. Strong [evidence fully supports the assumption]
- V. **Uses (Secondary Claims):** an acknowledgement of the uses that the program intends to support if the primary claim is valid

Table 11-1. Crosswalk of MSAA Evidence and Assumptions

Category	Assumption and Sources of Evidence in the 2025 MSAA Technical Report	
Primary Claim		
MSAA scores provide trustworthy information about the grade-level knowledge and skills of students with significant cognitive disabilities in the subjects of Mathematics and English Language Arts		
A. Construct Conceptualization and Representation: assumptions about the construct of measurement and how the MSAA assessments have been designed.		
Chapters 2-4	A1.	The assessed standards are appropriate for the target population.
Chapters 2-4, 9; Appendices C, I, J	A2.	MSAA consists of high-quality items aligned to the CCCs and grade-level standards.
Chapters 2-4, 9; Appendix I	A3.	Test items are appropriately targeted to the varied capabilities of the test takers.
B. Fairness: assumptions about student engagement and opportunity to demonstrate what they know and can do		
Chapters 3-4, 8-9; Appendices I, H, M	B1.	Choices made in the specific format or content of the item on the test do not interfere with students showing what they have learned.
C. Administration and Scoring: assumptions about the administration, scoring and security of MSAA tests		
Chapter 5	C1.	Testing follows prescribed administration procedures.
Chapters 2, 5; Appendix A	C2.	Administration procedures allow students to demonstrate what they have learned.
Chapters 3,5,6	C3.	Student responses are accurately scored.
Chapters 8-9	C4.	Students are engaged by the assessment.
D. Precision and Generalization: assumptions about score comparability (e.g., equating and linking), internal consistency (e.g., reliability), generalization (e.g., to performances or outcomes in the future) and the psychometric modeling used to support the assessment.		
Chapter 10; Appendices J, N	D1.	Test scores on the MSAA provide reliable information about student performance.
Chapters 8, 9; Appendices I, K, M	D2.	The psychometric model is appropriate for its intended uses.
Chapters 8,9,10; Appendices J-M	D3.	Student performance can be meaningfully interpreted and compared within and across administrations.
Not Available	D4.	MSAA test scores are predictive of success outside the context of the MSAA assessment
E. Score Reporting: assumptions about establishing performance standards and communicating clear and useful information to constituents.		
Chapters 3,7,9	E1.	Performance level descriptors accurately reflect the knowledge and skills required to support interpretations and use.
Chapter 7; Appendices G, P	E2.	Student information is reported in a timely and clear manner appropriate to the intended users.

11.2 Evaluating the Primary Score Claim

As identified in the MSAA Validation Framework step IV, supporting the primary claim involves synthesizing the evidence available for each of the claim's assumptions as detailed in Table 11-1. We have chosen to separate the synthesis of the support section into two parts. The first is a holistic synthesis in narrative form presenting the assumption and the supporting evidence assembled by Cognia and the MSAA partners. The second part is an independent holistic summary of the sufficiency of the available evidence in supporting the assumption. This step is performed by technical advisors independent of Cognia and the MSAA partners, reinforcing the impartiality of reviewer judgments.

To support this process, MSAA commissioned an independent evaluation to determine the extent to which the available evidence sufficiently supports the assumptions underlying the primary intended score interpretation and to identify priority areas for ongoing improvement. Two evaluators from the National Center for the Improvement of Educational Assessment, one of whom also serves on the MSAA TAC, conducted a systematic review of each assumption linked to the primary score interpretation, as presented in Figure 11-1 and detailed in Table 11-1. For each assumption, the evaluators considered the following guiding question: What conditions must hold for this claim to be warranted, and to what extent does the existing evidence substantiate the associated assumption?

Using this structure, evaluators classified the sufficiency of evidence into five categories:

1. Insufficient Evidence – Little to no credible evidence is available to substantiate the assumption.
2. Limited Evidence – Some relevant evidence exists, but it is incomplete, outdated, or leaves key aspects of the assumption unverified.
3. Partial Evidence – The evidence base covers most major components of the assumption, though some areas require additional support, clarification, or updating.
4. Substantial Evidence – A generally acceptable body of evidence supports the assumption. Any gaps are relatively minor, and further clarification or routine monitoring would strengthen confidence in the claim.
5. Strong Evidence – Extensive, well-aligned evidence fully supports the assumption. Recommendations, if any, focus primarily on ongoing monitoring or maintenance.

Through this process, the evaluators synthesized findings across assumptions, identified strengths and areas for further development, and provided feedback to inform continuous improvement of the MSAA validity argument. Their findings are presented following the synthesis of supporting evidence for each assumption. The section concludes with an overall summary of findings.

Primary Claim: MSAA scores provide trustworthy information about the grade-level knowledge and skills of students with significant cognitive disabilities in the subjects of Mathematics and English Language Arts.

11.2.1 Category A: Construct Conceptualization and Representation

Assumption A1. The assessed standards (i.e., the Core Content Connectors) are appropriate for the target population.

This assumption should be supported by the following evidence:

- The process of making the standards accessible for the target population (i.e., developing the MSAA Core Content Connectors [CCCs]) was systematic and defensible (e.g., included experts informed by research)
- Educators can and do provide students with instruction that gives them the opportunity to learn the knowledge and skills identified in the CCCs
- It is reasonable to expect all students in this population to demonstrate some degree of mastery of the target knowledge and skills in the CCCs

Supporting Evidence: Chapters 2-4

The subject of the MSAA Core Content Connectors is introduced in Chapter 2, where it states the assessment is intentionally aligned with the state content standards and CCCs. Chapter 3 notes that the CCCs were created during initial design of the assessment, where experts informed by research examined learning targets and grade-level expectations from the CCSS to identify guide grade-level instruction and assessment for students with significant cognitive disabilities. The process described in this chapter indicates that the CCCs were intentionally aligned with both Learning Progression Frameworks (LPFs) and the CCSS, providing evidence that the CCCs are aligned with expectations for students in this population as well as the expected learning opportunities necessary to demonstrate some degree of mastery.

Table 3-2 summarizes the studies conducted as evidence supporting this assumption, including confirmation that the content and skills in the CCCs represent an adequate and appropriate sample of the grade-level CCSS.

Chapter 4 indicates the role of stakeholders in standards alignment to the CCCs. The MSAA partners have confirmed alignment between the MSAA Core Content Connectors and each partner's academic content standards. Additionally, membership in MSAA requires adoption of the academic content standards assessed on the MSAA, and the evidence supports this alignment.

Evaluation Summary: Partial Evidence

While the original studies provide sufficient evidence that over a decade ago the CCCs were appropriate for the population, the studies should be replicated to confirm this holds true or that the population of students has not been impacted by the COVID19 pandemic (e.g. lowering IQ scores, increases in population of test takers) or changes in diagnostic criteria (more detailed and limiting). Also, since many states have revised expectations for the general assessment based on evidence that the targets may not align with college and career readiness, MSAA should conduct studies to determine if this trend holds for students performing at, above, and below proficient on MSAA.

Additionally, some assumptions necessitate more frequent investigations. For example, the claim that teachers provide instruction and the opportunity to learn can be done annually through a teacher survey, while being supplemented with more credible evidence (e.g., classroom observations). There is a reference to the NCSC theory of action, and it implies that being "career ready" is a goal, but the past studies cited in Table 3-2 do not provide evidence that the standards are aligned to these outcomes.

Assumption A2. MSAA consists of high-quality items aligned to the CCCs and grade-level standards.

This assumption should be supported by evidence:

- MSAA items are aligned to the CCCs, representing appropriate breadth, depth, and emphasis.
- Items are developed in a manner that is consistent with professionally accepted practices
 - Detailed specifications address the characteristics and features of high quality, aligned items and tasks
 - Items are written by qualified and well-trained item writers
 - Item review procedures that involve representative panels with subject matter expertise and experience with the student population
- Rules have been established to determine the collection of items (i.e., test forms) that students are administered
 - Appropriate form specifications have been developed and reviewed by qualified experts
 - Assessment forms appropriately represent the breadth, depth, and target emphasis
- Items have adequate psychometric properties (i.e., discrimination and difficulty)

Supporting Evidence: Chapters 2-4, 9; Appendices C, I and J

As in assumption A1, details of alignment studies confirming the coherence between test items and the CCCs, including appropriate breadth, depth, and emphasis are documented in Table 3.2 and Chapters 2-3.

Chapter 2 states the assessment is intentionally aligned with the state content standards and CCCs, and tables 3.2 documents the appropriate breadth, depth, and emphasis. Chapters 2-3 document the process of aligning content to state standards, including evidence that the development process follows industry standards, including alignment studies, as evidence that the process is designed to create high quality, aligned items.

Chapter 3 also documents the training and qualifications of item writers as well as the processes of writing and procedures for reviewing items using representative panels with subject matter expertise and experience with the student population as evidence supporting this assumption. Chapter 4 details the involvement of MSAA partners and other stakeholders in item and test development , indicating that industry accepted procedures of using SMEs with experience teaching the tested population are being followed, again supporting the assumption that these are high quality items with multiple reviews for quality and alignment.

Appendix C provides details regarding the test design and development process, and psychometric properties of the items and test forms are detailed in Appendices I and J, respectively and reflect industry best practices for test design and psychometric evaluation of high-quality items.

Evaluation Summary: Limited Evidence

Several claims are made about the quality of test development, and a detailed review process is outlined. However, limited evidence is provided to support the claims. For example, the process includes a bias review, but no criteria for the bias review are provided. Likewise, no qualifications for test developers are included (they are included for raters) despite claims that high quality item writers were used. While the original studies may provide substantial evidence, they are dated and need to be updated. See A1 for rationale for updating evidence.

Assumption A3. Test items are appropriately targeted to the varied capabilities of the test takers.

- Item scaffolding (in the form of items written to different ordinal levels of scaffolding) enhances accessibility
- Differences in item levels are associated with differences in empirical item difficulty
- The stage adaptive design of the test routes students to content aligned to their capabilities

Supporting Evidence: Chapters 2-4, 9; Appendix I

The MSAA test development process aims to maximize accessibility for all students by incorporating permissible accommodations (such as presentation and communication modes) and integrating accessibility features into the system, all guided by comprehensive accessibility standards. Stringers, a relevant learner trait in this context, might lead to repetitive responses in some students. To mitigate this, new item types have been introduced in the hope of diversifying their answers and improving overall accessibility to demonstration of KSAs in the MSAA. This aligns with the item development concept of levels. Section 3.3.1 describes the concept of item level development; level guidelines for items allow for test content of graduated complexity, with the lowest level addressing the EU that has the most decreased level of complexity, and (also as part of the item) the greatest level of scaffolding and support.

Additionally, the stage adaptive test design allows for targeting of items to three levels of empirical difficulty, which are referred to in this documentation as paths A, B and C (progressing from lower to greater difficulty). This and other aspects of the design are detailed in Chapters 2-4.

Evidence regarding the association between paths and empirical difficulty is documented in Chapter 9. Tables 9-5 and 9-6 show that the average difficulty of items progresses from Path A to Path C in all grades. Appendix I presents evidence in the form of box plots indicating the progression of empirical difficulty from items targeting Level 1 to Level 3.

Evaluation Summary: Partial Evidence

Data on item statistics and the scaffolded design criteria provide partial evidence that test items are appropriately targeted to the varied capabilities of the test takers. However, the original studies cited in Table 3-2 are dated and require updates. See A1 for the rationale for updating evidence. In addition, trend data on item difficulty should also be included here.

11.2.2 Category B: Fairness

Assumption B1. Choices made in the specific format or content of the item on the test do not interfere with students showing what they have learned.

This assumption should be supported by evidence:

- Item development procedures incorporate the principles of universal design and/or other appropriate accessibility frameworks
- Items are free of bias and sensitivity concerns. Minimal evidence of differential item functioning
- Item scaffolding enhances accessibility. Differences in levels are associated with differences in empirical item difficulty

Supporting Evidence: Chapters 3-4, 8-9; Appendices I, H, M

As described in assumption A3, the test design embraces the concepts of scaffolding to enhance accessibility. Chapter 3 details the test design process, including the intentional manipulation of item complexity and inherent scaffolding and supports to address the diverse accessibility needs of the target population. Chapter 9 and Appendix I present evidence that differences in item development levels are associated with empirical difficulty.

Chapter 3 also shows use of Universal Design and other accessibility frameworks including APIP in the design of MSAA tests. Chapter 8 and Appendix H present evidence regarding evidence of minimal DIF. Evidence of data, bias and sensitivity reviews is presented in Chapters 3 and 4. Additionally, Section 10.2 provides evidence to allow for comparison of test reliability over different groups and subgroup reliability is further detailed in Appendix M.

Evaluation Summary: Limited Evidence

There is some evidence that items do not exhibit DIF for race and gender. However, the text also claims DIF was computed for economically disadvantaged vs. non economically disadvantaged, but Appendix H does not report DIF statistics. Also, Appendix M includes sample sizes that indicate there may be errors. For example, Grade 3 shows more students with hearing loss than without hearing loss. Finally, MSAA should consider adding DIF analyses for secondary disability categories like vision, communication, and hearing since sample sizes appear sufficient. Other notes:

- There is a lack of evidence for the assumption that "Item development procedures incorporate the principles of universal design and/or other appropriate accessibility frameworks."
- Page 32 references UDL, but the assumptions reference Universal design more broadly and APIP.
- Chapter 3 and 4 refer to things like Bias and Fairness reviews but do not mention UDL (which is stated in the summary) or APIP (which is included in the assumption).
- Suggest providing the checklists and reviewer guidelines in an appendix. Also, list the qualifications and expertise of reviewers.
- Address procedures for what transpires when items do not pass reviews.
- Clarify what is included in the UDL check and how that is done. Prior presentations suggest that much of UDL implementation stems from the teacher adapting content; however, this aspect is not adequately conveyed.
- The text references "Accessible Portable Item Protocol (APIP) reviews," but no guidelines are provided.
- Suggest removing Appendix I as evidence relevant to fairness.

11.2.3 Category C: Administration and Scoring

Assumption C1. Testing follows prescribed administration procedures.

This assumption should be supported by evidence:

- Test administrators receive adequate training
- Test administrators follow established administration and security protocols
- Administration irregularities are documented and addressed in accordance with established procedures

Supporting Evidence: Chapter 5

Chapter 5 presents all available evidence supporting the assumption that testing follows prescribed administration procedures. The chapter describes the administration for the assessments, including administration modes, procedures, requirements, documentation, training, test security, irregularities, and support. The adequacy of administrator training and the evidence that administrators follow proper protocols is detailed in section 5.2. Test Administrators (TAs) are provided with multiple pieces of documentation, including a Test Administration Manual (TAM), Directions for Test Administration (DTA), and other guides, to aid in their training. Additionally, Table 5-1 documents annually updated and required training modules that TAs are required to view prior to administration. The TAs are also required to take a quiz on each module and receive a passing certification to access secure test materials. These training materials model potential student-TA interactions to better prepare administrators for actual administration processes.

The Monitoring and Quality Control section of the chapter provides evidence surrounding the safeguards put in place to ensure proper testing procedures (including security) are followed during administration. Procedures for reporting any violation or suspected violation of test security or confidentiality are detailed in the TAM. Additionally, the online MSAA system uses built-in measures to ensure testing procedures are followed. Observed irregularities are collected, including support calls and emails, and Cognia regularly synthesizes these for partners. In a survey last year, most respondents agreed that the TAM, User Guides and Training Modules were helpful in their understanding of how to perform their role as a TA.

Evaluation Summary: Partial Evidence

Evidence is sufficient that policies and procedures are in place.

Evidence from the post-test survey shows high participation rates and positive responses. This should be stated explicitly in the summary.

The following areas are insufficient due to lack of data on annual numbers of reported irregularities and calls for help that were resolved.

- Test administrators follow established administration and security protocols.
- Administration irregularities are documented and addressed in accordance with established procedures.

Areas for improvement are reporting call logs and resolutions. In addition, report annual data on the number of calls for test security and resolution rather than stating that a process is in place.

Assumption C2. Administration procedures allow students to demonstrate what they have learned.

This assumption should be supported by evidence:

- Students receive clear instructions about the testing process
- Students have an opportunity to gain familiarity with the testing process
- The testing platform makes it easy to elicit and record a student's response
- Items are rendered as intended
- Appropriate accommodations, assistive technology, and/or accessibility features are provided to meet student needs

Supporting Evidence: Chapters 2, 5; Appendix A

Chapter 5 presents most available evidence supporting the assumption that the administration procedures allow students to demonstrate what they have learned. For example, section 5.8.1 details the unique online or paper administration of the MSAA. Students taking the MSAA receive detailed instructions and supports as outlined in the TAM and modeled in the training videos and modules, although this evidence is not directly provided by this report. This training includes instructions, certainly, but also supports student answer entry and access to the test items, including but not limited to TA presentation of the item and recording of the response. As noted in Chapter 2, the test is “designed to be administered one-on-one” by the TA with myriad accommodations including the use of assistive technology, a scribe and/or sign language. Appendix A summarizes the use of accommodations with some detail.

Further evidence regarding the training and familiarizing of students to the testing process are detailed in Appendix E which demonstrates the type of materials teachers are provided to familiarize students with the assessment.

Evaluation Summary: Substantial Evidence

There is substantial evidence that the administration procedures allow students to demonstrate what they know and can do. Areas for improvement involve providing additional evidence that:

- The testing platform makes it easy to elicit and record a student’s response.
- Items are rendered as intended.

Assumption C3. Student responses are accurately scored.

This assumption should be supported by evidence:

- Quality assurance checks establish that selected-response item answers (e.g., option “A” was selected) are accurately scored (e.g., “A” = 1)
- For open-ended items:
 - Scorers are well-qualified and successfully meet training standards
 - Scoring checks demonstrate a high degree of accuracy.

Supporting Evidence: Chapters 3, 5, 6

As detailed in Chapter 6, machine-scored items are verified for accuracy through a key validation process, while constructed-response item scoring adheres to industry standards, employing the rigorous practice of double-blind scoring and expert read-behinds. The specific item types are detailed in Chapter 3.

Machine-scored items use key validation as a method of quality assurance. The key-validation process reports statistical anomalies to content and reporting resources indicating that an item may be systematically scored incorrectly, including key and distractor analyses that may indicate an item has an incorrect key. Open response items are all scored by at least two raters, with a percentage of items additionally read by leadership as a check on scorer accuracy and an arbiter of borderline judgments.

Evidence regarding the training and required consistency of scorers is detailed in Section 6.2, including Table 6-6 which reports Interrater Agreement.

Evaluation Summary: Substantial Evidence

There is substantial evidence that student responses are accurately captured. However, consider removing the text on crisis paper flagging from the technical manual since teachers are uploading responses. This should be in the test administrators’ guide or school guidelines.

Assumption C4. Students are engaged by the assessment.

This assumption should be supported by evidence:

- Examination of response time and process data
- Repetitive response (stringer) analyses (e.g., proportion of students that repeat the same response option some designated number of times)

Supporting Evidence: Chapters 8-9

Although there is a lack of traditional response process analyses or think aloud studies supporting evidence of thought processes for MSAA test takers, the unique one-on-one administration of the tests does provide some opportunity to understand the thought processes of students engaging with MSAA. As suggested in Chapter 2, TAs have substantial leeway in developing a testing schedule, including the ability to start and stop a test depending on the engagement of a student. The lack of time constraints and extended administration window allow for administrators to tailor the pace of item delivery to optimize engagement. This is specified further in Chapter 5, section 5.8, which confirms that “the MSAA is not a timed test. Testing time varies for each student, with testing paused and resumed based on a student’s needs. If a student becomes sick or exhibits frustration, lack of engagement, or refusal to participate,” TAs are empowered to pause the testing for minutes or days. “MSAA protocols allow the TA to pause and resume the administration as often as necessary” during the 7-week testing window.

Similarly, psychometricians evaluated repetitive responses as an analysis of engagement. Students identified as having repetitive responses that represent no relevance to the content or intended construct being measured are identified and removed from item calibrations (Chapter 9) and dimensionality analyses (Chapter 8). Evidence as to the efficacy of this detection is revealed in the dimensionality analysis with and without stringers. Table 8-3 shows the current percent of students by grade and subject identified with repetitive construct irrelevant responses for the current year, as well as a comparison to previous administrations.

Evaluation Summary: Insufficient Evidence

While there are references to the stringer analyses, no explicit connections between these (or other) analyses and engagement are addressed. Stringers are removed from some analyses (e.g., dimensionality), but they are not further examined to provide insights about engagement.

It would be necessary to reference previous studies or commission new investigations to better understand engagement and its potential impact on the primary MSAA claim.

11.2.4 Category D: Precision and Generalization

Assumption D1. Test scores on the MSAA provide reliable information about student performance.

This assumption should be supported by evidence:

- MSAA scores are sufficiently precise
 - Acceptable overall and conditional standard errors of measurement for all examinees
 - Acceptable reliability coefficients for each assessment and student group.
 - Test information functions are appropriately differentiated for each administration path
- Assignment to performance levels is supported by evidence of acceptable classification accuracy and consistency

Supporting Evidence: Chapter 10; Appendices J, N

The technical report Chapter 10 provides evidence to the precision of MSAA test scores, as demonstrated by the inverse of standard error, the Test Information Function (TIF) as well as in the form of Conditional Standard Error of Measurement (CSEM) and IRT Marginal Reliability. Marginal reliability is reported in Chapter 10 in Tables 10-1 and 10-2, providing evidence of internal consistency that meets or exceeds the professional standards for high quality tests. TIFs can be found for each grade and subject area in Appendix J. TIFs are indicators of test precision along the continuum of student performance. Appendix J organizes these by grade, subject, and path (described earlier as one of three forms available in the adaptive test, depending on preliminary in-test estimates of ability). This allows the reader to evaluate evidence of the differences in precision (and where precision is maximized) along the performance continuum, depending on routing during the test. Reliability coefficients for each assessment overall and subset by student group are provided in Appendix M.

Decision Accuracy and Consistency (DAC) and evidence related to the reliability of performance level classifications is detailed in Appendix N.

Evaluation Summary: Strong Evidence

There is substantial evidence of precision and score generalizability. Marginal reliabilities are within professionally accepted ranges. Decision accuracy and consistency studies show that students are reliably placed into performance levels. However, precision is less consistent for some student groups and scale regions. Ongoing monitoring and improvement should focus on increasing precision at the extremes of the score scale, which may be accomplished by developing new items in regions of need and evaluating approaches to provide more differentiation for stage adaptive paths.

Assumption D2. The psychometric model is appropriate for its intended uses.

This assumption should be supported by evidence:

- The IRT model is used to
 - Establish a common score scale on which student performance is interpreted
 - Route students appropriately to different sets of test items (i.e., paths)
 - Allow comparisons to be made across years
- Item response data shows adequate fit to the specification and functional form of the IRT model (e.g., analysis of item and person fit)
- The dimensionality of the IRT model is sufficiently consistent with the observed item covariance matrix (e.g., there is evidence that the data is “essentially” unidimensional)

Supporting Evidence: Chapter 8, 9; Appendices I, K, M

As described throughout this technical report, the test forms (paths) for each grade, subject and path were the same as used in 2024. As a result, this year's performance of students in grade 3 ELA (for example) is directly comparable to last year. In previous administrations, items were field tested and placed on a common scale using FCIP, holding previously calibrated items fixed to the MSAA scale, while freely calibrating field-tested items for best alignment with the MSAA scale and model. Each free calibration involved fit analyses not just of the calibrated items, but also of any possible drift in the fixed item values used for the fixed item parameter set. Details of the models used for each item type, linking and scaling methods, fit evaluation procedures, and related quality checks can all be found in Chapter 9. Appendix I presents the IRT item parameters. Appendix K includes the lookup tables used to calculate student scores based on their path and number correct, also comparing this year to the last year where the paths had different items (2023). Although not available in this report, artifacts from the original calibration and linking of current operational items (done in previous years) exist in those collected technical reports (available online) and are believed to have been faithfully produced and checked for quality using the

procedures described in this technical manual, as verified by technical advisors, partner review and peer review throughout the course of the program.

Essential unidimensionality, in terms of comparability of student performance over time, could also be thought of as stability of dimensionality. For this reason, evidence is provided not only of DETECT and DIMTEST measures, which help determine the effect size of multidimensionality, but also the differences year over year. The more similar those values are, the more comparable such comparisons become (comparisons of performance to prior years, placement of students on a similar scale, placement of students in performance levels). Dimensionality summaries are presented in Chapter 8, referenced in Chapter 9 and indicate that the effect size of multidimensionality has remained stable over administrations indicating the interpretations of the unidimensional composite score should be comparable between years.

Evaluation Summary: Substantial Evidence

Evidence provides substantial support that the model is suitable for establishing a common score scale, routing students to appropriate item sets, and supporting year-to-year comparability. This is supported through implementation of professionally accepted practices for calibration and linking, checks for item drift, and fit analyses, producing results that are within established ranges with some exception.

However, the claim of 'essential unidimensionality' requires additional support as moderate to strong effect sizes are found especially for paths and A and B, even with stringers removed. This suggests ongoing work is needed to expand item development and explore improvements to the stage adaptive design.

Assumption D3. Student performance can be meaningfully interpreted and compared within and across administrations.

This assumption should be supported by evidence:

- Content representation is stable across administrations
- Scaling procedures are appropriate and implemented with fidelity
- Equating procedures are appropriate and implemented with fidelity

Supporting Evidence: Chapters 8,9,10; Appendices J-M

Evidence supporting the assumption can be found throughout the psychometric chapters and appendices. Content representation has historically been stable across administrations, but even more in 2025, with the consideration of form reuse. The reliability of MSAA scores and categorizations into performance levels is supported by evidence of internal consistency, scaled score standard errors, and performance-level classification consistency and accuracy. Existing evidence strongly supports this element, with a possible challenge being the impact of local item dependence detected in dimensionality analysis. Evidence regarding reliability and decision accuracy is in Chapter 10, Appendix M, and Appendix N.

The item characteristics for the MSAA support intended interpretations about all students. Evidence includes DIF analyses, dimensionality assessment, calibration, and model fit evaluation. All items strongly represent a primary dimension, and this dimension is presumed by inference to be interpreted as the intended construct. Many checks are done, including DIF and stringer analysis, to support the claim that attempts are made to limit construct irrelevant variance in MSAA.

Test characteristics for paths A, B, and C support intended interpretations. Evidence includes dimensionality analysis and test information functions (Appendix J).

The fidelity of equating and scaling of MSAA test forms is supported by the quality of common items used in drift analysis and the stability indicated by those analyses.

Evidence supporting the correct routing of students in the multi-stage test is bolstered by the MSAA test construction process, which outlines how item and test information is tailored to each stage, including routing criteria. Performance-level distributions across different test paths (A, B, and C) are analyzed in Section 9.6 to gauge the psychometric characteristics of the stages, and these stages exhibit distinguishable and logical differences in performance-level distributions. These and similar efforts provide evidence that MSAA scores and their associated performance-level classification are accurate indicators of student achievement and comparable to prior administrations.

Evaluation Summary: Substantial Evidence

Evidence suggests that scaling and equating procedures, test construction procedures, quality checks, model fit analyses follow professional accepted practices and are implemented with fidelity to support score comparability. The same test forms and paths were used in 2025 as in 2023, allowing direct comparisons. More updated information about item and form alignment will strengthen evidence of comparability.

Assumption D4. MSAA test scores are predictive of success outside the context of the MSAA assessment.

This assumption should be supported by evidence:

- Association between MSAA test scores and external outcomes (e.g., grades, participation in post-secondary activities, employment)

Supporting Evidence: Limited Evidence

Evaluation Summary: Limited Evidence

Prior studies have examined the predictive relationships between MSAA scores and post-secondary outcomes. These studies are:

- *Exploring the Post-School Outcomes of Students Who Participated in MSAA Mathematics*
- *Exploring the Post-School Outcomes of Students Who Participated in MSAA ELA*

These studies provide limited evidence that achieving proficiency on the ELA and Math standards is associated with larger rates of postsecondary education and employment. Evidence could be strengthened by updating these studies and broadening them to include additional state partners regularly and examining a broader set of post-secondary outcomes for students across the full range of performance on the assessments.

11.2.5 Category E: Score Reporting

Assumption E1. Performance level descriptors accurately reflect the knowledge and skills required to support interpretations and use.

This assumption should be supported by evidence:

- Performance level descriptors accurately reflect the knowledge and skills required to support interpretations and use
- Standard setting procedures are appropriate and defensible.
 - Adequate representation and expertise of standard setting panelists
 - Documentation and review of procedures and results
 - Panelist evaluations indicate support for the determined cut-scores

Supporting Evidence: Chapters 3,9

Evidence regarding PLDs and their alignment with the KSAs necessary to support the primary claim can be found not only in the technical reports of the MSAA, but also in published reports and letters from USED as well as standard setting reports, advisor notes and memos, and other artifacts typical to a large-scale summative testing program. Within this report, evidence regarding the PLDs can primarily be found in Chapter 9. Section 9.6 describes the history of setting standards for the current cuts used in MSAA ELA and Mathematics. While both studies were completed in 2015, in July of 2018 Cognia and the MSAA Psychometric Subcommittee did conduct a standards validation to weigh the impact of adding. Part of the 2018 study was vertically articulating the cuts to improve the coherence of interpretation. The details of these efforts are not posted as evidence in this report, but the report does point to both the 2015 and 2018 MSAA technical reports as evidentiary support for the fidelity of those efforts. The cuts derived from those activities were approved by both the MSAA partners, Cognia psychometricians and eventually peer review. Table 9-8 presents the MSAA cuts on the theta metric by grade and subject.

Chapter 3, Table 3-2 documents a 2015 item mapping study designed to answer the research question “To what degree do the assessment tasks and items align to the performance-level descriptors?”, where researchers found adequate alignment between the KSAs required by the items and those described by the PLDs.

Evidence is presented in Chapter 7 related to score reports, including the reporting of performance levels.

Evaluation Summary: Partial Evidence

Work to establish performance standards for the MSAA was conducted in 2015 and validated in 2018. Procedures are well documented in the technical manual and in supplementary standards setting reports. Evidence suggests these activities are consistent with professional standards. However, as noted elsewhere, updated information should be produced. This may include updates to studies conducted in the summer of 2015 to address the extent to which items and tasks on the assessment align with the PLDs and ongoing data collection to evaluate the extent test information can be effectively used as intended.

Assumption E2. Student information is reported in a timely and clear manner appropriate to the intended users.

This assumption should be supported by evidence:

- Individual and summary reports convey relevant information about student performance in a manner that can be easily understood.
 - Documentation of usability testing
 - Feedback from constituents about the clarity and utility of reports and supporting resources.
- Resources and supports are provided to help constituents understand and use performance information as intended

Supporting Evidence: Chapter 7; Appendices G, P

Chapter 7 documents score reporting. Although the validity synthesis stops short of attempting to provide evidentiary support for the secondary claims of use, it does acknowledge that the student reports are used and that the student information is reported with the intention that it be used. As a result we present evidence regarding the primary claim that the test scores are reported in a way that is relevant enough to be used in expected ways and timely enough to be useful, including supports to aid interpretation. The chapter notes that the templates for report design were developed with a subcommittee of MSAA partners, who meet annually to refine the reports to ensure they remain meaningful to the intended users. The partners also produce a Guide for Score Report Interpretation. Again, these are updated annually and then further edited by each partner to ensure specific relevance to the users in each state. The guide is included in this report as Appendix G.

States additionally offer guidance to local districts for the purpose of professional development. An example is the document “How to Teach the State Standards to Students Who Take Alternate Assessments”
<https://www.azed.gov/sites/default/files/2016/12/How%20to%20Teach%20State%20Standards%20to%20Students%20Who%20Take%20Alternate%20Assessments%20030617%20Updated%20Links.pdf?id=5866dbe1aadebe085c4de5b4>

Appendix P does provide some information about usability. The referenced 2024 validity survey involved teachers use of student information from MSAA. The report noted that roughly 40% of them used the reports on an annual basis. Users gave feedback about where they thought reports could be changed to improve use.

Evaluation Summary: Partial Evidence

MSAA's collaborative and ongoing work to design and regularly evaluate score reports for various users at various levels and ancillary guidance (i.e., Guide for Score Report Interpretations) provides partial evidence in support of this claim. Information is not presented regarding the timing of score report distribution at each level. The 2024 validity report reveals that a majority of educators are not using reports on an annual basis, which elevates the need for further study to understand and address why and to identify strategies to potentially improve reports and supporting resources.

Validity evaluation summary of findings:

The validity evaluation revealed adequate evidence for most technical aspects of MSAA, including score reliability, comparability across years, and the appropriateness of the underlying measurement model. Moreover, there is good evidence that the administration and scoring procedures generally function as intended, and students can demonstrate what they know under established testing conditions.

However, several components of the validity argument rely on dated or incomplete evidence, especially in areas related to construct alignment, item quality, fairness, accessibility, and student engagement. Some assumptions received partial or limited evidence ratings because empirical studies have not been refreshed in many years or because documentation in the technical manual is not available, is incomplete, or does not fully substantiate the claims. One assumption, student engagement (C4), received Insufficient Evidence, indicating the need for new empirical work. This represents a high priority for improving the validity evidence of MSAA moving forward.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Baker, F. B., & Kim, S-H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York: Marcel Dekker, Inc.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Chapelle, C. A. (2021). *Argument-based validation in testing and assessment*. Sage.
- 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: Mantel-Haenszel and standardization. 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.
- Forte, E. (2013a). *Evaluating alignment for assessments developed using evidence-centered design*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.
- Forte, E. (2013b). *The next generation of alignment*. Paper presented at the National Conference on Student Assessment, Washington, DC.
- Hagge, S., & Keller, R. (2009). *Equating mixed-format tests: Examining the impact of intra-individual consistency in the IRT framework*. Presented at the American Educational Research Association, San Diego, CA.
- 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.
- Hambleton, R. K., & van der Linden, W. J. (1997). *Handbook of modern item response theory*. New York, NY: Springer-Verlag.
- Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, D.C.: National Council on Measurement in Education.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73.
- Keller, L., Keller, R., & Parker, P. (2008). *The effect of shifting content on the accuracy of equating*. Presented at the American Educational Research Association, New York, NY.
- Keller, R., Kim, W., Nering, M., & Keller, L. (2007). *What breaks the equating? A preliminary investigation into threats to a five-year equating chain*. Presented at the American Educational Research Association, Chicago, IL.
- 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. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of meta test scores*. Reading, MA: Addison-Wesley.
- Measured Progress Department of Psychometrics and Research (2015). *National Center and State Collaborative (NCSC) standard setting report*. Unpublished report.³
- Muraki, E., & Bock, R. D. (2003). PARSCALE 4.1 [Computer software]. Lincolnwood, IL: Scientific Software International.

³ See chapter 7: ncscpartners.org/Media/Default/PDFs/Resources/NCSC15_NCSC_TechnicalManualNarrative.pdf

- Parker, P., Keller, R., & von Davier, A. (2009). *The Examination of four equating methods: The effects of reclassifying students into performance categories and the population sensitivity assumption*. Presented at the National Council on Measurement in Education, San Diego, CA.
- Persuasiveness. (2019). In *Vocabulary.com* Retrieved from www.thesaurus.com/browse/persuasiveness
- 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.
- Rudner, L. M. (2001). Computing the expected proportions of misclassified examinees. *Practical Research & Evaluation*, 7 (14). Retrieved from PARE.online.net/getvn.asp?v=7&n=14
- Rudner, L. M. (2005). Expected classification accuracy. *Practical Research & Evaluation*, 10 (13). Retrieved from pareonline.net/pdf/v10n13.pdf
- Samejima, F. (1994). Estimation of reliability coefficients using the test information function and its modifications. *Applied Psychological Measurement*, 18, 229–244.
- Ferrara, S., & Qunbar, S. (2022). Validity Arguments for AI-Based Automated Scores: Essay Scoring as an Illustration. *Journal of Educational Measurement*, 59(3), 288-313.
- 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 Duign, & 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: Cambridge University Press.
- Wakeman, S., Lee, A., Browder, D. (2012, August 23). *National Center and State Collaborative approach to Content for Students with Significant Disabilities*. National Center and State Collaborative. ncscpartners.org/Media/Default/PDFs/Resources/Wakeman-et al-CEC2012-NCSC-Approach-to-Content-SWSCD.pdf
- Webb, N. L. (2005, November). *Alignment, depth of knowledge, and change*. Paper presented at the annual meeting of the Florida Educational Research Association, Miami, Florida.
- 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
ACCOMMODATION FREQUENCIES

Table A-1. Accommodation Frequencies—ELA

Accommodations	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade HS
LCI_Vision ¹	95	75	83	67	79	78	102
SAR_Assistive_Response_After ²	299	272	248	241	232	189	195
SAR_No_Accomm_Needed_After ³	325	336	345	305	333	315	329
SAR_Paper_Version_After ⁴	2	2	1	1	0	0	0
SAR_Scribe_After ⁵	125	115	109	108	89	87	55
SAR_Sign_Interpretation_After ⁶	389	353	328	264	267	241	166

1: LCI_Vision - Input could occur through alternate keyboards, eye-gaze, switch devices, speech-to-text, and other similar input devices. Students are also expected to access text using AT devices (e.g., screen readers), but refreshable Braille display is not supported for presentation of text-based content for the first operational year.

2: SAR_Assistive_Response_After - Assistive Technology (AT) for viewing, responding, or interacting with test items.

3: SAR_No_Accomm_Needed_After - No accommodations needed.

4: SAR_Paper_Version_After - Paper version of item/s.

5: SAR_Scribe_After - A scribe will enter in the MSAA Online Assessment System the student-indicated answer to a selected-response item. For the constructed-response writing item, the scribe will record the student's response to the writing prompt on the response templates in the MSAA Online Assessment System.

6: SAR_Sign_Interpretation_After - TA may communicate passages, items and response options using sign language to student.

Table A-2. Accommodation Frequencies—Mathematics

Accommodations	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade HS
LCI_Vision ¹	93	73	84	67	78	77	102
SAR_Assistive_Response_After ²	296	273	249	241	231	189	193
SAR_No_Accomm_Needed_After ³	326	334	345	305	336	316	330
SAR_Paper_Version_After ⁴	2	2	1	1	0	0	0
SAR_Scribe_After ⁵	126	118	108	108	88	87	55
SAR_Sign_Interpretation_After ⁶	387	356	327	264	264	241	165

1: LCI_Vision - Input could occur through alternate keyboards, eye-gaze, switch devices, speech-to-text, and other similar input devices. Students are also expected to access text using AT devices (e.g., screen readers), but refreshable Braille display is not supported for presentation of text-based content for the first operational year.

2: SAR_Assistive_Response_After - Assistive Technology (AT) for viewing, responding, or interacting with test items.

3: SAR_No_Accomm_Needed_After - No accommodations needed.

4: SAR_Paper_Version_After - Paper version of item/s.

5: SAR_Scribe_After - A scribe will enter in the MSAA Online Assessment System the student-indicated answer to a selected-response item. For the constructed-response writing item, the scribe will record the student's response to the writing prompt on the response templates in the MSAA Online Assessment System.

6: SAR_Sign_Interpretation_After - TA may communicate passages, items and response options using sign language to student.

Table A-3. Accommodation Frequencies—Science

Accommodations	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade HS
LCI_Vision ¹	0	0	64	0	0	59	86
SAR_Assistive_Response_After ²	0	0	207	0	0	165	167
SAR_No_Accomm_Needed_After ³	0	0	273	0	0	256	266
SAR_Paper_Version_After ⁴	0	0	1	0	0	0	0
SAR_Scribe_After ⁵	0	0	77	0	0	75	46
SAR_Sign_Interpretation_After ⁶	0	0	213	0	0	163	127

1: LCI_Vision - Input could occur through alternate keyboards, eye-gaze, switch devices, speech-to-text, and other similar input devices. Students are also expected to access text using AT devices (e.g., screen readers), but refreshable Braille display is not supported for presentation of text-based content for the first operational year.

2: SAR_Assistive_Response_After - Assistive Technology (AT) for viewing, responding, or interacting with test items.

3: SAR_No_Accomm_Needed_After - No accommodations needed.

4: SAR_Paper_Version_After - Paper version of item/s.

5: SAR_Scribe_After - A scribe will enter in the MSAA Online Assessment System the student-indicated answer to a selected-response item. For the constructed-response writing item, the scribe will record the student's response to the writing prompt on the response templates in the MSAA Online Assessment System.

6: SAR_Sign_Interpretation_After - TA may communicate passages, items and response options using sign language to student.

Table A-4. Accommodation Summary

Content Area	Grade	Number of Students Tested with Accommodations	Number of Students Tested Without Accommodations
ELA	3	953	646
ELA	4	919	620
ELA	5	900	609
ELA	6	810	668
ELA	7	786	592
ELA	8	735	547
ELA	HS	697	563
Mathematics	3	950	642
Mathematics	4	918	622
Mathematics	5	901	608
Mathematics	6	810	667
Mathematics	7	788	584
Mathematics	8	736	548
Mathematics	HS	695	559
Science	5	680	463
Science	8	575	415
Science	11	568	438

APPENDIX B
TEST PARTICIPATION

Table B.1 Summary of Participation by Demographic Category—ELA

Description	# Tested Complete	# No Tested Observable Mode of Communication ¹	Total Tested	Total Percent
All Students	9,393	652	10,045	100
Female	3,127	237	3,364	33
Male	6,266	415	6,681	67
Hispanic or Latino	3,054	180	3,234	32
American Indian or Alaska Native	658	43	701	7
Asian	294	22	316	3
Black or African American	933	66	999	10
Native Hawaiian or Pacific Islander	107	4	111	1
White (non-Hispanic)	3,379	286	3,665	36
Two or More Races (non-Hispanic)	482	40	522	5
No Primary race/Ethnicity Undefined	486	11	497	5
Currently receiving LEP services	1,188	79	1,267	13
Not receiving LEP services	1,521	145	1,666	17
LEP: All Other Students	6,684	428	7,112	71
Economically Disadvantaged Students	1,884	170	2,054	20
Non-economically Disadvantaged Students	560	61	621	6
SES: All Other Students	6,949	421	7,370	73
Migrant	53	9	62	1
Non- migrant	1,526	160	1,686	17
Undefined Migrant Status	7,814	483	8,297	83
Augmentative Communication	2,512	298	2,810	28
No Augmentative Communication	6,824	349	7,173	71
Undefined Augmentative Communications	57	5	62	1
Hearing Loss	248	93	341	3
Within Normal Limits	9,145	559	9,704	97
Undefined Hearing Loss	0	0	0	0
Visual Impairment	386	193	579	6
Within Normal Limits	8,979	456	9,435	94
Undefined Visual Impairment	28	3	31	0
Sensory Stimuli Response	760	461	1,221	12
Follow Directions	8,633	191	8,824	88
Undefined Receptive Language	0	0	0	0
Special School	1,049	142	1,191	12
Regular School Self-contained	5,467	447	5,914	59
Regular School Resource Room	1,510	42	1,552	15
Regular School Primarily Self-contained	903	16	919	9
Regular School General Education	464	5	469	5
Undefined Classroom Setting	0	0	0	0
Student Communicates Primarily Through Cries	680	431	1,111	11
Uses Intentional Communication	2,524	161	2,685	27
Uses Symbolic Language	6,189	60	6,249	62
Undefined Expressive Communication	0	0	0	0

¹ No Observable Mode of Communication indicates that the students' test was closed because they had no visible means of communication.

Table B-2. Summary of Participation by Demographic Category—Mathematics

Description	# Tested Complete	# Tested No Observable Mode of Communication ¹	Total Tested	Total Percent
All Students	9,376	652	10,028	100
Female	3,120	237	3,357	33
Male	6,256	415	6,671	67
Hispanic or Latino	3,048	180	3,228	32
American Indian or Alaska Native	654	43	697	7
Asian	294	22	316	3
Black or African American	929	66	995	10
Native Hawaiian or Pacific Islander	106	4	110	1
White (non-Hispanic)	3,377	286	3,663	37
Two or More Races (non-Hispanic)	480	40	520	5
No Primary race/Ethnicity Undefined	488	11	499	5
Currently receiving LEP services	1,185	79	1,264	13
Not receiving LEP services	1,510	145	1,655	17
LEP: All Other Students	6,681	428	7,109	71
Economically Disadvantaged Students	1,875	170	2,045	20
Non-economically Disadvantaged Students	557	61	618	6
SES: All Other Students	6,944	421	7,365	73
Migrant	52	9	61	1
Non- migrant	1,515	160	1,675	17
Undefined Migrant Status	7,809	483	8,292	83
Augmentative Communication	2,509	298	2,807	28
No Augmentative Communication	6,812	349	7,161	71
Undefined Augmentative Communications	55	5	60	1
Hearing Loss	251	93	344	3
Within Normal Limits	9,125	559	9,684	97
Undefined Hearing Loss	0	0	0	0
Visual Impairment	381	193	574	6
Within Normal Limits	8,967	456	9,423	94
Undefined Visual Impairment	28	3	31	0
Sensory Stimuli Response	751	461	1,212	12
Follow Directions	8,625	191	8,816	88
Undefined Receptive Language	0	0	0	0
Special School	1,045	142	1,187	12
Regular School Self-contained	5,459	447	5,906	59
Regular School Resource Room	1,503	42	1,545	15
Regular School Primarily Self-contained	906	16	922	9
Regular School General Education	463	5	468	5
Undefined Classroom Setting	0	0	0	0
Student Communicates Primarily Through Cries	680	431	1,111	11
Uses Intentional Communication	2,516	161	2,677	27
Uses Symbolic Language	6,180	60	6,240	62
Undefined Expressive Communication	0	0	0	0

¹ No Observable Mode of Communication indicates that the students' test was closed because they had no visible means of communication.

Table B-3. Summary of Participation by Demographic Category—Science

Description	# Tested Complete	# Tested No Observable Mode of Communication ¹	Total Tested	Total Percent
All Students	2,950	189	3,139	100
Female	1,019	73	1,092	35
Male	1,931	116	2,047	65
Hispanic or Latino	1,167	66	1,233	39
American Indian or Alaska Native	202	4	206	7
Asian	114	6	120	4
Black or African American	185	21	206	7
Native Hawaiian or Pacific Islander	38	0	38	1
White (non-Hispanic)	1,028	84	1,112	35
Two or More Races (non-Hispanic)	161	8	169	5
No Primary race/Ethnicity Undefined	55	0	55	2
Currently receiving LEP services	333	25	358	11
Not receiving LEP services	443	39	482	15
LEP: All Other Students	2,174	125	2,299	73
Economically Disadvantaged Students	241	22	263	8
Non-economically Disadvantaged Students	146	18	164	5
SES: All Other Students	2,563	149	2,712	86
Migrant	18	4	22	1
Non- migrant	393	42	435	14
Undefined Migrant Status	2,539	143	2,682	85
Augmentative Communication	726	84	810	26
No Augmentative Communication	2,207	102	2,309	74
Undefined Augmentative Communications	17	3	20	1
Hearing Loss	95	34	129	4
Within Normal Limits	2,855	155	3,010	96
Undefined Hearing Loss	0	0	0	0
Visual Impairment	137	72	209	7
Within Normal Limits	2,802	115	2,917	93
Undefined Visual Impairment	11	2	13	0
Sensory Stimuli Response	218	133	351	11
Follow Directions	2,732	56	2,788	89
Undefined Receptive Language	0	0	0	0
Special School	360	52	412	13
Regular School Self-contained	1,698	121	1,819	58
Regular School Resource Room	496	14	510	16
Regular School Primarily Self-contained	252	1	253	8
Regular School General Education	144	1	145	5
Undefined Classroom Setting	0	0	0	0
Student Communicates Primarily Through Cries	212	128	340	11
Uses Intentional Communication	744	39	783	25
Uses Symbolic Language	1,994	22	2,016	64
Undefined Expressive Communication	0	0	0	0

¹ No Observable Mode of Communication indicates that the students' test was closed because they had no visible means of communication.

Table B-4. Test Participation by Subgroup

Description	Total Tested	Invalidated	Did Not Test
ELA	10,045	103	496
Mathematics	10,028	113	503
Science	3,139	35	175

APPENDIX C
TEST DESIGN BLUEPRINTS

20-21 MSAA ELA Operational Blueprint

Notes:

- Measured Progress psychometricians have analyzed passage sets as a whole to show how well they differentiate between stages 2A, 2B, and 2C using IRT stats. For additional details about this process, please reference the Test Construction Process documentation here:
[T:\Contracts\MSAA\6027 - 2018\Program Management\Test Construction\TC process document](#)
 - Goal is to move toward:
 - 2A: difficulty range-low
 - 2B: difficulty range-medium
 - 2C: difficulty range-high
- Linking passage sets may occur in Session 2A, B & C, but they will vary based on how well they differentiate based on IRT stats.
- Writing standalones are included in Session One.
- Writing Prompt-SRs (Level 1) are administered in all Session 2 versions.
- Writing Prompt-OR WP Level 2 is administered in Session 2A. Writing Prompt -OR WP Level 3 is administered in Sessions 2B and 2C.
- Reading Foundational items are added to Session 1, Form 1 for grades 3 and 4 in Field Test. Grades 3-8 & High School will have FT writing items, L1 Writing Prompts or a shortened passage set.

Item Types:

SR- Selected Response: an independent item that is not connected to any other items.

- Two-Part SR: a two-part Selected Response item in which answering one item is not dependent on answering the previous item. Students can reference the previous item without impacting their score.
- MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.
- The CR writing prompts are scored based on 9 possible points, but score point 1 and 2 are collapsed for reporting purposes for a total of 6 possible points.

Blueprint Guidelines ELA

When the item pool allows, these are the blueprint guidelines that will inform test construction.

ELA Content Category	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	HS
Reading Literary	24-32%	24-32%	25-33%	21-30%	17-26%	17-26%	17-26%
Reading Informational	18-26%	18-26%	25-33%	26-34%	32-36%	32-36%	32-36%
Reading Vocabulary and Foundational (G3 and G4)	12-16%	12-16%	6-10%	9-11%	6-9%	6-9%	6-9%
Writing	36-38%	32-38%	31-40%	36-40%	36-40%	36-40%	36-40%

Grade 3 Targets by Standard
MSAA ELA Operational Test Blueprint Grade 3

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	24-32%	3.RL.h1** Answer questions related to the relationship between characters, setting, events, or conflicts (e.g., characters and events, characters and conflicts, setting and conflicts) NOT 2-PART	SR, MSR one or two-part item	12-16	12-16
		3.RL.i2 Answer literal questions and refer to text to support your answer	SR		
		3.RL.k2** Determine the central message, lesson, moral, and key details of a text read aloud or information presented in diverse media and formats, including visually, quantitatively, and orally	MSR, MSR two-part		
Reading: Informational	18-26%	3.RI.h1** Identify the purpose of a variety of text features NOT 2-PART	SR	9-13	9-13
		3.RI.h4 Use illustrations (e.g., maps, photographs, diagrams, timelines) in informational texts to answer questions	SR		
		3.RI.i2 Determine the main idea of text read or read aloud or information presented in diverse media and formats, including visually, quantitatively, and orally	SR		
		3.RI.k5** Determine the main idea of a text; recount the key details and explain how they support the main idea	SR, MSR two-part		
Reading: Vocabulary and Foundational	12-16%	3.RWL.i2 Use sentence context as a clue to the meaning of a new word, phrase, or multiple meaning word	SR	6-8	6-8
		3.RWL.i1 Use context to confirm or self-correct word recognition.	SR		
Writing	36-38%	3.WI.i4 Sort evidence (e.g., graphic organizer) collected from print and/or digital sources into provided categories	SR	3-4	3-4
		3.WI.p1 Include text features (e.g., numbers, labels, diagrams, charts, graphics) to enhance clarity and meaning	SR		
		3.WL.o1 With guidance and support from adults, produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (e.g., to entertain), or audience	MSR, CR	7	15
Total	100%			42 Total Items	50 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

Actual percentages by point value, not item count.

ELA Content Category	Gr 3
Reading Literary	24-32%
Reading Informational	18-26%
Reading Vocabulary and Foundational (G3 and G4)	12-16%
Writing	36-38%

Grade 4 Targets by Standard
MSAA ELA Operational Test Blueprint Grade 4

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	24-32%	4.RL.i1 Refer to details and examples in a text when explaining what the text says explicitly	SR	12-16	12-16
		4.RL.k2** Determine the theme of a story, drama, or poem; refer to text to support answer	SR, MSR one or two-part item		
		4.RL.l1** Describe character traits (e.g., actions, deeds, dialogue, description, motivation, interactions); use details from text to support description	SR, MSR two-part		
Reading: Informational	18-26%	4.RI.h4 Use information presented visually, orally, or quantitatively (e.g., in charts, graphs, diagrams, time lines, animations, or interactive elements on Web pages) to answer questions	SR	9-13	9-13
		4.RI.i3 Determine the main idea of an informational text	SR		
		4.RI.l1** Interpret information presented visually, orally, or quantitatively (e.g., in charts, graphs, diagrams, time lines, animations, or interactive elements on Web pages) and explain how the information contributes to an understanding of the text in which it appears	SR, two-part MSR		
Reading: Vocabulary Reading Foundational	12-16%	4.RWL.i2 Use context as a clue to determine the meaning of unknown words, multiple meaning words, or words showing shades of meaning	SR	6-8	6-8
		4.RWL.j1 Use general academic and domain specific words and phrases accurately	SR		
		4.RWL.i1 Use context to confirm or self-correct word recognition.	SR		
Writing	32-38%	4.WI.q1 Provide a concluding statement or section to support the information presented	SR	3-4	3-4
		4.WI.p1 Include formatting (e.g., headings, bulleted information), illustrations, and multimedia when useful to promote understanding	SR		
		4.WL.o1 Produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (e.g. to entertain), or audience	MSR, CR	5-7	13-15
Total	100%			42 Total Items	50 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

ELA Content Category	Gr 4
Reading Literary	24-32%
Reading Informational	18-26%
Reading Vocabulary and Foundational (G3 and G4)	12-16%
Writing	32-38%

Grade 5 Targets by Standard
MSAA ELA Operational Test Blueprint Grade 5

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	25-33%	5.RL.b1 Refer to details and examples in a text when explaining what the text says explicitly	SR	12-16	12-16
		5.RL.c2** Summarize a text from beginning to end in a few sentences	SR, MSR single or multi-part		
		5.RL.d1 Compare characters, settings, events within a story; provide or identify specific details in the text to support the comparison	SR		
Reading: Informational	25-33%	5.RI.d5** Compare and contrast the overall structure (e.g., chronology, comparison, cause/effect, problem/solution) of events, ideas, concepts, or information in two or more texts 2 Part	SR	12-16	12-16
		5.RI.c4** Determine the main idea, and identify key details to support the main idea 2 PART	SR, MSR two-part		
		5.RI.e2 Explain how an author uses reasons and evidence to support particular points in a text	SR		
Reading: Vocabulary	6-10%	5.RWL.a2 Use context to determine the meaning of unknown or multiple meaning words or phrases	SR	3-5	3-5
Writing	31-40%	5.WI.b3 Organize ideas, concepts, and information (using definition, classification, comparison/contrast, and cause/effect)	SR	2-4	2-4
		5.WI.d1 Support a topic with relevant facts, definitions, concrete details, quotations, or other information and examples	SR		
		5.WL.h1 Produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (e.g. to entertain), or audience	MSR, CR	5-7	13-15
Total	100%			40 Total Items	48 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

ELA Content Category	Gr 5
Reading Literary	25-33%
Reading Informational	25-33%
Reading Vocabulary	6-10%
Writing	31-40%

Grade 6 Targets by Standard
MSAA ELA Operational Test Blueprint Grade 6

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	21-30%	6.RL.b2 Refer to details and examples in a text when explaining what the text says explicitly	SR	10-14	10-14
		6.RL.b3 Use specific details from the text (words, interactions, thoughts, motivations) to support inferences or conclusions about characters including how they change during the course of the story	SR		
		6.RL.c3** Summarize a text from beginning to end in a few sentences without including personal opinions 3-PART	SR, SR two-part, MSR		
Reading: Informational	26-34%	6.RI.b4 Summarize information gained from a variety of sources including media or texts	SR	12-16	12-16
		6.RI.c2** Provide a summary of the text distinct from personal opinions or judgments 2 PART	SR, MSR single or multi-part		
		6.RI.g4 Determine how key individuals, events, or ideas are elaborated or expanded on in a text	SR		
		6.RI.g6 Evaluate the claim or argument; determine if it is supported by evidence	SR		
Reading: Vocabulary	9-11%	6.RWL.a1 Use context to determine the meaning of unknown or multiple meaning words or phrases	SR	4-5	4-5
		6.RWL.c1 Use general academic and domain specific words and phrases accurately	SR		
Writing	36-40%	6.WL.c1 Organize ideas and event so that they unfold naturally	SR	3-4	3-4
		6.WL.c3 Use a variety of transition words, phrases, and clauses to convey sequence and signal shifts from one time frame or setting to another	SR		
		6.WI.h2 Produce a clear, coherent, permanent product that is appropriate to the specific task (e.g., topic), purpose (e.g., to inform), and audience (e.g., reader)	MSR, CR	6-7	14-15
Total	100%			39 Total Items	47 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

ELA Content Category	GR 6
Reading Literary	21-30%
Reading Informational	26-34%
Reading Vocabulary	9-11%
Writing	36-40%

Grade 7 Targets by Standard
MSAA ELA Operational Test Blueprint Grade 7

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	17-26%	7.RL.i2** Use two or more pieces of textual evidence to support inferences, conclusions, or summaries of text	SR, SR two-part	8-12	8-12
		7.RL.j1 Analyze the development of the theme or central idea over the course of the text	SR		
Reading: Informational	32-36%	7.RI.i1** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR, SR two-part	15-17	15-17
		7.RI.j5 Analyze the interactions between individuals, events, and ideas in a text (e.g., how ideas influence individuals or events, or how individuals influence ideas or events)	SR		
		7.RI.l1** Compare/contrast how two or more authors write about the same topic	SR, SR two-part		
		7.RI.k4 Evaluate the claim or argument to determine if they are supported by evidence	SR		
Reading: Vocabulary	6-9%	7.RWL.g1 Use context as a clue to determine the meaning of a grade appropriate word or phrase	SR	3-4	3-4
Writing	36-40%	7.WL.o1 Select or provide a concluding statement or paragraph that follows from the narrated experiences or events.	SR	2-4	2-4
		7.WL.l1 Use precise words and phrases, relevant descriptive details, and sensory language to capture the action and convey experiences and events	SR		
		7.WI.o1 Produce a clear, coherent, permanent product (e.g. select/generate responses to form paragraph/essay) that is appropriate to the specific task (e.g., topic), purpose (e.g., to inform), and audience(reader)	MSR, CR	7	15
Total	100%			39 Total Items	47 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

ELA Content Category	GR 7
Reading Literary	17-26%
Reading Informational	32-36%
Reading Vocabulary	6-9%
Writing	36-40%

Grade 8 Targets by Standard
MSAA ELA Operational Test Blueprint Grade 8

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	17-26%	8.RL.i2** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR, SR two-part	8-12	8-12
		8.RL.j2 Analyze the development of the theme or central idea over the course of the text including its relationship to the characters, setting, and plot	SR		
Reading: Informational	32-36%	8.RI.j1** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text 2 PART	SR, SR two-part	15-17	15-17
		8.RI.l1 Analyze a case in which two or more texts provide conflicting information on the same topic and identify where the texts disagree on matters of fact or interpretation	SR		
		8.RI.k2 Determine how the information in each section contribute to the whole or to the development of ideas	SR		
		8.RI.k4 Identify an argument or claim that the author makes	SR		
Reading: Vocabulary	6-9%	8.RWL.g1 Use context as a clue to the meaning of a grade-appropriate word or phrase	SR	3-4	3-4
		8.RWL.i1 Use general academic and domain specific words and phrases accurately	SR		
Writing	36-40%	8.WP.k2 Create an organizational structure in which ideas are logically grouped to support the writer's claim	SR	2-4	2-4
		8.WP.j1 Gather relevant information (e.g., highlight in text, quote or paraphrase from text or discussion) from print and/or digital sources	SR		
		8.WI.o1 Produce a clear, coherent, permanent product (e.g. select/generate responses to form paragraph/essay) that is appropriate to the specific task (e.g., topic), purpose (e.g., to inform), and audience (e.g., reader)	MSR, CR	7	15
Total	100%			39 Total Items	47 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

ELA Content Category	Gr 8
Reading Literary	17-26%
Reading Informational	32-36%
Reading Vocabulary	6-9%
Writing	36-40%

High School Targets by Standard MSAA ELA Operational Test Blueprint--HS

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range	2021 Point Range
Reading: Literary	17-26%	1112.RL.b1** Use two or more pieces of evidence to support inferences, conclusions, or summaries of the plot, purpose, or theme within a text	SR, SR two-part	8-12	8-12
		1112.RL.d1 Analyze how an author's choices concerning how to structure specific parts of a text (e.g., the choice of where to begin or end a story, the choice to provide a comedic or tragic resolution) contribute to its overall structure and meaning	SR		
Reading: Informational	32-36%	1112.RI.b1** Use two or more pieces of evidence to support inferences, conclusions, or summaries or text	SR, SR two-part	15-17	15-17
		1112.RI.b5** Determine how key details support the development of the central idea of a text	SR, SR two-part, MSR		
		1112.RI.d1 Determine the author's point of view or purpose in a text	SR		
		1112.RI.e1 Integrate and evaluate multiple sources of information presented in different media or formats (e.g., visually, quantitatively) as well as in words in order to address a question or solve a problem	SR		
Reading: Vocabulary	6-9%	1112.RWL.b1 Use context (e.g., the overall meaning of a sentence, paragraph, or text; a word's position in a sentence) as a clue to the meaning of a word or phrase	SR	3-4	3-4
		1112.RWL.c3 Develop and explain ideas for why authors made specific word choices within text	SR		
Writing	36-40%	1112.WI.b2 Create an organizational structure for writing that groups information logically (e.g., cause/effect, compare/contrast, descriptions and examples) to support paragraph focus	SR	2-4	2-4
		1112.WI.b4 Select the facts, extended definitions, concrete details, quotations, or other information and examples that are most relevant to the focus and appropriate for the audience	SR		
		1112.WP.f1 Produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (to persuade), and audience	MSR, CR	7	15
Total	100%			39 Total Items	47 Total Points

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

***MSR- Multi-Select Response: for MSAA, this item type is a two-part Selected Response with a dependency between items, similar to an Evidence-Based Selected Response (EBSR). Students are not allowed to reference the previous item because the answer to the first item in the pair is included in the directions/stem of the second item.

ELA Content Category	HS
Reading Literary	17-26%
Reading Informational	32-36%
Reading Vocabulary	6-9%
Writing	36-40%

20-21 MSAA Mathematics Operational Blueprint

- * Standards with operational CR items in 2019
- ** Standards with operational CR items beginning in 2020 and 2021

Grade 3 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Operations and Algebraic Thinking	28-32%	3.NO.2d3 Solve multiplication problems with neither number greater than 5	SR	10 -11
		3.NO.2e1 Solve or solve and check one- or two-step word problems requiring addition, subtraction, or multiplication with answers up to 100		
		3.PRF.2d1 Identify multiplication patterns in a real word setting		
Number and Operations Base Ten	17-23%	3.NO.1j3 Use place value to round to the nearest 10 or 100	SR CR	7
		3.NO.2c1** Solve multi-step addition and subtraction problems up to 100		
Number and Operations Fractions	17-23%	3.NO.1l3 Identify the fraction that matches the representation (rectangles and circles; halves, fourths, thirds, and eighths)	SR	7
		3.SE.1g1 Use =, <, or > to compare 2 fractions with the same numerator or denominator		
Measurement and Data	17-23%	3.DPS.1g1* Collect data; organize into picture or bar graph	SR CR	7
		3.ME.1d2 Measure area of rectilinear figures by counting squares		
Geometry	9-11%	3.GM.1i1 Partition rectangles into equal parts with equal area	SR	3 -4
Total	100%			35

Grade 4 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Operations and Algebraic Thinking	28-32%	4.NO.2d7 Determine how many objects go into each group when given the total number of objects and groups where the number in each group or number of groups is not > 10	SR	10-11
		4.PRF.1e3 Solve multiplicative comparisons with an unknown using up to 2-digit numbers with information presented in a graph or word problem (e.g., an orange hat cost \$3. A purple hat cost 2 times as much. How much does the purple hat cost? [$3 \times 2 = p$])		
		4.NO.2e2 Solve or solve and check one or two step word problems requiring addition, subtraction, or multiplication with answers up to 100		
Number and Operations Base Ten	9-11%	4.NO.1j5 Use place value to round to any place (i.e., ones, tens, hundreds, thousands)	SR	3-4
Number and Operations Fractions	28-32%	4.NO.1m1 Determine equivalent fractions	SR	10-11
		4.NO.1n2 Compare up to 2 given fractions that have different denominators		
		4.SE.1g2 Use $=$, $<$, or $>$ to compare 2 fractions (fractions with a denominator of 10 or less)		
Measurement and Data	17-23%	4.ME.1g2 Solve word problems using perimeter and area where changes occur to the dimensions of a rectilinear figure	SR CR	7
		4.DPS.1g3* Collect data; organize in graph (e.g. picture graph, line plot, bar graph)		
Geometry	9-11%	4.GM.1h2* Classify two-dimensional shapes based on attributes (# of angles)	SR CR	3-4
Total	100%			35

Grade 5 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Operations and Algebraic Thinking	9-11%	5.PRF.2b1 Generate or select a comparison between two graphs from a similar situation	SR	3-4
Number and Operations Base Ten	34-40%	5.NO.1b1 Read, write, or select a decimal to the hundredths place	SR CR	14
		5.NO.1b4 Round decimals to the next whole number		
		5.NO.2c1 Solve one-step problems using decimals		
		5.NO.2a5** Solve word problems that require multiplication or division		
Number and Operations Fractions	17-23%	5.NO.2c2 Solve word problems involving the addition, subtraction, multiplication, or division of fractions	SR	7
		5.PRF.1a1 Determine whether the product will increase or decrease based on the multiplier		
Measurement and Data	17-23%	5.ME.1b2 Convert standard measurements of length	SR	7
		5.ME.2a1 Use a calculator to solve one-step problems involving conversions of standard measurement units of area, volume, time, mass in the same system		
Geometry	9-11%	5.GM.1c3* Use order pairs to graph given points	SR CR	3-4
Total	100%			35

Grade 6 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Ratio and Proportions	28-32%	6.PRF.1c1 Describe the ratio relationship between two quantities for a given situation	SR	10-11
		6.ME.2a2 Solve one-step real world measurement problems involving unit rates with ratios of whole numbers when given the unit rate (3 inches of snow falls per hour, how much in 6 hours?)		
		6.NO.1f1 Find a percent of a quantity as rate per 100		
Expressions and Equations	17-23%	6.PRF.1d1 Solve real world single-step linear equations	SR	7
		6.NO.2a6 Solve problems or word problems using up to three-digit numbers and any of the four operations		
The Number System	28-32%	6.NO.2c3 Solve one-step, addition, subtraction, multiplication, or division problems with	SR CR	10-11
		6.NO.1d4** Select the appropriate meaning of a negative number in a real world situation		
		6.NO.1d2* Locate positive and negative numbers on a number line		
Statistics and Probability	9-11%	6.DPS.1d3 Select the statement that matches mean, mode, and spread of data for 1 measure of central tendency for a given data set	SR	3-4
Geometry	9-11%	6.GM.1d1 Find area of quadrilaterals	SR	3-4
Total	100%			35

Grade 7 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Ratio and Proportions	34-40%	7.NO.2f1** Identify the proportional relationship between two quantities (use rules or symbols to show quantitative relationships)	SR CR	14
		7.NO.2f2 Determine if two quantities are in a proportional relationship using a table of equivalent ratios or points graphed on a coordinate plane		
		7.NO.2f6 Solve word problems involving ratios		
		7.PRF.1f1 Use proportional relationships to solve multistep percent problems in real world situations		
Expressions and Equations	9-11%	7.PRF.1g2 Use variables to represent quantities in a real-world or mathematical problem, and construct simple equations and inequalities to solve problems by reasoning about the quantities	SR	3-4
The Number System	17-23%	7.NO.2i1 Solve multiplication problems with positive/negative numbers	SR	7
		7.NO.2i2 Solve division problems with positive/negative numbers		
Statistics and Probability	9-11%	7.DPS.1k1* Analyze graphs to determine or select appropriate comparative inferences about two samples or populations	SR CR	3-4
Geometry	17-23%	7.ME.2d1 Apply formula to measure area and circumference of circles	SR	7
		7.GM.1h2 Find the surface area of three-dimensional figures using nets of rectangles or triangles		
Total	100%			35

Grade 8 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Functions	17-23%	8.PRF.2e2** Identify the rate of change (slope) and initial value (y-intercept) from graphs	SR	7
		8.PRF.1f2 Describe or select the relationship between the two quantities given a line graph of a situation		
Expressions and Equations	17-23%	8.PRF.1e2 Represent proportional relationships on a line graph	SR	7
		8.PRF.1g3 Solve linear equations with 1 variable		
The Number System	9-11%	8.NO.1k3* Use approximations of irrational numbers to locate them on a number line	SR CR	3-4
Statistics and Probability	17-23%	8.DPS.1h1* Graph bivariate data using scatter plots and identify possible associations between the variables	SR CR	7
		8.DPS.1k2 Analyze displays of bivariate data to develop or select appropriate claims about those data		
Geometry	28-32%	8.ME.1e1 Describe the changes in surface area, area, and volume when the figure is changed in some way (e.g., scale drawings)	SR	10-11
		8.GM.1g1 Recognize congruent and similar figures		
		8.ME.2d2 Apply the formula to find the volume of 3-dimensional shapes (i.e., cubes, spheres, and cylinders)		
Total	100%			35

High School Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2021 Item Range
Algebra And Functions	47-52%	H.PRF.2b1** Translate a real-world problem into a one-variable linear equation	SR CR	17-18
		H.PRF.2b2 Solve equations with one or two variables using equations or graphs		
		H.ME.1b2 Solve a linear equation to find a missing attribute given the area, surface area, or volume and the other attribute		
		H.PRF.1c1 Select the appropriate graphical representation of a linear model based on real world events		
		H.PRF.2c1 Make predictions based on a given model (for example, a weather model, data for athletes over years)		
Number and Quantity	17-23%	H.ME.1a2 Solve real world problems involving units of measurement	SR	7
		H.NO.1a1 Simplify expressions that include exponents		
Statistics and Probability	17-23%	H.DPS.1b1* Complete a graph given the data, using dot plots, histograms, or box plots	SR CR	7
		H.DPS.1c1 Use descriptive stats, range, median, mode, mean, outliers/gaps, to describe data set		
Geometry	9-11%	H.GM.1b1 Use definitions to demonstrate congruency and similarity in figures	SR	3-4
Total	100%			35



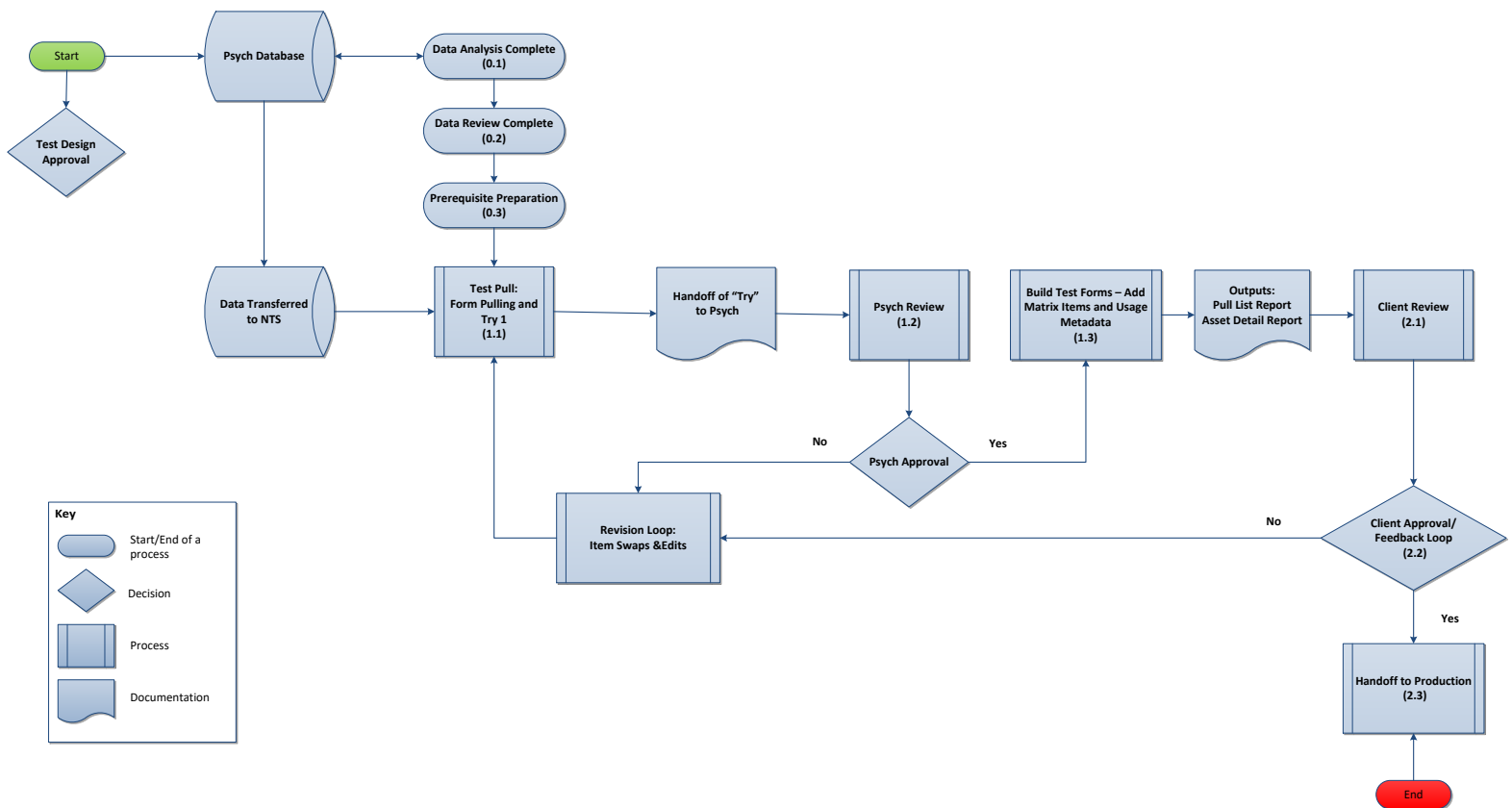
Cognia MSAA Test Construction Process

Purpose

The purpose of this resource is to document the MSAA test construction process beginning with approval of test design through final approval of forms. This resource is used as the guiding resource to replicate MSAA test construction processes across administration years. The process flow chart displayed below captures each individual step outlined in this document. Each process step identified in the flowchart is fully defined in this document to clarify each analysis performed, identify supporting documentation, and detail each step of the process.

Test Construction Process Flow Chart

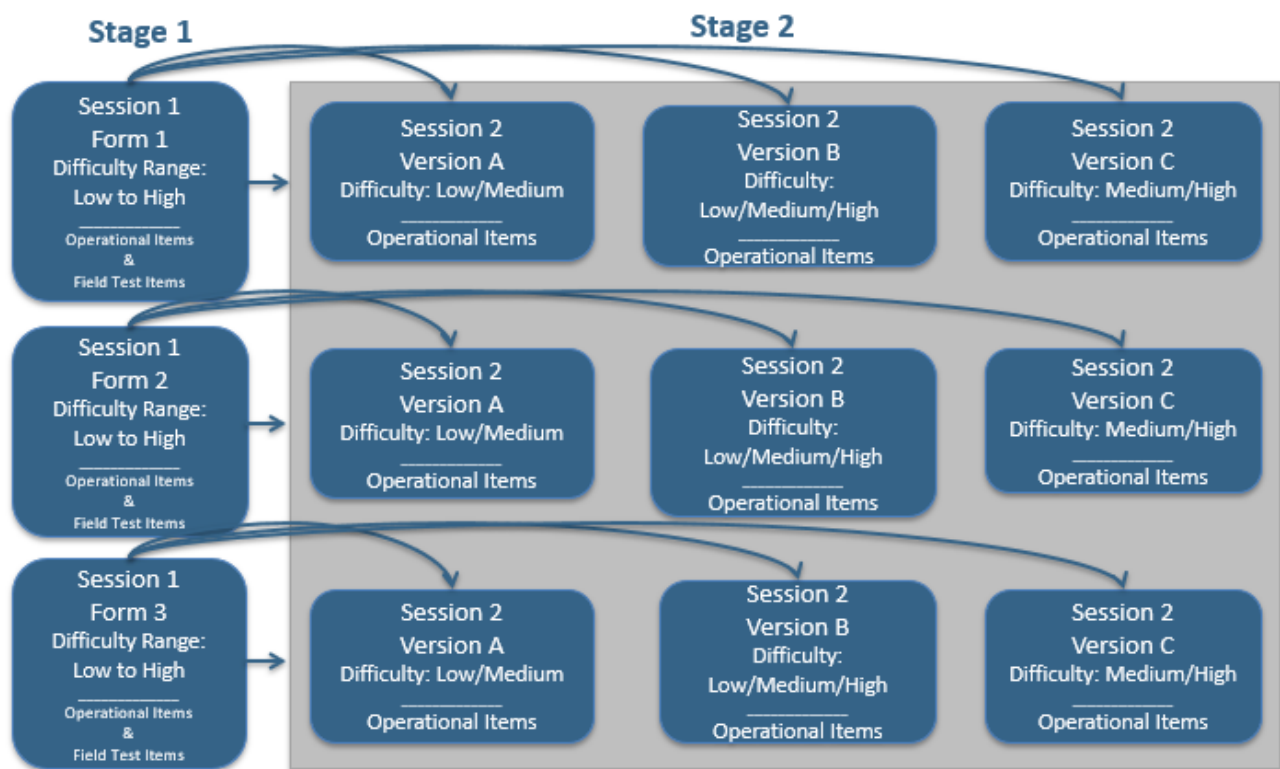
MSAA Test Construction Process



Pre-Test Construction Activities

Test Design

The MSAA ELA and Mathematics tests are multistage in design. Session 1 is administered to all students and includes items ranging in difficulty from low to high. Field test items are spiraled through forms 1-1, 1-2, and 1-3. Session 2 consists of three versions of the test targeting varied levels of difficulty: Version A (low/medium difficulty), Version B (low to high difficulty), and Version C (medium/high difficulty). Depending on performance in Session 1, students are then routed to either Session 2A, 2B, or 2C. Therefore, three possible paths exist for a student to move through the multistage test.



The MSAA test construction subcommittee is provided the opportunity to review and provide feedback on test design documentation annually. This review will occur prior to test construction activities and will include the following elements:

- Form design – states will confirm number of items per form including the positions of operational and embedded field test items.
- Assessment Blueprints – states will confirm assessment blueprints and category distribution for ELA and mathematics.
- Refreshment rate – states will confirm targeted annual rate of refreshment to each test/form.
- Item Types – states will confirm item types (i.e., SR and CR) targeted for each grade/form and content area.

➤ **Outcome:** Approval of Test Design

Data Analysis (0.1)

In preparation for the Data Review meeting (see 0.2), classical and Differential Item Functioning (DIF) statistics will be calculated on all the field-test items from the most recent operational administration. For dichotomously scored items, the classical statistics include the item p-value (the percentage of examinees who responded correctly) and the point-biserial (the correlation of item score with total score on the rest of the test). For polytomously scored items (for example, the trait scores on the writing prompts), in addition to the item p-value (i.e., the mean number of points achieved by examinees divided by the total possible points) and biserial, the number of students (and proportion of students) at each score point will be calculated. Any dichotomous or polytomous item with a negative point-biserial is designated as Do Not Use (DNU). Similarly, any dichotomous or polytomous item answered correctly by all students (or incorrectly by all students) is also designated as DNU. In the same way, if a score category for a writing prompt trait had no students, then the writing prompt would also be designated as DNU (or could be used with collapsed score categories). The DIF statistics will be calculated using the standardization procedure for dichotomously scored items and the standardized-mean-difference procedure for polytomously scored items. A DIF statistic value greater than 0.100 is considered large, while a value less than 0.050 is considered negligible.

In addition to classical statistics, we also look at the results of the Item Response Theory analyses (IRT). The most important results are the item fit plots. We compare the predicted item characteristic curve (ICC) (probability of a correct response, conditional on theta) with the empirical observed proportion of students who correctly responded at particular values of theta. If the ICC and the empirical data align closely to each other, the item is acceptable. If professional judgment indicates a strong mismatch, then the item is designated as DNU. When the mismatch is only partial, an item may be classified as Use With Caution (UWC). If the mismatch for an item is U-shaped or V-shaped, the item may be considered for recalibration using examinees only from Path B and/or Path C.

In addition, we study the IRT parameter estimates and their corresponding standard errors. For the discrimination parameter (the “a” parameter), an item with a value less than 0.25 will generally be flagged as DNU. Values greater than 3 are considered anomalies and require further investigation before being allowed in the common-eligible pool. In regard to the difficulty parameter (the “b” parameter), items with values outside of -4 to 4 are considered outliers and require further investigation before being allowed in the common-eligible pool. Also, if the standard error of the estimated “b” parameter for an item is greater than 0.30, the item is discouraged (generally DNU) for operational usage. For the polytomously scored writing prompt traits, if a trait score cannot be calibrated due to it having too few students (for example, fewer than 50), the writing prompt may be calibrated after collapsing score categories or may be designated as DNU.

Criteria	Use	Use w/ Caution	Do Not Use
Difficulty (<i>b</i>)	$b > -4 \& b < 4$	$b > -4 \& b < 4$	$b < -4$ or $b > 4$
Discrimination (<i>a</i>)	$a > .25$	$a > .25$	$a < .25$
SE of Difficulty [SE(<i>b</i>)]	SE(<i>b</i>) < .2	SE(<i>b</i>) > .2 & SE(<i>b</i>) < .3	SE(<i>b</i>) > .3

We summarize the finding of the IRT analyses in a Field-test Calibration Report (see Appendix A). In general, DNU items are typically items with large misfit, and UWC items are typically items with low discrimination but still having reasonable and appropriate mode fit.

- **Outcome:** Field Test Calibration Report for Data Review meeting

Data Review Meeting (0.2)

During the data review meeting, Cognia and the MSAA test construction subcommittee review the Field Test Calibration Report, which includes item statistics for each field test item that has been flagged by Psychometrics. The statistical parameters that trigger an item being flagged are also shared at this time. Referenced during data review are the IRT analyses summarized in the Field Test Calibration Report and DIF statistics. Data review attendees are also supplied with Asset Detail Reports for each of the flagged items. This allows for the content of the flagged items to be considered when determining future usage.

Some flagged items may be placed into different categories. The items might be flagged only for Form 2A (i.e., use only in forms 2B and 2C), flagged for 2A and 2B (i.e., use only in form 2C), or flagged for all three (i.e., flagged as DNU or UWC). The content of the item is reviewed along with the statistics. After each item is reviewed, the client representatives determine whether an item is accepted with the corresponding usage recommendation, rejected, or designated as revise and re-field test.

- **Outcome:** Future operational use of field-tested items determined

Prerequisite Preparation (0.3)

In preparation for test creation or “form pulling,” Cognia content specialists identify common eligible pools of items that can function as operational items. To meet common eligible status, content specialists must verify that the items and corresponding passages have been through development, field testing, and have been approved for use or have been used as an operational item on a previous form.

Using the common eligible pool, Psychometrics creates a Cut Stuff Report. The Cut Stuff Report indicates for each item its contribution to the value of the Test Characteristic Curve (TCC) and the Test Information Function (TIF) at the routing cut scores and at each of the three performance level cut points. In addition, for ELA, because items are administered in item sets, the Cut Stuff Report also provides the TCC and TIF contributions of the entire item set (by simply summing the contributions of each item in the set).

The TCC tells the expected value of a raw score on the test for a student at a specified level of performance. The TIF tells how much statistical information is provided by the test at each level of performance. The greater the TIF, the greater the test reliability and the smaller the measurement error in a student’s scaled score.

The TIF at the cuts for the proposed tests are compared to the TIFs in previous year(s). The comparison at Stage 1 is done at the routing cut scores. The comparison for Stage 2A focuses

primarily on Cut 1, with secondary emphasis on Cut 2. The comparison for Stage 2B is focused on Cut 2, with secondary emphasis on Cuts 1 and 3. The comparison for Stage 2C focuses on Cut 3, with secondary emphasis on Cut 2. If the TIF at the cut scores are at least 90% of the values in past years, the test is most likely providing adequate reliability as compared to past years.

In the Cut Stuff Report, Psychometrics will flag items with small TIF values as “Replace if possible.”

- **Outcome:** Cut Stuff Report

Test Pull: Form Pulling and Try 1

Form Pulling and Try 1 (1.1)

MSAA test creation (i.e., “form pulling”) begins with the forms from the previous year. The primary goal is to build new operational test forms, refreshing approximately one third of the items from the previous year. Content specialists focus their attention on replacing the items flagged by Psychometrics indicated in the Cut Stuff Report. Items are flagged with the goal of maximizing information at a performance level cut point and for the corresponding test form. To find a replacement for a flagged item, content specialists look for common eligible items that match the content standard of the flagged item to ensure alignment to the blueprint. The item with the largest TIF at the appropriate cut will be the best replacement item because it helps to maximize information for the form that corresponds to that performance level cut point and corresponding test form. When possible, efforts are also made to select items that have been most recently developed with the goal of gradually replacing older migrated NCSC items. Once the forms are constructed in NTS, the test forms are reviewed by another content specialist and by a special education specialist. After making any modifications, the forms are reviewed by Psychometrics.

- **Outcome:** Try 1 Proposed Operational Set

Psychometric Review (1.2)

The psychometric review of Try 1 results in the production of two files: the TIF file that shows graphs and tables for Stages 1, 2A, 2B, and 2C; and a Psychometric Review file that discusses the results and makes recommendations regarding the acceptability of the psychometric characteristics.

The graphs and tables in the TIF file compare the proposed test with the TIF that occurred the previous year. In particular, the tabulated results for Stage 1 report the TIF values at the routing cuts, while the tabulated results for Stage 2 report the TIF values at the performance-level cuts.

The Psychometric Review file gives a detailed summary and analysis of the results. It includes special notes and caveats to keep in mind for the particular test being analyzed.

- A review of the Stage 1 results focusing on the TIF at the routing cuts, comparing with the TIF of the previous year. The review also compares the TIF curves in general over the entire theta range of interest—that is, the scale score range for a test form—to Stage 1.
- A review of the Stage 2A results focusing on Cuts 1 and 2, again comparing the TIF with that of the previous year; while also comparing the TIF functions more generally over the theta range of interest to Stage 2A.
- A review of the Stage 2B results focusing on all three cuts, again comparing the TIF with that of the previous year; while also comparing the TIF functions more generally over the theta range of interest to Stage 2B.
- A review of Stage 2C results focusing on Cuts 2 and 3, again comparing the TIF with that of the previous year; while also comparing the TIF functions more generally over the theta range of interest to Stage 2C.
- A review of the TIF results for Paths A, B, and C (the three complete total test forms that make up the multistage test), focusing on the cuts of greatest interest for each Path (Cuts 1 and 2 for Path A; Cuts 1, 2, and 3 for Path B; and Cuts 2 and 3 for Path C). Also included in this section is a note on any restrictions in classification that may be present for a particular path (for example, Path A students sometimes cannot be categorized in the highest performance level).

Finally, an overall summary is provided giving an overall comparison of the proposed test with last year's test, noting where improvements have occurred and where improvements are still needed, either for the current test if the pool supports it or in the future.

The "Try 1" is either approved or one or more recommendations are provided to improve the technical quality of the forms. If any changes can be made given the depth of the item pool, the forms go through the review cycle again until they are approved.

When further tries beyond Try 1 are conducted, the Psychometric Review file is updated by adding further comments on any changes that were implemented for any of the stages. The original Try 1 comments are also included so that all the comments for all Tries are maintained in a single document.

- **Outcome:** Proposed Operational Forms (following additional Try 2 and Try 3 if needed), TIF and Psychometric Review files

Build Test Forms – Add Field Test Items and Usage Metadata (1.3)

When the operational test forms are approved by Psychometrics, the Cognia content specialists select items to slot into the field test positions in routing forms 1A, 1B, and 1C. These items are selected from the items that have been approved by MSAA as final following the post-IRC review revisions. Embedded field test items appear only in Session 1.

Once the forms are completed, content specialists designate test usage metadata. Content specialists then sequence the items following test construction best practices. The forms are then reviewed by a second content specialist and a special education specialist to ensure all field test

items are properly represented. A Pull List is then created for each test from NTS. A Pull List is then created for each test from NTS.

- **Outcome:** Proposed Test Forms including Field Test Items and Pull List for final MSAA approval

Client Review

Client Review (2.1)

All test forms are provided to the Test Construction Subcommittee for review. The subcommittee members are provided with the following resources to facilitate their review:

- Sequenced Asset Detail Reports (ADR) – The ADRs contain the test content presented for each item. The items will be sequenced in the proposed order for each test.
- Pull List – the pull list serves as the “test map” for each form of each test. The pull list includes metadata associated with each item (i.e., asset ID, CCC, item level, statistics). Reviewers are asked to provide their feedback on a tab on the pull list.
- TIF File (Core1_out) and Psychometric Review file (Core1_msg) – the IRT results and curves will be provided for each stage of each test. Psychometrics will also provide recommendations that summarize each “Try” and the characteristics of the final test.
- Cut Stuff Report – this file includes the routing cuts and cut points for each asset at each of the three performance levels on each form. This file also includes items recommended for replacement.
- Common Eligible File – This file identifies items that are eligible for operational usage.

- **Outcome:** Handoff of files for client review

Client Approval/Feedback Loop (2.2)

All test forms are reviewed by the client and feedback is provided. Client feedback may include requests to move or replace specific items. If items can be moved between forms and/or replaced, the client is provided with a revised pull list that reflects the changes. If there are any requested changes that cannot be made, documentation of the rationale is also provided. The revised test forms are returned to Psychometrics for approval and then handed off to the client for approval.

- **Outcome:** Approval of Test Forms or Revision Loop

Handoff to Production (2.3)

Once test forms have been approved by MSAA, the content specialists prepare the tests for Production.

- **Outcome:** Tests prepared for production process

Appendix

- A. Sample Field Test Calibration Report
- B. MSAA Psychometric Glossary

MSAA Sample FT Calibration Report

IRT STATISTICS

- a: item discrimination
 - closely tied to the slope of the blue curve (model-predicted probabilities)
 - Don't want below .25, and above 3 is an outlier
- b: item difficulty
 - lower value indicates easier item and vice versa
 - it is the inflexion point on the blue curve
 - generally falls between -3 and +3; if needed, we're willing to accept values between -4 and +4
 - outside -4 and +4 is a definite outlier
- se(b): standard error of b
 - indication of how sure the software is about the estimate of b (item difficulty)
 - we like seeing values less than .3

2017-18 Item Calibration Categorization Criteria

Criteria	Use	Use w/ Caution	DO NOT USE
Difficulty (b)	$b > -4 \& b < 4$	$b > -4 \& b < 4$	$b < -4$ or $b > 4$
Discrimination (a)	$a > .25$	$a > .25$	$a < .25$
SE of Difficulty [SE(b)]	SE(b) $< .2$	SE(b) $> .2 \& SE(b) < .3$	SE(b) $> .3$

Do Not Use Items

Subject	Grade	ItemID	a	SE(a)	b	SE(b)	Point Biserial r	Recommendation	a	SE(a)	b	SE(b)	Updated Recommendation
ELA	7	537384	0.244	0.034	0.429	0.150	0.150	USE					
ELA	11	532003	0.245	0.037	0.059	0.155	0.110	USE					
ELA	11	539001	0.246	0.036	1.639	0.270	0.110	2B/2C					
Math	3	442151	0.190	0.035	0.263	0.217	0.090	2B/2C					
Math	4	445822	0.294	0.047	2.425	0.390	0.110	2B/2C					
Math	4	454885	0.181	0.038	-0.156	0.221	-0.150	DNU					
Math	4	529973	0.227	0.039	0.647	0.191	0.090	2B/2C					
Math	4	531533	0.230	0.041	2.499	0.461	0.060	2B/2C					
Math	4	111138A	0.121	0.027	2.944	0.722	-0.050	DNU					
Math	4	111140A	0.135	0.030	4.392	1.013	-0.060	Check for 2C	0.474	0.097	1.984	0.265	2C Only-Use with Caution
Math	4	111146A	0.133	0.029	1.620	0.437	-0.030	DNU					
Math	4	111151A	0.236	0.041	1.668	0.319	0.070	Check for 2C	0.602	0.112	1.204	0.135	2C Only-USE
Math	5	450269	0.254	0.038	2.671	0.401	0.140	DNU					
Math	5	450289	0.076	0.017	0.000	0.479	-0.110	DNU					
Math	5	532879	0.237	0.034	0.459	0.168	0.160	2B/2C					
Math	5	532898	0.183	0.034	3.634	0.677	0.090	Check for 2C	0.416	0.075	2.486	0.301	DNU
Math	5	111305A	0.173	0.031	1.726	0.365	0.100	2B/2C					
Math	5	112349A	0.244	0.036	2.501	0.379	0.150	Check for 2C	0.415	0.072	2.215	0.246	2C Only-Use with Caution
Math	5	112370A	0.091	0.020	1.826	0.565	-0.070	Check for 2C	0.252	0.055	1.955	0.316	DNU
Math	5	112402A	0.183	0.032	0.654	0.222	0.110	2B/2C					
Math	6	442507	0.168	0.031	3.354	0.648	0.060	Check for 2C	0.370	0.069	2.408	0.302	DNU
Math	6	533985	0.188	0.032	1.502	0.307	0.080	2B/2C					
Math	6	534015	0.123	0.026	4.375	0.933	-0.010	Check for 2C	0.561	0.080	2.138	0.183	2C Only-USE
Math	6	110983A	0.217	0.035	2.378	0.401	0.080	2B/2C					
Math	7	446521	0.185	0.036	2.336	0.489	0.040	DNU					
Math	7	111755A	0.201	0.036	2.149	0.422	0.080	Check for 2C	0.353	0.070	1.599	0.231	DNU
Math	8	446908	0.081	0.018	0.000	0.428	-0.100	DNU					
Math	8	447116	0.360	0.047	-2.452	0.306	0.170	DNU					

Math	8	111282A	0.186	0.031	2.518	0.457	0.060	Check for 2C	0.337	0.055	2.242	0.286	2C Only-Use with Caution
Math	8	111611A	0.209	0.032	0.704	0.202	0.100	2B/2C					
Math	8	112507A	0.104	0.023	2.249	0.622	-0.050	2B/2C-Use with Caution					
Math	11	443403	0.339	0.055	2.083	0.342	0.110	2B/2C-Use with Caution					
Math	11	110856A	0.210	0.042	1.974	0.421	0.040	Check for 2C	0.387	0.078	1.574	0.238	DNU
Math	11	111107A	0.213	0.044	2.926	0.618	-0.020	DNU					

Use with Caution

Subject	Grade	ItemID	<i>a</i>	SE(<i>a</i>)	<i>b</i>	SE(<i>b</i>)	Point Biserial <i>r</i>	Recommendation	<i>a</i>	SE(<i>a</i>)	<i>b</i>	SE(<i>b</i>)	Updated Recommendation
ELA	6	540986	0.250	0.039	1.562	0.252	0.100	Checking for 2B/2C	0.334	0.056	1.451	0.191	2B/2C-USE
MAT	3	529844	0.310	0.047	1.703	0.296	0.130	Checking for 2B/2C	0.368	0.062	1.555	0.255	2B/2C- Use with Caution
									0.530	0.100	1.480	0.195	2C Only-USE*
MAT	3	111651A	0.316	0.045	1.700	0.281	0.180	Checking for 2B/2C	0.334	0.057	1.665	0.276	2B/2C- Use with Caution
									0.426	0.083	1.570	0.229	2C- Use with Caution
MAT	4	445528	0.314	0.046	1.428	0.231	0.140	Good for 2B/2C					
MAT	4	111691A	0.301	0.045	1.839	0.285	0.120	Good For 2B/2C					
MAT	5	450050	0.360	0.042	1.690	0.204	0.230	Checking for 2B/2C	0.396	0.053	1.678	0.194	2B/2C-USE
MAT	5	113837A	0.313	0.039	1.859	0.243	0.210	Checking for 2B/2C	0.453	0.054	1.636	0.168	2B/2C-USE
									0.705	0.095	1.590	0.121	2C Only-USE*
MAT	6	442625	0.324	0.038	1.840	0.224	0.180	Checking for 2B/2C	0.499	0.060	1.602	0.146	2B/2C-USE
MAT	6	450445	0.319	0.040	1.901	0.246	0.180	Checking for 2B/2C	0.430	0.060	1.719	0.186	2B/2C-USE
									0.583	0.085	1.742	0.153	2C Only-USE*
MAT	6	111023A	0.332	0.039	1.925	0.228	0.180	Checking for 2B/2C	0.472	0.059	1.736	0.164	2B/2C-USE
									0.693	0.088	1.649	0.117	2C Only-USE*
MAT	7	530662	0.261	0.041	1.633	0.289	0.120	Checking for 2B/2C	0.320	0.049	1.495	0.233	2B/2C- Use with Caution
									0.559	0.092	1.344	0.141	2C Only-USE*
MAT	7	111081A	0.334	0.045	1.370	0.210	0.230	Good For 2B/2C					
MAT	8	113936A	0.304	0.038	1.670	0.232	0.190	Good For 2B/2C					
MAT	8	113971A	0.278	0.036	1.654	0.244	0.180	Checking for 2B/2C	0.269	0.045	1.770	0.271	DNU
									0.276	0.051	1.735	0.263	

*Preference for use designation given analysis results; MSAA to make determination on preference for designation.

MSAA Sample FT Calibration Report (cont.)

FIT PLOTS

- Red dots: empirical proportion of students getting the item correct
 - We put students into buckets depending on their abilities (thetas), and compute the empirical proportion of students who got the particular item correct.
- Blue curve: model-predicted probabilities
 - The curve provides the model-predicted probability of students getting the particular item correct depending on their abilities.
- We like seeing the red dots and blue curves aligning closely to each other – means GOOD FIT!

IRT STATISTICS

- a: item discrimination
 - closely tied to the slope of the blue curve (model-predicted probabilities)
 - Don't want below .25, and above 3 is an outlier
- b: item difficulty
 - lower value indicates easier item and vice versa
 - it is the inflexion point on the blue curve
 - generally falls between -3 and +3; if needed, we're willing to accept values between -4 and +4
 - outside -4 and +4 is a definite outlier
- se(b): standard error of b
 - indication of how sure the software is about the estimate of b (item difficulty)
 - we like seeing values less than .3

2017-18 Item Calibration Categorization Criteria

Criteria	Use	Use w/ Caution	DO NOT USE
Difficulty (b)	$b > -4$ & $b < 4$	$b > -4$ & $b < 4$	$b < -4$ or $b > 4$
Discrimination (a)	$a > .25$	$a > .25$	$a < .25$
SE of Difficulty [SE(b)]	SE(b) $< .2$	SE(b) $> .2$ & SE(b) $< .3$	SE(b) $> .3$

Do Not Use Items

Subject	Grade	ItemID	a	$SE(a)$	b	$SE(b)$	Point Biserial r	Recommendation
ELA	7	537384	0.244	0.034	0.429	0.150	0.150	USE
ELA	11	532003	0.245	0.037	0.059	0.155	0.110	USE
ELA	11	539001	0.246	0.036	1.639	0.270	0.110	2B/2C
Math	3	442151	0.190	0.035	0.263	0.217	0.090	2B/2C
Math	4	445822	0.294	0.047	2.425	0.390	0.110	2B/2C
Math	4	454885	0.181	0.038	-0.156	0.221	-0.150	DNU
Math	4	529973	0.227	0.039	0.647	0.191	0.090	2B/2C
Math	4	531533	0.230	0.041	2.499	0.461	0.060	2B/2C
Math	4	111138A	0.121	0.027	2.944	0.722	-0.050	DNU
Math	4	111140A	0.135	0.030	4.392	1.013	-0.060	Check for 2C
Math	4	111146A	0.133	0.029	1.620	0.437	-0.030	DNU
Math	4	111151A	0.236	0.041	1.668	0.319	0.070	Check for 2C
Math	5	450269	0.254	0.038	2.671	0.401	0.140	DNU
Math	5	450289	0.076	0.017	0.000	0.479	-0.110	DNU
Math	5	532879	0.237	0.034	0.459	0.168	0.160	2B/2C
Math	5	532898	0.183	0.034	3.634	0.677	0.090	Check for 2C
Math	5	111305A	0.173	0.031	1.726	0.365	0.100	2B/2C
Math	5	112349A	0.244	0.036	2.501	0.379	0.150	Check for 2C
Math	5	112370A	0.091	0.020	1.826	0.565	-0.070	Check for 2C
Math	5	112402A	0.183	0.032	0.654	0.222	0.110	2B/2C
Math	6	442507	0.168	0.031	3.354	0.648	0.060	Check for 2C
Math	6	533985	0.188	0.032	1.502	0.307	0.080	2B/2C
Math	6	534015	0.123	0.026	4.375	0.933	-0.010	Check for 2C
Math	6	110983A	0.217	0.035	2.378	0.401	0.080	2B/2C
Math	7	446521	0.185	0.036	2.336	0.489	0.040	DNU
Math	7	111755A	0.201	0.036	2.149	0.422	0.080	Check for 2C
Math	8	446908	0.081	0.018	0.000	0.428	-0.100	DNU
Math	8	447116	0.360	0.047	-2.452	0.306	0.170	DNU
Math	8	111282A	0.186	0.031	2.518	0.457	0.060	Check for 2C
Math	8	111611A	0.209	0.032	0.704	0.202	0.100	2B/2C
Math	8	112507A	0.104	0.023	2.249	0.622	-0.050	2B/2C

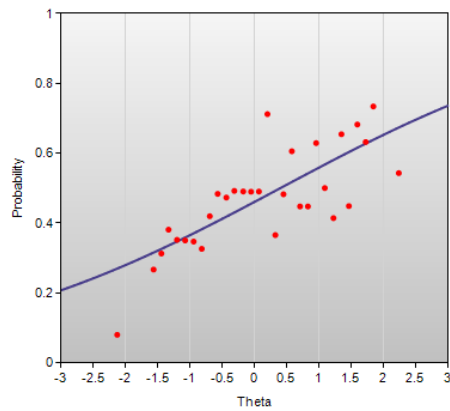
Math	11	443403	0.339	0.055	2.083	0.342	0.110	2B/2C
Math	11	110856A	0.210	0.042	1.974	0.421	0.040	Check for 2C
Math	11	111107A	0.213	0.044	2.926	0.618	-0.020	DNU

Use with Caution

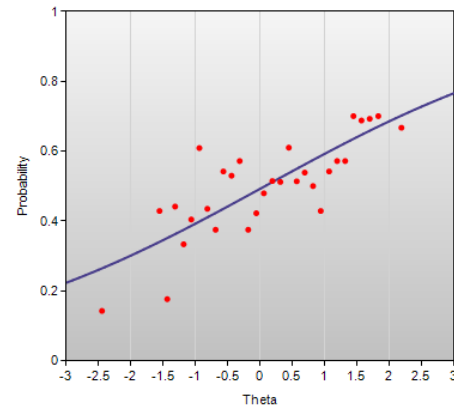
Subject	Grade	ItemID	α	SE(α)	b	SE(b)	Point Biserial r	FT Form
ELA	6	540986	0.250	0.039	1.562	0.252	0.100	Checking for 2B/2C
MAT	3	529844	0.310	0.047	1.703	0.296	0.130	Checking for 2B/2C
MAT	3	111651A	0.316	0.045	1.700	0.281	0.180	Checking for 2B/2C
MAT	4	445528	0.314	0.046	1.428	0.231	0.140	Good for 2B/2C
MAT	4	111691A	0.301	0.045	1.839	0.285	0.120	Good For 2B/2C
MAT	5	450050	0.360	0.042	1.690	0.204	0.230	Checking for 2B/2C
MAT	5	113837A	0.313	0.039	1.859	0.243	0.210	Checking for 2B/2C
MAT	6	442625	0.324	0.038	1.840	0.224	0.180	Checking for 2B/2C
MAT	6	450445	0.319	0.040	1.901	0.246	0.180	Checking for 2B/2C
MAT	6	111023A	0.332	0.039	1.925	0.228	0.180	Checking for 2B/2C
MAT	7	530662	0.261	0.041	1.633	0.289	0.120	Checking for 2B/2C
MAT	7	111081A	0.334	0.045	1.370	0.210	0.230	Good For 2B/2C
MAT	8	113936A	0.304	0.038	1.670	0.232	0.190	Good For 2B/2C
MAT	8	113971A	0.278	0.036	1.654	0.244	0.180	Checking for 2B/2C

Fit Plots – Do Not Use Items

English Language Arts Grade 7: 537384



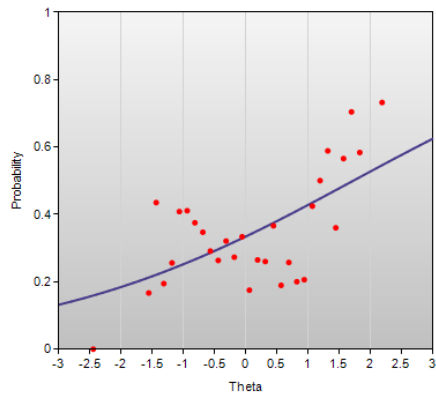
English Language Arts Grade 11: 532003



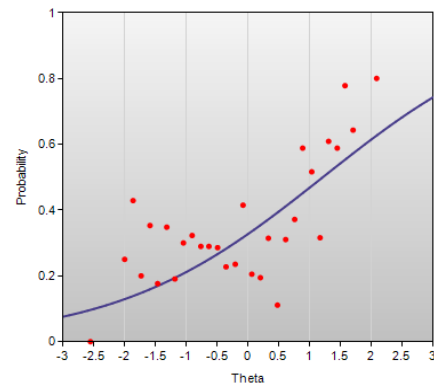
Recommend reclassifying as Use

Recommend reclassifying as Use

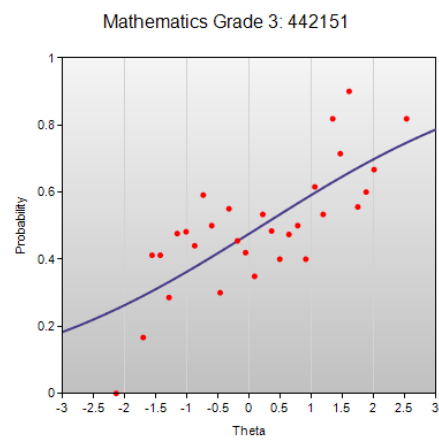
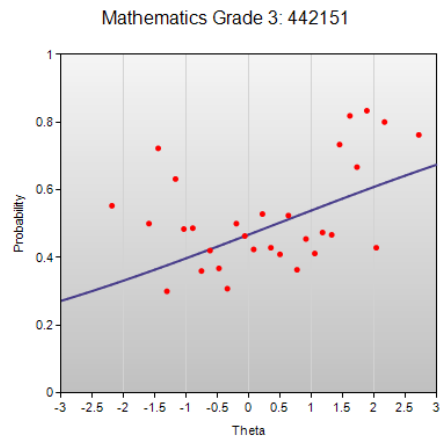
English Language Arts Grade 11: 539001



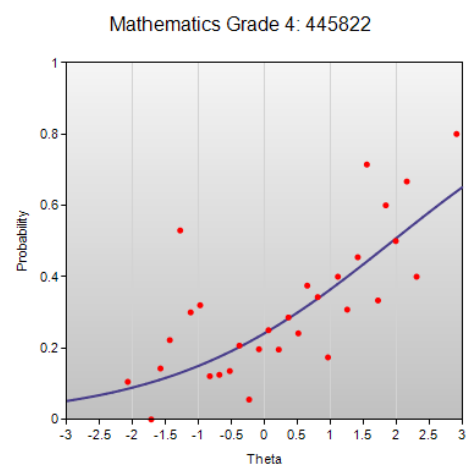
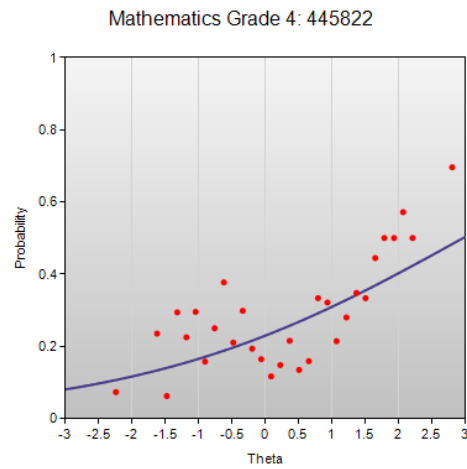
English Language Arts Grade 11: 539001



Recalibrated using 2B/2C. Went from DNU ($a=.246, b=1.64/se.27$) to use ($a=.39/b=1.4/se.17$)

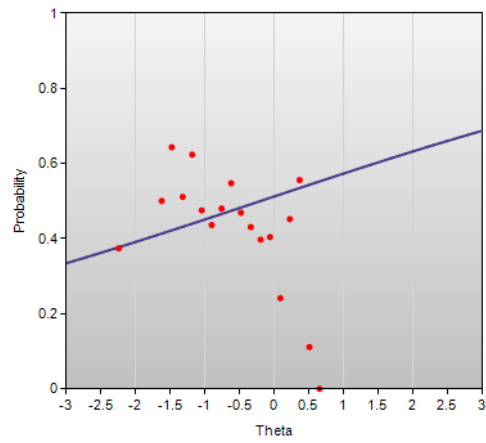


Recalibrated using 2B/2C. Went from DNU ($a=.19, b=.26/se.22$) to use ($a=.32/b=.40/se.15$)



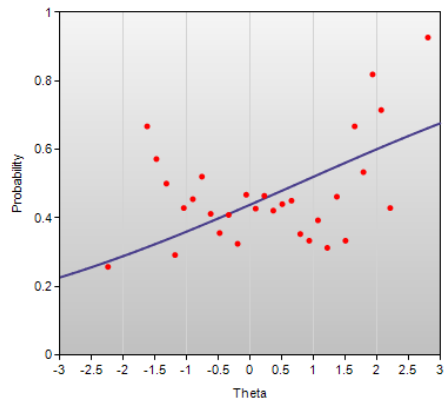
Recalibrated using 2B/2C. Went from DNU ($a=.29, b=2.43/se.39$) to UWC ($a=.43, b=1.92/se.26$)

Mathematics Grade 4: 454885

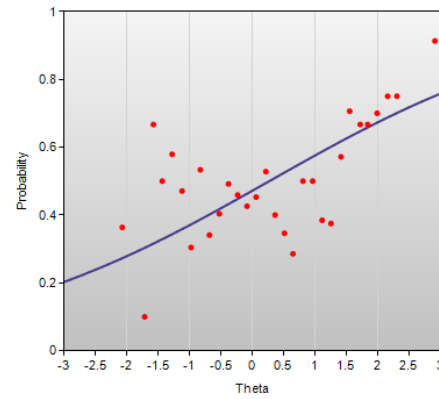


DNU

Mathematics Grade 4: 529973

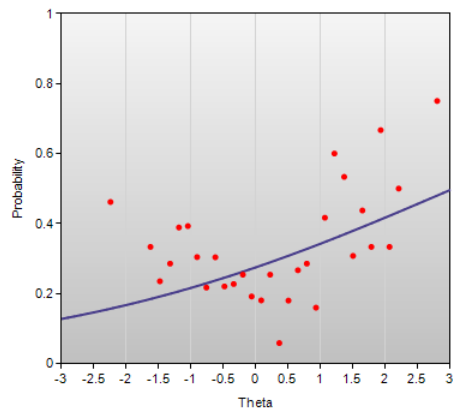


Mathematics Grade 4: 529973

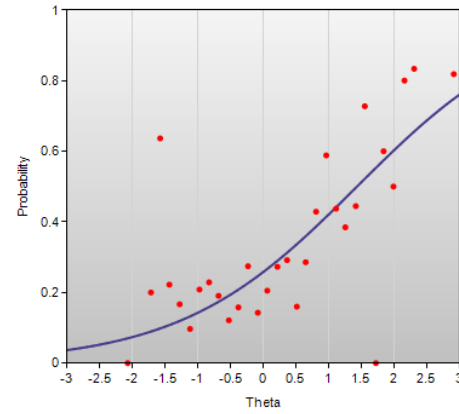


Recalibrated using 2B/2C. Went from DNU ($a=.23, b=.65/se.19$) to use ($a=.31, b=.59/se.15$)

Mathematics Grade 4: 531533

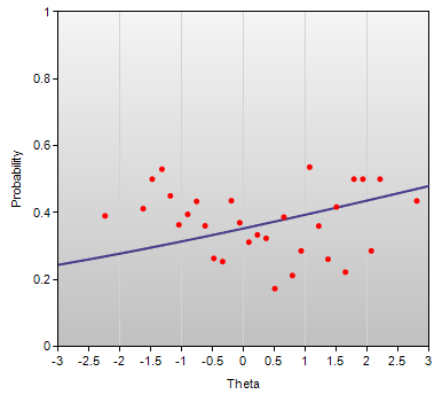


Mathematics Grade 4: 531533



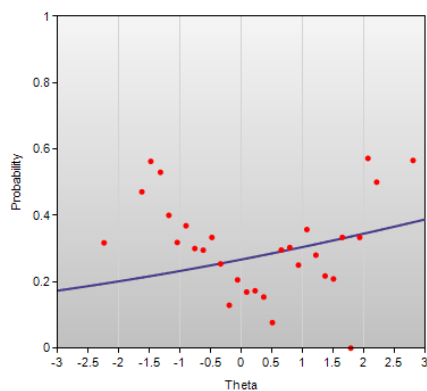
Recalibrated using 2B/2C. Went from DNU ($a=.23, b=2.50/se.46$) to use ($a=.48, b=1.63/se.198$)

Mathematics Grade 4: 111138A

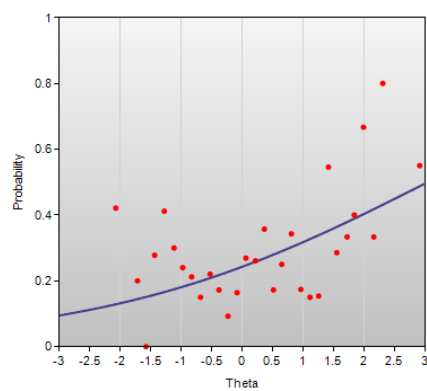


DNU

Mathematics Grade 4: 111140A

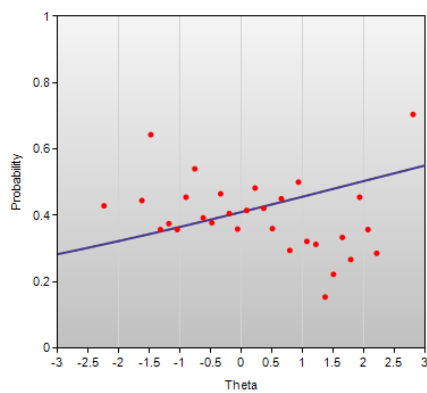


Mathematics Grade 4: 111140A

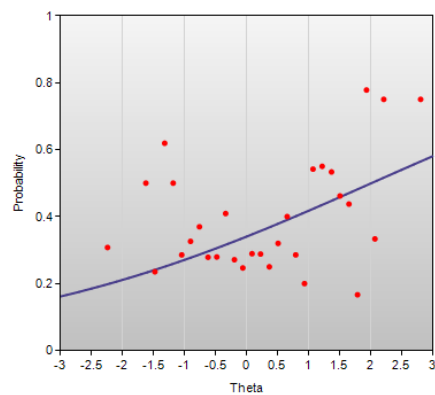


Recalibrated using 2B/2C. Still DNU (orig'l a=.14, b=4.4/se1.0) (new a=.31, b=2.5/se.43)

Mathematics Grade 4: 111146A



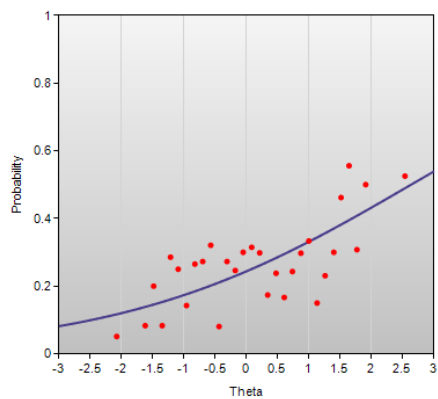
Mathematics Grade 4: 111151A



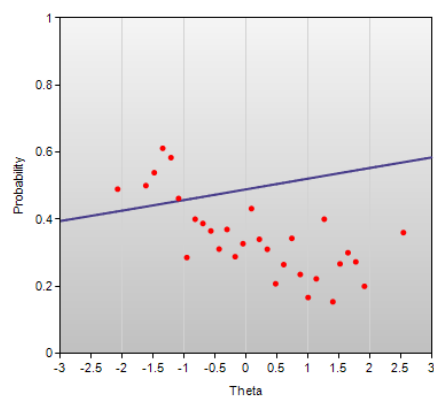
DNU

DNU

Mathematics Grade 5: 450269



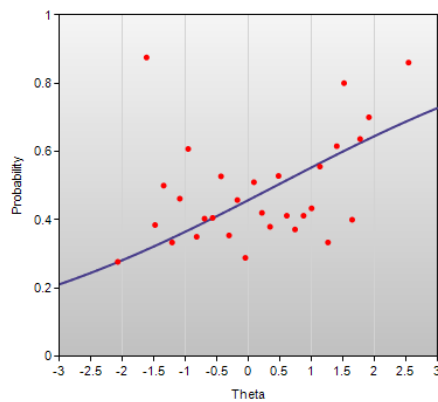
Mathematics Grade 5: 450289



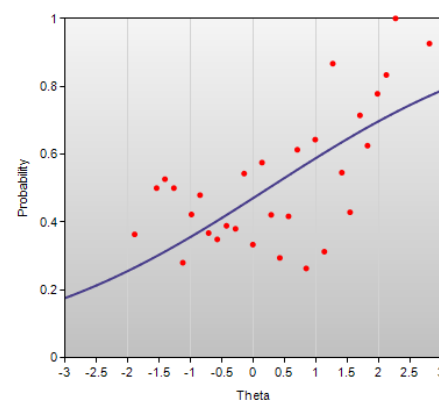
DNU

DNU

Mathematics Grade 5: 532879

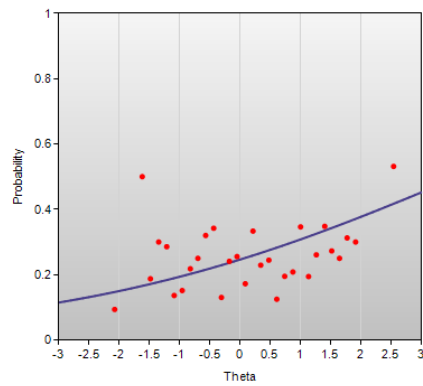


Mathematics Grade 5: 532879



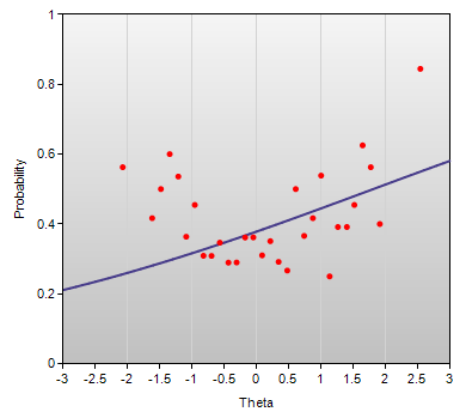
Recalibrated using 2B/2C. Went from DNU ($a=.236, b=.5/se=.17$) to use ($a=.34, b=.6/se=.13$)

Mathematics Grade 5: 532898

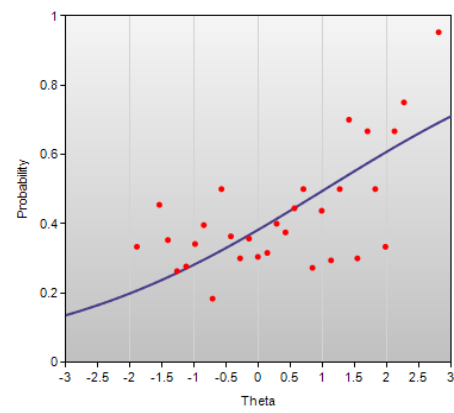


DNU

Mathematics Grade 5: 111305A

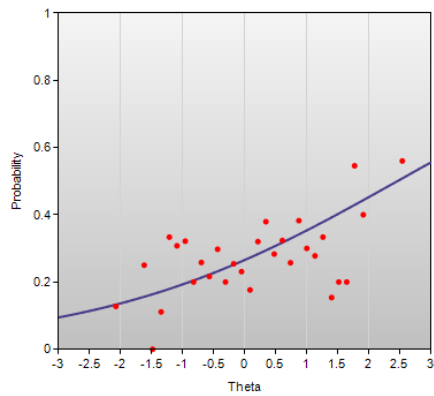


Mathematics Grade 5: 111305A

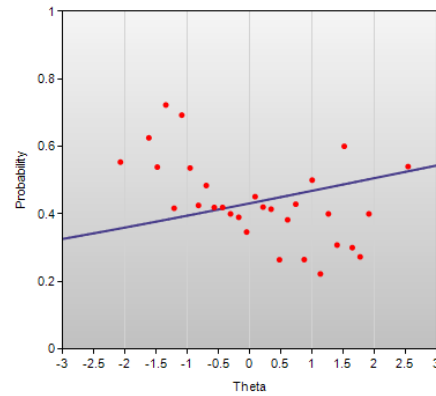


Recalibrated using 2B/2C. Went from DNU ($a=.17$, $b=1.7/se=.36$) to use ($a=.32$, $b=1.3/se=.19$)

Mathematics Grade 5: 112349A



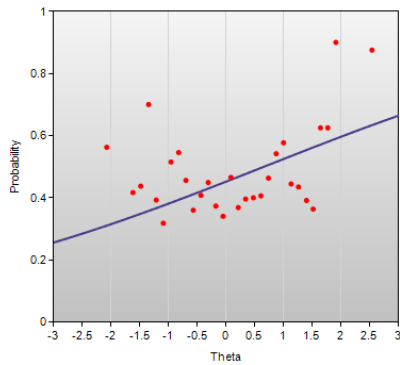
Mathematics Grade 5: 112370A



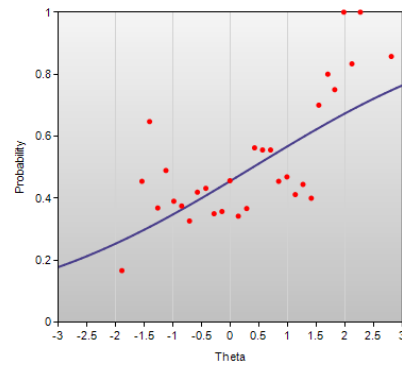
DNU

DNU

Mathematics Grade 5: 112402A

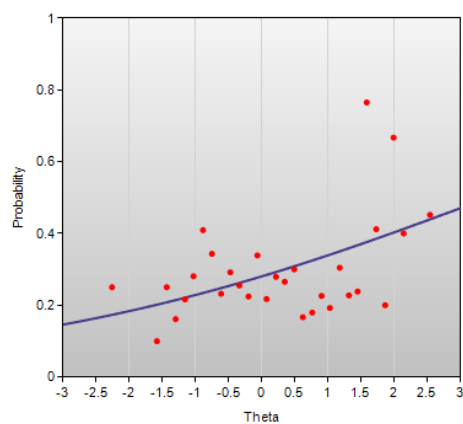


Mathematics Grade 5: 112402A



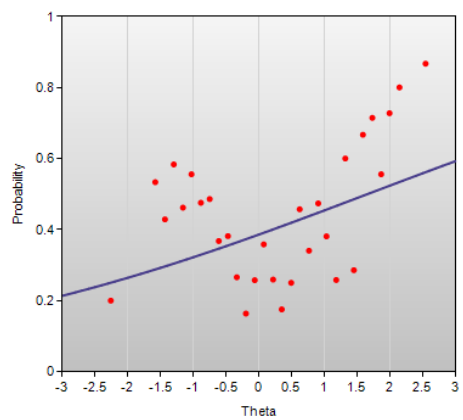
Recalibrated using 2B/2C. Went from DNU ($a=.18$, $b=.65/se=.22$) to use ($a=.32$, $b=.75/se.14$)

Mathematics Grade 6: 442507

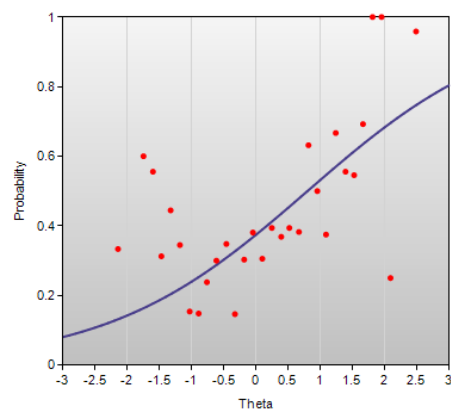


DNU

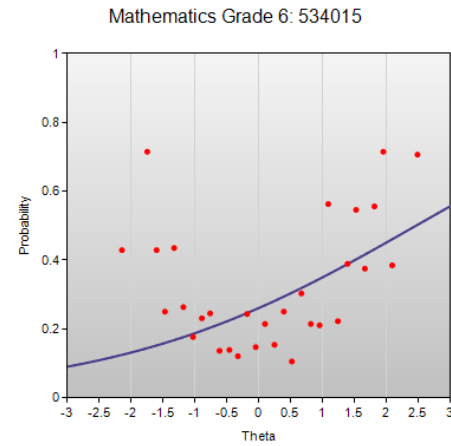
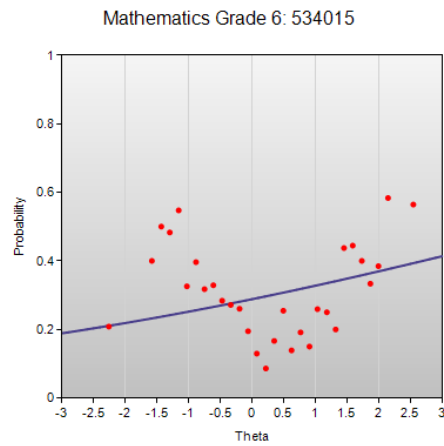
Mathematics Grade 6: 533985



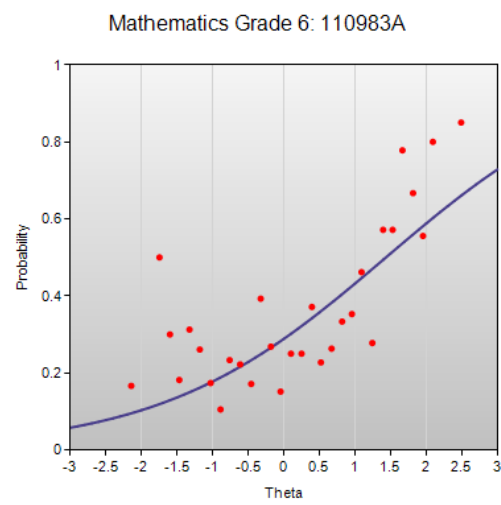
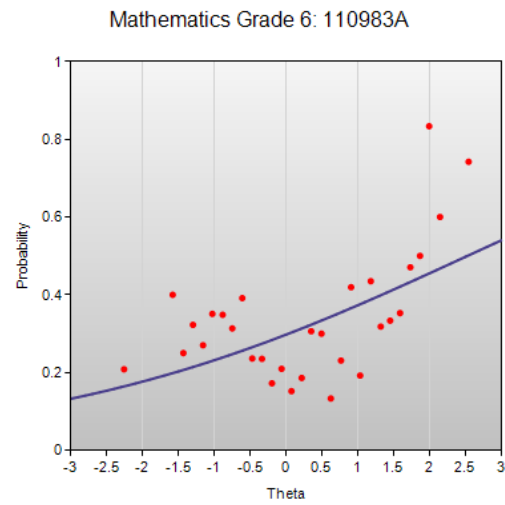
Mathematics Grade 6: 533985



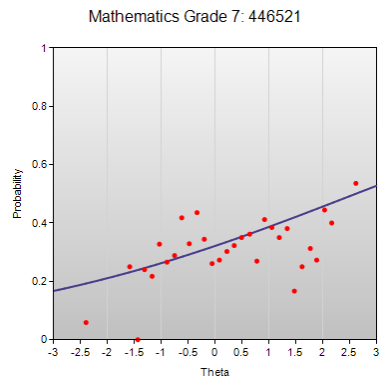
Recalibrated using 2B/2C. Went from DNU ($a=.17$, $b=1.5/se=.31$) to use ($a=.47$, $b=1.2/se.13$)



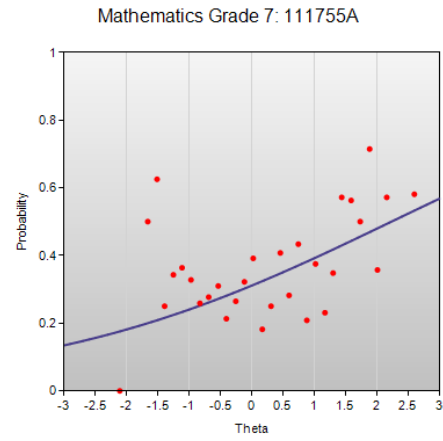
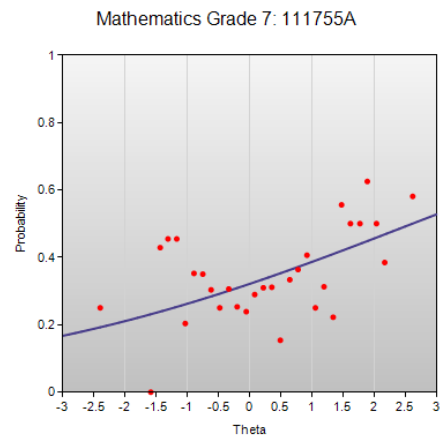
Recalibrated using 2B/2C. Still DNU (orig'l $a=.12$, $b=4.4/se.93$) (new $a=.31$, $b=2.6/se.35$)



Recalibrated using 2B/2C. Went from DNU ($a=.22$, $b=2.4/se=.40$) to use ($a=.44$, $b=1.78/se=.19$)

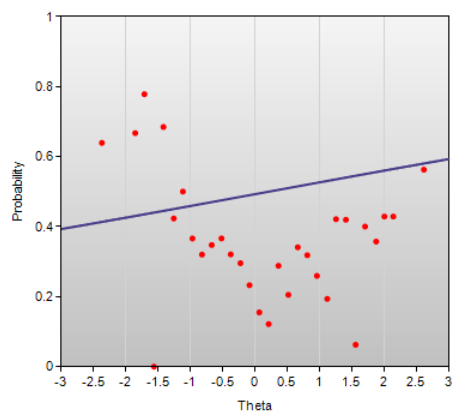


DNU

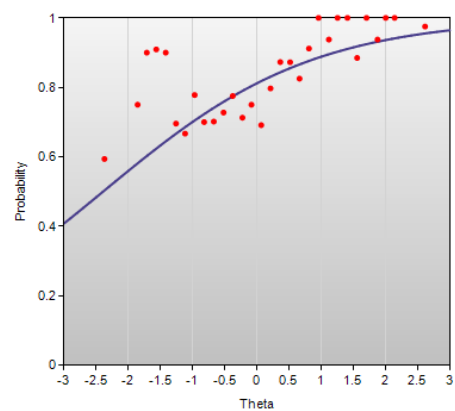


Recalibrated using 2B/2C. Still DNU (orig'l $a=.20$, $b=2.15/se.42$) (new $a=.25$, $b=1.89/se.33$)

Mathematics Grade 8: 446908



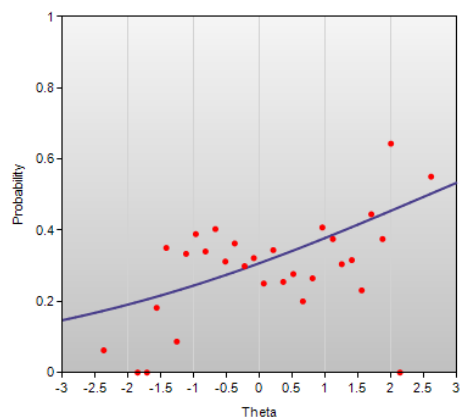
Mathematics Grade 8: 447116



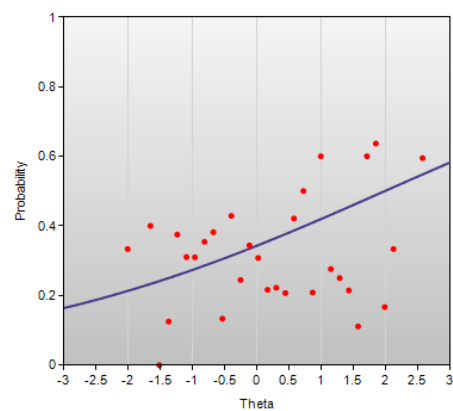
DNU

DNU

Mathematics Grade 8: 111282A

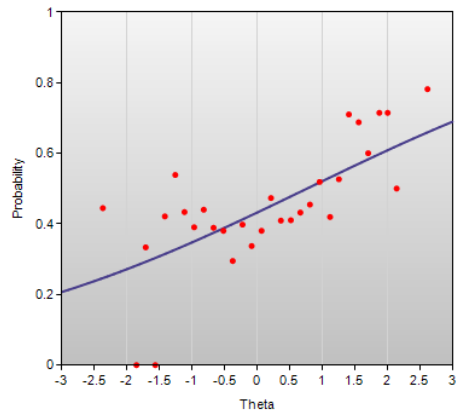


Mathematics Grade 8: 111282A

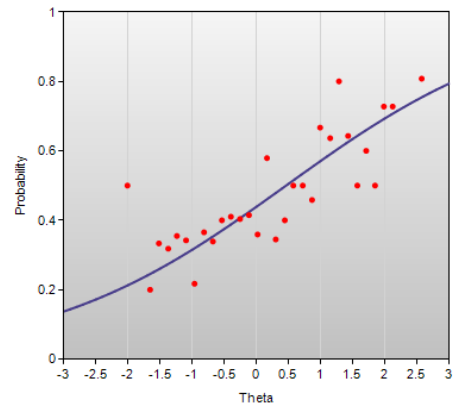


Recalibrated using 2B/2C. Still DNU (orig'l a=.19,b=2.5/se.46) (new a=.23,b=2.4/se.40)

Mathematics Grade 8: 111611A

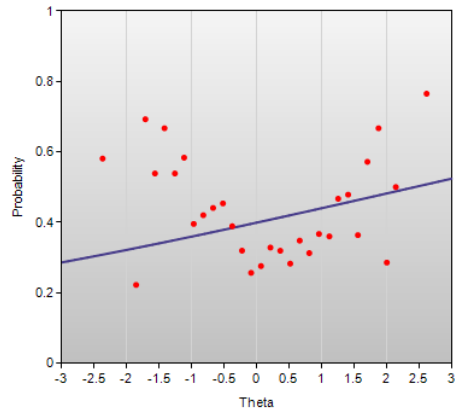


Mathematics Grade 8: 111611A

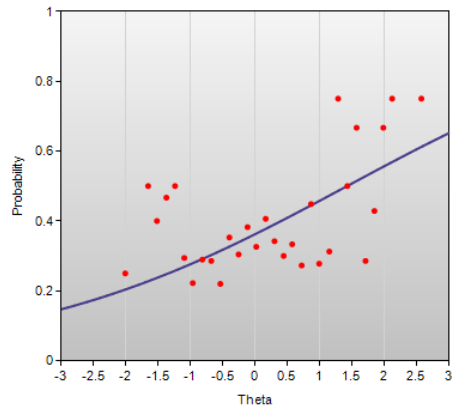


Recalibrated using 2B/2C. Went from DNU ($a=.21, b=.70/se.20$) to use ($a=.39, b=.73/se.12$)

Mathematics Grade 8: 112507A

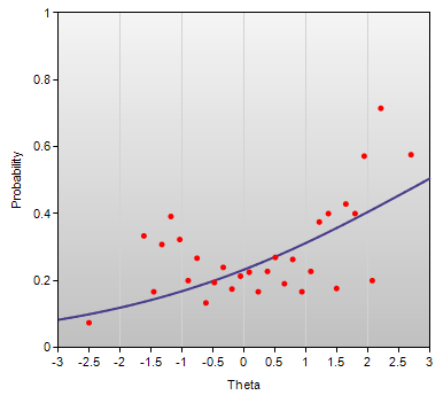


Mathematics Grade 8: 112507A

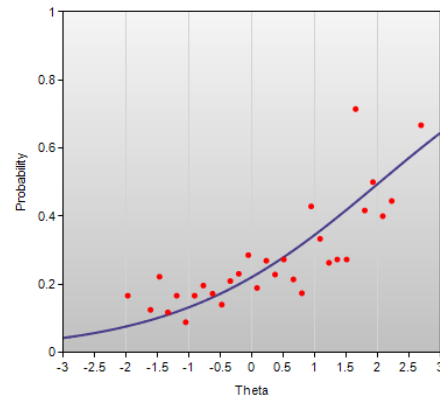


Recalibrated using 2B/2C. Went from DNU ($a=.10, b=2.25/se.62$) to UWC ($a=.31, b=1.5/se.24$)

Mathematics Grade 11: 443403

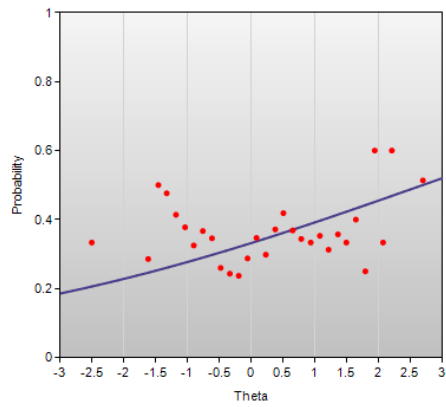


Mathematics Grade 11: 443403

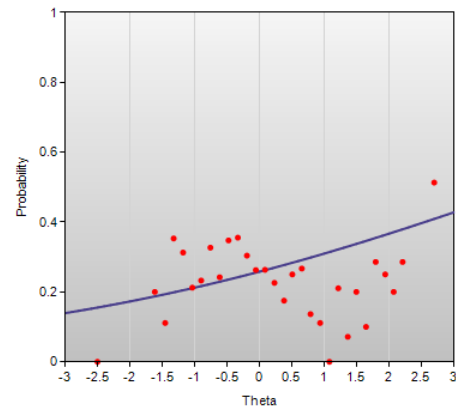


Recalibrated using 2B/2C. Went from DNU ($a=.34, b=2.08/se.34$) to UWC ($a=.50, b=1.7/se.22$)

Mathematics Grade 11: 110856A



Mathematics Grade 11: 111107A

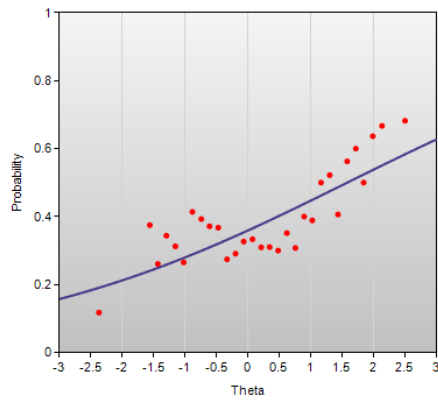


DNU

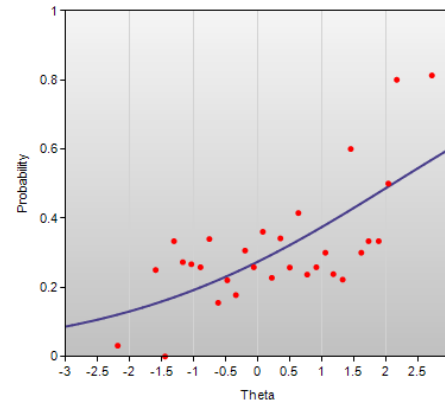
DNU

Fit Plots – Use With Caution Items

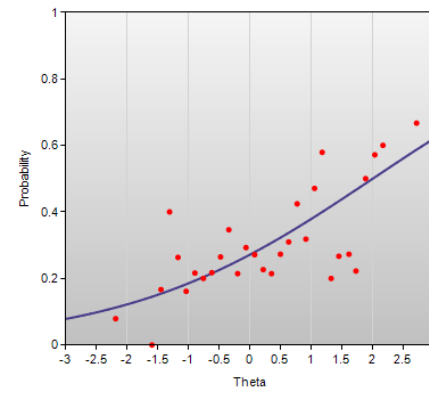
English Language Arts Grade 6: 540986



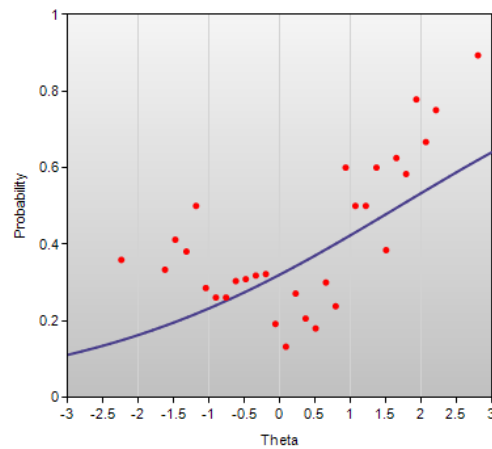
Mathematics Grade 3: 529844



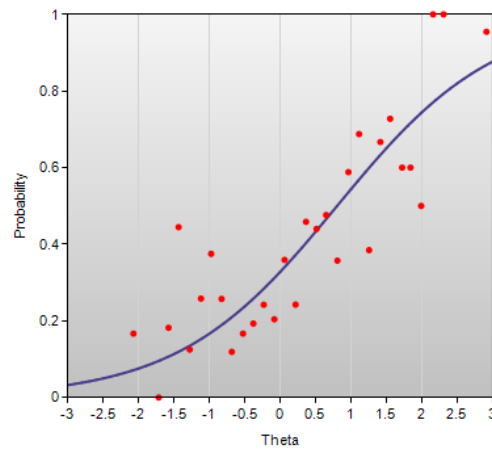
Mathematics Grade 3: 111651A



Mathematics Grade 4: 445528

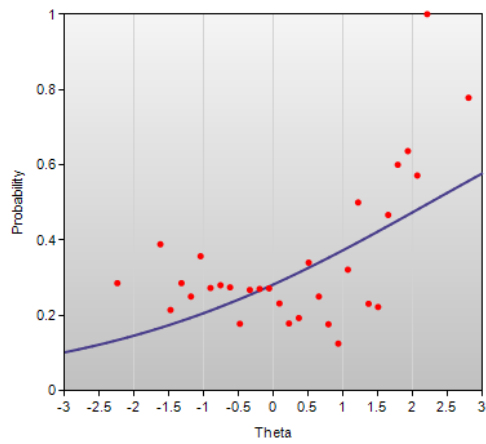


Mathematics Grade 4: 445528

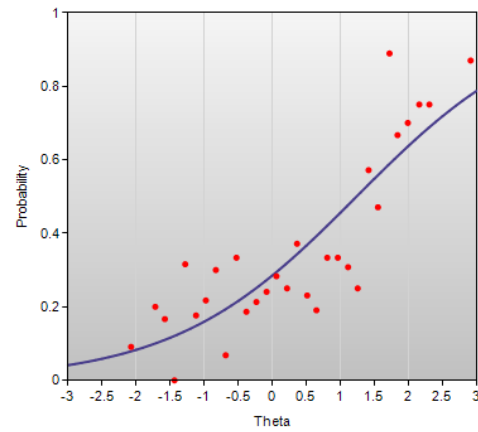


Recalibrated using 2B/2C. Went from UWC ($a=.31, b=1.43/se.23$) to use ($a=.65, b=1.02/se.11$)

Mathematics Grade 4: 111691A

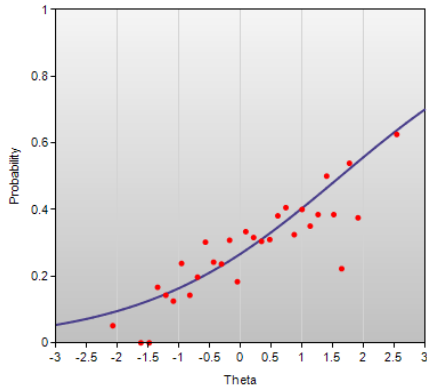


Mathematics Grade 4: 111691A

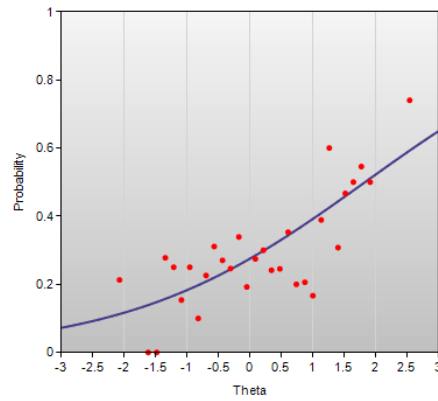


Recalibrated using 2B/2C. Went from UWC ($a=.30, b=1.84/se.28$) to use ($a=.53, b=1.4/se.15$)

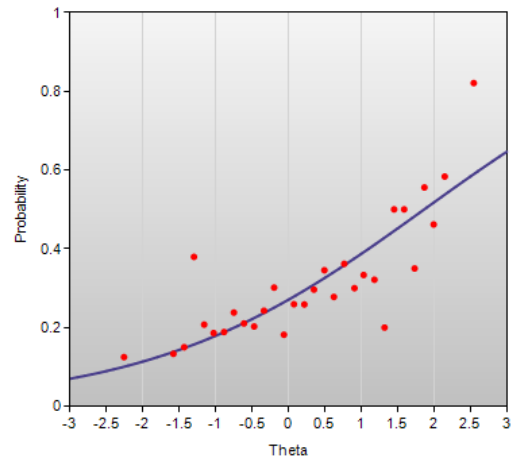
Mathematics Grade 5: 450050



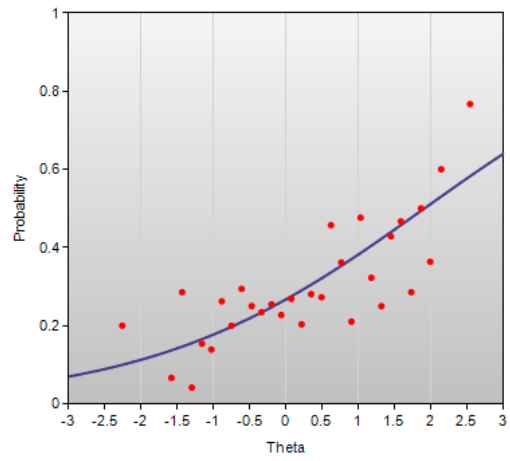
Mathematics Grade 5: 113837A



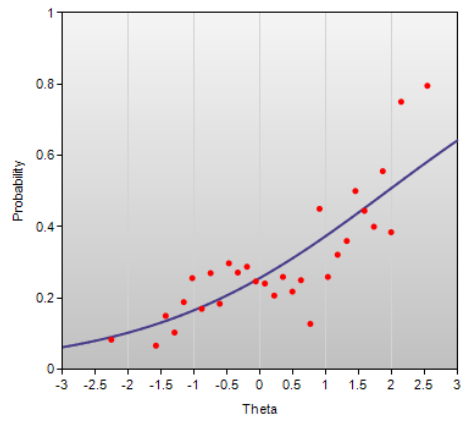
Mathematics Grade 6: 442625



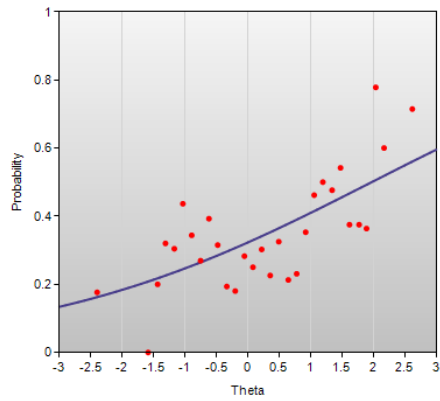
Mathematics Grade 6: 450445



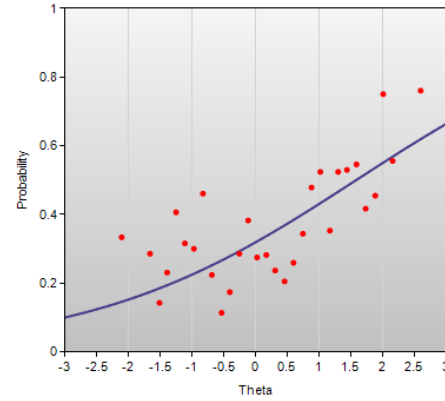
Mathematics Grade 6: 111023A



Mathematics Grade 7: 530662

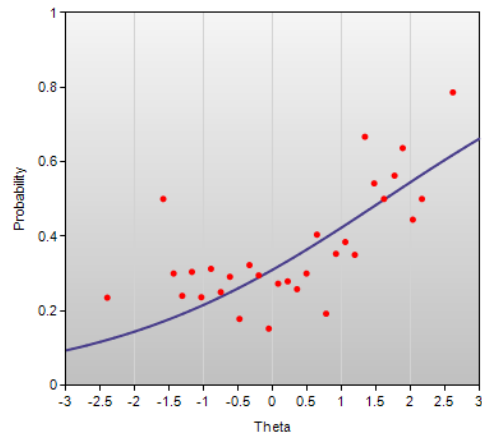


Mathematics Grade 7: 530662

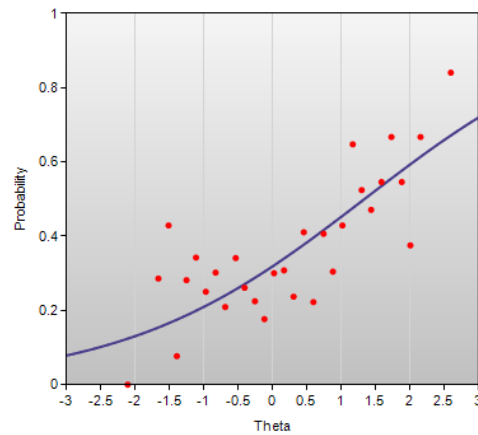


Recalibrated using 2B/2C. Still UWC (orig'l $a=.26, b=1.63/se.29$) to (new $a=.32, b=1.5/se.23$)

Mathematics Grade 7: 111081A

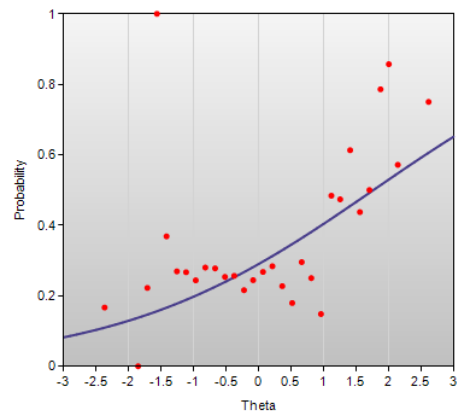


Mathematics Grade 7: 111081A

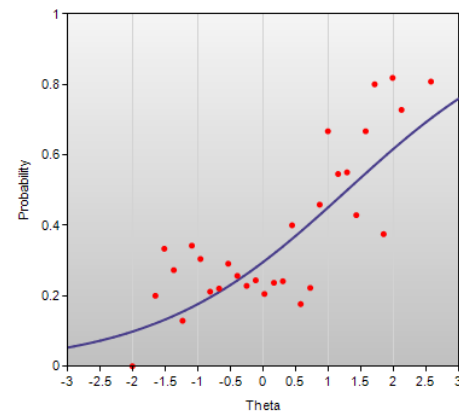


Recalibrated using 2B/2C. Went from UWC ($a=.33, b=1.37/se.21$) to use ($a=.40, b=1.26/se.17$)

Mathematics Grade 8: 113936A

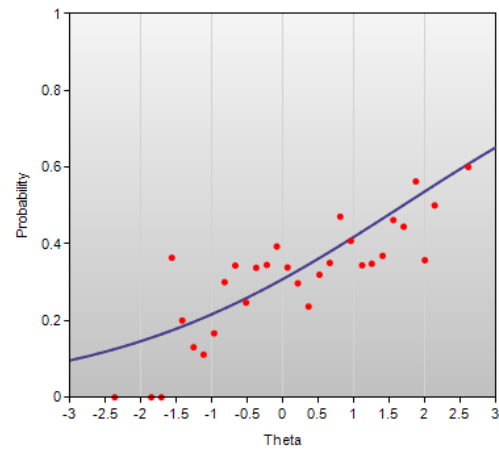


Mathematics Grade 8: 113936A



Recalibrated using 2B/2C. Went from UWC ($a=.30, b=1.67/se.23$) to use ($a=.48, b=1.37/se.15$)

Mathematics Grade 8: 113971A



MSAA Psychometrics Glossary

Term	Definition
a-parameter (Item Discrimination Index) "APARM"-	An estimate of the item discrimination (slope) parameter. The higher the a-parameter the more discrimination the item. A-parameters are typically above 0.50 and below 2.0 and values higher than 1.0 are desirable.
Biserial	A statistic that is used as an index of the extent to which an item discriminates between low and high scorers. The biserial is calculated by determining the correlation between an item with an underlying continuous distribution and a criterion measure, such as the total test score, and ranges in value from -1.0 to +1.0. Higher values, in general, are more desirable, and those with negative values are considered extremely undesirable. Indicates the extent to which the group's performance on an item is related to their performance on all other items on the test.
b-parameter (Item Difficulty Index) "BPARM"-	An estimate of the item difficulty parameter that can be interpreted similar to the p-value; however, the values typically range from -4.0 (easy) to +4.0 (hard) and values throughout this range are desirable.
CorrWTotal	Similar to Biserial (or Point Biserial), CorrWTotal describes the extent to which an item on a test differentiates between those examinees possessing much of the trait or skill being measured (high scorers) from those possessing little of the trait or skill being measured (low scorers). In Classical Test Theory, item discrimination indices generally range from -1.0 (little or no differentiation) to +1.0 (high differentiation), with higher values, in general, more desirable, and those with negative values considered extremely undesirable.
Criterion Score (Total Score)	An indication of student performance on the test overall. This is typically the raw score calculated across all operationally scored items. However, in certain instances it may be the proportion correct (i.e., when students are not all administered the same sets of items within an operational form).
Cuts (cut points)	When standards are set, the cut points are determined in terms of values of ability, or theta. These values do not change over time, although the raw scores associated with them <i>do</i> change.
"Cut Stuff"	The "Cut stuff" file gives for every item its contribution to the value of the TCC (Test Characteristic Curve) and the TIF (Test Information Function) at each of the three cut points. The "Cut stuff" file can be utilized to calculate these three cut scores. For a given set of items, the file can be used to find the contribution of each item to the TCC at each of the three cuts.

MSAA Psychometrics Glossary

DIF (Differential Item Functioning)	An item bias statistic indicating easiness of an item, favoring one student subgroup over another. Classifications with a negative sign (“–”) favor the reference group, while classifications with no sign favor the focal group. The ETS classification was used where A represents negligible DIF, B represents slight to moderate DIF, and C moderate to severe DIF. Contrasts can be made between white versus black, white versus Hispanic, and male versus female.																		
	Item Response Theory (IRT) models the response of each examinee of a given ability to each item in the test. It is based on the idea that the probability of a correct response to an item is a function of person (i.e., ability) and item parameters (i.e., difficulty) that are placed on the same underlying scale.																		
<table><tr><th>IRT parameters</th><th>Low</th><th>High</th><th>Description</th></tr><tr><td>a-parameter</td><td>< 0.25</td><td>> +3.00</td><td>Below .25 and above 3 will be flagged as outliers</td></tr><tr><td>b-parameter</td><td>< -4.00</td><td>> +4.00</td><td>Below -4.00 and above +4.00 will be flagged as outliers</td></tr><tr><td>b-standard error</td><td>N/A</td><td>> 0.30</td><td>Above .3 will be flagged as outliers</td></tr></table>				IRT parameters	Low	High	Description	a-parameter	< 0.25	> +3.00	Below .25 and above 3 will be flagged as outliers	b-parameter	< -4.00	> +4.00	Below -4.00 and above +4.00 will be flagged as outliers	b-standard error	N/A	> 0.30	Above .3 will be flagged as outliers
IRT parameters	Low	High	Description																
a-parameter	< 0.25	> +3.00	Below .25 and above 3 will be flagged as outliers																
b-parameter	< -4.00	> +4.00	Below -4.00 and above +4.00 will be flagged as outliers																
b-standard error	N/A	> 0.30	Above .3 will be flagged as outliers																
Item-Total Correlation (Item Discrimination Index)	A statistic that is used as an index of the extent to which an item discriminates between low and high scorers. The item-total correlation is a point biserial (for multiple-choice items) or a biserial (for constructed response items) and is calculated by determining the correlation between an item and the criterion score, and ranges in value from -1.0 to +1.0. Higher values, in general, are more desirable, and those with negative values are considered extremely undesirable. Indicates the extent to which the group's performance on an item is related to their performance on all other items on the test as measured by the criterion score.																		
P-Value (Item Difficulty Index)	An index that indicates an item's difficulty; calculated as the proportion of some specified group, such as grade, who answer a test item correctly. P-values range in value from 0.0 to 1.0, with lower values corresponding to more difficult items and higher values corresponding to easier items. Typically used in reference to multiple-choice items or other item types in which the item is only worth 1 point. For constructed-response items or item types in which the item is worth more than 1 point, a p-value can be estimated by dividing the resulting item mean by the maximum number of points possible for the item.																		

MSAA Psychometrics Glossary

<i>TCC (Test Characteristic Curve)</i>	The TCC tells the expected value of raw score on the test for a student who has a specified level of performance. In our case we are interested in the three levels of performance that correspond to the three performance level cuts. In particular, we want to know the raw score on the test that a student needs to achieve to be categorized into performance levels 2, 3, and 4 (i.e., TCC Cut 1/ICC Cut 1, TCC Cut 2/ICC Cut 2, TCC Cut 3/ICC Cut 3, etc.). The raw scores that determine the cut points for a given test are the scores associated through the test characteristic curve (TCC) for that test with the values of theta determined at standard setting.
<i>TIF (Test Information Function)</i>	How confident we are about the accuracy of determinations made at the cut points is described by the test information function (TIF). The TIF tells how much statistical information is being provided by the test at a specific level of performance. As with the TCC, we are interested in how much statistical information is being provided at the three performance level cuts. The greater the TIF, the greater the test reliability. The quality of inferences made at the cut points, in other words, how confident we are about the accuracy of determinations made at the cut points, is described by the test information function (TIF). Similar to TCC/ICC, corresponding information will be available for the following: TIF Cut 1/IIF Cut 1, TIF Cut 2/IIF Cut 2, TIF Cut 3/IIF Cut 3, etc.

APPENDIX D

PANELISTS AND COMMITTEE MEMBERS

Table D-1. MSAA 2023 Item Content and Bias Review Meeting & 2023 Passage Review Panelists by Grade and Content Area

ELA Content Grades 3–6		Mathematics Content Grades 3–6	
Name	State	Name	State
Rhonda Doris Camacho Gross	CNMI	Lizelle Torres Amirez	CNMI
Omar Tabb	AZ	Ashley Wilder	AZ
Nellisa Delgado	DC	Amy Dawson	AZ
Sonya Herbert	SD	Darla Stone	MT
Mary Ashes	SD	Sarah Weber	SD
Anntonelli Pola	AS	Dr. Melba A Martin	BIE
Edward Desiderio	AS	Lan Moi	AS
Racheal Ahearn	ME	Mark Dennett	ME
Abigail Trask	ME	Kelley Wallace	ME
		Zoe Milne	AZ
		Kailie Balcom	ME
		Cathrine Gilbertson	SD
ELA Content Grades 7, 8, HS		Mathematics Content Grades 7, 8, HS	
Name	State	Name	State
Tracy Lynn Del Rosario	CNMI	Brandon Bernard	AZ
Deborah Karpkala	AZ	Gloria Tuigamala	AS
Helene Cruz	GU	Becky Erickson	SD
Gaye McNeil	MT	Lynette Busch	SD
Christina Marino	DC	Andrea Bozzino	ME
Mary Fried	SD	Priscila Gomez	ME
Tagiilima Uikirifi	AS	Rebecca Dominguez	AZ
Tiffany Christman	SD	Atalina Coffin	AS
Gretchen Lehmann	BIE		
ELA Bias All Grades		Mathematics Bias All Grades	
Name	State	Name	State
Carolyn Boughton	AZ	Kimberly Aikins	AZ
Alaina Fisk	AZ	Jessica L. Wasisang-Mendez	CNMI
Thomas Vu	DC	Emma Haberman	SD
Christine Hernandez	GU	Lauren Thompson	DC
Monica Waltman	SD	Misty Favreau	ME
Senerita Kaleopa-Palauni	AS	Rucelio Vilar	BIE
Deanna Viola-Bennett	ME	Roberto Cruz	AZ
Richard Kisa Masiwemai	CNMI	Cassandra La'apui	AS
Melanie Bowlus	ME		
Jessica Fatina	AZ		
ELA Content Grades 3–8 & HS Passage Review			
Name	State		
Roberto Cruz	AZ		
Priscila Gomez	AZ		
Michael Petty	AZ		
Kristen Nash	AZ		
Margaret Dodson	AZ		
Andrew Sumner	ME		
John Reinhart	ME		
Tiffany Perry	ME		
Jestine Mayberry	DC		
Brittany Grow	DC		
Amy Hague	ME		
Elizabeth Neuts	ME		

Table D-2. MSAA 2023 Technical Advisory Committee Members

Name	Organization	Expertise
Derek Briggs	University of Colorado	• Assessment • Growth • Psychometrics
Chris Domaleski	Center for Assessment	• Accountability Systems • Psychometrics
Rachel Quenemoen	National Center on Educational Outcomes	• Students with Significant Cognitive Difficulties • NCSC Awareness
Mike Russell	Boston College	• Technology • Accessibility
Martha Thurlow	University of Minnesota/NCEO	• Special Education • Accessibility

APPENDIX E
TEACHER GUIDE—SAMPLE ITEMS

Sample Items 1 & 2		
Alignment	Core Content Connector (CCC): 3.RI.k5 Determine the main idea of a text; recount the key details and explain how they support the main idea.	
Learning Targets	Instructional Strategies	Scaffolds and Supports
I can determine the topic of an informational text presented in diverse media. I can identify a supporting detail of the topic in a text. I can identify a supporting detail in diverse media that supports the topic in the medium.	Graphic Organizer - List the topic of a text or multimedia and note events and/or details that support the topic. - Use a System of Least Prompts when selecting a supporting detail. Topic Board/Display - Identify pictures that represent the topic(s) of a given text. - Include illustrations or sentences from the text; include events and details that support the topic in a topic board/display or graphic organizer. Interactive Story Reading - Choose and pre-read a text prior to instruction. - Read the text aloud to students, stopping at predetermined points. - At each stopping point, ask students to share their thoughts and respond to text. Group Think Tell the students what the topic is prior to reading a text or watching multimedia. After reading the text, ask the students to identify sentences that tell you the topic and supporting details about the topic. Think Aloud Model the thought processes that occur while reading the text. This may include asking questions while reading the text, identifying important details, identifying the topic, and identifying the main idea.	- Pictures, objects, or tactile representations to illustrate the topic, events, or details - Sentence strips that reflect supporting details about the topic - Videos or storyboards/ cards of the story for visual supports. - Technology (e.g., interactive whiteboard, informational texts read by the computer that highlights text)

Item 1*

What is the main idea in this passage?



A. The Sun helps trees to grow big and tall.



B. People can guess a riddle and win a prize.



C. The Lantern Festival is important to families.

Would you like to read this question again, yes or no?

***Please note:** passage may be accessed in the sample items PDF and Directions for Test Administration.

APPENDIX F
REPORTING SERVICES DELIVERABLES DECISION RULES



MSAA Assessments Reporting Services Deliverables Decision Rules

2024-2025

01/25/2025

Tara LaPierre

Table of Contents

Overview	3
Partners	3
Assessment and Administrations	3
Reporting Services Deliverables List	4
Change Log	4
Pre-Test Administration Data Preparation	5
Organizational Data	5
MSAA Reporting Organizational Data Descriptions	6
MSAA ICORE Data Store	6
Test Meta Data	6
Source	6
Session Forms	6
Test Session & Position	7
Special Processing of Form Meta Data	7
Item Number	7
Item Types	7
Item Role on Test Form	7
Stage-Adaptive Requirements	7
ELA Reading and Writing Items	8
Test Administration Validation	8
Student Registration Data	8
Post-Test Administration Data Clean-up	8
Student Data Sources	8
Student Item Data	9
Student Item Response: Format	9
Student Item Response: Scoring Method	9
Item Excluded: Identify Student Modified Test Form	9
Writing Prompt: Raw Trait Dimension Scores	10
Writing Prompt: Valid Trait Dimension Score Combinations	10
Writing Prompt: Dimension Score Adjustment	10
Single-Select Choice Response: Response Adjustment	11
Student Item Attempt	11
Student Test Data	11
Student Test Status	11
Student Reporting Status (Participation Status)	12

Post-Test Administration Student Data Clean-Up	12
Post-Test Administration Psychometric Data	13
Post-Test Administration Reporting Calculations	14
<u>Student Data</u>	14
<u>Aggregate Data</u>	15
<u>Aggregation Level</u>	15
<u>Aggregation Formulas</u>	15
<u>Aggregation Suppression Rule</u>	16
Post-Test Administration Data File Deliverables	16
<u>Student Demographics Datafile (for Test Clean-up)</u>	16
<u>Organization Datafile (for Test Clean-up)</u>	16
<u>Test Materials Download Count Datafile</u>	16
<u>Writing Score Off-Topic Datafile</u>	16
<u>Billable Records Datafile</u>	17
<u>Scaled Score Lookup Datafile</u>	17
<u>Student Results School, District, State Datafile</u>	17
<u>Duplicate/Void Student Datafile</u>	18
<u>Stringer Report Datafile</u>	18
Post-Test Administration Report Deliverables	19
<u>Student Report</u>	19
<u>Report Delivery</u>	19
<u>Data Visualization</u>	19
<u>School and District Roster Report</u>	21
<u>Report Delivery</u>	21
<u>Data Visualization</u>	21
<u>School, District, and State Summary Report</u>	23
<u>Report Delivery</u>	23
<u>Data Visualization</u>	23
eMetric Data Interaction	24
Focal Point Reporting	25
Parental Rescore Request	25

Overview

This document describes the Reporting Services administration analysis and reporting requirements for the Multi-State Alternate Assessment (MSAA) administered during the **2024-2025** academic school year. For each Reporting Services responsibility, information needed to produce accurate and timely deliverables is included throughout this document.

Partners

MSAA is a consortium of Partners. Each Partner may select various analysis and reporting deliverable options. The active Partner for the current school year is included in the table below.

Partner	Partner Abbreviation
American Samoa	AS
Arizona	AZ
Bureau of Indian Education (BIE)	BI
District of Columbia	DC
Department of Defense Education Activity (DoDEA)	DD
Guam	GU
Maine	ME
Commonwealth of Northern Mariana Islands (CNMI)	MP
Montana	MT
South Dakota	SD
Vermont	VT
US Virgin Islands	VI

Assessment and Administrations

The MSAA contract consists of ELA and mathematics assessments administered during the school year to grades 03-08, and 11. As a Partner option, science may also be administered during the school year to grades 05, 08, and 11 students. Breakthrough's system will be used for registration and administration of the assessments. Student test data will be collected online only; there will be no scannable documents.

Assessment Content Area	Assessment Grade*	Brief Description	Start Date	End Date	Partner
ELA	03 - HS	Stage-Adaptive test that includes operational and embedded field test items consisting of Single Select Choice Items and Writing Prompt item types	3/10/2025	04/25/2025	All
Mathematics	03 - HS	Stage-Adaptive test that includes operational and embedded field test items consisting of Single Select Choice Items	3/10/2025	04/25/2025	All
Science	05, 08, 11	Operational Field test consisting of Single Select Choice Items	3/10/2025	04/25/2025	AS, AZ, BI, MP, GU, ME, VI, VT

(*) VT will administer the grade HS ELA and mathematics tests to grade 09 students only. All other partners administer the HS ELA and mathematics tests to grade 11 students only.



Reporting Services Deliverables List

Reporting Services will produce various data files and static report deliverables included in the table below. This document details the data preparation, processing, and formatting rules.

Post-Test Administration Deliverable	Delivery Method	File Type	Partners
Student Demographics Datafile (for Test Clean-Up)	sFTP Test Clean-up	CSV	All
Organization Datafile (for Test Clean-Up)	sFTP Test Clean-up	EXCEL	All
Test Materials Download Count	sFTP Final	EXCEL	All
Writing Score Off-Topic	sFTP Final	EXCEL	All
Billable Records Datafile (True-Up File)	sFTP Final	EXCEL	All
Scaled Score Lookup Datafiles	sFTP Final	EXCEL	All
Student Results School, District, and State Datafiles	sFTP Preliminary (State Only)	EXCEL	All
Student Results School, District, and State Datafiles	BT Online Final (School, District, State)	EXCEL	All
Student Results School, District, and State Datafiles	Focal Point/Emetric Online Final (State)	EXCEL	VT, BI
Student Results School, District, and State Datafiles	sFTP Final (State Only)	EXCEL	All
Duplicate/Void State Student Test Datafiles	sFTP Final	EXCEL	All
Stringer Report	sFTP Final	EXCEL	All
Student Report	BT Online	PDF	All
Student Report	eMetric Online	PDF	BI
Student Report	Print	PDF	BI, GU
School and District Roster Report	BT Online	PDF	All
School, District, and State Summary Report	BT Online	PDF	All
Student Report Text Lookups	FocalPoint	XLSX	VT
eMetric PDF Download	eMetric Online	CSV	BI
Parental Rescore Request	As Required	PDF/EXCEL	SD

Change Log

Administration	Description
• 2020-2021	• Datafile deliverables will be in EXCEL format instead of CSV
• 2020-2021	• ELA/Math Student Roster will be created at the district level in addition to school
• 2020-2021	• Any Partner choosing the Student Report Print Option will receive two copies of the report
• 2020-2021	• When both/all tests for a student are not launched/started but are closed (due to TA/TC misadministration) it will be reported as ESM. These tests were previously reported as DNT.
• 2020-2021	• Science will be administered to the Partners who select the science option
• 2020-2021	• Science Participation file will be created after test clean-up
• 2020-2021	• Administration window extended from 04/30/2021 to 05/14/2021 • Note: SD admin ends 05/07/2021
• 2020-2021	• Student Demographic test clean-process modified by combining bull-pen and demographic process • Partners can provide information for Cognia to add, remove, merge student data to be included in analysis and reporting • Partners can provide information for Cognia to update demographics, test status, and reporting status (participation status) • The process is outlined in the requirements document MSAA 2122 Student Demographic Instructions.pdf • Final reporting status values will be calculated as part of the demographic clean-up process and detailed in the requirements document MSAA 2122 Student Demographic Instructions.pdf
• 2020-2021	• DC does not plan to administer MSAA in 2021
• 2020-2021	• Do not print the Scaled Score Low/High sentence on the student report for students with a reporting status of ESR
• 2020-2021	• WRP Reporting Status will stay in the student results file. However, the rules will be to submit a value if different from ELA reporting status and blank if the same. Change the valid values to remove those that would not apply (remove TES, ESR, ESM, INC, ELL, EXE, DNT, WDR, and NLE). • A few supports/accommodation fields were removed from the student results layout since they no longer exist

Administration	Description
2021-2022	- Science will be operational. Two major rounds of reports and datafiles: Pre-Standard Setting and Post-Standard Setting. Pre-Standard Setting will include ELA/Math Results and science Participation as outlined in this document. Post-Standard Setting will include ELA/mathematics/science results for Partners who participated in the science assessment. - DoDEA joined MSAA - BI will have eMetric Data Interaction Reporting - Science Student Report design - School and District Student Roster re-designed such that one subject is reported on a single page. - Print Ready Student Report PDFs for Partners who opted in - Student Results Layout for ELA/mathematics/science new – modeled from ELA/mathematics layout - Added calculation rule for “Ethnic” using the individual Race/Ethnicity variables to create one Ethnic variable
2022-2023	- VT joined MSAA ELA, mathematics, and science: Testing at grade 09 instead of grade 11 for ELA and Math. The grade 09 ELA and Math tests will be analyzed as test grade 11 for psychometric analyses. - Additional partners are participating in science: AS, AZ, BI, MP, GU, ME, VI, VT - BT organization ID management /assignment and creation changed. BT Org ID should not change across years. (No impact on Reporting.) - Standards setting in science - No more Print Ready PDFs - Two new accommodations (SAR_Braille_Before and SAR_Braille_After) added to student results layout - Added Grade 09 as a valid value to the student results layout - Reporting will only use the final BT extract and ignore the initial BT extract for analysis reporting - Static Reporting: All reports - “Grade 11” will be replaced with “High School”; Remove science provisional score footnote; Student Report – parent letter edits and What to work on next text updates - Focal Point will receive VT state student results file for reporting and summary aggregation files for QC. File names for eMetric changed.
2023-2024	- TN is no longer participating in MSAA - Test Design – There will be two instead of three session 1 forms per test - DP Demographic Clean-up process – file type will be CSV instead of XLSX - Science will be reported at the same time as ELA/Mathematics - Scaled Score Low and High added to state student results layout - For completeness, Stringer Report Excel Datafile added to decision rules. No changes to the rules for the file. - ESM Test Report Status - do not report any scores
2024-2025	- Closure text updated in BT platform, no changes to BT data handoff - Student Results Layout: modified valid values description for LCI_PrimaryDisability field - ESM – student results – do not suppress student item responses in student results file

Pre-Test Administration Data Preparation

Organizational Data

Partners Cognia Operational Services department district and school data following a standardized layout. Cognia will load the data into an internal database referred to as ICORE. The requirements for district and school organizational handoff, load into ICORE, and data maintenance is out of scope for this document. However, the data will be used to support reporting assessment results. Internal use only school and district organizations are added to ICORE to support quality assurance. The fields and value descriptions used for MSAA reporting are detailed below.

MSAA Reporting Organizational Data Descriptions

Field	Field Description
ReportCode1	- Partner Abbreviation - Partner code DEMO is for internal use only
BT Org ID	Unique code assigned by the Breakthrough Portal to identify the Partners, Districts, and Schools
District Code	- Unique code (within Partner) to identify districts - District Code values of DEMOA and DEMOB are for internal use only - Length and Pattern of Values Varies
District Name	- District name used for reporting - ASCII Text field - Maximum allowable length 30
School Code	- Unique code (within Partner) when combined with District Code identifies a unique School - Schools associated with District Code values DEMOA and DEMOB are for internal use only - Length Varies and Pattern of Values
School Name	- School Name used for reporting - ASCII Text field - Maximum allowable length 30

MSAA ICORE Data Store

ICORE contract code is used to identify the set of organizational data used to support analysis and reporting.

Administration	ICORE Contract Code	Partners
Spring 2025	603400, 603452	All

Test Meta Data

The information in this section describes the test meta data needed to support data student test data validation as well as analysis and reporting activities. Test meta data includes information about tests, forms, and items being administered. Test meta data impacting analysis and reporting include Test Form ID, Test Form Session & Position, Item Number, Item Type, Item Points, Item Subject, Count Towards Student Score, Item Role on Test Form, Equating Eligible Status.

Source

NTS is the primary test meta data source support for MSAA analysis and reporting. Test meta data will be extracted from NTS after Content Development and Publications Cognia department (CDP) completes test clean up.

Session Forms

MSAA is designed to be stage adaptive. A student's score on the first session determines what form will be administered in the second session. Therefore, Forms will be constructed at the session level. Each eligible student is expected to take one form for session 1 and one form for session 2 for an assessment



content area (also referred to as test). All forms will be available in English only. Note: science is not stage-adaptive in 2024-2025.

Test Session & Position

Within the NTS data, for Stage-Adaptive tests, each form consists of one session where each session consists of a collection of items. The NTS form name includes the session. The position field indicates the order items are presented to students. Position should be unique in test form.

Special Processing of Form Meta Data

Session Form data will be used to create Test Form data by combining all possible combinations of Session 1 Form and Session 2 Forms.

Item Number

Item number (NTS AssetID) is used to support various psychometric analyses as well linking student test data to NTS data.

Item Types

Each item is characterized by its type. The item type identifies student response and score data formats. The table below lists the item types administered by MSAA. MSAA tests consist of single select choice items and a writing prompt (ELA only). Writing prompts are scored on three trait dimensions: Organization, Idea Development, and Conventions.

Item Type Label	NTS Identification	Reporting Abbreviation	CDP Abbreviation
Single-Select Choice	Interaction Type: choiceInteraction, and Correct Response: Exactly one option is the correct response	MC	SR
Writing Prompt: Scored on Three Dimensions/Traits	PointValue = 9 (Note: Each Dimension scored on 3 points)	WP	WP

Item Role on Test Form

Each item on a form is characterized as operational or field test. An item's role on a test form impacts various analyses including calculating student test scores.

Role	Abbreviation	Rule
Operational	OP	- Included in calculating student test scores - CountsTowardStudentScore = Yes
Field Test	FT	- Excluded from calculating student test scores - CountsTowardStudentScore = No

Stage-Adaptive Requirements

Reporting provides Psychometrics session 1 scaling items item lists. Psychometrics provides the routing lookups to Cognia CDP department to be incorporated in test production. Psychometrics determines the

raw scores for each session 1 form required for session 2 form assignment. Since science is not stage-adaptive in 2024-2025, routing item lists will not be produced.

ELA Reading and Writing Items

Every ELA item is assigned a Subject value of Reading or Writing in NTS. The Subject code is used for calculating Reading Percent of Points Earned and Writing Percent of Points Earned.

Test Administration Validation

Reporting participates in validating Breakthrough MSAA Testing System prior to the system going live for administration.

Student Registration Data

Student registration occurs with each Partner utilizing the Breakthrough MSAA Systems Portal. Registration requirements are outside the scope of this document. Each student will be associated with a Partner, district within the Partner, and school within the Partner in the portal.

Post-Test Administration Data Clean-up

Report Services receives data from various sources, validates the data, and applies processing rules to prepare data for psychometrics, analysis, and report generation. This section provides a general overview of the various sources and a detailed description of student item responses and scores as well as test status. In-depth detail on the data processing rules and data sources are out of scope of this document.

Student Data Sources

Student Online Test Data	BT Systems Portal
Description	- Breakthrough will provide Cognia data related to student online testing following and agreed upon schedule. - The data includes - student demographics at the time of testing, - student accommodation, - LCI data, - student response check data, - student test data including not tested reasons, student test session data, test date time stamp, student item responses item evidence, and scores, - test meta data - test proctor data - organization data - student's at the '9999' organization are excluded
General Rules	- Cognia Reporting will import and validate the files - Cognia Reporting will provide item evidence counts to Cognia Client Services for conformation that all evidence files have been received for scoring
File Layout	BT provides Cognia standard CSV files following an agreed upon format



Demographic File	Partner Updated
Description	- Partners provide an updated student demographic data file - Cognia will incorporate updates as part of post-test administration student test cleanup
General Rules	Refer to MSAA 2425 Student Demographic Instructions.pdf
File Layout	Refer to MSAA 2425 Student Demographic Instructions.pdf

Student Human Item	Scores
Description	Cognia Scoring Services will provide Reporting Services student level item scores and non-scorable scores
General Rules	Refer to section “Student Item Response: Human Score Type”
File Layout	Scoring Specifications

Student Item Data

The purpose of this section is to describe in detail the data associated with items on student tests necessary for analysis and reporting and student data clean-up activities.

Student Item Response: Format

Student item responses are captured, formatted, and stored as described below. Item type is used to categorize the response formats.

Item Type	Student Response Description	Sample Value
Single-Select Choice	Single alpha character	A
Writing Prompt	N/A	N/A

Student-Item Response: Scoring Method

Each student response to an item is assigned a score value. An item score is assigned either by machine scored or human scored. Student responses collected online are either machine scored by the testing platform or human scored.

Item Type	CBT	PBT	Scoring
Single-Select Choice	Testing Platform	N/A	Exact Match: 1 = student response match correct response; 0 otherwise
Writing Prompt	Human	N/A	Refer to sections Student Item Response: Human Score Type, Writing Prompt: Valid Dimension Score Combinations and Writing Prompt: Score Adjustment sections below

Item Excluded: Identify Student Modified Test Form

Rarely an administration issue may lead to excluding an item from a student test form during test clean-up. To exclude the item from scoring a particular student's test, the item response is set to X and score set to blank. Student test scores will be based on all core items administered to the student where the response is not X.

Writing Prompt: Raw Trait Dimension Scores

Student responses requiring a human score will have a final score of record, scorer 1, scorer 2, and scorer 3 scores as defined by scoring procedures. The final score of record value is used to calculate official student test scores and used to determine if a student attempted an item. Refer to the Writing Prompt: Score Adjustment section for more information on the writing prompt score. Scoring rubrics and procedures are out of scope for this document. Each student response requiring a human score will be assigned a final score of record score value for each rubric dimension as outlined in the table below.

Human Score	Interpretation	Raw iScore Value	Valid*	Item Attempt**
Numeric	Valid numeric score (an integer greater than or equal to 0 and less than or equal maximum allowed item score as defined in the rubric)	0,1,2,3	OP, FT	Yes
Blank	No deliberate marks in the answer space; No evidence submitted	B	OP, FT	No
Unreadable	Faint handwriting or otherwise obstructed student response	U	FT	Yes
Non-English	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	F	OP, FT	Yes
Off Topic	A response that is not related to the task/prompt administered or is not a valid attempt at answering any task/prompt on the test	5	OP, FT	Yes
Repeats the Prompt	The response copies the prompt or portions of it and offers no attempt to respond to the task/prompt	P	OP, FT	Yes
No Score	Any other response that cannot receive a numeric score	N	OP, FT	Yes
Insufficient Amount to Score	The response contains an insufficient amount of writing to score	A	N/A	Yes
Refusal	The response clearly indicates a refusal on the part of the student to address the prompt or participate in the test	R	N/A	Yes
Illegible	Tiny or poor handwriting, spelling that cannot be deciphered, or other conditions that render the student work indecipherable	I	N/A	Yes
Wrong Location	Item response inconsistent with student form	W	N/A	Yes
Response Not Scored	Field test item where students' response was not selected for scoring	# or blank	FT	Unknown

(*) Valid: OP = Human score value is valid for operational items

FT = Human score value is valid for field test items

N/A = Not applicable for project. If value is provided, resolution is needed.

Note: In 2024-2025, all Writing Prompts are OP.

(**) Item Attempt: Yes = Human score value indicates student attempted the item

No = Human score value indicates student did not attempt the item

Unknown = Not enough information to determine if the student attempted the item

Writing Prompt: Valid Trait Dimension Score Combinations

Writing prompts are scored on three trait dimensions: Organization, Idea Development, and Conventions. Each trait is assigned a score listed in the "Raw Score Value" column in "Writing Prompt: Raw Trait Dimension Scores." Off Topic is not a valid score for the Conventions trait. If one dimension score is scored a B, then all dimension scores must be a B.

Writing Prompt: Dimension Score Adjustment

The raw iScore dimension score values are translated as indicated below to support analysis and reporting requirements. During test cleanup, the raw iScore value is translated to the Student Results



value except Z will be set to B to be consistent with standard processes. “B” will be translated to “Z” when producing the student results and void/duplicate files.

Human Score	Raw iScore Value	Psychometric Score Value	Student Results
Rubric Score	0	0	0
Rubric Score	1	1	1
Rubric Score	2	1	1
Rubric Score	3	2	2
Blank	B	0	Z
Unreadable	U	N/A	U
Non-English	F	0	F
Off Topic	5	0	O
Repeats the Prompt	P	0	P
No Score	N	0	N
Item Excluded: Identify Student Modified Test Form during Clean Up	0-3,5, B, U, F, P, N	N/A	X

Single-Select Choice Response: Response Adjustment

Student responses to single-select choice items are translated below to support analysis and reporting.

Raw Response	Raw Value	Psychometric Score Value	Student Results
Raw Response	A, B, C, or D	0 = response does not match item key	A, B, C or D
Raw Response	+	1 = response matches item key	+
Raw Response	blank	0	Z
Item Excluded: Identify Student Modified Test Form during Clean Up	A, B, C, D, or blank	N/A	X

Student Item Attempt

Item Type	Item Attempt Rule
Single-Select Choice	If the student raw response is not blank or X, the student attempted the item
Writing Prompt	If the student's earned score value for one or more dimensions is listed as a "Yes" in "Item Attempt" column in "Writing Prompt: Raw Trait Dimension Scores" table, the student attempted the item.

Student Test Data

Test data applies at the ELA, mathematics, and science levels. Science test data will only exist for Partners who selected the option to administer the science test. The purpose of this section is to describe in detail the data associated with student tests necessary for analysis and reporting and student data Clean-Up activities.

Student Test Status

Each student test is assigned a test status in the Breakthrough Portal and adjusted during student data Clean-Up when necessary. This field will be updated during demographic clean-up.

Final Test Status	Condition
InProgress	- BT Portal value Paused value is changed to InProgress during test Clean-Up - Provided by field using BT Portal
Cancelled	- Provided by field using BT Portal - Canceled test status is also referred to as Closed Tests
Completed	- Provided by field using BT Portal - Completed test status value is also referred to as Submitted
[Blank]	- Final Test Status will be blank for science if a Partner does not participate in science - Final Test Status will be blank for students who were added during demographic clean up

Student Reporting Status (Participation Status)

Each student is assigned an ELA Reporting Status, Mathematics Reporting Status, a Writing Reporting Status, and Science Reporting Status during test cleanup. The allowed values are detailed in the table below. If a partner does not participate in science, the Science Reporting Status will be blank. The rules for assigning the final reporting status are out of scope of this document. Refer to student demographic clean-up instructions for reporting status assignment rules.

Test Reporting Status	Code	Description
Administration Irregularity	IRR	Administration irregularity reported, but does not necessitate an invalidation
Invalidated	INV	Student-based or administration-based irregularity resulting in invalidation
Parental Refusal	PRF	Parental refusal
ELL Exempt (ELA Only)	ELL	Student meets the ELA ELL 1 st Year in U.S. exemption requirements
Exempt	EXE	Student meets test exemption requirements
Withdrew	WDR	Student withdrew
No Longer Eligible	NLE	Student is no longer eligible for testing
Tested	TES	Submitted test, regardless of number of item responses
Tested-Incomplete	INC	In-Progress Test, with at least one item response
Early Stopping Rule	ESR	Closed Test – with no item response
Early Stopping Rule – Misadministration	ESM	Closed Test – with at least one item response Closed Test – both/all content area tests not launched or started
Did Not Test	DNT	No Test, or In-Progress Test with no item response

Post-Test Administration Student Data Clean-Up

Various data sources, including Test Meta Data, Organization Data, Online Student Test Data, Scores for Human scored items, and Demographic Clean-Up are used to conduct student data clean-up to produce student test data ready for analysis and reporting. The table below describes relevant details related to the clean-up process and requirements.

Data	Guidelines
General Information	- Cognia will update student data using the updated demographic files returned by each Partner as outlined in the Demographic Clean-Up Instructions for additional details - Updates include modifying demographic, test status, preliminary reporting (participation status), item responses/scores data as well as adding and removing student tests - After the updates are incorporated, Cognia will perform additional clean up as outlined below
Organization Data	- All student test records associated with the same student ID must have the same School, District, and State - State, District, and School codes associated with student tests must exist in ICORE and Breakthrough Organization file. - New or revised Organization data will be updated in both ICORE and Breakthrough reporting platforms - Cognia will work with Partners to identify the complete set of schools and district organizations, along with the names for reporting, during the demographic file acceptance and organization Clean-Up process with each Partner
Student Test Grade	- Test grade is expected to match Student Enrolled Grade. - If a student's enrolled grade level is provided in the final demographic data does not match the student's tested grade, the test is considered off-grade and will be marked as "Void/Duplicate"
Duplicate Test	- After Off-Grade tests have been resolved, duplicate tests are tests in the same Assessed Content Area and State Student ID within a State Partner - The final test used for analysis and reporting is determined used the following hierarchy - Submitted/Completed - Closed - In Progress - If two or more tests have the same status, the test associated with the latest date will be used, determined by the datetime stamp of the test record. Additionally, the larger TestID is used if still duplicate. - The duplicate test(s) not selected for analysis and reporting will marked as "Void/Duplicate"
Student Test Status	Final ELA, Mathematics, and Science Test Status will be audited based on MSAA 2425 Student Demographic Instructions.pdf
Student Test Reporting Status	Final ELA, Mathematics, and Science Test Reporting Status (Participation Status) will be calculated based on MSAA 2425 Student Demographic Instructions.pdf
Student Writing Prompt Reporting Status	Final Writing Prompt Reporting Status (Participation Status) will be calculated based on MSAA 2425 Student Demographic Instructions.pdf
Student Test Item Responses	Item responses could be removed based on Student Test Reporting Status as detailed in the demographic clean up instructions
Ethnic	- For DIF and eMetric DI the algorithm below is applied to assign one Ethnic value as follows: - If Hispanic is indicated, then "Hispanic" - Else, if DemographicRaceTwoOrMoreRaces is indicated then "Multi" - Else if AmericanIndianOrAlaskaNative is indicated then "AIAN" - Else if Asian is indicated then "Asian" - Else if BlackorAfricanAmerican is indicated, then "BAA" - Else if NativeHawaiianOthPacificIslander is indicated, then "NHOPi," - Else if White is indicated, then "White"

Post-Test Administration

Psychometric Data

Reporting Services will provide Cognia Psychometric team test meta data and student test administration data consisting of demographics, student test status, student test form, and student item level responses and scores. Psychometrics will conduct statistical key checks, Stringer Analyses, CTT, and IRT. The specifications for such activities are out of scope for this document. Psychometrics will provide Reporting Services pre-equated test scaling information and raw-score, to scaled-score lookup tables as described in this section to support creation of data file and report deliverables.

Psychometrics Assigned Scores

ELA Cut Scores by Test Grade	• Proficiency Level Scale Score Ranges
ELA Scaled Score Lookup by Test Grade	• Scale form • Raw Score • Scale Score • Proficiency Level • Scale Score Low/High
Mathematics Cut Scores by Test Grade	• Proficiency Level Scale Score Ranges
Mathematics Scaled Score Lookup by Test Grade	• Scale form • Raw Score • Scale Score • Proficiency Level • Scale Score Low/High
Science Cut Scores by Test Grade	• Proficiency Level Scale Score Ranges
Science Scaled Score Lookup by Test Grade	• Scale form • Raw Score • Scale Score • Proficiency Level

Post-Test Administration Reporting Calculations

This section details calculations and formatting applied after test clean-up is complete.

Student Data

The data listed below details student level data used to support various analysis and reporting tasks. It does not include a complete list of student data fields available. Student data prepared for psychometrics is merged with student scores calculated by psychometrics. [Test] Refers to ELA, mathematics, and science tests. Science test fields will be blank for Partners who did not participate in science.

Field	Description
[Test] Form	• Two letter test form identification where the first letter identifies the session1 form and the second letter identifies the session 2 form • Students without a test form who need to be reported are defaulted to form AA or 01
[Test] Scale Form	• Identifies the unique set of scaling and equating items based on Test Form and "Item Excluded: Identify Student Modified Test Form during Clean Up"
[Test] Form Modified	• If during test clean up the student test was identified as "Item Excluded: Identify Student Modified Test Form during Clean Up" the field will be set to a "1"; otherwise, it will be "0"
[Test] Raw Score	• Sum of final non-flawed item scores classified as "counts toward student score" items for the student test
[Test] Scaled Score	• Using calculated Test Scale Form, Test Raw Score and Psychometric Raw Score to scale score lookup, assign a Test Scaled Score • Apply Reporting Status test score rules as appropriate for a specific deliverable
[Test] Performance Level	• Using calculated Test Scale Form, Test Raw Score and Psychometric Raw Score to scale score lookup, assign a Test Performance Level • Apply Reporting Status test score rules as appropriate for a specific deliverable
[Test] Scaled Score Low/High	• Using calculated Test Scale Form, Test Raw Score and Psychometric Raw Score to scale score lookup, assign a Test Scaled Score Low/High • Apply Reporting Status test score rules as appropriate for a specific deliverable

Field	Description
[Test] State Compare	- Calculate by comparing the student's [test] scaled score with the state average scaled score and the student's scaled score SEM - Below (-): state average scaled score – student's scaled score SEM > student's scaled Score - At (=): state average scaled score – student's scaled score SEM ≤ student's scaled Score ≤ state average scaled score + student's scaled score SEM - Above (+): < student's scaled Score > state average scaled score + student's scaled score SEM
[Test] Item Score String	- Test Item Score/Response String - Apply Reporting Status test score rules formatting as appropriate for a specific deliverable - Each column in the string represents a core item (counts toward student score) - Selected Response: + = Correct Response - A,B,C,D = Incorrect Response - Z = No Response - X = Item Excluded from Student's form - Writing Prompt: 0,1,2 = Response Score - Z (blank), F (Foreign Language), P (Copy of Prompt), N (No Score), O (Off Topic) Non-Scorable Codes - X = Item Excluded from Student's form
[Test] Field Item	If at least one field test item is attempted on the test then "1", otherwise "0"
ELA Reading Percent of Points Earned	- Percentage of possible points correct for reading items - Values: 0-100, N/A - Apply Reporting Status test score formatting rules as appropriate for a specific deliverable - Include all core items administered to the student
ELA Writing Percent of Points Earned	- Percentage of possible points correct for writing items - Values: 0-100, N/A - Apply Reporting Status test score rules as appropriate for a specific deliverable - Include all core items administered to the student
WR Trait Scores	- Student level writing trait scores are included part of overall ELA test - Apply Reporting Status test score formatting rules as appropriate for a specific deliverable - Refer to Writing Prompt: Dimension Score Adjustment table Student results column 0,1,2 = Response Score - Z (blank), F (Foreign Language), P (Copy of Prompt), N (No Score), O (Off Topic) Non-Scorable Codes

Aggregate Data

Aggregation Level

Each student is assigned one State, District, and School code to use for aggregations as described in the table below.

Aggregation Organizational Level	Aggregation Code
State	Partner Abbreviation
District	Combined Partner Abbreviation and District Code
School	Combined Partner Abbreviation, District and School Code

Aggregation Formulas

The aggregations below are calculated to support various datafiles and reports. The calculations are aggregated by state, school, and district. *Student tests identified as Void/Duplicate or Remove are excluded from all aggregations.*



Aggregation	Calculation
Number Enrolled	Number of student tests that have at least one test assigned one of the final reporting status values other than WDR or NLE for the aggregation level
Number Tested	Number of student tests assigned TES, ESR, or IRR final reporting status for the aggregation level
Number of Did Not Test	Number Enrolled minus Number Tested
Average Scale Score	Average test scale score for students included in the "Number Tested" aggregation rounded to the nearest whole number for the aggregation level
Number of Students at each Performance Level	Number of student tests included in the "Number Tested" count with the specific Performance Level Value for the aggregation level
Percent of Students at each Performance Level	Divide the "Number of Students at each Performance Level" by the Number Tested for the aggregation level. Multiply by 100 and round to the nearest whole number.

Aggregation Suppression Rule

Aggregations with less than 10 students included in the denominator will be suppressed from state level reports only.

Post-Test Administration Data File Deliverables

Student Demographics Datafile (for Test Clean-up)

Description	Cognia provides each participating Partner an excel file containing raw student data to support data cleanup
Generation Rules	Refer to MSAA 2425 Student Demographic Instructions.pdf
File Layout	Refer to MSAA 2425 Student Demographic Instructions.pdf
File Name	Refer to MSAA 2425 Student Demographic Instructions.pdf

Organization Datafile (for Test Clean-up)

Description	Cognia provides each participating Partner an excel file containing organization data to support data cleanup
Generation Rules	ICORE organization data are used directly to create the file as detailed in the layout
File Layout	MSAA2425OrgDataLayout.xlsx
File Name	MSAA2425_ICORE_[state abbreviation].xlsx

Test Materials Download Count Datafile

Description	- Breakthrough provides test materials data table - Cognia uses the data table to create a data file for each state containing the relevant state data
File Name	MSAA2425_tblFilddownloads_[state abbreviation].xlsx

Writing Score Off-Topic Datafile

Description	The writing off-topic datafile lists students and their writing prompt trait scores.
Generation Rules	Raw ISCORE scores are provided in the file except Off Topic is O and B, F, N are translated to 0
File Layout	State, DistrictCode, SchoolCode, DistrictName, SchoolName, Lname, Fname, StateStudentID, Grade, ItemNumber, Trait1score, Trait2Score, Trait3Score
File Name	WritingDelivareble-[state abbreviaton].xlsx



Billable Records Datafile

Description	MSAA States shall be billed out based on record results. Billable results shall be delivered to Cognia's Finance Department for true up and final billing.
Generation Rules	- Each tested student is considered a billable record - Each student test shall be considered a valid billable record when a test is launched and In Progress, Closed or Submitted. - A billable record does not include where a student does not have a test record, is no longer enrolled or is withdrawn - Records with a blank nap_delivery_id will be highlighted - The datafile will include two tabs: one for Reporting records (included in results datafile) and Not Reported Records (included in Duplicate/Void datafile) - The records will be reported in the file with their SSID
File Name	Billing_[state abbreviation].xlsx

Scaled Score Lookup Datafile

Description	The rawscore to scaled score lookup will be created and provided as an option to Partners
Generation Rules	- One EXCEL file for each Test Subject will be created containing the psychometric raw score to scale score lookup data - Each EXCEL file will contain a worksheet for each test grade
File Layout	Each worksheet will contain columns: Grade, Subject, ScaleForm, RawScore, ScaledScore, LowScaledScore, HighScaledScore, and PerfLevel
File Name	- MSAA2425ScaledScoreLookups_mat.xlsx - MSAA2425ScaledScoreLookups_ela.xlsx - MSAA2425ScaledScoreLookups_sci.xlsx

Student Results School, District, State Datafile

Description	- The student results data file will contain all data for student tests not identified as Void/Duplicate during test Clean-Up as well as students tests added during test clean-up following the file layout - State files will be produced and provided on the sFTP - State, District, and School files will be provided to Breakthrough
Generation Rules	- The student results data file is sliced by state, district, and school. - Student tests are included in the specific version of the file based on the Aggregation Organization Level of State, District, and School assignment rules - Refer to table "Final Report Status Formatting of Student Scores table" - The file layout defines each field and valid values - The file will be exported to EXCEL.
File Layout	- MSAA2425StudentResultsLayout.xlsx - District and School files will contain a subset of variables as indicated in the layout "District, School Files" column - For Partners not participating in science: worksheet StuResults_ELAMat will be used to generate the files - For Partners participating in science: worksheet StuResults_ELAMATSCI will be used to generate the files
Preliminary State File Name	2025_[Partner abbreviation]_PreliminaryStudentResults.xlsx
State, eMetric, Focal Point File Name	2025_[Partner abbreviation]_StateStudentResults.xlsx
BT State File Name	2025_[BT Org ID]_StateStudentResults.xlsx
BT District File Name	2025_[BT Org ID]_DistrictStudentResults.xlsx
BT School File Name	2025_[BT Org ID]_SchoolStudentResults.xlsx

Student Results Datafile: Final Reporting Status Formatting of Student Scores

Final Test Reporting Status	Code	State File All Scores*	District & School File Scaled Score	District & School File Perf Level	District & School File R/W Percent
Administration Irregularity	IRR	Yes	Yes	Yes	Yes: 0-100
Invalidated	INV	Yes	No	No	N/A
Parental Refusal	PRF	No	No	No	No
ELL Exempt (ELA Only)	ELL	No	No	No	N/A
Exempt	EXE	No	No	No	N/A
Withdrew	WDR	No	No	No	N/A
No Longer Eligible	NLE	No	No	No	N/A
Tested	TES	Yes	Yes	Yes	Yes: 0-100
Tested-Incomplete	INC	Yes	Yes	No	Yes: 0-100
Early Stopping Rule	ESR	Yes	Yes	Yes	N/A
Early Stopping Rule – Misadministration	ESM	No**	No	No	N/A
Did Not Test	DNT	No	No	No	N/A

(*) All Scores: State student results file includes item responses, WP trait scores, raw scores, scaled scores, and performance levels.

(**) Do not suppress item responses.

- Yes = Include score in data file; No = Leave column blank in data file; N/A = Put N/A in the data file

Duplicate/Void Student Datafile

Description	• The file contains the student tests identified as Void/Duplicate, including Off-Grade test records during test Clean-Up process • Data within the datafile shall be interpreted with caution since minimal Clean-Up has been applied
Generation Rules	• The file will follow the same layout and rules as the ELA/mathematics student results file, except only include student tests identified as Void/Dup • A file will be created for each Partner if there is at least one student test identified as Void/Dup • The file will be exported to EXCEL.
File Layout	• MSAA2425StudentResultsLayout.xlsx • District and School files will contain a subset of variables as indicated in the layout “District, School Files” column • For Partners not participating in science: worksheet StuResults_ELAMat will be used to generate the files • For Partners participating in science: worksheet StuResults_ELAMATSCI will be used to generate the files
State File Name	• 2025_[Partner abbreviation]_VoidDupResults.xlsx

Stringer Report Datafile

Description	• The file contains the student tests identified Stringers by Psychometrics for at least one test
Generation Rules	• The file will contain one row per student if they were identified as Stringer for at least one test. • The file will be exported to EXCEL.
File Layout	• The file will contain fields: State Barcode LastName FirstName disCode schCode DistrictID SchoolID State_Student_ID Grade ELA_Stringer MAT_Stringer SCI_Stringer • [Test]_Stringer column(s) will have a value of ‘yes’ if the student is identified as stringer
State File Name	• MSAA2425StringerReport_[Partner abbreviation].xlsx



Post-Test Administration Report Deliverables

Student Report

Report Delivery

- Students who have an ELA or mathematics final reporting status of TES, ESR, or IRR will receive an ELA/mathematics Student Report.
- Students who have a science final reporting status of TES, ESR, or IRR will receive a Science Student Report

Print

- Only Partners who selected the Print option will receive two printed copies of the student report.
- A print report package will be created by the school.
- Slip sheets will be created at the start of each new report pack. The slip sheet identifies the appropriate shipping information and provides a way to track the secure shipment.
- ELA/mathematics Student Reports will be printed and shipped.
- Science Reports will be printed and shipped.
- Printed student reports will be gray-scaled.

Online

- A PDF will be generated for each Partner and school containing all student reports for the school regardless of test grade.
- Student reports will be sorted by Test Grade, Student Last Name, Student First Name, and Student ID.
- The PDFs will contain ELA/mathematics Student Reports. For Partners participating in science, the online PDFs will include Science Student Reports at the end of each PDF.
- Online student reports will be in color.

Data Visualization

This section details the data visualizations for the ELA/Mathematics and Science Student Report. Each ELA/mathematics student report is a two-page report (front and back). The ELA/mathematics report is designed to display both ELA and mathematics results side by side. The Science Student Report is a two-page report (front and back) one- subject report. The front page of every student report is noted as "Confidential."

- Print Student First name possessive, when appropriate. Throughout the student report, the student's first name appears embedded in text, it will appear as is or modified to be possessive as follows:
 - If student first name ends in 's' append apostrophe to student first name

- Otherwise, print [Student First Name]'s in section introduction sentence
- First Page Header
 - Name: [Student First Name] [Student Last Name]
 - ID: [State Student ID]
 - School: Print School Name
 - Test Date: Spring [Year] (example: Spring 2025)
 - Grade: if Test Grade 11 then High School else [Two-Digit Test Grade]
- First Page Performance Summary

Format Performance Summary section based on the student's final test reporting status as detailed in the table below.

Test Final Reporting Status	Test Result Section	Visualization
TES or IRR	Performance Level	• Print formatted earned student performance level ○ Level 1 ○ Level 2 ○ Level 3 ○ Level 4
TES or IRR	Score	• Print the student earned scaled score
TES or IRR	Score Graphic	• Place arrow in the relative location of the graphic for the student's scaled score with score printed above the arrow • Print scaled score ranges in each performance level
TES or IRR	Score Low/High	• Print the student's lower and upper scaled score
TES or IRR	R/W Percent of Points Earned (ELA Only)	• Print the student's earned percent of points
ESR	Performance Level	• Print formatted student performance with an asterisk • Level 1*
ESR	Score	• Print the student scaled score provided by psychometrics. It is expected to be 1200.
ESR	Score Graphic	• Place arrow in the relative location of the graphic for the student's scaled score with score printed above the arrow • Print scaled score ranges in each performance level
ESR	Score Low/High	• Leave blank. Do not print the Low/High Scaled Score sentence.
ESR	R/W Percent of Points Earned (ELA Only)	• Print N/A
All Other Values	Leave blank under the Test Subject header except print the note:	• Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.

- First Page Performance Level Descriptors

Format Performance Level Descriptors section based on the student's final test reporting status detailed in the table below.

Test Final Reporting Status	Visualization
TES or IRR	• Print formatted performance level descriptors based on student test grade, test subject, and earned student performance level • The performance level descriptors were provided to Reporting during report design after standard setting. The text is carry forward from year to year. • Each statement starts with a checkmark
ESR	• Print the text under the Test Header: * Your child did not show an observable response mode during the test; therefore, the test was not administered by the teacher. If you have additional questions, please contact your child's teacher.
Other	• Leave section under Test header blank

- First Page Footer
 - Left Justified: Copyright information
 - Right justified: Page 1
- Second Page Header
 - 2025 Results for [Student First Name] [Student Last Name] ([State ID]) | “High School” or Grade [2-digit test grade] | [School Name]
 - Example: 2025 for Jane Smith (12345678) | Grade 04 | Demonstration School A
 - Example: 2025 for Jane Smith (12345678) | High School | Demonstration School A
- Second Page Letter to Parents and Guardians
 - Letter is provided by the Partner and one letter for all ELA/mathematics Student Reports
 - Letter is provided by the Partner and one letter for all Science Student Reports
- Second Page: What skills can be worked on next?

Format “What skills can be worked on next?” section based on the student’s final test reporting status as detailed in the table below.

Test Final Reporting Status	Visualization
TES or IRR	• Print the specific skills text provided during report design based on the students test grade and subject • Each statement starts with a plus symbol
ESR	• Print the text under the Test Header. • Revisit IEP communication goals in collaboration with the speech language pathologist, AT specialist, and others who assist the student in developing a consistent mode of communication.
Other	• Leave section under Test header blank

- Second Page What now?
 - Print the questions and suggestions developed during report design with student’s first name embedded in the statements and questions
- Second Page Footer
 - Left Justified: Copyright information
 - Right justified: Page 1

School and District Roster Report

Report Delivery

- A School Roster Report will be produced when a school has at least one student assigned an ELA, mathematics, or science reporting status value other than WDR or NLE.
- A District Roster Report will be produced when a district has at least one student assigned an ELA, mathematics, or science reporting status value other than WDR or NLE.
- Static PDFs will be generated to be posted online. The report is not printed.

Data Visualization

This section details the data visualizations for the School and District Roster Report.



- **District Roster Report**

- Header
 - Print: CONFIDENTIAL
 - Print: [Formatted State Name]
 - Print: [Formatted District Name]
 - Print: If test grade =11 then High School else Grade [Two Digit Test Grade]
- Summary Data Rows:
 - Each row will contain the state and district aggregated test results
 - Do not suppress aggregations
- Student Roster
 - Header: Spring 2025
 - Student Name [Student Last Name], [Student First Name]
 - Student ID [State Student ID]
 - Test Status Impact on Report of Student Test Results

Final Test Reporting Status	Code	Print Test Status	Print State Compare	Print Scale Score	Print Performance Level
Administration Irregularity	IRR	Yes	Yes	Yes	Yes
Invalidated	INV	Yes	No	No	No
Parental Refusal	PRF	Yes	No	No	No
ELL Exempt (ELA Only)	ELL	Yes	No	No	No
Exempt	EXE	Yes	No	No	No
Withdrew	WDR	Yes	No	No	No
No Longer Eligible	NLE	Yes	No	No	No
Tested	TES	No	Yes	Yes	Yes
Tested-Incomplete	INC	Yes	Yes	Yes	No
Early Stopping Rule	ESR	Yes	Yes	Yes	Yes
Early Stopping Rule – Misadministration	ESM	Yes	No	No	No
Did Not Test	DNT	Yes	No	No	No

Print Test Status: Yes – print the three-letter code; No – Leave blank

Print State Compare: Yes – print -, +, or = based on student score; No – Leave blank

Print Scale Score: Yes – print student scale score; No – Leave blank

Print Performance Level: Yes: Print “Level 1”, “Level 2”, “Level 3”, or “Level 4” student performance level; No – Leave blank

- Footer
 - State Comparison Key
 - Copyright
 - Page X (Restart page count at 1 for each test grade)

School, District, and State Summary Report

Report Delivery

- Each participating Partner with at least one student included the “Number Enrolled” calculation will receive a State Summary Report.
- Each district with at least one student included the “Number Enrolled” calculation will receive a District Summary Report.
- Each school with at least one student included in the “Number Enrolled” will receive a School Summary Report.
- Static PDFs will be generated to be posted online. The report is not printed.
- For Partners who are not participating in science, each static PDF will contain a page for ELA and a page for mathematics.
- For Partners who participated in science, each static PDF will contain a page for ELA, a page for mathematics, and a page for science.

Data Visualization

This section details the data visualizations for the State, District, and School Summary Report

- **State Summary Report**
 - Title: [Formatted Subject]
 - Right Justified Header: Print [Formatted State Name]
 - Summary Data Rows:
 - Each row will contain the state aggregated test results for each grade
 - If the “Number Tested” is less than 10, then suppress the Number and Percent at each Performance Level and Average Scale Score
 - Footnote: Copyright statement
- **District Summary Report**
 - Title:
 - Print CONFIDENTIAL
 - Print [Formatted Subject]
 - Right Justified Header:
 - Print [Formatted State Name]
 - Print [District Name]
 - Summary Data Rows:
 - Each grade row will contain the state and district aggregated test results
 - Only grades with at least one student enrolled in the district will be included on the district roster
 - Do not suppress aggregations
 - Footnote: Copyright statement
- **School Summary Report**



- Title:
 - Print: CONFIDENTIAL
 - Print: [Formatted Subject]
- Right Justified Header:
 - Print: [Formatted State Name]
 - Print: [District Name]
 - Print: [School Name]
- Summary Data Rows:
 - Each grade row will contain the state, district, and school aggregated test results
 - Only grades with at least one student enrolled in the school will be included on the district roster
 - Do not suppress aggregations
- Footnote: Copyright statement

eMetric Data Interaction

Student & Summary Results

Description	- Cognia will provide eMetric data to support eMetric Data Interaction reporting for Partners who opted into this option - eMetric will receive two types of files: Student Results, Summary Results
Generation Rules	- Exclude Void/Duplicate Student Tests from the student results file (and subsequently aggregations in summary files) - The file layouts define each field and valid values - The student files will be exported to EXCEL - The summary files will be exported to EXCEL.
File Layouts	- Student Results: MSAA2425StudentResultsLayout.xlsx worksheet StuResults_ELAMATSCI - Summary: MSAA2425eMetricSummaryDataTransfer.xlsx
File Names	- MSAA2425_[Partner abbreviation]_StudentResults.xlsx - MSAA2425_[Partner abbreviation]_[test grade]SummaryData.xlsx

PDF Metadata

Description	Cognia will provide eMetric data to support the eMetric PDF download hub for Partners who opted into this option
Generation Rules	- Each school student report PDF will be included in the CSV - The first row will contain field names: ProgramName, ReportName, Org_Num, PDF_Name
File Layouts	- Program Name: MSAA (Alternate Assessment) - Year: 2025 - ReportName: Individual Student Report - Org_Num: <client district code>-<client school code> - PDF_Name: <school student report PDF name>
File Names	MSAA2425_[Partner abbreviation]_eMetricPDFMetaData.csv

Focal Point Reporting

Student & Summary Results

Description	• Cognia will provide Focal Point data to support Focal Point reporting for Partners who opted into this option • Focal Point will receive two types of results files: Student Results, Summary Results • Focal Point will also receive data needed for recreating the student report
Generation Rules	• Exclude Void/Duplicate Student Tests from the student results file (and subsequently aggregations in summary files) • The file layouts define each field and valid values • The student files will be exported to EXCEL • The summary files will be exported to EXCEL.
File Layouts	• Student Results: MSAA2425StudentResultsLayout.xlsx worksheet StuResults_ELAMATSCI • Summary: MSAA2425eMetricSummaryDataTransfer.xlsx
File Names	• MSAA2425_[Partner abbreviation]_StudentResults.xlsx • MSAA2425_[Partner abbreviation]_[test grade]SummaryData.xlsx

Parental Rescore Request

For Partners selecting the Parental Rescore Request option, if one or more students require a score update as part of the parental rescore request the following deliverables will be updated with the corrected student scores and provided to Client Services Program Management to be delivered to each Partner. Aggregate data will not be re-calculated as part of the parental rescore request.

- Student Results Datafile
- Student Report
- School and District Roster Report

APPENDIX G
GUIDE FOR SCORE REPORT INTERPRETATION



2025 Guide for Score Report Interpretation

All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Multi-State Alternate Assessment (MSAA). 2025 Guide for Score Report Interpretation.

MSAA Service Center

Phone: 866-834-8879

Email: MSAAServiceCenter@cognia.org

MSAA Online Assessment System: www.msaaassessment.org

State-Specific Information

State MSAA Coordinator contact information is listed below.

American Samoa	Arizona	BIE
(refer to contact information below)	www.azed.gov/assessment/alternate-assessments-dtcs	www.bie.edu
Thor Tinitali: 684-633-1323 ext. 226 thort@doe.as Anntonelli Pola: 684-633-1323 ext. 233 nelly.pola@doe.as	AlternateAssessment@azed.gov Main line: 602-542-8239	Donald Griffin: 703-282-3316 Donald.Griffin@bie.edu Aurelia Shorty: 505-274-3746 Aurelia.Shorty@bie.edu
CNMI	District of Columbia	DoDEA
(refer to contact information below)	osse.dc.gov/service/alternate-assessments	www.dodea.edu/education/assessment/assessment-calendar/comprehensive-assessment-program-sy24-25
Fasefulu Tigilau: 670-789-8739 fasefulu.tigilau@cnmipss.org	Assessment Team OSSE.Assessment@dc.gov Stephanie Snyder: 202-765-7158 Stephanie.Snyder@dc.gov	Dr. Blessing Mupanduki: 571-372-7983 blessing.mupanduki@dodea.edu Nanci Rangelhelli-Bohlke nanci.rangelhelli-bohlke@dodea.edu
Guam	Maine	Montana
(refer to contact information below)	www.maine.gov/doe/Testing_Accountability/MECAS/ela_math_materials/msaa	opi.mt.gov/Leadership/Assessment-Accountability/MontCas/Participation-Eligibility
Michelle M. Camacho: 671-475-0479 ext. 1347 mmcamacho@gdoe.net Joshua C. Blas: 671-475-0479 ext. 1254 jcblas@gdoe.net	Mechelle Ganglfinger: 207-242-4244 mechelle.ganglfinger@maine.gov	Assessment Help Desk: 844-867-2569 OPIAssessmentHelpDesk@mt.gov
South Dakota	USVI	Vermont
doe.sd.gov/assessment/alternate.aspx	(refer to contact information below)	education.vermont.gov/student-learning/assessments/alternate-assessments
Sabrina Johnson: 605-773-6156 Sabrina.Johnson@state.sd.us Stacy Holzbauer: 605-295-3441 Stacy.Holzbauer@state.sd.us	Uilez Semper uilez.sempervide.vi James Richardson: 340-998-2633 james.richardson@vide.vi	Amanda Hass: 802-828-5513 Amanda.Hass@vermont.gov AOE Assessment Team aoe.statewideassessment@vermont.gov

Table of Contents

MSAA Service Center.....	1
State-Specific Information	1
Introduction to the MSAA	3
Purpose	3
Student Participation.....	4
Overview of the MSAA Format.....	5
Scoring.....	5
MSAA Score Reports.....	6
Overview	6
Interpreting and Using the MSAA Scores	7
Talking to Parents and Guardians	7
Reporting Codes and Descriptions	8
Types of Score Reports	9
Testing Participation	9
Reports for the District.....	10
District Summary Report	10
District Roster Report	11
Reports for the School.....	12
School Summary Report.....	12
School Roster Report.....	13
Individual Student Report	14
Appendix A: Writing Scoring Rubrics.....	16
Appendix B: Performance Level Descriptors	31
Performance Level Descriptors for ELA, Mathematics, and Science	31
Appendix C: Scale Score Ranges.....	61
Appendix D: Individual Student Report Samples.....	62

Introduction to the MSAA

Purpose

The Multi-State Alternate Assessment (MSAA) is a comprehensive assessment system, designed to promote increasing higher academic outcomes for students with the most significant cognitive disabilities, in preparation for a broader array of post-secondary outcomes. The MSAA is designed to assess students with the most significant cognitive disabilities and measures academic content that is aligned to and derived from each participating state's content standards. This assessment contains many built-in supports that allow students to use materials they are most familiar with, and communicate what they know and can do as independently as possible. The MSAA is administered in the content areas of English Language Arts (ELA) and mathematics in grades 3–8 and high school. American Samoa, Arizona, the Bureau of Indian Education (BIE), Commonwealth of the Northern Mariana Islands (CNMI), Guam, Maine, the United States Virgin Islands (USVI), and Vermont also administered science in grades 5, 8, and high school.

This assessment was developed with Cognia through the research and development done by the National Center and State Collaborative (NCSC), and is now carried forward by the MSAA Partners, including American Samoa, Arizona, BIE, CNMI, Department of Defense Education Activity (DoDEA), District of Columbia, Guam, Maine, Montana, South Dakota, USVI, and Vermont.

This guide provides information regarding the administration and results of the spring 2025 MSAA to district and school personnel.

Student Participation

The criteria for student participation in the MSAA reflect the pervasive nature of a significant cognitive disability. All content areas should be considered by the Individualized Education Program (IEP) team when determining who should participate in this assessment. The table below shows the participation criteria and the descriptors used to determine eligibility for participation for each student. Students must meet the following eligibility criteria:

Participation Criteria	Participation Criteria Descriptors
1. The student has a significant cognitive disability.	Review of student records indicates a disability or multiple disabilities that significantly impact intellectual functioning and adaptive behavior.* *Adaptive behavior is defined as essential for someone to live independently and to function safely in daily life.
2. The student is learning content linked to grade-level content standards.	Goals and instruction listed in the IEP for this student are linked to the enrolled grade-level content standards and address knowledge and skills that are appropriate and challenging for this student.
3. The student requires extensive, direct individualized instruction and substantial supports to achieve measurable gains in a grade- and age-appropriate curriculum.	The student (a) requires extensive, repeated, individualized instruction and support that is not of a temporary or transient nature, and (b) uses substantially adapted materials and individualized methods of accessing information in alternative ways to acquire, maintain, generalize, demonstrate, and transfer skills across multiple settings.

Assessments for students with the most significant cognitive disabilities rely on a foundation of communicative competence. Students who do not have receptive and expressive communication are unlikely to be able to demonstrate what they know and can do on an assessment. Students who do not have a mode of communication are identified during the assessment process.

Post assessment, teachers may use the Communication Tool Kit developed by NCSC to help these students develop a mode of communication. The tool kit can be found here:

wiki.ncscpartners.org/index.php/Communication_Tool_Kit.

Overview of the MSAA Format

The MSAA assesses ELA (reading and writing) and mathematics at grades 3–8 and high school, and is aligned to the state’s content standards and the MSAA Core Content Connectors (CCCs). The MSAA is a computer-based, on-demand, stage-adaptive assessment consisting mostly of selected-response items and some constructed-response items. The test items are written at three levels of complexity that represent different levels of skill acquisition by students.

Students with the most significant cognitive disabilities often need materials and instructional strategies that are substantially adapted, are scaffolded, and have built-in supports to meet their individual needs.

The MSAA levels of complexity are designed to follow instructional practices. When students begin to learn a new skill, or acquire new knowledge, they need more support. As students learn and develop mastery of that skill or knowledge, they need less support. The test items on the MSAA are developed with many scaffolds and supports embedded within the items. Supports not embedded in the test items may be provided as accommodations, as well as other allowable ways to present the item to a student, based on their individual requirements.

The assessment is a computer-based test and is administered one-on-one. Based on the needs of the student, the assessment may also be delivered in a paper-pencil format. The needs of the student may also be addressed through other supports and accommodations, such as reading the test aloud, having a scribe, using manipulatives, using object replacement, translating the test into American Sign Language, among others. Test administrators (TAs) have substantial leeway in developing a testing schedule, with the ability to start and stop a test depending on the engagement of the student.

Each content area consists of 45–55 items across two test sessions. These are primarily selected-response items with some constructed-response items. The writing portion of the ELA test contains a scaffolded writing prompt at each grade level.

American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont also administered science in grades 5, 8, and high school. The science test is aligned to the state’s content standards and the Extended Performance Expectations (EPEs). The science assessment is also computer based and consists of selected-response items.

Scoring

Scoring of most items is accomplished within the online test platform. The selected-response items are scored as correct or incorrect by the test platform based on the answer keys programmed into the system. Constructed-response items are scored by the TA and then marked correct or incorrect in the test platform. Items without responses receive a score of zero. Student responses to writing prompts are hand scored by trained scorers utilizing the rubrics in [Appendix A](#).

Overview

This guide describes the types of score reports provided for the 2024–25 MSAA administration. The data in the sample reports are for illustrative purposes only and are not intended to reflect performance of any student(s).

Information included on the score reports:

- **Performance levels** describe how the student performed in relation to the knowledge and skills of that content area and grade level. Each performance level has two components: the scale scores that make up each level and the performance level descriptors (PLDs). The PLDs are broad and general statements regarding skills and abilities of students who have attained each level.
 - o Performance levels for ELA and mathematics for the MSAA were established by committees of educators after the first NCSC administration of the assessment in 2015 and were updated in 2018. PLDs for each grade level of ELA and mathematics can be found in [Appendix B](#).
 - o Content and Accessibility specialists collaborated with MSAA Science Partners to develop PLDs for science in 2022. Science PLDs consist of policy PLDs and range PLDs. Policy PLDs provide high-level student performance expectations, and range PLDs describe the knowledge, skills, and abilities that students must demonstrate to be classified into a performance level. PLDs for grades 5, 8, and high school science can be found in [Appendix B](#).
- **Scale scores** are numerical values that summarize student performance. The scale score allows for an appropriate comparison across test forms and administration years within a grade and content area. MSAA Individual Student Reports (ISRs) provide overall scale scores for ELA, mathematics, and science, which determine a student's performance level for each content area. Scale scores range from 1200–1290 for all grades and content areas.

For example, a student who earns an overall scale score of 1250 on one form of the grade 8 mathematics assessment would be expected to earn an overall scale score of 1250 on any other form of the grade 8 mathematics assessment. Furthermore, the student's overall scale score would be comparable to a student who took the same assessment the previous year or the following year.

In [Appendix C](#), scale score ranges for each performance level are shown by content area and grade.
- **Descriptive and informative reports.** In addition to including student demographic information, performance level, and scale scores, the ISR contains supportive information about student performance and MSAA measures.
 - o **Reading and writing scores.** The percentage of items answered correctly for reading and writing separately. The writing items consisted of selected response (multiple choice) and constructed response (the writing prompt).
 - o **What skills can be worked on next?** Skills related to the standards in the following grade.
 - o **What now?** Conversation starters for parents when talking with teachers about instruction for their child.

Interpreting and Using the MSAA Scores

The MSAA tests student performance based on the state's content standards at the student's enrolled grade level. The student's performance level is based on alternate academic achievement standards. Results for the MSAA are reported by a scale score and performance level for each content area.

MSAA scores should be used in conjunction with the IEP progress reports, student work, diagnostic assessments, district-required assessments, and report cards in order to place the student's performance on academic content and skills in context and to provide a complete picture of the student's progress across a wide range of categories.

It is helpful to read the PLDs to understand the expectations for the performance level and grade level for each student. This information can provide a concrete link from the test to instructional planning.

Talking to Parents and Guardians

MSAA parent overviews are available for parents to introduce and describe the assessment. To view the parent guides, visit www.msaastates.com and select the "Resources for Families" tab. You may also contact your State MSAA Coordinator to locate these materials. Contact information can be found at the beginning of this document.

When talking to parents and guardians about their child's score, it may be helpful to keep the following in mind:

- MSAA test results should be used along with local assessment results and other information to determine what changes in curriculum and instruction may be needed to support their student's learning.
- MSAA scores alone should not be used to make placement or eligibility decisions.

Reporting Codes and Descriptions

A complete list of reporting codes and their associated descriptions is provided below. It is important to note that not all codes are used in all entities. For additional information on reporting codes, contact your State MSAA Coordinator. Contact information can be found at the beginning of this document.

Code	Test Status	Description
TES	Tested	The student's test was submitted by the close of testing. Please note that students with a status code of TES (Tested), ESR (Early Stopping Rule), or IRR (Administration Irregularity) are all included in the total number of students tested for reporting.
ESR	Early Stopping Rule	If the TA did not observe a student response after the presentation of four items, the test was closed by the test coordinator (TC). Students with this code are included in the overall count for students tested.
ESM	Early Stopping Rule Misadministration	Testing may have ended early on the basis that a consistent mode of communication was not observed or the student was not administered the first four items in at least one content area. At least one response was recorded for the student, but the student may not have had the opportunity to complete the entire test.
INC	Tested – Incomplete	The student's test was not submitted by the close of testing. The student may not have had the opportunity to complete the entire test.
IRR	Administration Irregularity	An administration irregularity not necessitating an invalidation of scores was reported for the student's test. Students with this code are included in the overall count for students tested.
INV	Invalidated	The results of the student's test have been invalidated.
PRF	Parental Refusal	The student did not test due to a parent/guardian refusal.
ELL	ELL Exempt (ELA Only)	The student was exempt from ELA testing due to being a first year English Language Learner.
EXE	Exempt (Emergency, Medical, Other)	The student was exempt from testing.
DNT	Did Not Test	The student did not test via the MSAA.
WDR	Withdrew	The student withdrew from the school.
NLE	No Longer Eligible	The student is not eligible to test via the MSAA.

Types of Score Reports

Below are the types of MSAA score reports that will be available on the MSAA Reporting Portal. Only district TCs using their current MSAA username and password may access the MSAA reports at www.msaaassessment.org under the Reporting tab. Reports are only available during the online reporting window. All MSAA score reports are confidential documents.

- Reports for the District
 - o District Summary Report (DSR)
 - o District Roster Report (DRR)
 - o Student Results File
- Reports for the School
 - o School Summary Report (SSR)
 - o School Roster Report (SRR)
 - o Individual Student Report (ISR)
 - o Student Results File

An Excel file of all student results at the district and school levels will be available to district TCs through the MSAA Reporting Portal. For information regarding this file or questions about accessing the reports, contact your State MSAA Coordinator. Contact information can be found at the beginning of this document.

Testing Participation

All students in grades 3–8 and high school are required to be assessed in ELA and mathematics. Participation status is assigned independently for ELA and mathematics.

All submitted tests receive a participation status, regardless of the number of responses to items.

For additional information regarding the reported test status, contact your State MSAA Coordinator. Contact information can be found at the beginning of this document.

Reports for the District

District Summary Report

The DSR provides district staff with a summary of student participation and performance by district and school. State-level data is taken from the individual participating state. See Figure 1 below.

Figure 1. Sample District Summary Report

CONFIDENTIAL

msaa

Multi-State Alternate Assessment

1

English Language Arts

2

SUMMARY REPORT

Demonstration State

Demonstration District B

		3		Did Not Test	4		5							
		Enrolled	Tested		Average Scale Score	Performance Level								
						Level 1		Level 2		Level 3		Level 4		
						N	%	N	%	N	%	N	%	
03	State	22	12	10	1225	5	42	6	50	1	8	0	0	
	District	22	12	10	1225	5	42	6	50	1	8	0	0	
04	State	17	9	8	1219	7	78	1	11	1	11	0	0	
	District	17	9	8	1219	7	78	1	11	1	11	0	0	
05	State	28	14	14	1223	8	57	3	21	2	14	1	7	
	District	28	14	14	1223	8	57	3	21	2	14	1	7	
06	State	17	10	7	1222	5	50	2	20	3	30	0	0	
	District	17	10	7	1222	5	50	2	20	3	30	0	0	
07	State	17	10	7	1223	5	50	4	40	1	10	0	0	
	District	17	10	7	1223	5	50	4	40	1	10	0	0	
08	State	21	12	9	1221	5	42	3	25	4	33	0	0	
	District	21	12	9	1221	5	42	3	25	4	33	0	0	
High School	State	26	15	11	1223	8	53	3	20	4	27	0	0	
	District	26	15	11	1223	8	53	3	20	4	27	0	0	

The DSR contains the following features, highlighted above:

1. The content area of the report
2. The state and district included in the report
3. The number of students by grade who were enrolled,* tested, and did not test by state and district
4. The average scale score for each grade by state and district
5. The number and percentage of students at each performance level by grade in the state and district

* Number of students in this content area who have a reporting status other than WDR or NLE

District Roster Report

The DRR provides student performance information at the district level for each grade, including each student's test status, scale score, and performance level. See Figure 2 below.

Figure 2. Sample District Roster Report



Multi-State Alternate Assessment

CONFIDENTIAL

DISTRICT ROSTER REPORT

1

Demonstration State
Demonstration District B

Grade 03

Mathematics

2	Enrolled	Tested	Average Scale Score	Level 1 (%)	Level 2 (%)	Level 3 (%)	Level 4 (%)
State	22	13	1234	31	0	62	8
District	22	13	1234	31	0	62	8

Spring 2025

4	Student Name Student ID	3 Mathematics			
		Test Status*	State Compare	Scale Score	Performance Level
	LASTNAME56, FIRST56 D056	DNT			
	LASTNAME67, FIRST67 D067	ESR	-	1200	Level 1
	LASTNAME77, FIRST77 D077	ESR	-	1200	Level 1
	LASTNAME79, FIRST79 D079	ESR	-	1200	Level 1
	LASTNAME81, FIRST81 D081	ESR	-	1200	Level 1
	LASTNAME83, FIRST83 D083		+	1245	Level 3
	LASTNAME91, FIRST91 D091		+	1253	Level 3

The DRR contains the following features, highlighted above:

1. The state and district included in the report
2. The number of students who were enrolled and tested, the average scale score, and the percentage of students at each performance level by state and district
3. The test status, state comparison, scale score, and performance level by student for the specified content area. Refer to the Reporting Codes and Descriptions for information regarding test status.
4. The students who tested in the district for the specified grade
5. The symbols used in the "State Compare" column

Reports for the School

School Summary Report

The SSR provides summarized performance information at the state, district, and school levels for each grade, including the number of students who were enrolled, tested, and did not test, as well as average scale score and performance level. See Figure 3 below.

Figure 3. Sample School Summary Report



msaa
Multi-State Alternate Assessment

CONFIDENTIAL

1 English Language Arts

2 SUMMARY REPORT

Demonstration State
Demonstration District B
Demonstration School 4

		3		Did Not Test	4 Average Scale Score	5 Performance Level							
		Enrolled	Tested			Level 1		Level 2		Level 3		Level 4	
						N	%	N	%	N	%	N	%
03	State	22	12	10	1225	5	42	6	50	1	8	0	0
	District	22	12	10	1225	5	42	6	50	1	8	0	0
	School	22	12	10	1225	5	42	6	50	1	8	0	0
04	State	17	9	8	1219	7	78	1	11	1	11	0	0
	District	17	9	8	1219	7	78	1	11	1	11	0	0
	School	17	9	8	1219	7	78	1	11	1	11	0	0
05	State	28	14	14	1223	8	57	3	21	2	14	1	7
	District	28	14	14	1223	8	57	3	21	2	14	1	7
	School	28	14	14	1223	8	57	3	21	2	14	1	7
06	State	17	10	7	1222	5	50	2	20	3	30	0	0
	District	17	10	7	1222	5	50	2	20	3	30	0	0
	School	17	10	7	1222	5	50	2	20	3	30	0	0
07	State	17	10	7	1223	5	50	4	40	1	10	0	0
	District	17	10	7	1223	5	50	4	40	1	10	0	0
	School	17	10	7	1223	5	50	4	40	1	10	0	0
08	State	21	12	9	1221	5	42	3	25	4	33	0	0
	District	21	12	9	1221	5	42	3	25	4	33	0	0
	School	21	12	9	1221	5	42	3	25	4	33	0	0
High School	State	26	15	11	1223	8	53	3	20	4	27	0	0
	District	26	15	11	1223	8	53	3	20	4	27	0	0
	School	26	15	11	1223	8	53	3	20	4	27	0	0

© 2025 MSAA. All Rights Reserved.

© 2025 MSAA. All Rights Reserved.

The SSR contains the following features, highlighted above:

1. The content area of the report
2. The state, district, and school included in the report
3. The number of students by grade who were enrolled, * tested, and did not test by state, district, and school
4. The average scale score for each grade by state, district, and school
5. The number and percentage of students at each performance level by grade in the state, district, and school

* Number of students in this content area who have a reporting status other than WDR or NLE

School Roster Report

The SRR provides student performance information at the school level for each grade, including each student's test status, scale score, and performance level. See Figure 4 below.

Figure 4. Sample School Roster Report



Multi-State Alternate Assessment

CONFIDENTIAL

SCHOOL ROSTER REPORT

1 Demonstration State
Demonstration District B
Demonstration School 4
Grade 05

English Language Arts

2	Enrolled	Tested	Average Scale Score	Level 1 (%)	Level 2 (%)	Level 3 (%)	Level 4 (%)
State	28	14	1223	57	21	14	7
District	28	14	1223	57	21	14	7
School	28	14	1223	57	21	14	7

Spring 2025

		3 English Language Arts			
Student Name Student ID	Test Status*	State Compare	Scale Score	Performance Level	
4 LASTNAME10, FIRST10 D010		+	1230	Level 1	
LASTNAME11, FIRST11 D011	ESR	-	1200	Level 1	
LASTNAME12, FIRST12 D012	ESR	-	1200	Level 1	
LASTNAME133, FIRST133 D133		+	1240	Level 3	
LASTNAME137, FIRST137 D137	ESM	-	1200		
LASTNAME141, FIRST141 D141	DNT				
LASTNAME144, FIRST144 D144	DNT				
LASTNAME148, FIRST148 D148	ESM	-	1200		
LASTNAME149, FIRST149 D149	ESM	-	1200		
LASTNAME152, FIRST152 D152	DNT				
LASTNAME155, FIRST155 D155	DNT				
LASTNAME157, FIRST157 D157	DNT				
LASTNAME160, FIRST160 D160	INV				
LASTNAME161, FIRST161 D161	INV				
LASTNAME21, FIRST21 D021	ESM	-	1200		

5	State Comparison Key
-	Performance is lower than state average
=	Performance is similar to state average
+	Performance is greater than state average

* For descriptions of the Test Statuses, see your State's Guide for Score Report Interpretation.

© 2025 MSAA. All Rights Reserved.

Page 1

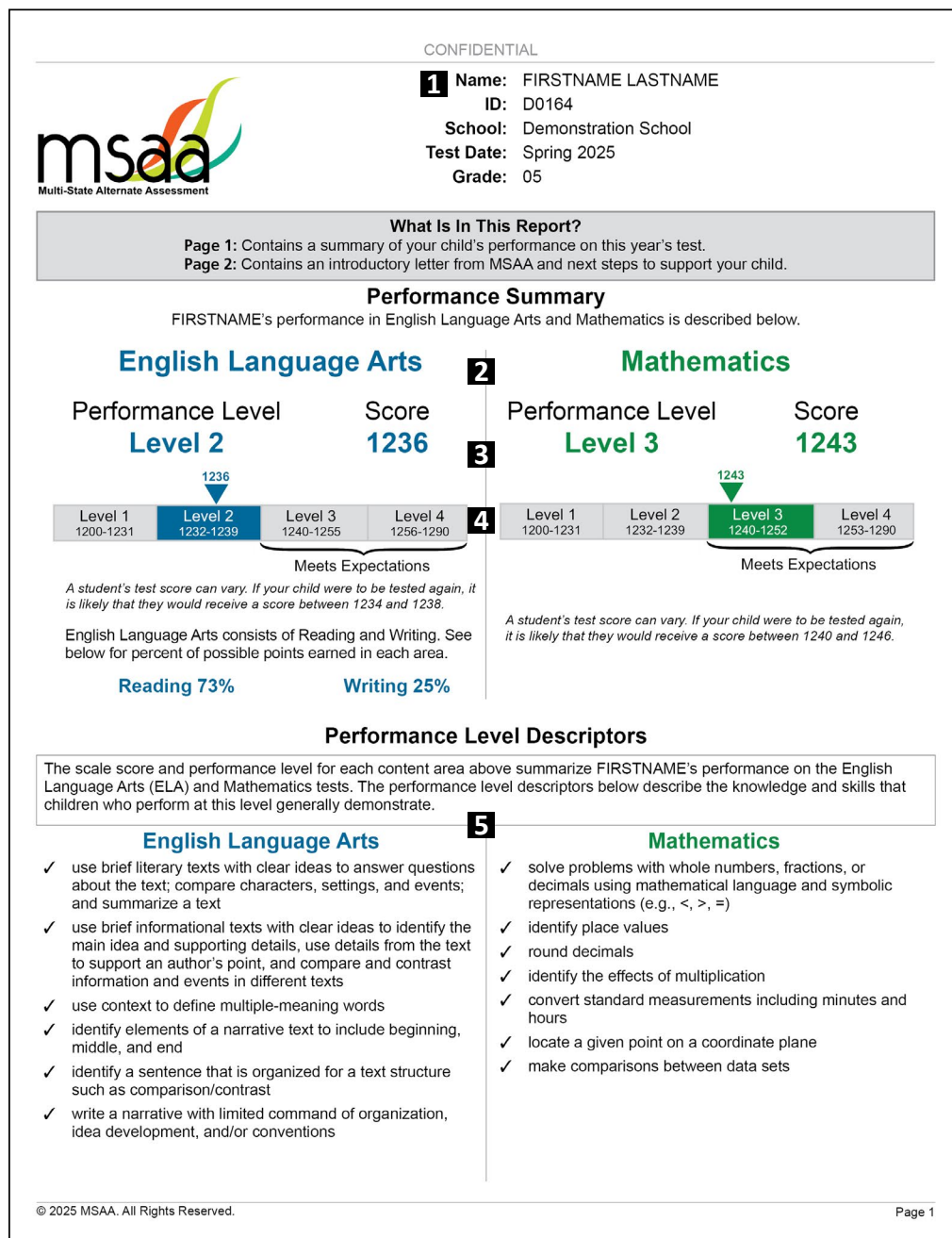
The SRR contains the following features, highlighted above:

1. The state, district, and school included in the report
2. The number of students who were enrolled and tested, the average scale score, and the percentage of students at each performance level by state, district, and school
3. The test status, state comparison, scale score, and performance level by student for the specified content area. Refer to the Reporting Codes and Descriptions for more information regarding test status.
4. The students who tested in the school for the specified grade
5. The symbols used in the "State Compare" column

Individual Student Report

The ISR provides scale score and performance level information for a specific student. Figure 5 shows page 1 of the ISR. Full samples of the ISR are included in [Appendix D](#).

Figure 5. Sample Individual Student Report



The ISR contains the following features, highlighted above:

1. The student's full name, student ID, school, and grade
2. The results for each content area
3. The student's scale score and performance level for each content area
4. The student's score compared to the performance level scale
5. The PLDs for the student's performance level

[Page is intentionally blank]

Appendix A: Writing Scoring Rubrics

Grade 3 Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The narrative establishes a situation (activity and setting) and includes a character with relevant descriptive statements. The response provides a conclusion.	The narrative includes at a minimum: • character and situation (activity and setting) • a conclusion that follows from the narrated experiences or events	The narrative includes at a minimum: • character and situation (activity or setting) • a conclusion that may not follow from the narrated experiences or events	The narrative includes at a minimum: • some evidence related to a character, situation (activity or setting), or conclusion	0 • no evidence of organization	5 • evidence is off topic
Idea Development – The narrative includes a sequence of events that unfold naturally and develops a story using temporal words.	The narrative includes at a minimum: • a sequence of two events related to the situation (activity or setting) • both events include a detail	The narrative includes at a minimum: • two events related to the situation (activity or setting) • one of the events includes a detail	The narrative includes at a minimum: • one event related to the situation (activity or setting)	0 • no evidence of idea development	5 • evidence is off topic
Conventions – Students use standard English conventions (e.g., end punctuation, subject-verb agreement).	The narrative includes more than one sentence and at a minimum: • end punctuation for more than one thought unit • one simple sentence that contains a complete thought with subject-verb agreement (e.g., “Dog runs” or “dog runs”)	The narrative includes at a minimum: • end punctuation for one thought unit • one thought unit with or without subject-verb agreement	The narrative includes at a minimum: • one use of standard English conventions (end punctuation for one thought unit or one thought unit with or without subject-verb agreement)	0 • no evidence of standard English conventions	

Grade 3 Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The narrative establishes a situation (activity and setting) and includes a character with relevant descriptive statements. The response provides a conclusion.	The narrative includes at a minimum: • character and situation (activity or setting) • two descriptions related to a character • a conclusion that follows from the narrated experiences or events	The narrative includes at a minimum: • character and situation (activity or setting) • one description related to a character • a conclusion that may not follow from the narrated experiences or events	The narrative includes at a minimum: • some evidence related to a character, situation (activity or setting), or conclusion OR • descriptive words related to a character or situation (activity or setting)	0 • no evidence of organization	5 • evidence is off topic
Idea Development – The narrative includes a sequence of events that unfold naturally and develops the story using temporal words (e.g., first, then, next).	The narrative includes at a minimum: • two sequenced events related to the situation (activity or setting) • both events include a detail • appropriate use of temporal words that signal order of events	The narrative includes at a minimum: • two events related to the situation (activity or setting) • one of the events includes a detail • one temporal word that may or may not be used appropriately	The narrative includes at a minimum: • one event related to the situation (activity or setting)	0 • no evidence of idea development	5 • evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The narrative includes more than one sentence and at a minimum: • capitalization at the beginning of the majority of thought units • end punctuation for more than one thought unit • one simple sentence that contains a complete thought with subject-verb agreement (e.g., “Dog runs” or “dog runs”)	The narrative includes at a minimum two of the following: • capitalization at the beginning of one thought unit • end punctuation for one thought unit • one simple sentence with or without subject-verb agreement	The narrative includes at a minimum: • one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 • no evidence of standard English conventions	

Grade 4 Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The narrative establishes a situation (activity <u>or</u> setting) and includes a character. The response provides a conclusion.	The narrative includes at a minimum: • character <u>and</u> situation (activity <u>or</u> setting) • a conclusion that follows from the narrated experiences <u>or</u> events	The narrative includes at a minimum: • character <u>and</u> situation (activity <u>or</u> setting) • a conclusion that <u>may not</u> follow from the narrated experiences <u>or</u> events	The narrative includes at a minimum: • <u>some</u> evidence related to a character, situation (activity <u>or</u> setting), <u>or</u> conclusion	0 • <u>no</u> evidence of organization	5 • evidence is <u>off topic</u>
Idea Development – The narrative includes a description of events using concrete words or sensory details (e.g., how things look, sound, taste, smell, or feel) related to the events.	The narrative includes at a minimum: • <u>two</u> events related to the situation (activity <u>or</u> setting) • <u>both</u> of the events include a detail related to a character’s action <u>or</u> response to a situation (activity <u>or</u> setting)	The narrative includes at a minimum: • <u>two</u> events related to the situation (activity <u>or</u> setting) • <u>one</u> of the events includes a detail related to a character’s action <u>or</u> response to a situation (activity <u>or</u> setting)	The narrative includes at a minimum: • <u>one</u> event related to the situation (activity <u>or</u> setting)	0 • <u>no</u> evidence of idea development	5 • evidence is <u>off topic</u>
Conventions – Students use standard English conventions (e.g., end punctuation, subject-verb agreement).	The essay includes <u>more than one sentence and</u> at a minimum: • end punctuation for <u>more than one</u> thought unit • <u>one</u> complex thought unit that expresses a complete idea <u>with</u> subject-verb agreement (e.g., “The dog runs” or “the dog runs”)	The narrative includes at a minimum: • end punctuation for <u>one</u> thought unit • <u>one</u> complex thought unit <u>with or without</u> subject-verb agreement	The narrative includes at a minimum: • <u>one</u> use of standard English conventions (end punctuation for <u>one</u> thought unit <u>or one</u> thought unit <u>with or without</u> subject-verb agreement)	0 • <u>no</u> evidence of standard English conventions	

Grade 4 Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The narrative establishes a situation (activity and setting) and includes a character. The response provides a conclusion.	The narrative includes at a minimum: - character and situation (activity and setting) - description of character and situation (activity or setting) - a conclusion that follows from the narrated experiences or events	The narrative includes at a minimum: - character and situation (activity or setting) - description of the character or the situation (activity or setting) - a conclusion that may not follow from the narrated experiences or events	The narrative includes at a minimum: some evidence related to a character, situation (activity or setting), or conclusion OR - descriptive words related to a character or situation (activity or setting)	0 no evidence of organization	5 evidence is off topic
Idea Development – The narrative includes a description of events using concrete words or sensory details (e.g., how things look, sound, taste, smell, or feel) related to the events.	The narrative includes at a minimum: two events related to the situation (activity or setting) both events include a detail related to a character’s action or response to a situation (activity or setting)	The narrative includes at a minimum: two events related to the situation (activity or setting) one of the events includes a detail related to a character’s action or response to a situation (activity or setting)	The narrative includes at a minimum: one event related to the situation (activity or setting)	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The narrative includes at a minimum: - capitalization at the beginning of the majority of thought units - end punctuation for more than one thought unit one complex thought unit that expresses a complete idea with subject-verb agreement (e.g., “The dog runs” or “the dog runs”)	The narrative includes at a minimum: - capitalization at the beginning of one thought unit - end punctuation for one thought unit one complex thought unit with or without subject-verb agreement	The narrative includes at a minimum: one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

Grade 5 Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The narrative establishes a situation (activity and setting) for the story and includes characters. The response provides a conclusion.	The narrative includes at a minimum: • two characters unchanged through the narrative • an established situation (activity and setting) • a conclusion that follows from the narrated experiences or events	The narrative includes at a minimum: • two characters • a situation (activity or setting) • a conclusion that may not follow from the narrated experiences or events	The narrative includes at a minimum: • some evidence related to a character, situation (activity or setting), or conclusion	0 • no evidence of organization	5 • evidence is off topic
Idea Development – The narrative includes dialogue and events supported with relevant details and descriptive statements.	The narrative includes at a minimum: • two events that connect to the narrative • both of the events include a detail related to a character's action or response to a situation (activity or setting) • one dialogue statement from one character to the other character relevant to the narrative (e.g., I said "No, I want to play.")	The narrative includes at a minimum: • two events related to a character's action or response to a situation (activity or setting) • one of the events includes a detail related to a character's action or response to a situation (activity or setting) • one dialogue statement from one character to the other character that may not be relevant to the narrative	The narrative includes at a minimum: • one event related to the situation (activity or setting)	0 • no evidence of idea development	5 • evidence is off topic
Conventions – Students use standard English conventions (e.g., end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: • end punctuation for more than one thought unit • one complete sentence that expresses an idea with subject-verb agreement (e.g., "The dog runs.")	The narrative includes at a minimum: • end punctuation for one thought unit • one complete sentence with or without subject-verb agreement	The narrative includes at a minimum: • one use of standard English conventions (end punctuation for one thought unit or one thought unit with or without subject-verb agreement)	0 • no evidence of standard English conventions	

Grade 5 Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The narrative establishes a situation (activity and setting) for the story and includes characters. The response provides a conclusion.	The narrative includes at a minimum: • two characters unchanged through the narrative • identification of the situation (activity and setting) • a conclusion that follows from the narrated experiences or events	The narrative includes at a minimum: • two characters • identification of the setting or the activity • a conclusion that may not follow from the narrated experiences or events	The narrative includes at a minimum: • some evidence related to a character or conclusion	0 • no evidence of organization	5 • evidence is off topic
Idea Development – The narrative includes dialogue and events supported with relevant details and descriptive statements.	The narrative includes at a minimum: • two sequenced events related to the situation (activity or setting) • both events include a detail related to a character's action or response to a situation (activity or setting) • one relevant conversation between two characters (e.g., I said "No! I don't want to go to bed." Mom said "OK.")	The narrative includes at a minimum: • two events related to a character's action or response to a situation (activity or setting) • one event that includes a detail related to a character's action or response to a situation (activity or setting) • one relevant piece of dialogue showing what one character said to the other	The narrative includes at a minimum: • one event related to the situation (activity or setting)	0 • no evidence of idea development	5 • evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The narrative includes more than one sentence and at a minimum: • capitalization at the beginning of the majority of thought units • end punctuation for the majority of thought units • one complete sentence that expresses an idea with subject-verb agreement (e.g., "The dog runs.")	The narrative includes at a minimum: • capitalization at the beginning of one thought unit • end punctuation for one thought unit • one complete sentence with subject-verb agreement	The narrative includes at a minimum: • one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 • no evidence of standard English conventions	

Grade 6 Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses a specified topic and is organized to describe two opposing conditions (e.g., compare/contrast).	The essay includes at a minimum: - an introduction that states the essay is about two opposing conditions - a body that includes: one activity for each of the two opposing conditions and one activity common to both conditions - a conclusion that states two opposing conditions or summarizes the content	The essay includes at a minimum: - an introduction that states one activity or topic - a body that relates two conditions with activities - a conclusion that states one activity or the topic	The essay includes at a minimum: some evidence related to the specified topic (i.e., introduction, compare/contrast relationship, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The essay develops a topic, and includes relevant facts and details to promote meaning and create clarity.	The essay includes at a minimum: three activities, each with relevant details (the same detail may be used for all activities if relevant to each)	The essay includes at a minimum: one activity with a relevant detail	The essay includes at a minimum: one detail that describes an activity	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - end punctuation for more than one thought unit one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - end punctuation for one thought unit one complete sentence with or without subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (end punctuation for one thought unit or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

Grade 6 Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses a specified topic and is organized to describe two opposing conditions (e.g., compare/contrast). The response provides a conclusion.	The essay includes at a minimum: - an introduction that presents the two opposing conditions - a body that includes: one activity common to both conditions one activity related to each of the two opposing conditions - a conclusion that states the two opposing conditions	The essay includes at a minimum: - an introduction that presents the topic - a body that includes: one activity common to both conditions one activity related to one of the two opposing conditions - a conclusion that states the topic	The essay includes at a minimum: some evidence related to the specified topic (i.e., introduction, compare/contrast relationship, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The essay develops a topic, and includes relevant facts and details to promote meaning and create clarity.	The essay includes at a minimum: one activity related to both conditions with a relevant detail one activity related to each of the two opposing conditions, each with relevant details	The essay includes at a minimum: two activities each with a relevant detail	The essay includes at a minimum: one activity OR one detail that describes an activity	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - capitalization at the beginning of the majority of thought units - end punctuation for the majority of thought units one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - capitalization at the beginning of one thought unit - end punctuation for one thought unit one complete sentence with subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

Grade 7 Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses a specified topic and is organized with an effect related directly to a cause (e.g., cause/effect).	The essay includes at a minimum: - an introduction that states the topic/cause - a body that relates the effect to the provided cause - a conclusion that states the essay is about a cause and its effect	The essay includes at a minimum: - an introduction that states the topic/cause - a body that includes an effect that may not relate to the provided cause - a conclusion that states a cause or the effect	The essay includes at a minimum: some evidence related to the specified topic (i.e., introduction, cause/effect relationship, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The essay develops a topic, and includes details to promote meaning and create clarity.	The essay includes at a minimum: one relevant detail to describe the effect	The essay includes at a minimum: one effect with no relevant detail	The essay includes at a minimum: one idea related to the topic	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - end punctuation for more than one thought unit one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - end punctuation for one thought unit one complete sentence with or without subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (end punctuation for one thought unit or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

Grade 7 Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses a specified topic and is organized with an effect related directly to a cause (e.g., cause/effect).	The essay includes at a minimum: - an introduction that presents the cause and its effects - a body that includes two effects and refers them to the cause - a conclusion that states the essay is about a cause and its effects	The essay includes at a minimum: - an introduction that presents a topic - a body that includes one effect and refers it to the cause - a conclusion that states the topic	The essay includes at a minimum: some evidence related to the specified topic (i.e., introduction, on-topic cause/effect relationship, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The essay develops a topic, and includes details and transitional words to promote meaning and create clarity.	The essay includes at a minimum: two effects, each with a relevant detail - transitional words to connect the cause to each of the two effects	The essay includes at a minimum: one effect with a relevant detail - transitional word to connect one cause/effect relationship	The essay includes at a minimum: one detail that describes the cause or effect OR one transition word	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - capitalization at the beginning of the majority of thought units - end punctuation for the majority of thought units one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - capitalization at the beginning of one thought unit - end punctuation for one thought unit one complete sentence with subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

Grade 8 Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses the specified topic and is organized with a solution related directly to the problem (e.g., problem/solution).	The essay includes at a minimum: - an introduction that states both parts of the problem - a body that relates how the solution can be applied to the problem - a conclusion that states the problem and the solution	The essay includes at a minimum: - an introduction that states the problem one solution that may not relate to the problem - a conclusion that states the problem or the solution	The essay includes at a minimum: some evidence related to the specified topic (i.e., introduction, on-topic problem/solution relationship, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The essay develops a topic, and includes details to promote meaning and create clarity.	The essay includes at a minimum: one relevant detail to describe the problem one relevant detail to describe the solution	The essay includes at a minimum: one relevant detail to describe the problem or the solution	The essay includes at a minimum: one detail or word that describes the problem or the solution	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - end punctuation for more than one thought unit one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - end punctuation for one thought unit one complete sentence with or without subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (end punctuation for one thought unit or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

Grade 8 Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses the specified topic and is organized with a solution related directly to the problem (e.g., problem/solution).	The essay includes at a minimum: - an introduction that states both parts of the problem - a body that includes a solution and refers to the problem - a conclusion that states the problem and its solution	The essay includes at a minimum: - an introduction that states one part of the problem - a body that includes a related solution - a conclusion that states the problem or the solution	The essay includes at a minimum: some evidence related to the specified topic (i.e., introduction, on-topic problem/solution relationship, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The essay develops a topic, and includes details and transitional words to promote meaning and create clarity.	The essay includes at a minimum: one problem with a relevant detail one solution with a relevant detail one transitional word that connects the problem to the solution	The essay includes at a minimum: one problem or solution with a relevant detail one transitional word that is in relation to the problem or the solution	The essay includes at a minimum: one detail or word that describes the problem or the solution	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - capitalization at the beginning of the majority of thought units - end punctuation for the majority of thought units one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - capitalization at the beginning of one thought unit - end punctuation for one thought unit one complete sentence with subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

High School Writing Scoring Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses a specified claim supported with organized complex ideas.	The essay includes at a minimum: - an introduction that states the claim and a rational reason - a conclusion that states the claim and the rational reason	The essay includes at a minimum: - an introduction that states the claim or a reason - a conclusion that states the claim or the reason	The essay includes at a minimum: some evidence related to the specified claim/topic (i.e., introduction, claim/topic, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The defended claim includes relevant evidence, and uses words, phrases, and clauses to clarify the relationship among claim, reasons, and evidence.	The essay includes at a minimum: - a body with two relevant facts or examples - words or phrases to connect the reason with one relevant fact or example	The essay includes at a minimum: - a body with one relevant fact or example one word or phrase to connect the reason with one fact or example	The essay includes at a minimum: one word related to the reason	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - end punctuation for more than one thought unit one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - end punctuation for one thought unit one complete sentence with or without subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (end punctuation for one thought unit or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

High School Writing Scoring Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0 or 5	
Organization – The essay addresses a specified claim supported with organized complex ideas.	The essay includes at a minimum: - an introduction that states the claim and is supported by two rational reasons - a body that includes two reasons related to the claim - a conclusion that states the claim and is supported by two rational reasons	The essay includes at a minimum: - an introduction that states the claim - a body that includes one reason related to the claim - a conclusion that states the claim with one rational reason or relevant evidence	The essay includes at a minimum: some evidence related to the specified claim/topic (i.e., introduction, claim/topic, or conclusion)	0 no evidence of organization	5 evidence is off topic
Idea Development – The defended claim includes relevant evidence, and uses words, phrases, and clauses to clarify the relationship among claim, reasons, and evidence.	The essay includes at a minimum: one piece of relevant evidence that follows each of the two provided reasons - words or phrases that connect each of the two reasons with relevant evidence	The essay includes at a minimum: - a body with one reason and one piece of relevant evidence - a word or phrase that connects one reason with one piece of relevant evidence	The essay includes at a minimum: one word related to the reason or a connecting word or phrase	0 no evidence of idea development	5 evidence is off topic
Conventions – Students use standard English conventions (e.g., capitalization, end punctuation, subject-verb agreement).	The essay includes more than one sentence and at a minimum: - capitalization at the beginning of the majority of thought units - end punctuation for the majority of thought units one complete sentence that expresses an idea with subject-verb agreement (e.g., “The dog runs.”)	The essay includes at a minimum: - capitalization at the beginning of one thought unit - end punctuation for one thought unit one complete sentence with subject-verb agreement	The essay includes at a minimum: one use of standard English conventions (capitalization at the beginning of one thought unit, end punctuation for one thought unit, or one thought unit with or without subject-verb agreement)	0 no evidence of standard English conventions	

[Page is intentionally blank]

Appendix B: Performance Level Descriptors

Performance Level Descriptors for ELA, Mathematics, and Science

MSAA developed PLDs for ELA and mathematics at grades 3–8 and high school through an iterative process involving multiple stakeholder groups. Content and Accessibility specialists also collaborated with MSAA Science Partners to develop PLDs for science in grades 5, 8, and high school. The MSAA partnership developed grade-level PLDs to summarize the knowledge, skills, and abilities (KSAs) prioritized for the MSAA that students need to attain at each level of achievement (Level 1–Level 4). Each performance level is understood to include the KSAs of the preceding performance levels.

The PLDs included in this appendix provide a detailed description for teachers, parents, and the public to see not only what grade-level content a student should know and be able to do in order to meet high expectations, but also the depth, breadth, and complexity of that content.

By using the PLDs, test results become multi-dimensional. Test results in the form of scale scores are one way educators, parents, and guardians find out where a student’s performance is in relation to other students. The PLDs provide another dimension that completes the description of how a student interacts with the standards the test measures. Both the scale score and the PLDs provide information that helps teachers, schools, parents, and guardians build a path to student learning.

Grade 3 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: - identify the topic of a literary text - identify a detail from a literary text - identify a character or setting in a literary text - identify the topic of an informational text - identify a title, caption, or heading in an informational text - identify an illustration related to a given topic - identify a topic presented by an illustration - identify the meaning of words (i.e., nouns)	In reading, the student is able to: - determine the central idea and supporting details in literary text - determine the main idea and identify supporting details in informational text - determine the main idea of visually presented information - identify the purpose of text features in informational text - use information from charts, graphs, diagrams, or timelines in informational text to answer questions - use context to identify the meaning of multiple-meaning words	In reading, the student is able to: - determine the central idea and supporting details in literary text - determine the main idea and identify supporting details in informational text - determine the main idea of visually presented information - identify the purpose of text features in informational text - use information from charts, graphs, diagrams, or timelines in informational text to answer questions - use context to identify the meaning of multiple-meaning words	In reading, the student is able to: - determine the central idea and supporting details in literary text - determine the main idea and identify supporting details in informational text - determine the main idea of visually presented information - identify the purpose of text features in informational text - use information from charts, graphs, diagrams, or timelines in informational text to answer questions - use context to identify the meaning of multiple-meaning words
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	- use details from a literary text to answer specific questions - describe the relationship between characters, and character and setting in literary text	- use details from a literary text to answer specific questions - describe the relationship between characters, and character and setting in literary text	
	AND with accuracy, the student is able to: identify simple words (i.e., words with a consonant at the beginning, a consonant at the end, and a short vowel in the middle)	AND with accuracy, the student is able to: identify grade-level words	
AND in writing, the student is able to: - identify a statement related to an everyday topic - use the writing process to create a narrative product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify elements of a narrative text to include beginning, middle, and end - identify the category related to a set of facts - use the writing process to create a narrative product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify a text feature (e.g., captions, graphs, or diagrams) to present information in explanatory text - use the writing process to create a narrative product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: use the writing process to create a narrative product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 4 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: - identify a topic of a literary text - identify a detail from a literary text - identify a character in a literary text - identify charts, graphs, diagrams, or timelines in an informational text - identify a topic of an informational text - use context to identify the meaning of multiple-meaning words - identify general academic words	In reading, the student is able to: - determine the theme of literary text and identify supporting details - describe character traits using text-based details in literary text - determine the main idea of informational text - locate information in charts, graphs, diagrams, or timelines - use information from charts, graphs, diagrams, or timelines in informational text to answer questions - use general academic words	In reading, the student is able to: - determine the theme of literary text and identify supporting details - determine the main idea of informational text - explain how the information provided in charts, graphs, diagrams, or timelines contributes to an understanding of informational text - use information from charts, graphs, diagrams, or timelines in informational text to answer questions - use general academic words	In reading, the student is able to: - determine the theme of literary text and identify supporting details - determine the main idea of informational text - explain how the information provided in charts, graphs, diagrams, or timelines contributes to an understanding of informational text - use information from charts, graphs, diagrams, or timelines in informational text to answer questions - use general academic words
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	- use details from a literary text to answer specific questions - use context to identify the meaning of multiple-meaning words	- use details from a literary text to answer specific questions - describe character traits using text-based details in literary text - use context to identify the meaning of multiple-meaning words	
	AND with accuracy, the student is able to: identify simple words (i.e., words with a consonant at the beginning, a consonant at the end, and a short vowel in the middle)	AND with accuracy, the student is able to: identify grade-level words	
AND in writing, the student is able to: - identify the concluding sentence in a short explanatory text - use the writing process to create a narrative product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify elements of a narrative text to include beginning, middle, and end - identify a concluding sentence related to information in explanatory text - use the writing process to create a narrative product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify a text feature (e.g., headings, charts, or diagrams) to present information in explanatory text - use the writing process to create a narrative product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: use the writing process to create a narrative product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 5 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: - identify an event from the beginning of a literary text - identify a detail from a literary text - identify a character, setting, and event in a literary text - identify the topic of an informational text - identify the main idea of an informational text - identify the difference in how information is presented in two sentences	In reading, the student is able to: - compare characters, settings, and events in literary text - determine the main idea and identify supporting details in informational text - use details from the text to support an author's point in informational text - compare and contrast how information and events are presented in two informational texts - use context to identify the meaning of multiple-meaning words	In reading, the student is able to: - compare characters, settings, and events in literary text - determine the main idea and identify supporting details in informational text - use details from the text to support an author's point in informational text - compare and contrast how information and events are presented in two informational texts - use context to identify the meaning of multiple-meaning words	In reading, the student is able to: - compare characters, settings, and events in literary text - determine the main idea and identify supporting details in informational text - use details from the text to support an author's point in informational text - compare and contrast how information and events are presented in two informational texts - use context to identify the meaning of multiple-meaning words
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	- summarize a literary text from beginning to end - use details from a literary text to answer specific questions	- summarize a literary text from beginning to end - use details from a literary text to answer specific questions	
AND in writing, the student is able to: - identify the category related to a set of common nouns - use the writing process to create a narrative product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify elements of a narrative text to include beginning, middle, and end - identify a sentence that is organized for a text structure such as comparison/contrast - use the writing process to create a narrative product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: - support an explanatory text topic with relevant information - use the writing process to create a narrative product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: use the writing process to create a narrative product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 6 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: - identify an event from the beginning or end of a literary text - identify a detail from a literary text - identify a character in a literary text - identify the topic of an informational text - identify the main idea of an informational text - identify a fact from an informational text - identify a description of an individual or event in an informational text - use context to identify the meaning of multiple-meaning words - identify the meaning of general academic words	In reading, the student is able to: - summarize a literary text from beginning to end without including personal opinions - support inferences about characters using details in literary text - use details from the text to elaborate a key idea in informational text	In reading, the student is able to: - summarize a literary text from beginning to end without including personal opinions - support inferences about characters using details in literary text - summarize an informational text without including personal opinions - use details from the text to elaborate a key idea in informational text - use evidence from the text to support an author's claim in informational text - summarize information presented in two informational texts - use domain-specific words accurately	In reading, the student is able to: - summarize a literary text from beginning to end without including personal opinions - use details from a literary text to answer specific questions - support inferences about characters using details in literary text - use details from the text to elaborate a key idea in an informational text - use evidence from the text to support an author's claim in informational text - use domain-specific words accurately
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	- use details from a literary text to answer specific questions - use context to identify the meaning of multiple-meaning words	- use details from a literary text to answer specific questions - use context to identify the meaning of multiple-meaning words	
AND in writing, the student is able to: - identify an everyday order of events - use the writing process to create an explanatory product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify elements of an explanatory text to include introduction, body, and conclusion - identify the next event in a brief narrative - use the writing process to create an explanatory product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify transition words and phrases to convey a sequence of events in narrative text - use the writing process to create an explanatory product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: use the writing process to create an explanatory product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 7 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: - identify a theme from a literary text - identify an inference from a literary text - identify a conclusion from an informational text - identify a claim the author makes in an informational text - compare and contrast two statements related to the same topic - use context to identify the meaning of words	In reading, the student is able to: - identify the relationship between individuals or events in an informational text - use evidence from the text to support an author's claim in informational text	In reading, the student is able to: - use details to support a conclusion from informational text - use details to explain how the interactions between individuals, events, or ideas in informational texts are influenced by each other - use evidence from the text to support an author's claim in informational text - compare and contrast how two authors write about the same topic in informational texts - use context to identify the meaning of grade-level phrases	In reading, the student is able to: - use details to support a conclusion from informational text - use details to explain how the interactions between individuals, events, or ideas in informational texts are influenced by each other - use evidence from the text to support an author's claim in informational text - compare and contrast how two authors write about the same topic in informational texts - use context to identify the meaning of grade-level phrases
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	- use details to support themes from literary text - use details to support inferences from literary text	- use details to support themes from literary text - use details to support inferences from literary text	
AND in writing, the student is able to: - identify a graphic that includes an event as described in a text - use the writing process to create an explanatory product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify elements of an explanatory text to include introduction, body, and conclusion - identify the next event in a brief narrative - use the writing process to create an explanatory product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify a sentence that provides a conclusion in narrative text - use the writing process to create an explanatory product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: use the writing process to create an explanatory product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 8 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: • identify a theme from a literary text • identify an inference from a literary text • identify a fact related to a presented argument in informational text • identify a similar topic in two informational texts • use context to identify the meaning of multiple-meaning words • identify the meaning of general academic words	In reading, the student is able to: • use details to support a conclusion from literary text • identify an inference drawn from an informational text • identify the portion of text that contains specific information • identify an argument the author makes in informational text • examine parts of two informational texts to identify where the texts disagree on matters of fact or interpretation • use domain-specific words or phrases accurately	In reading, the student is able to: • use details to support a conclusion from literary text • use details to support an inference from informational text • identify the information (e.g., facts or quotes) in a section of text that contributes to the development of an idea • identify an argument the author makes in informational text • examine parts of two informational texts to identify where the texts disagree on matters of fact or interpretation • use domain-specific words and phrases accurately	In reading, the student is able to: • use details to support a conclusion from literary text • use details to support an inference from informational text • identify the information (e.g., facts or quotes) in a section of text that contributes to the development of an idea • identify an argument the author makes in informational text • examine parts of two informational texts to identify where the texts disagree on matters of fact or interpretation • use domain-specific words and phrases accurately
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	• analyze the development of a theme including the relationship between a character and an event in literary text • use context to identify the meaning of grade-level words and phrases	• analyze the development of a theme including the relationship between a character and an event in literary text • use context to identify the meaning of grade-level words and phrases	
AND in writing, the student is able to: • identify a writer’s opinion • use the writing process to create an explanatory product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: • identify elements of an explanatory text to include introduction, body, and conclusion • identify an idea relevant to a claim • use the writing process to create an explanatory product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: • identify relevant information to support a claim • use the writing process to create an explanatory product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: • use the writing process to create an explanatory product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

High School ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Low text complexity – <i>Brief text with straightforward ideas and relationships; short, simple sentences</i>	Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>
In reading, the student is able to: - identify a summary of a literary text - identify an event from a literary text - identify the central idea of an informational text - identify facts from an informational text - identify what an author tells about a topic in informational text - use context to identify the meaning of multiple-meaning words - identify a word used to describe a person, place, thing, action, or event	In reading, the student is able to: - use details to support a summary of literary text - identify a conclusion from an informational text - identify key details that support the development of a central idea of an informational text - use details presented in two informational texts to answer a question - explain why an author uses specific word choices within texts	In reading, the student is able to: - use details to support a summary of literary text - use details to support a conclusion presented in informational text - identify key details that support the development of a central idea of an informational text - use details presented in two informational texts to answer a question - explain why an author uses specific word choices within texts	In reading, the student is able to: - use details to support a summary of literary text - use details to support a conclusion presented in informational text - identify key details that support the development of a central idea of an informational text - use details presented in two informational texts to answer a question - explain why an author uses specific word choices within texts
	AND with Moderate text complexity – <i>Text with clear, complex ideas and relationships and simple, compound sentences</i>	AND with High text complexity – <i>Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words</i>	
	- evaluate how the author’s use of specific details in literary text contributes to the text - determine an author’s point of view about a topic in informational text - use context to identify the meaning of grade-level phrases	- evaluate how the author’s use of specific details in literary text contributes to the text - determine an author’s point of view about a topic in informational text - use context to identify the meaning of grade-level phrases	
AND in writing, the student is able to: - identify information that is unrelated to a given topic - use the writing process to create an argumentative product and demonstrate minimal (or no) command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify elements of an argument to include introduction, claim, evidence, and conclusion - identify how to group information for a specific text structure - use the writing process to create an argumentative product and demonstrate limited command of organization, idea development, and/or conventions	AND in writing, the student is able to: - identify relevant information to address a given topic and support the purpose of a text - use the writing process to create an argumentative product and demonstrate partial command of organization, idea development, and/or conventions	AND in writing, the student is able to: use the writing process to create an argumentative product and demonstrate overall command of organization, idea development, and/or conventions

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

[Page is intentionally blank]

Grade 3 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: • solve addition problems • identify growing number patterns • identify an object showing a specified number of parts shaded • identify which object has the greater number of parts shaded • identify an object equally divided into two parts • identify the number of objects to be represented in a pictograph	The student is able to: • solve addition and subtraction word problems • identify an arrangement of objects that represents factors in a problem • solve multiplication equations in which both numbers are equal to or less than 5 • identify multiplication patterns • identify a set of objects as nearer to 1 or 10 • identify a representation of the area of a rectangle	The student is able to: • solve addition and subtraction word problems • check the correctness of an answer in the context of a scenario • solve multiplication equations in which both numbers are equal to or less than 5 • identify multiplication patterns • match fraction models to unitary fractions • compare fractions with different numerators and the same denominator • transfer data from an organized list to a bar graph	The student is able to: • solve addition and subtraction word problems • check the correctness of an answer in the context of a scenario • solve multiplication equations in which both numbers are equal to or less than 5 • identify multiplication patterns • match fraction models to unitary fractions • compare fractions with different numerators and the same denominator • transfer data from an organized list to a bar graph
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	• identify geometric figures that are divided into equal parts	• round numbers to the nearest 10 • identify geometric figures that are divided into equal parts • count unit squares to compute the area of a rectangle	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 4 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: - identify an array with the same number of objects in each row - identify values rounded to the nearest tens place - identify equivalent representations of a fraction (e.g., shaded diagram) - compare representations of a fraction (e.g., shaded diagram) - identify a rectangle with the larger or smaller perimeter - identify a given attribute of a shape - identify the data drawn in a bar graph that represents the greatest value	The student is able to: - match a model to a multiplication expression using two single-digit numbers - identify a model of a multiplicative comparison - show division of objects into equal groups - round numbers to the nearest 10, 100, or 1,000 - differentiate parts and wholes - compute the perimeter of a rectangle	The student is able to: - solve multiplication word problems - show division of objects into equal groups - round numbers to the nearest 10, 100, or 1,000 - compare two fractions with different denominators - sort a set of two-dimensional shapes - compute the perimeter of a rectangle - transfer data to a graph	The student is able to: - solve multiplication word problems - show division of objects into equal groups - round numbers to the nearest 10, 100, or 1,000 - compare two fractions with different denominators - sort a set of two-dimensional shapes - compute the perimeter of a rectangle - transfer data to a graph
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	- identify equivalent fractions - select a two-dimensional shape with a given attribute	- solve a multiplicative comparison word problem using up to two-digit numbers - check the correctness of an answer in the context of a scenario - identify equivalent fractions	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 5 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: • solve one-step subtraction word problems • divide sets (no greater than six) into two equal parts • identify values in the tenths place • identify a number in the ones, tens, or hundreds place • identify a given axis of a coordinate plane • match the conversion of 3 feet to 1 yard to a model • calculate elapsed time (i.e., hours) • identify whether the values increase or decrease in a line graph	The student is able to: • identify if the total will increase or decrease when combining sets • perform operations with decimals • identify a symbolic representation of the addition of two fractions • identify place values to the hundredths place • convert standard measurements	The student is able to: • solve multiplication and division word problems • perform operations with decimals • solve word problems involving fractions • identify place values to the hundredths place • locate a given point on a coordinate plane when given an ordered pair • convert standard measurements • convert between minutes and hours • make quantitative comparisons between data sets shown as line graphs	The student is able to: • solve multiplication and division word problems • perform operations with decimals • solve word problems involving fractions • identify place values to the hundredths place • locate a given point on a coordinate plane when given an ordered pair • convert standard measurements • convert between minutes and hours • make quantitative comparisons between data sets shown as line graphs
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	• compare the values of two products based upon multipliers • round decimals to the nearest whole number	• compare the values of two products based upon multipliers • round decimals to the nearest whole number	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 6 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: - identify a model of a given percent - match a given unit rate to a model - identify a representation of two equal sets - identify a number less than 0 on a number line - identify the meaning of an unknown in a modeled equation - count the number of grids or tiles inside a rectangle to find the area of a rectangle - identify the object that appears most frequently in a set of data (mode) - identify a representation of a set of data arranged into even groups (mean)	The student is able to: - match a given ratio to a model - recognize a representation of the sum of two halves - solve real-world measurement problems involving unit rates - identify a representation of a value less than 0 - identify the median or the equation needed to determine the mean of a set of data	The student is able to: - perform operations using up to three-digit numbers - solve real-world measurement problems involving unit rates - identify positive and negative values on a number line - determine the meaning of a value from a set of positive and negative integers - solve word problems with expressions including variables - compute the area of a parallelogram - identify the median or the equation needed to determine the mean of a set of data	The student is able to: - solve real-world measurement problems involving unit rates - identify positive and negative values on a number line - solve word problems with expressions including variables - compute the area of a parallelogram - identify the median or the equation needed to determine the mean of a set of data
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	- perform one-step operations with two decimal numbers - solve word problems using a percent	- perform one-step operations with two decimal numbers - solve word problems using a percent - solve word problems using ratios and rates	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 7 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: - identify a representation that represents a negative number and its multiplication or division by a positive number - identify representations of area and circumference of a circle - identify representations of surface area - make qualitative comparisons when interpreting a data set presented on a bar graph or in a table	The student is able to: - match a given ratio to a model - identify the meaning of an unknown in a modeled equation - describe a directly proportional relationship (i.e., increases or decreases) - find the surface area of a three-dimensional right prism	The student is able to: - solve division problems with positive/negative whole numbers - solve word problems involving ratios - use a proportional relationship to solve a percentage problem - identify proportional relationships between quantities represented in a table - identify unit rate (constant of proportionality) in tables and graphs of proportional relationships - compute the area of a circle - find the surface area of a three-dimensional right prism	The student is able to: - solve division problems with positive/negative whole numbers - solve word problems involving ratios - identify proportional relationships between quantities represented in a table - compute the area of a circle - find the surface area of a three-dimensional right prism
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	- solve multiplication problems with positive/negative whole numbers - interpret graphs to qualitatively contrast data sets	- solve multiplication problems with positive/negative whole numbers - evaluate variable expressions that represent word problems - interpret graphs to qualitatively contrast data sets	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Grade 8 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: - locate a given decimal number on a number line - identify the relatively larger data set when given two data sets presented in a graph - identify congruent rectangles - identify similar rectangles - identify an attribute of a cylinder - identify a rectangle with the larger or smaller area as compared to another rectangle - identify an ordered pair and its point on a graph	The student is able to: - identify the solution to an equation that contains a variable - identify the y-intercept of a linear graph - match a given relationship between two variables to a model - identify a data display that represents a given situation - interpret data presented in graphs to identify associations between variables	The student is able to: - locate approximate placement of an irrational number on a number line - solve a linear equation that contains a variable - identify the relationship shown on a linear graph - calculate slope of a positive linear graph - compute the change in area of a figure when its dimensions are changed - solve for the volume of a cylinder - plot provided data on a graph	The student is able to: - locate approximate placement of an irrational number on a number line - solve a linear equation that contains a variable - identify the relationship shown on a linear graph - compute the change in area of a figure when its dimensions are changed - plot provided data on a graph
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	- identify congruent figures - use properties of similarity to identify similar figures - interpret data tables to identify the relationship between variables	- interpret data presented in graphs to identify associations between variables - interpret data tables to identify the relationship between variables - use properties of similarity to identify similar figures - identify congruent figures	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

High School Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Low task complexity – <i>Simple problems using common mathematical terms and symbols</i>	Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	High task complexity – <i>Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</i>
The student is able to: • arrange a given number of objects into two sets in multiple combinations • match an equation with a variable to a provided real-world situation • determine whether a given point is or is not part of a data set shown on a graph • identify an extension of a linear graph • use a table to match a unit conversion • complete the formula for area of a figure	The student is able to: • identify the model that represents a square number • identify variable expressions that represent word problems • identify the hypotenuse of a right triangle • identify the greatest or least value in a set of data shown on a number line • identify the missing label on a histogram • calculate the mean and median of a set of data	The student is able to: • compute the value of an expression that includes an exponent • identify variable expressions that represent word problems • solve real-world measurement problems that require unit conversions • find the missing attribute of a three-dimensional figure • determine two similar right triangles when a scale factor is given • make predictions from data tables and graphs to solve problems • plot data on a histogram • calculate the mean and median of a set of data	The student is able to: • identify variable expressions that represent word problems • solve real-world measurement problems that require unit conversions • determine two similar right triangles when a scale factor is given • make predictions from data tables and graphs to solve problems • plot data on a histogram • calculate the mean and median of a set of data
	AND with Moderate task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	AND with High task complexity – <i>Common problems presented in mathematical context using various mathematical terms and symbols</i>	
	• identify the linear representation of a provided real-world situation • use an equation or a linear graphical representation to solve a word problem	• identify the linear representation of a provided real-world situation • use an equation or a linear graphical representation to solve a word problem • identify a histogram that represents a provided data set	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

[Page is intentionally blank]

Grade 5 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Policy				
Standards	Level 1 (Beginning – in need of additional support) <i>Students at Level 1 are beginning to access the science content and can be expected to need additional support to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students attempt to perform basic science tasks but will require additional support in order to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations by using disciplinary core ideas, practices, and/or crosscutting concepts to address more basic and concrete science phenomena and problems in Level 1.</i>	Level 2 (Approaching Expectations) <i>Students at Level 2 can be expected to demonstrate developing knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate developing knowledge and skills in some disciplinary core ideas together with some aspects of the practices and crosscutting concepts from the Extended Performance Expectations to address primarily basic and concrete science phenomena and problems in Level 2.</i> <i>At Level 2, students are expected to have the knowledge and skills of Level 1 and may be able to demonstrate some of the knowledge and skills described in Level 3.</i>	Level 3 (Meeting Expectations) <i>Students at Level 3 can be expected to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate knowledge and skills in the majority of disciplinary core ideas, practices, and crosscutting concepts from the K–12 science framework Extended Performance Expectations to address moderately complex science phenomena and problems, some concrete and some abstract in Level 3.</i> <i>At Level 3, students are expected to have the knowledge and skills of Level 2 and may be able to demonstrate some of the knowledge and skills described in Level 4.</i>	Level 4 (Exceeding Expectations) <i>Students at Level 4 can be expected to demonstrate understanding and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate understanding and skills in the disciplinary core ideas, practices, and crosscutting concepts from the K–12 science framework Extended Performance Expectations in more sophisticated ways than students in Level 3 to address science phenomena and problems that are complex, more abstract, and/or multi-factorial. Students are expected to describe, explain, and/or respond to phenomena and problems using reasonably complex evidence, analysis, and inference in Level 4.</i> <i>At Level 4, students are expected to have the knowledge and skills described in Level 3.</i>
Range				
PS-1 Matter and Its Interactions • 5-PS1-2 SEP • Using Mathematics and Computational Thinking CCC • Scale, Proportion, and Quantity	Attempt to identify the appropriate tools or units of measurement (for weight, time, temperature, or volume) for a scientific task.	Identify the appropriate tools or units of measurement (for weight, time, temperature, or volume) for a scientific task.	Compare the weight of matter before and after heating, cooling, or mixing by using data.	Show that the weight of matter does not change when substances are heated, cooled, or mixed by measuring, graphing, or using mathematical relationships.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 5 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
PS-2 Motion and Stability: Forces and Interactions 3-PS2-2 5-PS2-1 SEP - Planning and Carrying Out Investigations (3-PS2-2, Supporting 5-PS2-1) - Engaging in Argument from Evidence (5-PS2-1) - Analyzing and Interpreting Data (Supporting 3-PS2-2) - Developing and Using Models (Supporting 5-PS2-1) CCC - Patterns (3-PS2-2, Supporting 5-PS2-1) - Cause and Effect (5-PS2-1)	Attempt to identify patterns in the motion of an object by using observations or data. Attempt to identify patterns in the motion of falling objects on Earth by using observations.	Identify patterns in the motion of an object by using observations or data. Identify patterns in the motion of falling objects on Earth by using observations.	Predict the future motion of an object by using observations or data. Show the direction objects move when released on Earth (downward toward Earth's center) by identifying or developing a model.	Determine predictable patterns in the motion of an object by describing observations or measurements that can be made in an investigation. Support the claim that Earth's gravity pulls objects downward (toward Earth's center) by describing evidence (observations, data, or a model).
PS-3 Energy 4-PS3-4 5-PS3-1 SEP - Constructing Explanations and Designing Solutions (4-PS3-4) - Developing and Using Models (5-PS3-1) CCC - Energy and Matter (4-PS3-4, 5-PS3-1) - Patterns (Supporting 5-PS3-1)	Attempt to identify various forms of energy present in a system. Attempt to identify that the Sun is a source of energy for ecosystems.	Identify the various forms of energy involved in energy transfers that occur in an everyday object or device. Identify the Sun as a source of energy for ecosystems by using patterns in food chains or drawings of ecosystems.	Describe the various ways that energy transfer can occur between everyday objects or devices. Describe the direction of energy transfer between two organisms (e.g., plant-animal, animal-animal) or between the Sun and a plant by using a model.	Identify which design or improvement will maximize energy transfer from one form to another by designing or modifying a device. Describe how the energy animals obtain from food comes from the Sun by using a model.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 5 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
LS-1 From Molecules to Organisms: Structures and Processes • 4-LS1-1 SEP • Engaging in Argument from Evidence • Developing and Using Models (Supporting) • Analyzing and Interpreting Data (Supporting) CCC • Systems and System Models • Structure and Function (Supporting)	Attempt to identify the parts of plants or animals that have a specific function by using evidence from data and/or a model.	Identify the parts of plants or animals that have specific functions by using evidence from data and/or a model.	Describe how parts of plants or animals have specific functions that help them survive, grow, or reproduce by using data and/or a model.	Describe evidence to support a claim that parts of plants and/or animals have specific functions that help them survive, grow, or reproduce by using evidence from data and/or a model.
LS-3 Heredity: Inheritance and Variation of Traits • 3-LS3-1 SEP • Analyzing and Interpreting Data • Obtaining, Evaluating, and Communicating Information (Supporting) CCC • Patterns	Attempt to identify patterns in trait variations between parents and their baby/babies by using data or observations.	Identify patterns in trait variations between parents and their offspring by using data or observations.	Describe patterns in trait variation between groups of organisms (e.g., parents and their offspring, siblings, populations of similar organisms) by using data or observations.	Describe how patterns in trait variation between groups of organisms (e.g., parents and their offspring, siblings, populations of similar organisms) provide evidence of inheritance between parents and their offspring and that there are differences in these traits by analyzing and interpreting data.
LS-4 Biological Evolution: Unity and Diversity • 3-LS4-1 SEP • Analyzing and Interpreting Data • Obtaining, Evaluating, and Communicating Information (Supporting) CCC • Scale, Proportion, and Quantity	Attempt to recognize that there was life on Earth long ago by using fossils and/or data.	Identify that plants and/or animals lived on Earth long ago by using information about fossils and/or data.	Describe how modern-day plants or animals compare to their ancestors by using observations of fossils and/or data.	Describe the type of environment in which plants and/or animals lived on Earth long ago by using observations of fossils and/or data.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 5 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
ESS-1 Earth's Place in the Universe 5-ESS1-2 SEP - Analyzing and Interpreting Data - Developing and Using Models (Supporting) CCC - Patterns - Systems and System Models (Supporting)	Attempt to identify the positions of the Sun, the Moon, and Earth in the solar system by using data or a model.	Identify the positions of the Sun, the Moon, and Earth in the solar system by using data or a model.	Identify patterns concerning the rotation of Earth, Earth's orbit around the Sun, or the Moon's orbit around Earth by analyzing data (e.g., length and direction of shadows, day and night, seasonal appearance of stars) or a model.	Predict or infer patterns concerning the rotation of Earth, Earth's orbit around the Sun, or the Moon's orbit around Earth by analyzing data (e.g., length and direction of shadows, day and night, seasonal appearance of stars) or a model.
ESS-2 Earth's Systems 3-ESS2-1 5-ESS2-1 SEP - Analyzing and Interpreting Data (3-ESS2-1) - Planning and Carrying Out Investigations (Supporting 3-ESS2-1) - Developing and Using Models (5-ESS2-1) CCC - Patterns (3-ESS2-1) - Systems and Systems Models (5-ESS2-1)	Attempt to describe weather conditions by using observations of weather data. Attempt to identify parts of an Earth system (e.g., geosphere, hydrosphere, atmosphere, biosphere) by using data or a model.	Describe weather conditions by using observations of weather data. Identify parts of an Earth system (e.g., geosphere, hydrosphere, atmosphere, biosphere) by using data or a model.	Describe patterns of weather conditions for a particular season by analyzing weather data. Describe the interaction between two Earth systems (e.g., geosphere, hydrosphere, atmosphere, biosphere) by using a model.	Predict weather conditions for a particular season by analyzing patterns in weather data. Represent the interaction between two Earth systems (e.g., geosphere, hydrosphere, atmosphere, biosphere) by developing a model.
ESS-3 Earth and Human Activity 5-ESS3-1 SEP - Obtaining, Evaluating, and Communicating Information CCC - Cause and Effect (Supporting) - Systems and System Models	Attempt to identify a natural or human impact on the environment by using data.	Identify a natural or human impact on the environment by using data.	Describe an effect (positive or negative) of human activities on the environment by using data.	Describe how humans are using science to protect Earth's resources and/or the environment by using data.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 8 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Policy				
Standards	Level 1 (Beginning – in need of additional support) <i>Students at Level 1 are beginning to access the science content and can be expected to need additional support to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students attempt to perform basic tasks but will require additional support in order to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations by using disciplinary core ideas, practices, and/or crosscutting concepts to address more basic and concrete science phenomena and problems in Level 1.</i>	Level 2 (Approaching Expectations) <i>Students at Level 2 can be expected to demonstrate developing knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate developing knowledge and skills in some disciplinary core ideas together with some aspects of the practices and crosscutting concepts from the K–12 science framework Extended Performance Expectations to address primarily basic and concrete science phenomena and problems at Level 2.</i> <i>At Level 2, students are expected to have the knowledge and skills of Level 1 and may be able to demonstrate some of the knowledge and skills described in Level 3.</i>	Level 3 (Meeting Expectations) <i>Students at Level 3 can be expected to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate knowledge and skills in the majority of disciplinary core ideas, practices, and crosscutting concepts from the K–12 science framework Extended Performance Expectations to address moderately complex science phenomena and problems, some concrete and some abstract at Level 3.</i> <i>At Level 3, students are expected to have the knowledge and skills of Level 2 and may be able to demonstrate some of the knowledge and skills described in Level 4.</i>	Level 4 (Exceeding Expectations) <i>Students at Level 4 can be expected to demonstrate understanding and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate understanding and skills in the disciplinary core ideas, practices, and crosscutting concepts from the K–12 science framework Extended Performance Expectations in more sophisticated ways than students in Level 3 to address science phenomena and problems that are complex, more abstract, and/or multi-factorial. Students are expected to describe, explain, and/or respond to phenomena and problems using reasonably complex evidence, analysis, and inference at Level 4.</i> <i>At Level 4, students are expected to have the knowledge and skills described in Level 3.</i>
Range				
PS-1 Matter and Its Interactions • MS-PS1-2 SEP • Analyzing and Interpreting Data • Planning and Carrying Out Investigations (Supporting) CCC • Patterns • Scale, Proportion, and Quantity (Supporting)	Attempt to identify properties of a substance by using data or observations.	Identify properties of a substance by using data or observations.	Determine the identities of substances by using data or observations on the properties of substances.	Determine whether a chemical reaction occurred by using data or observations on the properties of substances before and after an interaction.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 8 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
PS-2 Motion and Stability: Forces and Interactions - MS-PS2-2 SEP - Planning and Carrying Out Investigations - Analyzing and Interpreting Data (Supporting) CCC - Stability and Change - Cause and Effect (Supporting)	Attempt to identify the effects of pushes and pulls on objects by using data from an investigation.	Identify the effects of pushes and pulls on objects by using data from an investigation.	Identify the change in an object's motion when the mass of the object or the force on the object is changed by using data from an investigation.	Describe how the mass of an object or the force on an object will change the motion of the object by using data from an investigation.
PS-3 Energy - MS-PS3-5 SEP - Engaging in Argument from Evidence - Asking Questions and Defining Problems (Supporting) - Analyzing and Interpreting Data (Supporting) CCC - Energy and Matter	Attempt to determine whether energy is being transferred in a system by asking questions or by using data.	Determine whether energy is being transferred in a system by asking questions or by using data.	Identify the forms of energy that increase or decrease when the kinetic energy of an object changes by using data as evidence.	Make or support a claim that a transfer of energy occurs when the kinetic energy of an object changes by using data as evidence.
PS-4 Waves and Their Applications in Technologies for Information Transfer - MS-PS4-2 SEP - Developing and Using Models - Planning and Carrying Out Investigations (Supporting) CCC - Structure and Function	Attempt to identify whether a wave is being reflected, absorbed, or transmitted through a material by using data or a model.	Identify whether a wave is being reflected, absorbed, or transmitted through a material by using data or a model.	Describe the path of a wave that is reflected, absorbed, or transmitted through different materials by using a model.	Represent what happens to waves when they are reflected, absorbed, or transmitted through different materials by developing a model.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 8 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
LS-1 From Molecules to Organisms: Structures and Processes - MS-LS1-3 - MS-LS1-5 SEP - Engaging in Argument from Evidence (MS-LS1-3) - Developing and Using Models (Supporting MS-LS1-3) - Obtaining, Evaluating, and Communicating Information (Supporting MS-LS1-3) - Constructing Explanations and Designing Solutions (MS-LS1-5) - Analyzing and Interpreting Data (Supporting MS-LS1-5) - Asking Questions and Defining Problems (Supporting MS-LS1-5) CCC - Systems and System Models (MS-LS1-3) - Cause and Effect (MS-LS1-5)	Attempt to identify structures that are part of human body systems and those that are not by using charts, diagrams, or graphic organizers. Attempt to identify factors that could be affecting the growth of an organism by asking questions.	Identify structures that are part of human body systems and those that are not by using charts, diagrams, or graphic organizers. Identify factors that could be affecting the growth of an organism by asking questions.	Identify those parts that belong to a particular body system and the organization of those parts by using a model. Determine whether a particular factor is affecting the growth of organisms by analyzing data.	Make a claim about two body systems (e.g., circulatory, respiratory, muscular, digestive, nervous, excretory) working together to carry out various functions by using evidence. Explain how the growth of organisms is influenced by various environmental and/or genetic factors by using data.
LS-2 Ecosystems: Interactions, Energy, and Dynamics - MS-LS2-1 - MS-LS2-3 SEP - Analyzing and Interpreting Data (MS-LS2-1) - Developing and Using Models (MS-LS2-3) CCC - Cause and Effect (MS-LS2-1) - Energy and Matter (MS-LS2-3)	Attempt to identify resources (e.g., food, water, nutrients, space) that are necessary for the growth or survival of organisms or populations of organisms by using data. Attempt to identify the role of organisms (e.g., producer, consumer, decomposer) or nonliving things (e.g., the Sun, water, minerals, air) in cycling energy or matter in an ecosystem by using a model.	Identify resources (e.g., food, water, nutrients, space) that are necessary for the growth or survival of organisms or populations of organisms by using data. Identify the role of organisms (e.g., producer, consumer, decomposer) or nonliving things (e.g., the Sun, water, minerals, air) in cycling energy or matter in an ecosystem by using a model.	Describe the effects of resource availability on organisms and/or populations of organisms by using data or observations. Identify how energy is transferred or that matter is cycled from one specific part of an ecosystem to another specific part by using a model.	Identify evidence of a cause-effect relationship between resource availability and growth of organisms and/or populations of organisms by analyzing data. Describe how energy is transferred or how matter is cycled among living and nonliving parts of ecosystems by developing a model.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Grade 8 Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
ESS-1 Earth's Place in the Universe - MS-ESS1-1 SEP - Developing and Using Models CCC - Patterns - Systems and System Models (Supporting)	Attempt to show the positions of Earth (with its tilt), the Sun, and the Moon as Earth orbits the Sun and the Moon orbits Earth in the solar system by identifying a model.	Show the positions of Earth (with its tilt), the Sun, and the Moon as Earth orbits the Sun and the Moon orbits Earth in the solar system by identifying a model.	Describe or compare the positions of the Sun, the Moon, and Earth or the amount or path of light in the cyclic patterns of seasons, lunar phases, or eclipses by using a model.	Compare or show patterns in seasons, lunar phases, or eclipses by using or developing a model of the Earth-Sun-Moon system.
ESS-2 Earth's Systems - MS-ESS2-2 - MS-ESS2-4 SEP - Constructing Explanations (MS-ESS2-2) - Obtaining, Evaluating, and Communicating Information (Supporting MS-ESS2-2) - Developing and Using Models (MS-ESS2-4) CCC - Scale, Proportion, and Quantity (MS-ESS2-2) - Cause and Effect (Supporting MS-ESS2-2) - Energy and Matter (MS-ESS2-4)	Attempt to identify the process or agent that causes a particular change to Earth's surface by using observations as evidence. Attempt to trace the path of water through Earth's systems by using a model.	Identify the process or agent that causes a particular change to Earth's surface by using observations as evidence. Trace the path of water through Earth's systems by using a model.	Identify whether a geological process or event on Earth was small/ large scale and/or whether a process or event happened gradually/rapidly by using information in charts, diagrams, or graphic organizers. Describe the state of water or how water changes state in various parts of the water cycle by using a model.	Explain how geological processes on Earth have caused changes to Earth's surface at various times or spatial scales by using evidence to support an explanation. Describe how the Sun's energy or the force of gravity move water through the water cycle by developing a model.
ESS-3 Earth and Human Activity - MS-ESS3-3 SEP - Constructing Explanations and Designing Solutions - Engaging in Argument from Evidence (Supporting MS-ESS3-3) - Asking Questions and Defining Problems (Supporting MS-ESS3-3) CCC - Cause and Effect	Attempt to identify an environmental problem caused by human activities/impact by using data.	Identify an environmental problem caused by human activities/impacts by using data.	Make a claim about how a particular method would work to reduce a human impact on the environment by using data.	Select or evaluate a design for a method that is intended to minimize a human impact on the environment by using data.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

High School Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Policy				
Standards	Level 1 (Beginning – in need of additional support) <i>Students at Level 1 are beginning to access the science content and can be expected to need additional support to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students attempt to perform basic tasks but will require additional support in order to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations by using disciplinary core ideas, practices, and/or crosscutting concepts to address more basic and concrete science phenomena and problems in Level 1.</i>	Level 2 (Approaching Expectations) <i>Students at Level 2 can be expected to demonstrate developing knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate developing knowledge and skills in some disciplinary core ideas together with some aspects of the practices and crosscutting concepts from the K–12 science framework Extended Performance Expectations to address primarily basic and concrete science phenomena and problems at Level 2.</i> <i>At Level 2, students are expected to have the knowledge and skills of Level 1 and may be able to demonstrate some of the knowledge and skills described in Level 3.</i>	Level 3 (Meeting Expectations) <i>Students at Level 3 can be expected to demonstrate knowledge and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate knowledge and skills in the majority of disciplinary core ideas, practices, and crosscutting concepts from the K–12 science framework Extended Performance Expectations to address moderately complex science phenomena and problems, some concrete and some abstract at Level 3.</i> <i>At Level 3, students are expected to have the knowledge and skills of Level 2 and may be able to demonstrate some of the knowledge and skills described in Level 4.</i>	Level 4 (Exceeding Expectations) <i>Students at Level 4 can be expected to demonstrate understanding and skills of the K–12 science framework Extended Performance Expectations.</i> <i>Students can be expected to demonstrate understanding and skills in the disciplinary core ideas, practices, and crosscutting concepts from the K–12 science framework Extended Performance Expectations in more sophisticated ways than students in Level 3 to address science phenomena and problems that are complex, more abstract, and/or multi-factorial. Students are expected to describe, explain, and/or respond to phenomena and problems using reasonably complex evidence, analysis, and inference at Level 4.</i> <i>At Level 4, students are expected to have the knowledge and skills described in Level 3.</i>
Range				
PS-1 Matter and Its Interactions - HS-PS1-2 SEP - Constructing Explanations and Designing Solutions - Obtaining, Evaluating, and Communicating Information (Supporting) - Developing and Using Models (Supporting) CCC - Patterns - Energy and Matter (Supporting)	Attempt to show how substances react in a chemical reaction by using provided information to complete an incomplete chemical reaction model.	Show how substances react by using provided information to complete an incomplete chemical reaction model.	Identify or classify elements that will react similarly in chemical reactions by using a periodic table model.	Construct an explanation for why specific chemical reactions occur by using a periodic table.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

High School Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
PS-2 Motion and Stability: Forces and Interactions - HS-PS2-3 - HS-PS2-5 SEP - Constructing Explanations and Designing Solutions (HS-PS2-3) - Engaging in Argument from Evidence (Supporting HS-PS2-3) - Developing and Using Models (Supporting HS-PS2-3) - Planning and Carrying Out Investigations (HS-PS2-5) - Analyzing and Interpreting Data (Supporting HS-PS2-5) CCC - Cause and Effect (HS-PS2-3, HS-PS2-5) - Systems and System Models (Supporting HS-PS2-3) - Stability and Change (Supporting HS-PS2-5)	Attempt to identify how forces are acting on a macroscopic object during a collision in a model. Attempt to identify examples of electric current producing magnetic fields or magnetic fields producing electric current by using data or observations.	Identify how forces are acting on a macroscopic object during a collision in a model. Identify examples of electric current producing magnetic fields or magnetic fields producing electric current by using data or observations.	Construct a claim for how a familiar device functions to minimize the forces on a macroscopic object during a collision. Predict or draw conclusions about how a change to a system affects how electric current produces magnetic fields or how magnetic fields produce electric current by using data.	Select, evaluate, or revise the design of a familiar device that minimizes the forces on a macroscopic object during a collision. Plan or conduct an investigation to determine cause-and-effect relationships between magnetic fields and electric current.
PS-3 Energy - HS-PS3-2 SEP - Developing and Using Models - Asking Questions and Defining Problems (Supporting) CCC - Energy and Matter	Attempt to identify questions that would determine whether an object's kinetic or potential energy is changing in a system.	Identify questions that would determine whether an object's kinetic or potential energy is changing in a system.	Show how kinetic and potential energy change in a system when an object's position changes or when the particles making up an object change their motion by using a model.	Develop or use models to describe how energy is conserved at the macroscopic or particle level when kinetic and/or potential energy are transferred or converted from one form to another in a system.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

High School Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
LS-2 Ecosystems: Interactions, Energy, and Dynamics - HS-LS2-2 SEP - Using Mathematics and Computational Thinking - Obtaining, Evaluating, and Communicating Information (Supporting) CCC - Scale, Proportion, and Quantity - Cause and Effect (Supporting)	Attempt to identify factors that affect population size or biodiversity by using provided information.	Identify factors that affect population size or biodiversity by using provided information.	Describe how a factor affects population size or biodiversity in an ecosystem by interpreting data.	Explain how a factor affects population size or biodiversity in an ecosystem at different scales (e.g., habitat size compared to population size) by using mathematical representations of data.
LS-3 Heredity: Inheritance and Variation of Traits - HS-LS3-1 SEP - Asking Questions and Defining Problems - Developing and Using Models (Supporting) - Obtaining, Evaluating, and Communicating Information (Supporting) CCC - Cause and Effect - Structure and Function (Supporting) - Patterns (Supporting)	Attempt to identify the function of DNA or chromosomes by using provided information.	Identify the function of DNA or chromosomes by using provided information.	Describe how genes and traits are inherited from parents to offspring by using a model.	Ask questions that will provide information about the cause-and-effect relationships among DNA/chromosomes and/or traits that are inherited from parents to offspring.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

High School Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
LS-4 Biological Evolution: Unity and Diversity - HS-LS4-1 - HS-LS4-3 SEP - Obtaining, Evaluating, and Communicating Information (HS-LS4-1, Supporting HS-LS4-3) - Analyzing and Interpreting Data (HS-LS4-3, Supporting HS-LS4-1) CCC - Patterns (HS-LS4-1, HS-LS4-3) - Stability and Change (Supporting HS-LS4-1)	Attempt to identify how organisms have changed over time by using provided information. Attempt to identify physical traits that can vary in an organism by using provided information.	Identify how organisms have changed over time by using provided information. Identify physical traits that can vary in an organism by using provided information.	Draw conclusions about patterns of relatedness among organisms by using data (e.g., DNA sequences, amino acid sequences, structures found in organisms, embryos, fossils). Describe changes in the distribution of physical traits that can vary in a population by using data.	Describe how comparing patterns in data (e.g., DNA sequences, amino acid sequences, structures found in organisms, embryos, fossils) provide evidence for evolution and common ancestry of living things. Demonstrate that organisms with helpful traits increase in proportion to organisms lacking those traits by using data as evidence.
ESS-1 Earth's Place in the Universe - HS-ESS1-6 SEP - Constructing Explanations and Designing Solutions - Asking Questions and Defining Problems (Supporting) - Analyzing and Interpreting Data (Supporting) CCC - Stability and Change - Patterns (Supporting)	Attempt to identify patterns in data about ancient Earth materials, meteorites, or other planetary surfaces by using data.	Identify patterns in data about ancient Earth materials, meteorites, or other planetary surfaces by using data.	Describe Earth's formation and early history by asking questions about ancient Earth materials, meteorites, and other planetary surfaces.	Explain Earth's formation and early history by using data about ancient Earth materials, meteorites, or other planetary surfaces.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

High School Science Performance Level Descriptors

(for American Samoa, Arizona, BIE, CNMI, Guam, Maine, USVI, and Vermont)

Range				
ESS-2 Earth's Systems - HS-ESS2-4 - HS-ESS2-5 SEP - Developing and Using Models (HS-ESS2-4) - Planning and Carrying Out Investigations (HS-ESS2-5) - Analyzing and Interpreting Data (Supporting HS-ESS2-5) - Asking Questions and Defining Problems (Supporting HS-ESS2-5) CCC - Cause and Effect (HS-ESS2-4, Supporting HS-ESS2-5) - Energy and Matter (Supporting HS-ESS2-4) - Structure and Function (HS-ESS2-5)	Attempt to identify how energy flows between two Earth systems by using a model. Attempt to identify testable questions about how water affects Earth's materials and surface processes.	Identify how energy flows between two Earth systems by using a model. Identify testable questions about how water affects Earth's materials and surface processes.	Describe how energy from the Sun drives Earth's climate system by using a model. Use data or observations to draw conclusions about how water affects Earth's materials and surface processes.	Predict or draw conclusions about how various factors (e.g., large volcanic eruptions, human activity, solar output, changes to Earth's orbit and axis, changes to atmospheric composition, etc.) cause changes to Earth's climate (measured as changes in surface temperatures, precipitation patterns, glacial ice volumes, sea levels, biosphere distribution) by using models. Plan or conduct an investigation of the properties of water and its effects on Earth materials and surface processes (e.g., stream transportation and deposition using a stream table, frost wedging by the expansion of water as it freezes, or chemical weathering and recrystallization by testing the solubility of different materials).
ESS-3 Earth and Human Activity - HS-ESS3-4 SEP - Constructing Explanations and Designing Solutions - Engaging in Argument from Evidence (Supporting) - Analyzing and Interpreting Data (Supporting) CCC - Stability and Change - Cause and Effect (Supporting)	Attempt to identify the impact of positive or negative local human activities on natural systems by using data.	Identify the positive or negative impacts of local human activities on natural systems by using data.	Construct a claim about how a local technological solution reduces the negative impact of human activities on natural systems.	Evaluate or refine the design of a local technological solution that reduces the negative impact of human activities on natural systems.

SEP—Science and Engineering Practices; CCC—Crosscutting Concepts; ESS—Earth and Space Science; PS—Physical Science; LS—Life Science

Appendix C: Scale Score Ranges

Table 1. 2025 Performance-Level Scale Score Ranges for ELA and Mathematics

Performance Level	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	High School
English Language Arts							
Level 4	1254–1290	1259–1290	1256–1290	1251–1290	1255–1290	1250–1290	1255–1290
Level 3	1240–1253	1240–1258	1240–1255	1237–1250	1240–1254	1238–1249	1240–1254
Level 2	1234–1239	1234–1239	1232–1239	1231–1236	1236–1239	1230–1237	1236–1239
Level 1	1200–1233	1200–1233	1200–1231	1200–1230	1200–1235	1200–1229	1200–1235
Mathematics							
Level 4	1254–1290	1251–1290	1253–1290	1251–1290	1254–1290	1251–1290	1250–1290
Level 3	1242–1253	1239–1250	1240–1252	1239–1250	1240–1253	1240–1250	1240–1249
Level 2	1235–1241	1232–1238	1232–1239	1233–1238	1234–1239	1234–1239	1235–1239
Level 1	1200–1234	1200–1231	1200–1231	1200–1232	1200–1233	1200–1233	1200–1234

Table 2. 2025 Performance-Level Scale Score Ranges for Science

Performance Level	Grade 5	Grade 8	High School
Level 4	1247–1290	1254–1290	1251–1290
Level 3	1240–1246	1240–1253	1240–1250
Level 2	1236–1239	1236–1239	1236–1239
Level 1	1200–1235	1200–1235	1200–1235

Appendix D: Individual Student Report Samples

CONFIDENTIAL



Name: FIRSTNAME LASTNAME
ID: D0164
School: Demonstration School
Test Date: Spring 2025
Grade: 05

What Is In This Report?

Page 1: Contains a summary of your child's performance on this year's test.

Page 2: Contains an introductory letter from MSAA and next steps to support your child.

Performance Summary

FIRSTNAME's performance in English Language Arts and Mathematics is described below.

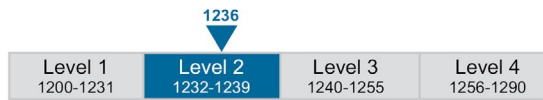
English Language Arts

Performance Level

Level 2

Score

1236



Meets Expectations

A student's test score can vary. If your child were to be tested again, it is likely that they would receive a score between 1234 and 1238.

English Language Arts consists of Reading and Writing. See below for percent of possible points earned in each area.

Reading 73%

Writing 25%

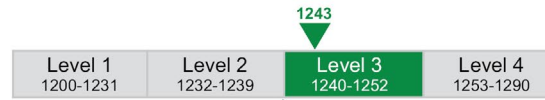
Mathematics

Performance Level

Level 3

Score

1243



Meets Expectations

A student's test score can vary. If your child were to be tested again, it is likely that they would receive a score between 1240 and 1246.

Performance Level Descriptors

The scale score and performance level for each content area above summarize FIRSTNAME's performance on the English Language Arts (ELA) and Mathematics tests. The performance level descriptors below describe the knowledge and skills that children who perform at this level generally demonstrate.

English Language Arts

- ✓ use brief literary texts with clear ideas to answer questions about the text; compare characters, settings, and events; and summarize a text
- ✓ use brief informational texts with clear ideas to identify the main idea and supporting details, use details from the text to support an author's point, and compare and contrast information and events in different texts
- ✓ use context to define multiple-meaning words
- ✓ identify elements of a narrative text to include beginning, middle, and end
- ✓ identify a sentence that is organized for a text structure such as comparison/contrast
- ✓ write a narrative with limited command of organization, idea development, and/or conventions

Mathematics

- ✓ solve problems with whole numbers, fractions, or decimals using mathematical language and symbolic representations (e.g., $<$, $>$, $=$)
- ✓ identify place values
- ✓ round decimals
- ✓ identify the effects of multiplication
- ✓ convert standard measurements including minutes and hours
- ✓ locate a given point on a coordinate plane
- ✓ make comparisons between data sets

Dear Parents and Guardians,

This report summarizes your child's performance on the online 2025 Multi-State Alternate Assessment (MSAA). This report shows the scaled score and performance levels in English Language Arts (ELA) and Mathematics. Also shown is the percent of possible points earned in Reading and Writing. The performance level descriptors describe the knowledge and skills that children who perform at this level generally demonstrate.

The MSAA is designed to assess students in grades 3-8 and High School with significant cognitive disabilities and measures academic content that is aligned to and derived from your state's content standards. The test contains many built-in supports that allow students to take the test using materials they are most familiar with and to communicate what they know and can do. These are some of the built-in supports found in the MSAA:

- shortened ELA reading passages
- pictures, charts, tables, and maps to help students understand the reading passages
- models and examples that explain important ideas and concepts
- smaller numbers on the mathematics tests

To support communication independence to the greatest extent possible, the MSAA is designed to work with different communication modes and systems. Please discuss the supports your child used on the MSAA with your child's teacher.

More information and resources for helping your child are available at your state's alternate assessment web page or by talking with your child's teacher. If you require this letter or your child's report in a different format, please contact your state's department of education.

What skills can be worked on next?

English Language Arts

- + Summarize a text
- + Summarize a text and use inferences
- + Use content vocabulary
- + Use transition words in writing

Mathematics

- + Use mathematical terms and symbols (<, >, =)
- + Solve problems related to percent, rates, and ratios
- + Find the area of a parallelogram
- + Identify numbers on a number line
- + Solve word problems
- + Identify mean, median, and mode
- + Solve equations with decimals

What now?

Bring this report to your next conference with FIRSTNAME's teachers.

You can ask FIRSTNAME's teachers:

- What is FIRSTNAME learning in ELA and Mathematics this year?
- How is FIRSTNAME doing?
- How can I use this information to work with FIRSTNAME this year?
- What resources should I use to support FIRSTNAME?



Name: FIRSTNAME LASTNAME
ID: D137
School: Demonstration School
Test Date: Spring 2025
Grade: 05

What Is In This Report?

Page 1: Contains a summary of your child's performance on this year's test.
Page 2: Contains an introductory letter from MSAA and next steps to support your child.

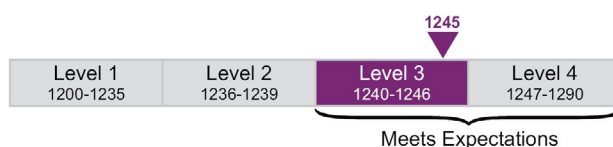
Performance Summary

FIRSTNAME's performance in Science is described below.

Science

Performance Level
Level 3

Score
1245



A student's test score can vary. If your child were to be tested again, it is likely that they would receive a score between 1243 and 1247.

Performance Level Descriptors

The scale score and performance level for the content area above summarizes FIRSTNAME's performance on the Science test. The performance level descriptors below describe the knowledge and skills that children who perform at this level generally demonstrate.

Science

- ✓ Compare the weight of matter before and after a change
- ✓ Predict the future motion of an object
- ✓ Show the direction objects move when released on Earth
- ✓ Describe energy transfer between everyday objects or devices
- ✓ Describe the direction of energy transfer between two organisms
- ✓ Describe how parts of plants or animals have specific functions
- ✓ Describe patterns in trait variations between groups of organisms
- ✓ Describe how modern-day plants or animals compare to their ancestors
- ✓ Identify patterns of rotation or Earth's orbit around the Sun or the Moon's orbit around Earth
- ✓ Describe patterns of weather conditions for a particular season
- ✓ Describe the interaction between two Earth systems
- ✓ Describe an effect of human activities on the environment

Dear Parents and Guardians,

This report summarizes your child's performance on the online 2025 Multi-State Alternate Assessment (MSAA). This report shows the scaled score and performance levels in Science. The performance level descriptors describe the knowledge and skills that children who perform at this level generally demonstrate.

The MSAA Science is designed to assess students in grades 5, 8, and High School with significant cognitive disabilities and measures academic content that is aligned to and derived from your state's science content standards. The test contains many built-in supports that allow students to participate using materials they are most familiar with and to communicate what they know and can do. These are some of the built-in supports found in the MSAA Science:

- pictures, charts, tables, and diagrams to help students understand the science concept
- models and examples that explain important ideas and concepts
- use of concrete science terminology and scenarios

To support communication independence to the greatest extent possible, the MSAA is designed to work with different communication modes and systems. Please discuss the supports your child used on the MSAA with your child's teacher.

More information and resources for helping your child are available at your state's alternate assessment web page or by talking with your child's teacher. If you require this letter or your child's report in a different format, or if you have questions about provisional performance levels and scaled scores, please contact your state's department of education.

What skills can be worked on next?

Science

- + Use charts, graphs, and models to answer questions
- + Focus on physical science concepts such as,
 - the identities of substances
 - motion
 - changes in forms of energy
 - paths of waves
- + Focus on life science concepts such as,
 - body systems
 - organism growth
 - effects of resource availability
 - energy transfer in an ecosystem
- + Focus on Earth and space science concepts such as,
 - positions of the Sun, the Moon, and Earth
 - seasons, lunar phases, and eclipses
 - geological processes or events on Earth
 - the water cycle
 - human impacts on the environment

What now?

Bring this report to your next conference with FIRSTNAME's teachers.

You can ask FIRSTNAME's teachers:

- What is FIRSTNAME learning in Science this year?
- How is FIRSTNAME doing?
- How can I use this information to work with FIRSTNAME this year?
- What resources should I use to support FIRSTNAME?

APPENDIX H
DIFFERENTIAL ITEM FUNCTIONING RESULTS

Table H-1. DIF—ELA

Grade	Group Reference	Group Focal	Number “Low” Favoring Focal	Number “Low” Favoring Reference	Number “High” Favoring Focal	Number “High” Favoring Reference
3	White	BAA	2	4	NA	1
3	White	Hispanic	6	6	NA	NA
3	M	F	8	5	1	2
4	White	BAA	3	4	2	NA
4	White	Hispanic	9	8	NA	NA
4	M	F	8	6	1	2
5	White	BAA	7	1	NA	NA
5	White	Hispanic	5	9	2	NA
5	M	F	4	8	1	1
6	White	BAA	5	2	1	NA
6	White	Hispanic	6	4	NA	1
6	M	F	4	3	1	2
7	N	Y	NA	1	2	1
7	White	BAA	6	6	1	NA
7	White	Hispanic	10	2	NA	1
7	M	F	4	10	1	NA
8	White	BAA	3	1	1	NA
8	White	Hispanic	5	3	1	NA
8	M	F	5	8	1	2
11	White	BAA	4	3	NA	NA
11	White	Hispanic	8	4	1	NA
11	M	F	5	6	2	3

Table H-2. DIF—Mathematics

Grade	Grade	Group Reference	Group Focal	Number “Low” Favoring Focal	Number “Low” Favoring Reference	Number “High” Favoring Focal
3	White	BAA	3	NA	NA	NA
3	White	Hispanic	10	9	NA	1
3	M	F	4	6	1	1
4	White	BAA	2	3	NA	1
4	White	Hispanic	10	6	3	2
4	M	F	5	6	NA	1
5	White	BAA	2	2	NA	NA
5	White	Hispanic	6	5	NA	1
5	M	F	10	10	1	NA
6	White	BAA	3	4	NA	NA
6	White	Hispanic	10	9	1	NA
6	M	F	6	7	1	NA
7	N	Y	2	3	NA	2
7	White	BAA	1	3	1	1
7	White	Hispanic	5	8	3	2
7	M	F	9	4	2	2
8	White	BAA	2	3	NA	NA
8	White	Hispanic	4	8	4	4
8	M	F	8	9	1	NA
11	White	BAA	4	2	1	NA
11	White	Hispanic	6	5	1	1
11	M	F	9	6	NA	NA

APPENDIX I
ITEM RESPONSE THEORY PARAMETERS

Image I-1. Boxplot of Item Difficulty by Level for ELA—Grade 3

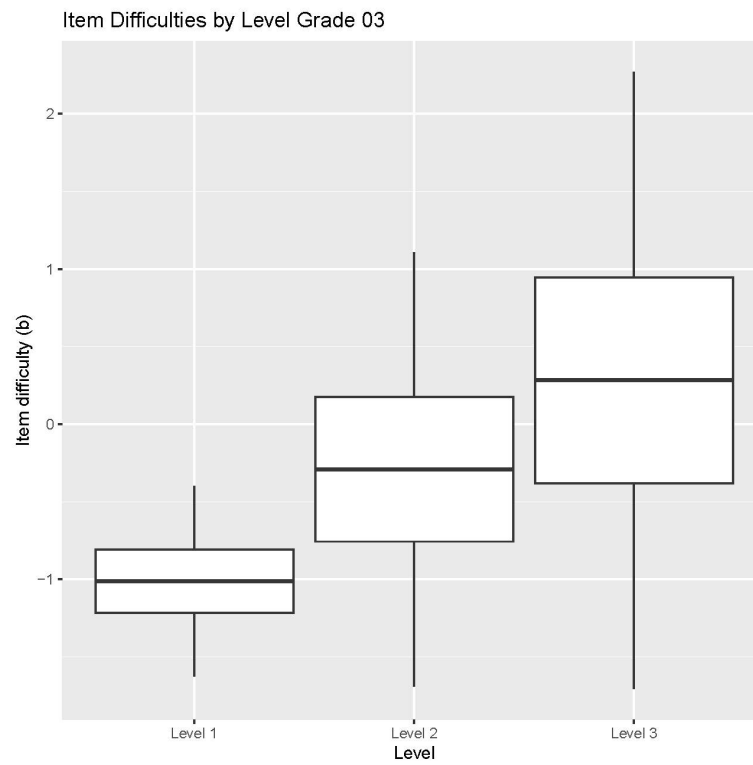


Image I-2. Boxplot of Item Difficulty by Level for ELA—Grade 4

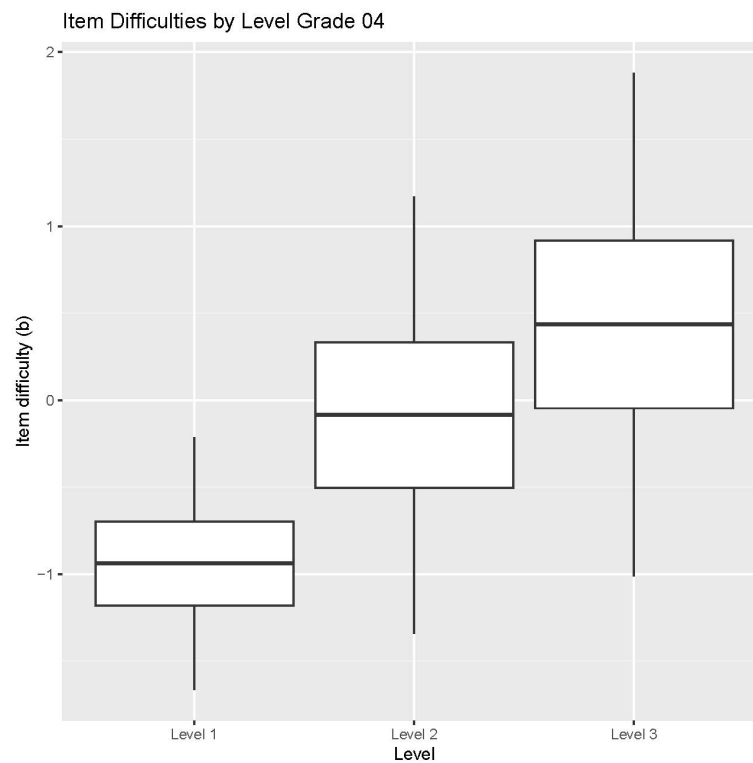


Image I-3. Boxplot of Item Difficulty by Level for ELA—Grade 5

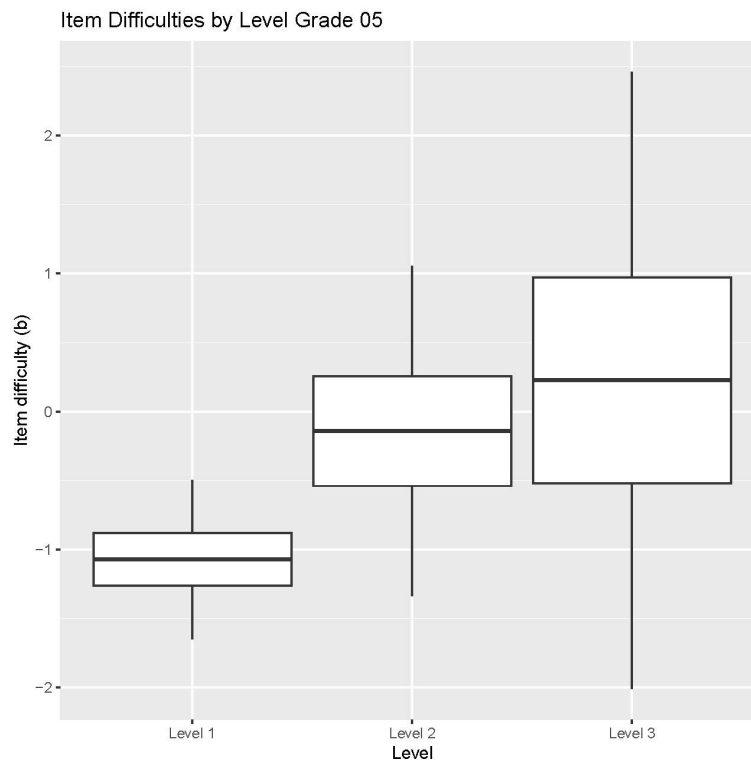


Image I-4. Boxplot of Item Difficulty by Level for ELA—Grade 6

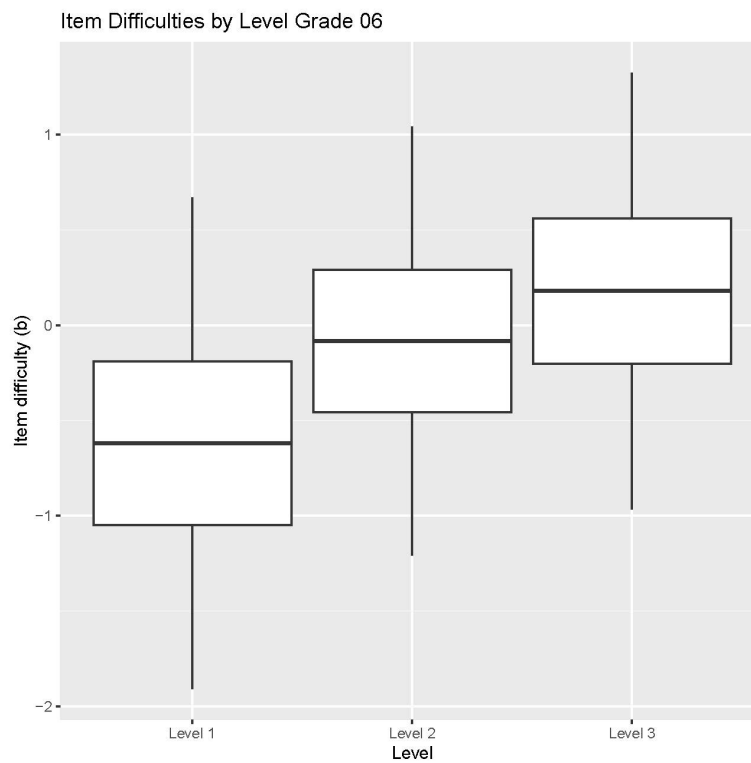


Image I-5. Boxplot of Item Difficulty by Level for ELA—Grade 7

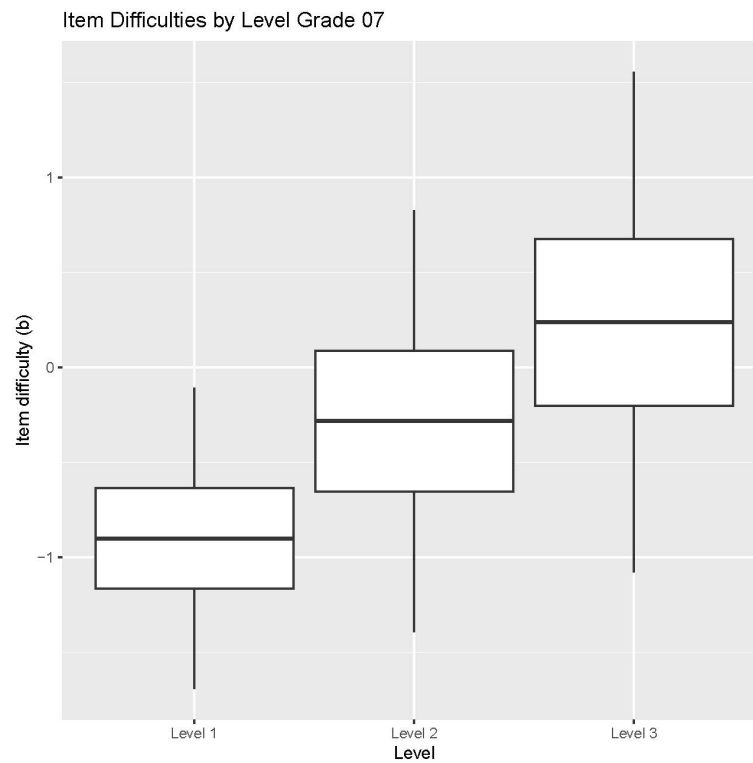


Image I-6. Boxplot of Item Difficulty by Level for ELA—Grade 8

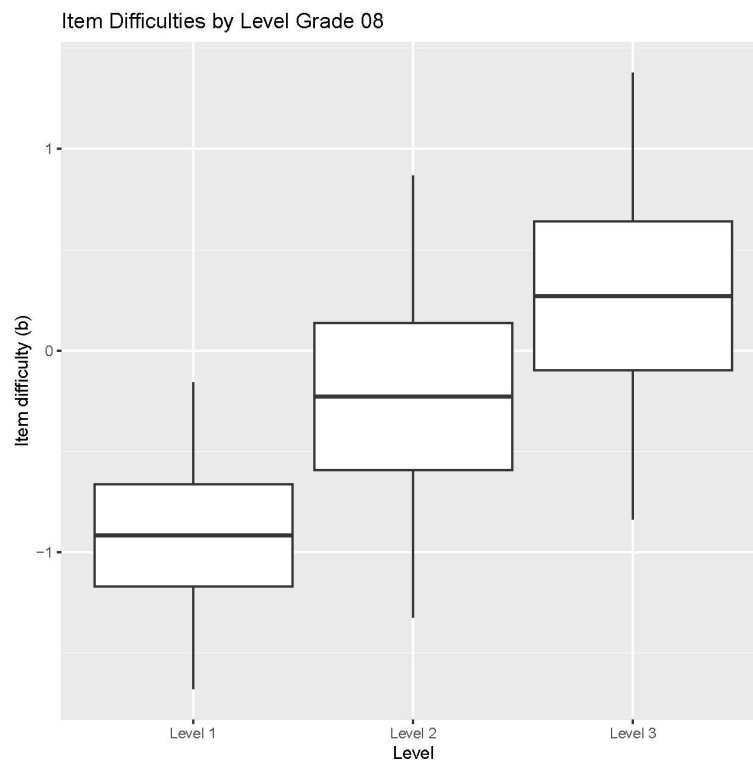


Image I-7. Boxplot of Item Difficulty by Level for ELA—Grade 11

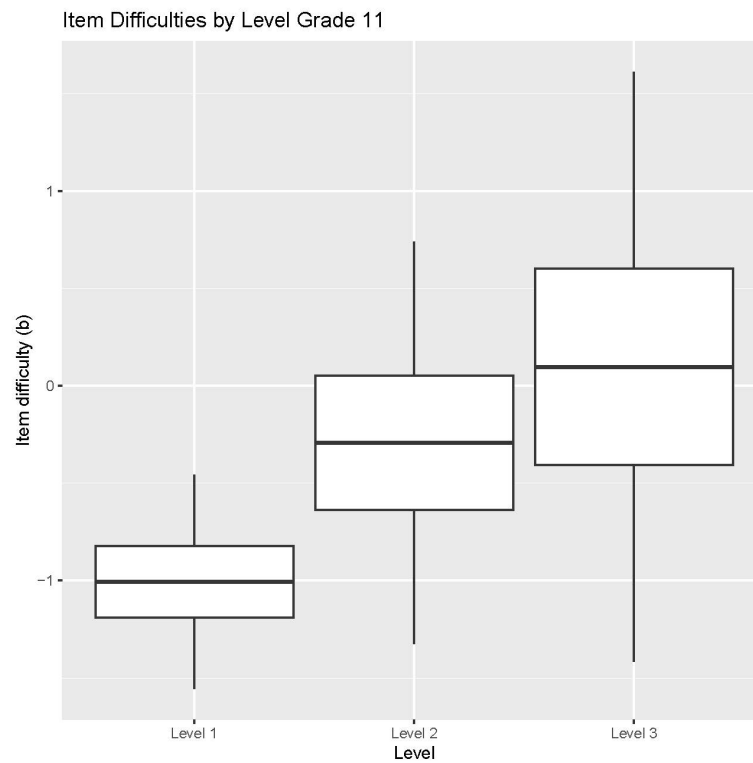


Image I-8. Boxplot of Item Difficulty for Mathematics—Grade 3

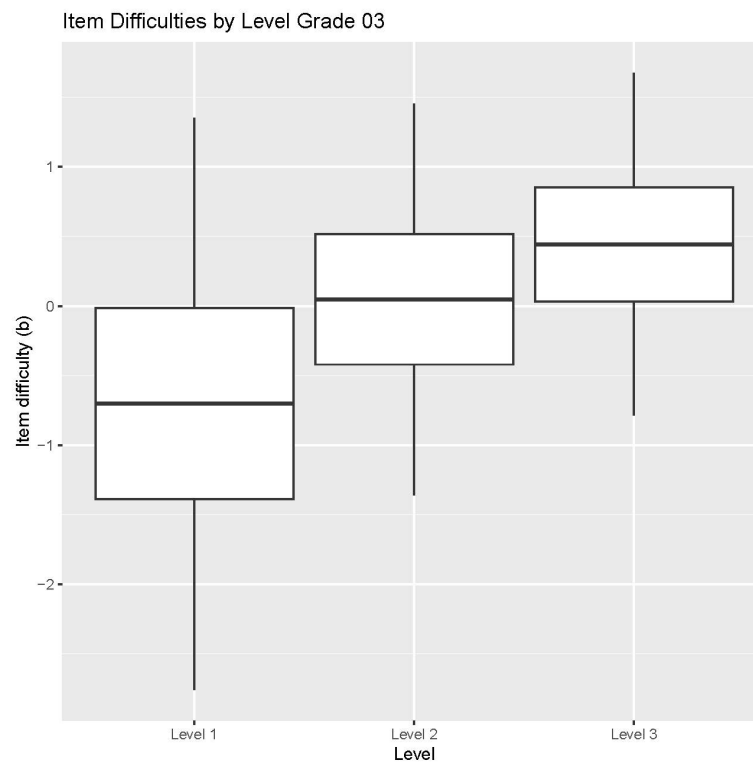


Image I-9. Boxplot of Item Difficulty for Mathematics—Grade 4

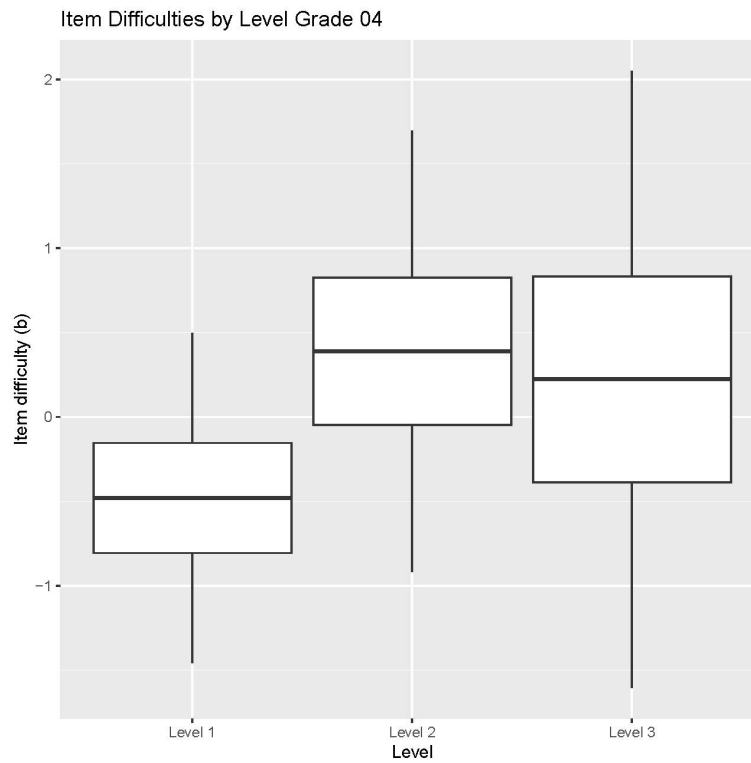


Image I-10. Boxplot of Item Difficulty for Mathematics—Grade 5

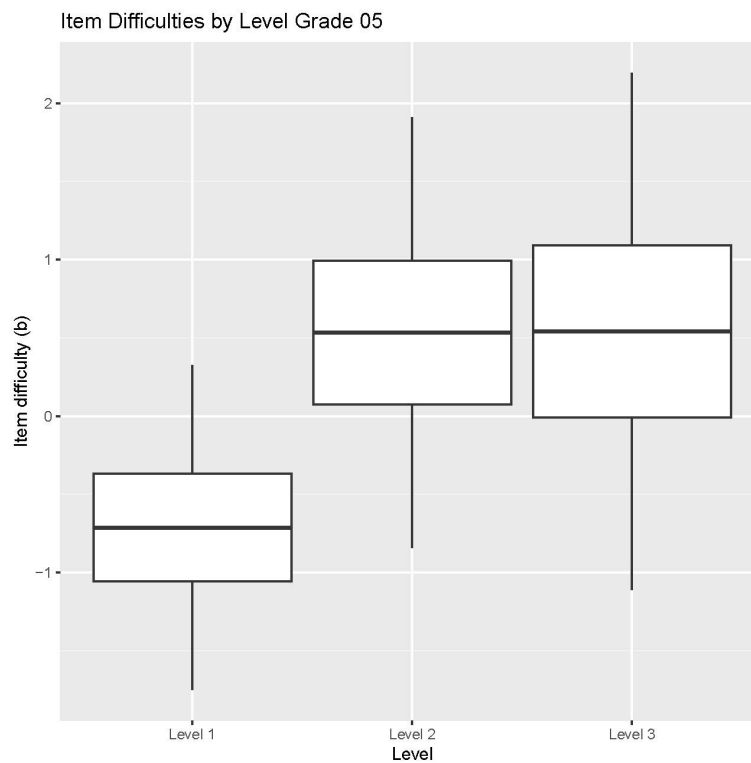


Image I-11. Boxplot of Item Difficulty for Mathematics—Grade 6

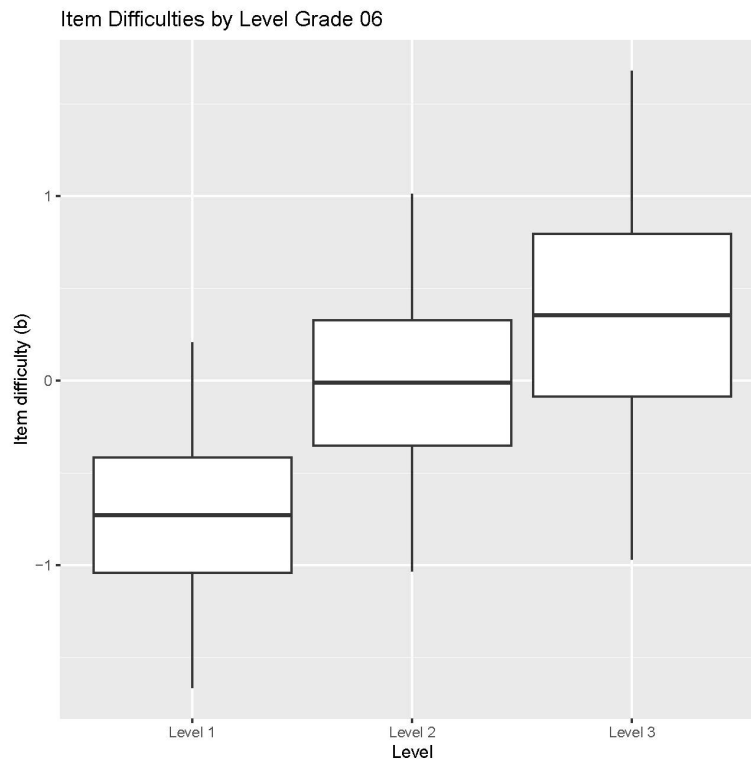


Image I-12. Boxplot of Item Difficulty for Mathematics—Grade 7

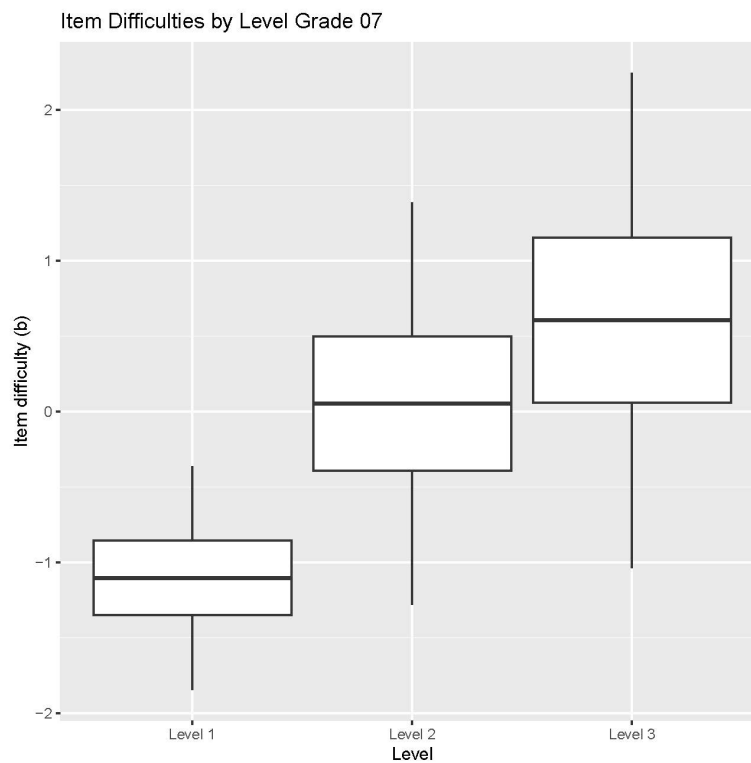


Image I-13. Boxplot of Item Difficulty for Mathematics—Grade 8

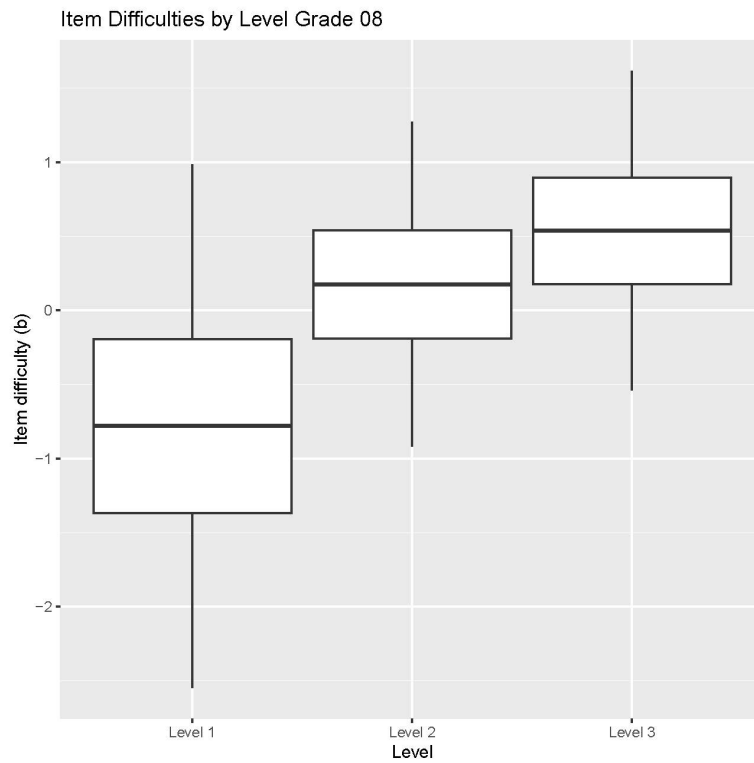


Image I-14. Boxplot of Item Difficulty for Mathematics—Grade 11

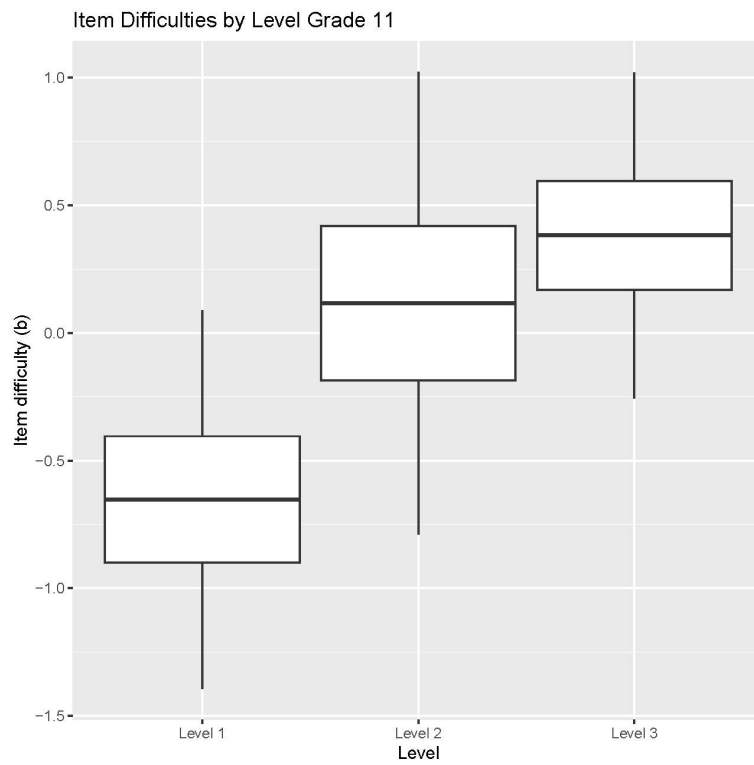


Table I-1. IRT Parameters for ELA Grade 3

IREF	a	SE (a)	b	SE (b)
451136	0.74	0.04	-0.41	0.04
451474	1.28	0.17	-0.59	0.06
451486	0.64	0.13	-0.33	0.09
451534	0.76	0.13	-0.14	0.07
451521	0.52	0.11	0.92	0.25
528721	0.76	0.04	-0.99	0.05
451172	0.72	0.04	-0.35	0.04
451186	0.86	0.05	-0.42	0.04
451498	1.08	0.15	-0.2	0.05
538024	1.05	0.12	-1.25	0.05
538036	1.45	0.11	-0.99	0.03
538060	1.34	0.13	-1.12	0.04
451148	0.65	0.04	-0.86	0.05
451160	0.73	0.04	-0.70	0.05
538072	1.45	0.11	-1.07	0.03
607566	0.69	0.08	0.43	0.08
607601	0.97	0.05	-0.83	0.03
607611	0.24	0.03	-1.29	0.15
538048	1.04	0.12	-1.07	0.04
607609	1.07	0.05	-0.69	0.03
659071	0.91	0.19	1.40	0.13
659143	0.77	0.09	0.44	0.08
607604	0.85	0.06	-1.02	0.05
772618	0.92	0.09	-0.96	0.06
772657	1.5	0.24	-0.61	0.10
607598	0.97	0.05	-0.99	0.04
607607	1.03	0.05	-0.68	0.03
658842	1.09	0.12	0.35	0.06
658921	1.3	0.15	-1.02	0.06
658945	1.01	0.14	-1.47	0.12
659002	0.82	0.09	0.61	0.07
659083	0.87	0.11	-0.01	0.11

IREF	a	SE (a)	b	SE (b)
659095	0.64	0.09	-0.06	0.14
786892	0.73	0.09	0.61	0.08
772611	0.77	0.14	-1.00	0.23
772646	0.94	0.09	-0.86	0.06
772753	0.83	0.07	-0.64	0.04
772774	0.74	0.07	-0.43	0.05
815047	1.27	0.26	0.34	0.11
815061	0.88	0.17	1.24	0.12
815131	0.75	0.12	-0.60	0.15
815159	0.63	0.13	0.86	0.14
658854	1.75	0.19	0.06	0.06
658878	0.85	0.19	1.55	0.16
658909	1.14	0.12	-0.96	0.07
658933	1.19	0.14	-1.18	0.08
675582	0.79	0.07	-0.18	0.06
675606	1.5	0.18	-0.44	0.04
658866	0.73	0.10	0.13	0.11
658957	1.02	0.13	-1.32	0.10
772632	1.14	0.12	-1.02	0.05
772767	0.88	0.07	-0.85	0.04
815040	0.95	0.19	0.60	0.11
815054	0.88	0.17	1.10	0.11
815138	1.19	0.10	-0.50	0.05
815145	1.11	0.21	0.44	0.11
772760	1.03	0.08	-0.64	0.04
824816	0.95	0.10	-0.92	0.07
827729	0.6	0.13	0.64	0.16
786899	0.87	0.10	0.46	0.07
815152	0.5	0.06	-0.25	0.10
824802	0.62	0.13	0.96	0.14
824823	0.94	0.09	-0.63	0.06
824849	1.28	0.10	-0.38	0.05

Table I-2. IRT Parameters for ELA Grade 4

IREF	a	SE (a)	b	SE (b)
113055A	1.38	0.13	-0.21	0.04
113095A	0.98	0.12	-0.45	0.07
113280A	0.74	0.04	-0.64	0.05
113283A	0.81	0.04	-0.57	0.04
113281A	0.68	0.04	-0.29	0.04
113096A	1.55	0.14	-0.19	0.04
121550A	0.58	0.03	-0.61	0.05
121569A	1.09	0.11	-0.12	0.05
121551A	0.39	0.03	-0.28	0.07
451679	1.10	0.13	0.63	0.07
451634	1.07	0.14	0.35	0.10
451646	0.79	0.11	0.51	0.11
451694	0.68	0.10	0.47	0.12
449675	0.95	0.07	0.07	0.05
451663	0.63	0.11	-0.01	0.20
615093	0.77	0.11	0.54	0.10
615160	1.43	0.13	-1.15	0.04
615184	1.35	0.12	-0.87	0.03
608475	0.71	0.04	-0.31	0.04
615079	1.27	0.18	0.12	0.11
615137	0.71	0.11	0.19	0.16
615172	1.44	0.13	-1.03	0.03
773120	0.73	0.10	0.79	0.09
659377	0.87	0.05	-0.84	0.04
659401	0.91	0.11	0.62	0.08
615196	1.38	0.13	-0.93	0.03
615065	1.67	0.27	-0.06	0.12
789723	0.94	0.08	-0.30	0.06
772948	0.98	0.11	-1.25	0.06
773099	0.79	0.11	0.71	0.09
773113	0.46	0.05	-0.48	0.10
815225	1.20	0.11	-0.85	0.06

IREF	a	SE (a)	b	SE (b)
815232	1.86	0.18	-0.86	0.04
659365	0.42	0.03	0.24	0.07
659389	0.59	0.03	0.12	0.05
823550	0.83	0.10	-1.27	0.11
789734	0.83	0.06	0.54	0.05
659413	0.43	0.03	0.14	0.07
772927	0.98	0.11	-1.03	0.05
772955	0.84	0.10	-1.29	0.07
773106	0.81	0.14	-0.19	0.21
815178	0.54	0.09	0.20	0.12
815199	1.09	0.22	1.00	0.10
677265	0.57	0.10	0.23	0.18
772934	1.20	0.12	-1.22	0.05
772941	1.20	0.12	-0.98	0.04
772969	1.23	0.20	-0.65	0.14
773092	1.04	0.15	0.11	0.13
824863	1.16	0.12	-1.04	0.07
824905	1.07	0.08	-1.03	0.05
815246	1.42	0.15	-1.07	0.06
815185	0.82	0.18	1.03	0.13
815192	0.74	0.17	1.38	0.14
815213	0.86	0.20	0.70	0.16
815253	1.20	0.11	-0.90	0.06
823578	1.54	0.22	-0.35	0.08
823585	1.50	0.24	-0.53	0.11
824856	0.85	0.09	-1.09	0.09
823557	1.03	0.10	-0.78	0.06
823564	1.18	0.11	-0.71	0.05
824898	0.86	0.07	-0.29	0.06
823571	1.06	0.19	-0.70	0.16
830881	0.86	0.05	0.59	0.05

Table I-3. IRT Parameters for ELA Grade 5

IREF	a	SE (a)	b	SE (b)
114072A	0.91	0.18	-0.73	0.27
114331A	0.64	0.10	-0.21	0.07
114332A	1.26	0.13	-0.26	0.04
114073A	0.88	0.15	-0.39	0.19
114074A	0.71	0.10	0.56	0.09
114075A	0.88	0.11	0.75	0.06
114329A	0.79	0.11	-0.37	0.06
121733A	0.62	0.10	0.02	0.08
121735A	0.65	0.11	-0.13	0.07
452025	1.54	0.18	-0.92	0.07
452001	0.44	0.09	-0.67	0.14
452013	0.87	0.12	0.18	0.07
540501	0.77	0.11	-0.82	0.06
530809	0.49	0.09	0.94	0.11
451036	0.68	0.11	-0.05	0.07
452038	1.40	0.14	-0.40	0.04
540418	1.57	0.15	-1.09	0.03
540443	1.28	0.13	-0.87	0.04
530777	0.74	0.11	0.12	0.08
540430	0.85	0.12	-1.51	0.09
659523	1.60	0.15	-1.06	0.03
659626	1.43	0.18	0.12	0.07
659643	0.98	0.12	0.41	0.07
659667	0.78	0.04	-0.46	0.04
659715	0.62	0.04	-0.39	0.05
677579	1.29	0.07	-0.80	0.03
659468	0.75	0.04	-0.12	0.04
659535	2.04	0.17	-1.11	0.03
659679	0.55	0.04	-0.03	0.05
659691	0.97	0.05	-0.49	0.03
677531	0.81	0.05	-0.92	0.05
677555	0.62	0.04	-1.33	0.08
791161	0.88	0.07	-0.17	0.05
816592	1.00	0.08	-0.44	0.05

IREF	a	SE (a)	b	SE (b)
816599	1.32	0.12	-0.91	0.05
659426	0.75	0.04	-0.35	0.04
659480	0.69	0.04	-0.70	0.05
659492	0.90	0.12	-0.29	0.12
659547	1.46	0.18	-1.35	0.07
659559	1.53	0.14	-0.93	0.03
659602	1.22	0.16	0.01	0.10
659614	1.71	0.21	0.07	0.07
677515	0.72	0.05	-1.28	0.07
791200	0.53	0.09	1.14	0.11
659511	1.26	0.14	-1.27	0.05
659703	0.67	0.04	-0.41	0.04
659727	0.55	0.09	1.06	0.16
677543	1.13	0.06	-0.93	0.03
677567	0.76	0.05	-1.08	0.05
816468	1.40	0.16	-1.29	0.07
816585	0.76	0.11	0.03	0.09
816606	0.62	0.09	0.55	0.11
824952	1.29	0.13	-1.19	0.06
816512	0.92	0.21	0.68	0.11
816620	0.45	0.06	-0.07	0.11
824919	1.09	0.11	-0.98	0.06
824959	1.50	0.15	-1.07	0.05
816440	2.12	0.22	-1.05	0.04
816447	1.77	0.17	-0.95	0.04
816454	1.01	0.10	-0.85	0.06
816475	1.89	0.18	-0.92	0.04
816491	0.88	0.19	1.40	0.14
816498	0.51	0.14	1.09	0.19
816505	0.65	0.16	0.90	0.15
816519	0.53	0.14	1.55	0.23
816526	0.70	0.07	-0.25	0.08
816613	1.00	0.14	-0.38	0.09
824926	0.85	0.12	-0.26	0.09

Table I-4. IRT Parameters for ELA Grade 6

IREF	a	SE (a)	b	SE (b)
114439A	1.43	0.19	-0.91	0.04
121328A	1.36	0.19	-1.00	0.05
121329A	1.66	0.20	-0.80	0.04
121327A	1.79	0.21	-0.88	0.04
452231	0.71	0.04	0.06	0.04
452243	0.47	0.04	0.14	0.06
452219	0.83	0.11	-0.11	0.12
452269	1.87	0.09	-0.36	0.02
452311	0.99	0.07	0.13	0.04
452348	1.19	0.08	-0.08	0.03
452299	1.01	0.07	-0.26	0.05
530858	0.90	0.09	0.62	0.05
452257	0.83	0.05	0.22	0.04
452360	1.49	0.09	-0.16	0.03
608268	1.33	0.18	-0.84	0.05
530821	1.42	0.19	-0.83	0.04
612778	1.83	0.21	-0.06	0.06
612816	0.69	0.10	-0.07	0.13
616086	1.44	0.19	-0.84	0.04
616098	1.30	0.18	-0.96	0.05
616136	1.54	0.20	-0.85	0.04
616272	0.83	0.09	0.72	0.05
616309	0.91	0.11	0.03	0.09
616347	0.61	0.08	0.29	0.09
612701	0.89	0.10	0.26	0.08
661565	0.94	0.09	-0.65	0.07
661838	1.37	0.18	-0.61	0.04
616110	1.39	0.19	-0.68	0.04
608280	1.31	0.15	-0.20	0.04
612751	1.13	0.14	-0.15	0.10
612841	1.51	0.16	0.04	0.06
775660	0.88	0.11	-0.09	0.07
775847	0.59	0.04	0.63	0.06

IREF	a	SE (a)	b	SE (b)
612725	1.43	0.15	0.19	0.05
661589	0.72	0.11	-0.27	0.12
616074	1.31	0.19	-1.14	0.07
620737	1.08	0.10	0.81	0.04
816657	0.75	0.08	0.12	0.07
816664	0.90	0.17	0.78	0.08
816671	0.98	0.12	0.19	0.06
816748	1.03	0.12	0.02	0.06
661577	0.64	0.07	-0.39	0.09
661790	1.81	0.21	-0.85	0.03
661802	1.09	0.17	-0.88	0.06
661826	1.48	0.20	-0.96	0.05
775651	1.00	0.09	-0.18	0.05
775862	1.22	0.06	-0.24	0.03
677939	1.12	0.11	0.41	0.05
791241	1.12	0.15	0.45	0.08
661553	1.15	0.11	-0.71	0.06
661601	1.12	0.12	0.33	0.06
661814	1.47	0.19	-0.80	0.04
816650	0.47	0.06	0.51	0.11
775635	1.00	0.09	-0.27	0.05
775806	0.61	0.04	0.09	0.05
775840	0.45	0.06	-0.14	0.10
775617	0.93	0.08	-0.16	0.05
791213	1.46	0.16	-0.32	0.04
775669	0.76	0.08	-0.39	0.07
775824	0.85	0.11	0.74	0.09
775831	0.78	0.05	-0.43	0.04
816643	0.79	0.10	0.17	0.07
816755	0.95	0.11	0.19	0.06
816762	0.87	0.11	0.30	0.07
816769	1.47	0.13	-0.49	0.05
791220	1.44	0.12	-0.48	0.04

Table I-5. IRT Parameters for ELA Grade 7

IREF	a	SE (a)	b	SE (b)
114594A	0.82	0.05	-0.46	0.04
115432A	1.18	0.13	-0.17	0.04
114593A	1.23	0.06	-0.70	0.03
114596A	0.99	0.05	-0.61	0.04
115433A	1.20	0.13	-0.32	0.04
121313A	1.27	0.14	-0.21	0.04
121315A	0.53	0.10	-0.03	0.10
121997A	0.75	0.04	-0.31	0.04
452456	0.63	0.04	-0.43	0.05
452480	1.23	0.06	-0.55	0.03
452492	0.92	0.05	-0.62	0.04
531700	1.35	0.17	-1.10	0.05
537321	1.13	0.16	-0.62	0.06
452468	0.68	0.04	0.01	0.04
537337	0.65	0.05	-0.15	0.07
537369	1.60	0.19	-0.94	0.04
531750	0.54	0.05	0.32	0.06
537297	1.50	0.18	-0.83	0.04
537309	1.51	0.07	-0.72	0.02
610008	0.96	0.09	0.35	0.06
662045	0.59	0.05	-0.06	0.06
662069	1.48	0.09	-0.36	0.03
662130	1.35	0.17	-0.98	0.04
616799	0.78	0.06	-0.20	0.05
609676	1.60	0.19	-1.15	0.04
775944	1.74	0.14	-0.81	0.03
775981	0.61	0.04	-0.16	0.05
775988	0.44	0.03	-0.06	0.06
662081	1.02	0.07	-0.18	0.04
662106	1.23	0.16	-0.95	0.05
662118	0.74	0.13	-1.11	0.08
616811	0.46	0.07	0.48	0.10
618684	0.77	0.05	0.05	0.05
616749	0.77	0.06	-0.07	0.05
616773	0.84	0.09	0.26	0.07

IREF	a	SE (a)	b	SE (b)
616785	0.85	0.06	-0.13	0.05
793091	1.37	0.17	-0.71	0.05
816839	0.82	0.15	0.55	0.10
816992	0.75	0.10	-0.04	0.09
662033	0.49	0.10	0.69	0.20
775937	1.03	0.11	-1.29	0.08
776002	0.58	0.04	-0.19	0.05
678083	0.78	0.14	-0.81	0.07
678095	0.78	0.06	-0.10	0.05
662021	1.41	0.14	-0.28	0.04
662057	0.62	0.05	0.00	0.06
662142	1.19	0.16	-1.21	0.06
662154	0.90	0.14	-0.87	0.06
823707	1.47	0.14	-0.80	0.05
816860	0.53	0.11	1.09	0.14
816867	0.54	0.08	0.50	0.10
816999	0.72	0.09	0.63	0.09
831412	1.37	0.15	-0.17	0.05
775958	1.46	0.12	-0.86	0.04
776009	0.61	0.04	-0.03	0.05
823700	0.65	0.11	-0.88	0.19
823728	1.03	0.11	-0.95	0.08
831419	1.01	0.11	-0.02	0.06
775922	1.21	0.10	-0.79	0.04
775974	0.87	0.05	-0.40	0.04
776016	0.63	0.04	0.59	0.05
816985	0.96	0.16	0.80	0.08
793031	0.73	0.11	0.02	0.07
823714	0.63	0.07	-0.63	0.10
823721	1.29	0.12	-0.81	0.06
831433	1.20	0.14	-0.29	0.07
816846	0.72	0.13	0.94	0.10
817006	0.65	0.09	-0.20	0.11
823693	0.76	0.10	-0.13	0.09

Table I-6. IRT Parameters for ELA Grade 8

IREF	a	SE (a)	b	SE (b)
114876A	0.82	0.05	-0.59	0.04
114877A	0.39	0.04	0.19	0.07
114879A	1.08	0.06	-0.60	0.03
122082A	0.49	0.04	-0.25	0.06
538809	1.24	0.16	-0.44	0.05
537726	0.92	0.10	0.36	0.06
537740	1.01	0.11	0.06	0.07
538821	0.56	0.12	0.59	0.18
538833	1.57	0.18	-0.28	0.04
538845	1.21	0.16	-0.27	0.05
538857	2.00	0.20	-0.18	0.03
610340	0.91	0.07	-0.63	0.05
537758	1.26	0.13	0.02	0.06
537774	1.36	0.14	-0.02	0.06
662173	1.44	0.12	-0.43	0.04
662197	0.37	0.06	0.39	0.15
662221	0.75	0.08	-0.44	0.08
662233	1.28	0.12	-0.71	0.05
662379	1.28	0.14	-0.81	0.04
662391	1.05	0.11	-0.92	0.08
617023	1.56	0.15	-0.99	0.03
617956	0.97	0.07	-0.03	0.04
776676	2.04	0.18	-0.93	0.03
776683	1.63	0.16	-0.99	0.03
617007	1.84	0.20	-1.38	0.05
617038	1.09	0.14	-1.24	0.06
617920	0.95	0.07	-0.03	0.04
617052	1.29	0.14	-0.83	0.04
617908	0.93	0.07	0.25	0.04
617932	0.99	0.08	-0.18	0.04
617944	0.69	0.06	0.48	0.05
662355	2.01	0.19	-1.19	0.04
817079	0.68	0.14	1.40	0.17

IREF	a	SE (a)	b	SE (b)
817093	0.84	0.11	0.55	0.08
817144	0.88	0.18	0.44	0.11
817188	0.95	0.12	0.02	0.07
817202	0.83	0.11	0.06	0.08
776697	1.88	0.17	-0.94	0.03
678299	1.39	0.09	-0.58	0.03
662209	0.84	0.08	-0.04	0.07
662343	1.99	0.18	-0.98	0.03
662367	1.78	0.17	-1.10	0.03
678251	1.58	0.11	-0.25	0.03
823736	0.81	0.08	-0.62	0.07
817151	1.06	0.13	0.27	0.06
817158	0.98	0.14	-0.03	0.07
817195	0.35	0.07	0.48	0.16
776669	1.68	0.17	-1.14	0.04
793123	0.79	0.05	0.20	0.05
823771	1.38	0.22	-0.62	0.11
776690	1.35	0.14	-1.02	0.04
835193	0.92	0.07	-0.87	0.06
817123	0.53	0.12	0.97	0.16
817137	0.59	0.13	0.87	0.14
817209	0.46	0.06	0.12	0.10
823743	1.09	0.10	-0.64	0.05
835200	1.16	0.11	-0.03	0.04
835280	1.24	0.10	0.14	0.04
817086	0.57	0.07	0.15	0.09
817100	0.78	0.11	0.28	0.08
817174	0.69	0.07	-0.08	0.07
817181	0.77	0.11	0.22	0.08
823750	1.29	0.13	-0.93	0.06
823757	1.19	0.15	-0.20	0.07
823764	1.64	0.16	-0.94	0.05

Table I-7. IRT Parameters for ELA Grade HS

IREF	a	SE (a)	b	SE (b)
114166A	0.95	0.10	1.01	0.06
114193A	0.70	0.11	-0.57	0.08
121714A	1.16	0.12	0.28	0.05
121702A	0.59	0.10	0.05	0.09
121703A	1.72	0.16	-0.38	0.03
121711A	1.19	0.13	0.22	0.05
121875A	1.14	0.13	-0.28	0.04
121695A	0.82	0.10	0.55	0.06
124328A	1.08	0.12	0.28	0.06
453060	1.10	0.13	-0.48	0.05
453074	1.34	0.15	-0.57	0.04
453111	0.69	0.04	0.26	0.04
453087	1.53	0.15	-0.33	0.03
453099	0.74	0.11	0.15	0.08
539072	1.01	0.06	-0.66	0.04
453138	1.21	0.14	-0.59	0.05
538950	1.27	0.14	0.03	0.06
539001	0.83	0.10	1.14	0.08
538962	0.75	0.11	-0.01	0.10
538986	0.91	0.11	0.10	0.08
539058	0.85	0.05	-0.26	0.04
531943	1.37	0.14	0.23	0.05
618234	0.92	0.10	1.23	0.08
618275	0.70	0.10	0.23	0.09
538974	1.34	0.15	-0.05	0.07
539013	1.02	0.06	-0.47	0.03
539025	0.84	0.05	-0.45	0.04
662568	0.65	0.05	-1.04	0.07
678432	1.10	0.13	-0.30	0.04
612940	1.34	0.14	-0.22	0.03
778000	1.56	0.20	-1.15	0.04
778023	1.24	0.17	-1.04	0.05
778030	1.00	0.16	-1.08	0.06
618247	1.03	0.11	0.55	0.05
618287	1.09	0.13	-0.01	0.08
618316	0.66	0.10	-0.20	0.14
793157	1.17	0.09	-0.40	0.03

IREF	a	SE (a)	b	SE (b)
793199	1.52	0.19	-1.04	0.04
817032	1.89	0.20	-1.02	0.05
817304	1.21	0.11	-0.47	0.05
817370	1.03	0.10	-0.68	0.06
817405	0.88	0.12	0.04	0.07
662580	0.47	0.04	-0.55	0.07
662628	0.61	0.04	0.13	0.05
778037	2.07	0.22	-1.02	0.03
678469	1.49	0.14	-0.24	0.03
793192	1.49	0.12	-0.59	0.04
662604	0.52	0.04	-0.35	0.06
678420	0.93	0.11	0.45	0.06
823779	1.12	0.11	-0.80	0.06
823800	0.54	0.07	-0.76	0.11
817039	1.94	0.20	-0.98	0.05
817377	1.34	0.16	-0.23	0.05
817398	0.67	0.10	0.19	0.09
836411	1.38	0.16	-1.05	0.07
778051	2.09	0.23	-1.01	0.03
823793	1.63	0.17	-0.97	0.05
817046	2.36	0.27	-1.12	0.05
817283	0.78	0.08	-0.03	0.07
817311	1.21	0.11	-0.57	0.05
777951	2.20	0.24	-1.14	0.03
777958	2.02	0.23	-1.17	0.04
777967	2.38	0.25	-1.02	0.03
777983	1.81	0.22	-1.21	0.04
778044	1.67	0.21	-1.25	0.05
793164	0.96	0.13	-0.27	0.12
823786	1.09	0.12	-1.09	0.08
823807	1.62	0.15	-0.82	0.05
817060	1.60	0.18	-1.12	0.07
817276	1.12	0.13	-0.19	0.06
817384	0.94	0.14	-0.45	0.09
817391	0.94	0.14	-0.52	0.10
823814	1.87	0.18	-0.87	0.04

Table I-8. IRT Parameters for Mathematics Grade 3

IREF	a	SE (a)	b	SE (b)
110855A	0.73	0.03	0.13	0.03
110975A	0.55	0.11	0.16	0.12
110923A	0.95	0.08	-0.64	0.04
111387A	0.79	0.08	0.61	0.06
111405A	0.78	0.11	-0.90	0.06
111425A	0.81	0.11	-0.60	0.07
111430A	0.76	0.05	-0.49	0.05
111649A	0.64	0.08	-1.30	0.14
111386A	0.73	0.04	-0.20	0.04
111420A	0.83	0.09	0.34	0.06
111426A	0.68	0.05	0.31	0.05
111391A	0.70	0.08	0.24	0.08
112552A	1.17	0.14	-0.27	0.09
112600A	0.70	0.11	-0.47	0.10
112559A	0.56	0.10	-0.60	0.10
112595A	0.82	0.04	-0.24	0.04
112601A	1.39	0.13	0.11	0.05
112569A	0.67	0.10	-0.78	0.08
112579A	0.67	0.06	-0.98	0.09
112615A	0.60	0.07	0.99	0.09
442035	0.70	0.07	0.42	0.07
442166	0.71	0.09	0.86	0.09
442037	1.07	0.05	-0.39	0.03
441796	1.01	0.08	-0.50	0.05
442416	0.83	0.12	-0.44	0.09
529146	0.53	0.10	-0.08	0.10
451116	0.71	0.08	0.89	0.07
528663	1.36	0.10	-0.14	0.04
598100	0.57	0.07	0.80	0.08
604164	0.91	0.08	-0.53	0.06
528813	0.73	0.04	-0.55	0.04
656884	0.80	0.11	-1.06	0.06
597535	0.84	0.10	0.12	0.08
597576	0.69	0.08	0.38	0.07

IREF	a	SE (a)	b	SE (b)
604123	1.19	0.06	0.73	0.04
604155	0.84	0.10	0.26	0.07
603958	1.83	0.23	0.50	0.07
604382	1.62	0.22	0.64	0.10
595744	1.09	0.11	0.34	0.05
597774	0.93	0.13	-0.82	0.09
598213	0.60	0.10	-0.71	0.18
773532	0.71	0.04	-0.15	0.04
773546	0.61	0.12	-1.06	0.16
773558	0.68	0.07	-0.88	0.09
773560	0.57	0.08	-0.98	0.12
656872	0.50	0.09	-1.18	0.11
656860	0.72	0.12	-1.02	0.14
656876	1.03	0.13	-0.11	0.05
656824	0.46	0.09	-0.82	0.11
656836	0.66	0.04	-0.67	0.05
656868	1.83	0.18	0.08	0.04
656874	0.85	0.12	-0.11	0.06
656880	0.51	0.07	-0.11	0.11
822319	0.75	0.08	-0.86	0.08
822326	0.52	0.07	-0.88	0.11
822350	0.43	0.06	-1.00	0.14
822376	0.76	0.07	-0.06	0.06
773488	0.52	0.06	1.05	0.16
773552	0.40	0.07	0.84	0.18
773562	0.77	0.05	-0.83	0.05
822316	1.31	0.10	0.23	0.05
822339	0.56	0.09	1.62	0.22
822402	0.56	0.08	0.60	0.11
822354	0.92	0.16	1.08	0.10
822356	0.97	0.09	-0.81	0.06
822394	0.49	0.08	0.75	0.15
822322	0.89	0.07	-0.12	0.06

Table I-9. IRT Parameters for Mathematics Grade 4

IREF	a	SE (a)	b	SE (b)	IREF	a	SE (a)	b	SE (b)
111136A	0.87	0.07	-0.07	0.04	656928	0.84	0.17	0.91	0.12
111166A	0.86	0.05	0.24	0.04	656946	1.43	0.13	0.25	0.04
111122A	0.96	0.15	-0.41	0.08	656954	1.38	0.07	0.22	0.02
111123A	0.85	0.07	-0.02	0.05	599462	0.81	0.13	-0.28	0.14
111162A	1.18	0.18	-0.06	0.10	599574	0.60	0.09	1.04	0.10
111685A	0.94	0.14	-0.79	0.06	599952	1.21	0.17	-0.27	0.07
111663A	1.20	0.07	-0.71	0.04	600620	0.87	0.06	-0.63	0.04
111681A	2.13	0.19	0.49	0.03	774536	0.96	0.11	-0.34	0.05
111658A	1.22	0.12	-0.56	0.05	774540	1.33	0.18	-0.30	0.07
111678A	1.53	0.15	0.37	0.04	774542	1.05	0.13	0.69	0.09
111712A	0.52	0.05	0.73	0.08	774562	0.91	0.09	-0.59	0.07
111730A	0.77	0.10	0.33	0.07	656932	0.64	0.12	-0.53	0.10
111731A	0.99	0.11	0.30	0.06	656960	0.95	0.14	-1.18	0.08
111686A	0.64	0.04	0.38	0.05	656904	1.10	0.13	-0.17	0.05
111687A	1.27	0.13	0.44	0.04	656920	0.84	0.13	-0.78	0.07
111716A	0.76	0.07	0.16	0.06	656952	1.48	0.14	0.15	0.04
111717A	0.73	0.05	-0.15	0.04	656926	0.90	0.05	-0.37	0.04
112817A	1.05	0.11	0.90	0.05	774510	1.03	0.10	-0.79	0.07
112838A	0.55	0.09	0.86	0.09	774522	0.78	0.05	0.32	0.04
112794A	0.67	0.07	0.19	0.07	774534	0.86	0.14	-0.40	0.09
112783A	1.45	0.15	0.28	0.05	774548	0.92	0.14	-0.88	0.06
112812A	0.85	0.10	1.13	0.08	822452	0.84	0.12	0.16	0.06
120551A	0.83	0.10	1.15	0.08	822472	0.91	0.17	0.94	0.10
121750A	0.93	0.16	-0.11	0.11	822478	1.26	0.23	1.05	0.10
454751	0.94	0.05	0.26	0.03	822492	0.48	0.06	0.75	0.14
445558	1.63	0.15	0.55	0.03	774558	0.54	0.09	0.86	0.14
445572	0.78	0.05	-0.58	0.05	774506	0.65	0.04	0.15	0.04
446443	0.63	0.05	-0.64	0.06	822448	0.57	0.10	0.28	0.09
599071	0.93	0.06	-0.90	0.05	822498	1.08	0.14	0.09	0.05
600647	0.99	0.16	-0.19	0.10	822424	1.31	0.11	0.29	0.04
455024	1.06	0.06	-0.54	0.03	822466	1.13	0.10	-0.34	0.05
600654	0.72	0.06	-0.27	0.06	822488	0.51	0.07	0.49	0.11
529949	0.56	0.06	0.21	0.08	822496	0.81	0.16	1.18	0.13
656896	0.69	0.10	-0.39	0.07	822412	1.19	0.13	0.94	0.08
656924	0.60	0.08	0.28	0.10	822430	0.69	0.07	0.36	0.07

Table I-10. IRT Parameters for Mathematics Grade 5

IREF	a	SE (a)	b	SE (b)
111242A	1.00	0.14	-0.69	0.06
111257A	1.14	0.14	-1.05	0.06
111294A	1.35	0.12	0.25	0.05
111243A	0.93	0.05	0.24	0.03
111277A	0.56	0.05	0.41	0.06
112358A	0.92	0.06	0.45	0.04
112335A	0.96	0.13	-0.80	0.06
112348A	0.78	0.09	0.63	0.06
112353A	0.75	0.07	0.19	0.06
112368A	0.89	0.11	0.23	0.06
112377A	0.73	0.05	-0.65	0.05
112385A	0.77	0.06	0.21	0.04
112346A	0.73	0.05	-0.53	0.05
112354A	0.77	0.09	1.02	0.07
112364A	0.88	0.11	0.44	0.08
112373A	0.72	0.09	0.22	0.08
112392A	0.64	0.08	1.26	0.10
112384A	1.22	0.07	0.18	0.03
112408A	1.02	0.11	0.27	0.06
112410A	0.88	0.10	0.31	0.06
113843A	0.75	0.13	-0.12	0.12
113884A	0.48	0.04	-0.44	0.07
113889A	0.88	0.13	-0.79	0.07
113862A	0.88	0.11	0.44	0.08
120737A	0.98	0.06	1.14	0.06
120724A	0.99	0.16	0.75	0.21
450187	0.47	0.03	-0.93	0.08
450352	0.56	0.09	-0.41	0.09
450153	0.98	0.06	-0.70	0.04
450194	1.05	0.10	0.55	0.05
450200	0.89	0.14	-0.28	0.09
449994	0.66	0.06	0.09	0.06
450339	1.02	0.10	0.51	0.05
449975	1.32	0.13	0.16	0.05
450055	0.95	0.11	0.31	0.07
450210	1.00	0.14	-0.84	0.06

IREF	a	SE (a)	b	SE (b)
462181	1.07	0.08	0.78	0.05
450296	0.56	0.06	1.24	0.11
601619	0.43	0.09	-0.89	0.13
602060	0.66	0.11	-0.97	0.09
602432	0.61	0.11	-0.98	0.10
532547	1.01	0.12	0.46	0.07
656974	0.97	0.12	-1.10	0.10
657022	0.95	0.11	-0.70	0.06
657030	0.84	0.10	1.10	0.13
601957	0.66	0.12	-0.41	0.10
601986	1.41	0.16	-0.87	0.05
602486	0.47	0.05	0.59	0.11
601633	1.02	0.11	-1.01	0.08
774378	0.95	0.09	0.68	0.05
657002	0.82	0.11	0.52	0.09
657038	0.92	0.15	1.40	0.23
656968	0.66	0.05	-1.05	0.08
657010	0.51	0.09	0.26	0.10
774368	0.61	0.09	1.22	0.16
774372	0.43	0.08	1.29	0.21
774404	1.59	0.30	0.49	0.18
774436	1.13	0.11	-0.85	0.06
822552	0.59	0.09	0.12	0.09
822568	0.57	0.13	0.28	0.16
822586	0.95	0.09	0.93	0.09
774414	0.51	0.10	-1.06	0.13
774376	1.16	0.11	-0.61	0.05
774422	0.85	0.05	-0.11	0.03
822578	0.96	0.16	-0.75	0.14
822592	0.83	0.12	-0.22	0.08
822530	0.68	0.13	0.88	0.12
822580	0.55	0.09	0.95	0.14
822538	0.99	0.09	0.00	0.05
822556	0.76	0.08	1.55	0.16
822574	0.75	0.10	0.34	0.08

Table I-11. IRT Parameters for Mathematics Grade 6

IREF	a	SE (a)	b	SE (b)
110982A	0.82	0.09	0.75	0.06
111025A	0.83	0.07	-0.32	0.04
111035A	0.55	0.10	0.13	0.09
111445A	1.44	0.13	0.25	0.05
111465A	0.63	0.08	0.51	0.08
111455A	1.48	0.14	0.29	0.04
111479A	1.27	0.13	-0.03	0.04
111630A	1.07	0.06	-0.81	0.04
111514A	0.72	0.05	-0.36	0.04
111518A	0.91	0.06	0.34	0.04
112655A	0.93	0.09	0.81	0.05
112699A	1.15	0.11	0.58	0.04
112653A	0.65	0.10	-0.11	0.06
112656A	1.00	0.15	-0.70	0.06
112672A	0.92	0.06	-0.09	0.04
112673A	0.98	0.08	-0.24	0.04
112666A	1.10	0.13	-0.64	0.06
112697A	1.88	0.16	-0.33	0.03
120854A	1.10	0.11	0.24	0.06
442631	0.75	0.11	0.12	0.07
450365	0.73	0.07	-0.01	0.06
453780	1.18	0.16	-0.80	0.05
442538	1.21	0.11	0.63	0.04
442634	1.37	0.17	-0.87	0.04
442683	0.65	0.14	0.12	0.20
442356	0.96	0.06	0.08	0.04
442566	1.41	0.07	-0.44	0.03
442628	0.67	0.04	0.48	0.05
453755	1.08	0.11	0.43	0.05
453771	0.77	0.05	0.01	0.04
450436	1.12	0.16	-0.89	0.05
453675	0.96	0.12	-0.74	0.05
534783	1.06	0.12	0.03	0.04
603458	0.72	0.13	-0.98	0.08
534823	1.19	0.18	-0.16	0.09

IREF	a	SE (a)	b	SE (b)
657092	1.28	0.17	-1.05	0.06
603290	1.15	0.08	-0.23	0.04
603418	1.36	0.15	-0.12	0.08
603462	0.80	0.07	-0.29	0.05
603743	0.99	0.10	-0.95	0.08
602872	1.67	0.15	-0.46	0.03
603188	0.86	0.08	-0.53	0.06
603329	0.73	0.09	0.10	0.10
773661	0.58	0.08	0.69	0.08
773681	1.22	0.06	0.41	0.03
773716	1.40	0.17	-0.94	0.05
603349	1.31	0.18	-0.89	0.05
603712	1.31	0.17	-0.68	0.05
657042	0.72	0.14	-0.35	0.11
657088	1.08	0.11	-0.47	0.06
657054	0.85	0.05	-0.40	0.04
773655	0.66	0.04	0.37	0.05
657074	1.18	0.06	-0.28	0.03
822648	0.75	0.09	-0.91	0.10
822656	0.90	0.09	-0.45	0.06
822658	0.90	0.10	0.17	0.06
773619	0.92	0.12	-0.73	0.08
773639	0.64	0.08	-0.86	0.11
773720	0.77	0.05	0.08	0.04
773697	1.24	0.12	0.35	0.05
822594	0.87	0.10	0.34	0.06
822610	1.10	0.11	0.22	0.05
822632	0.82	0.09	-0.67	0.08
822640	1.01	0.12	-0.25	0.07
822668	1.97	0.17	-0.56	0.03
822678	0.79	0.07	0.06	0.06
822652	0.85	0.15	1.07	0.11
822660	0.97	0.10	-0.95	0.08
822664	0.69	0.07	0.09	0.07

Table I-12. IRT Parameters for Mathematics Grade 7

IREF	a	SE (a)	b	SE (b)
111104A	1.18	0.12	-0.40	0.04
111054A	0.98	0.13	-0.74	0.06
111071A	1.20	0.11	-0.05	0.05
111074A	1.04	0.13	-0.95	0.05
111106A	0.88	0.08	1.21	0.07
111055A	0.64	0.04	0.09	0.05
111075A	1.34	0.14	-0.31	0.07
111130A	0.55	0.10	-1.20	0.10
111092A	0.68	0.11	-1.03	0.08
111099A	1.54	0.13	0.12	0.04
111090A	0.93	0.09	-1.12	0.05
111745A	0.48	0.10	0.20	0.26
111769A	0.77	0.10	-0.36	0.06
111765A	0.72	0.05	0.23	0.05
111778A	0.71	0.11	-1.47	0.10
111795A	0.81	0.05	0.31	0.04
111783A	1.27	0.16	-1.28	0.10
111796A	0.97	0.09	0.36	0.05
111764A	1.01	0.10	0.09	0.06
111766A	1.25	0.11	0.25	0.04
112523A	0.73	0.08	0.65	0.06
112882A	0.60	0.03	0.07	0.04
112909A	0.75	0.05	-1.08	0.06
113101A	1.47	0.12	0.29	0.04
112881A	0.57	0.04	0.39	0.06
112887A	1.16	0.10	0.33	0.04
446491	0.96	0.06	-1.09	0.05
446589	1.08	0.09	0.41	0.05
446722	1.33	0.12	-0.01	0.05
446901	0.94	0.12	-1.16	0.06
452111	0.49	0.06	0.92	0.11
454054	0.99	0.11	-0.56	0.05
451883	1.03	0.06	-1.35	0.06
454049	1.11	0.10	0.39	0.04
605075	0.48	0.04	-0.43	0.06
605241	0.80	0.13	-0.16	0.13

IREF	a	SE (a)	b	SE (b)
605361	0.77	0.11	0.59	0.14
606474	0.59	0.05	0.19	0.08
606510	0.89	0.07	-0.57	0.05
530662	0.56	0.07	1.22	0.11
531427	1.84	0.19	-0.53	0.04
531446	0.82	0.06	-0.02	0.05
605348	0.89	0.15	0.12	0.17
657152	0.81	0.11	0.22	0.08
657166	0.83	0.08	0.93	0.06
657168	0.81	0.12	-1.20	0.13
605643	1.15	0.14	-1.35	0.06
657170	0.54	0.04	-0.05	0.05
605009	1.18	0.06	-0.35	0.03
605373	0.98	0.05	-0.06	0.03
606208	0.58	0.04	0.16	0.05
657178	0.65	0.09	0.66	0.11
657138	0.76	0.11	-0.82	0.07
774651	0.48	0.06	-0.09	0.10
774676	0.51	0.04	-0.62	0.06
657158	0.41	0.09	0.74	0.24
774637	1.39	0.15	-1.22	0.04
774645	0.89	0.09	-0.85	0.06
774659	0.97	0.06	-1.30	0.06
774719	2.05	0.21	-0.48	0.04
774735	0.71	0.08	1.08	0.08
822682	1.15	0.11	-1.04	0.06
822700	0.67	0.07	-0.22	0.07
822718	0.49	0.09	1.74	0.25
822730	0.63	0.07	-0.24	0.07
822752	0.64	0.08	-1.47	0.14
822714	0.97	0.09	-0.95	0.06
822722	0.57	0.08	0.73	0.12
822760	1.82	0.14	-0.40	0.03

Table I-13. IRT Parameters for Mathematics Grade 8

IREF	a	SE (a)	b	SE (b)
111339A	0.73	0.08	0.73	0.06
111286A	0.96	0.13	0.15	0.07
111335A	0.84	0.12	-0.31	0.10
111588A	1.06	0.14	-0.16	0.10
111622A	1.24	0.07	0.07	0.03
111352A	0.96	0.05	-0.79	0.04
112474A	1.24	0.14	-1.36	0.06
112464A	0.73	0.05	-0.80	0.05
112475A	1.16	0.08	-0.51	0.04
112499A	1.34	0.14	-0.93	0.04
112516A	0.88	0.05	0.06	0.03
113922A	0.76	0.07	-0.01	0.06
113932A	0.77	0.06	0.29	0.04
113937A	0.67	0.04	0.56	0.04
113963A	0.80	0.04	0.16	0.04
446958	0.83	0.09	0.26	0.06
447063	1.08	0.13	-0.62	0.06
447166	1.73	0.16	-1.09	0.04
447054	0.69	0.03	0.02	0.03
454122	0.87	0.12	-0.82	0.06
446979	0.64	0.11	-0.18	0.07
447047	1.09	0.10	0.44	0.04
532251	0.97	0.10	0.01	0.06
532339	1.00	0.17	0.34	0.19
532355	1.40	0.14	0.10	0.04
454205	0.69	0.04	-0.06	0.04
607045	0.59	0.12	-1.55	0.27
657200	0.62	0.07	0.04	0.08
657204	0.61	0.10	-0.46	0.14
657234	0.82	0.14	-0.63	0.14
608118	0.79	0.10	1.29	0.10
608123	0.92	0.13	-0.12	0.05
609804	0.53	0.05	0.13	0.08
609162	0.87	0.08	0.62	0.05
607023	0.51	0.06	0.89	0.08

IREF	a	SE (a)	b	SE (b)
774991	0.64	0.13	0.70	0.32
608186	0.72	0.08	-1.24	0.12
609908	1.00	0.12	-0.76	0.05
610058	1.15	0.10	0.21	0.04
657222	0.62	0.08	-1.15	0.14
774987	0.47	0.05	0.32	0.11
774999	0.67	0.08	0.44	0.07
775013	0.83	0.09	0.23	0.06
775015	0.86	0.14	-0.59	0.12
657232	0.93	0.09	0.59	0.05
657260	0.69	0.08	0.77	0.11
822834	0.53	0.08	-1.57	0.23
774917	0.56	0.09	0.06	0.10
775061	0.93	0.05	-0.27	0.03
774921	0.50	0.06	-1.76	0.19
774923	0.57	0.04	-0.02	0.05
774995	0.84	0.05	0.03	0.04
822770	1.02	0.11	-0.88	0.08
822786	0.79	0.08	-0.50	0.06
822798	0.51	0.06	0.08	0.10
822816	0.69	0.08	-0.38	0.08
822818	0.80	0.10	0.13	0.07
822820	0.62	0.12	0.85	0.13
822772	0.67	0.07	0.01	0.07
822776	1.01	0.10	-0.71	0.06
822796	1.00	0.11	-1.01	0.09
822830	0.60	0.09	0.79	0.12
822780	1.50	0.22	0.34	0.06
822812	0.89	0.15	0.76	0.09
822814	0.61	0.09	0.29	0.09
822822	0.79	0.14	0.75	0.10
822824	0.92	0.12	0.09	0.06
822832	0.71	0.10	0.59	0.09
822838	0.74	0.13	0.81	0.11

Table I-14. IRT Parameters for Mathematics Grade HS

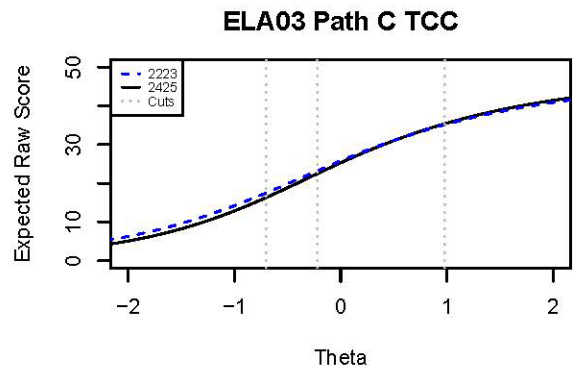
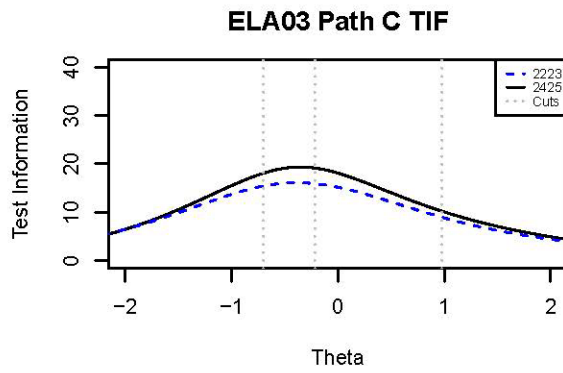
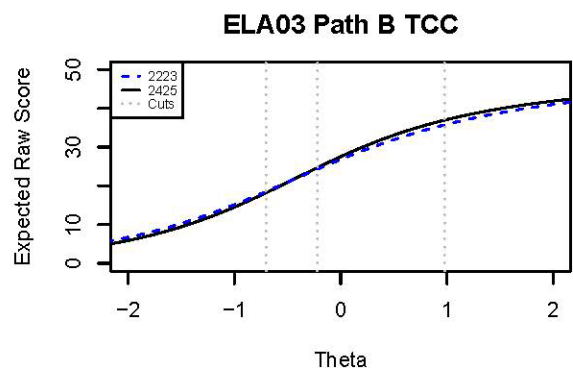
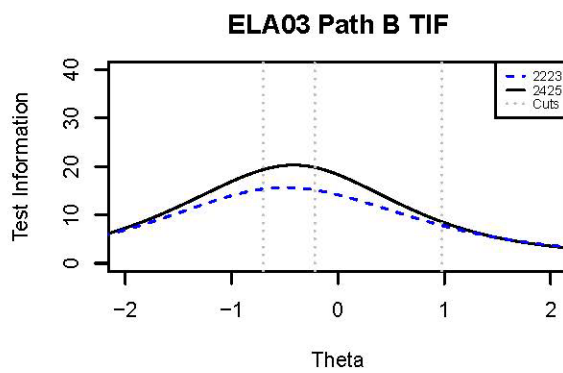
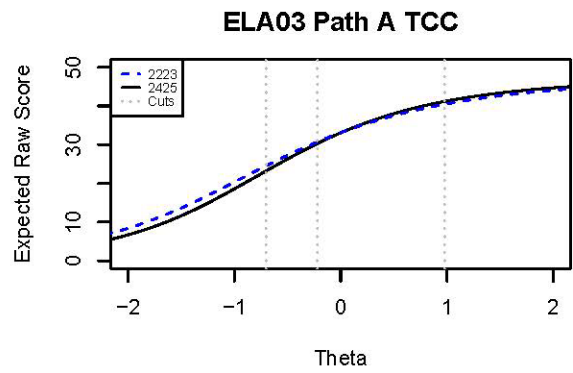
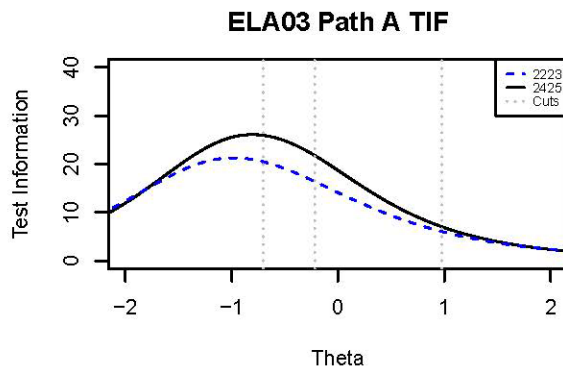
IREF	a	SE (a)	b	SE (b)
110914A	1.61	0.13	0.17	0.03
111002A	1.59	0.17	-0.10	0.03
110846A	1.04	0.06	0.05	0.03
110936A	0.77	0.05	-0.36	0.04
111000A	1.67	0.14	0.12	0.03
110913A	1.03	0.07	0.33	0.04
110921A	1.73	0.18	-0.08	0.03
111016A	0.84	0.10	0.10	0.07
111534A	1.38	0.11	0.48	0.03
111829A	1.36	0.11	0.41	0.03
111533A	1.60	0.19	-0.72	0.04
111537A	1.06	0.10	0.73	0.05
111824A	1.17	0.17	-0.31	0.07
112701A	1.35	0.17	-0.76	0.04
111834A	1.10	0.17	-0.30	0.07
112727A	1.21	0.06	-0.15	0.03
111830A	1.05	0.15	-0.80	0.06
112940A	1.19	0.11	0.38	0.04
112706A	1.19	0.16	-0.57	0.05
112924A	1.14	0.15	-0.51	0.05
112718A	1.04	0.07	0.45	0.04
112744A	0.92	0.10	0.16	0.05
112717A	0.60	0.12	0.24	0.13
112934A	0.86	0.14	-0.79	0.10
443381	0.87	0.14	-0.63	0.06
443575	1.29	0.07	-0.49	0.03
454987	1.13	0.11	-0.07	0.04
533370	0.61	0.08	0.21	0.07
533908	0.62	0.08	1.31	0.13
613648	0.80	0.05	-0.19	0.04
614219	1.01	0.11	-0.17	0.07
533128	1.05	0.10	0.00	0.04
657300	0.98	0.15	-0.97	0.07
613955	0.98	0.06	-0.01	0.03
613392	0.82	0.09	-0.84	0.09

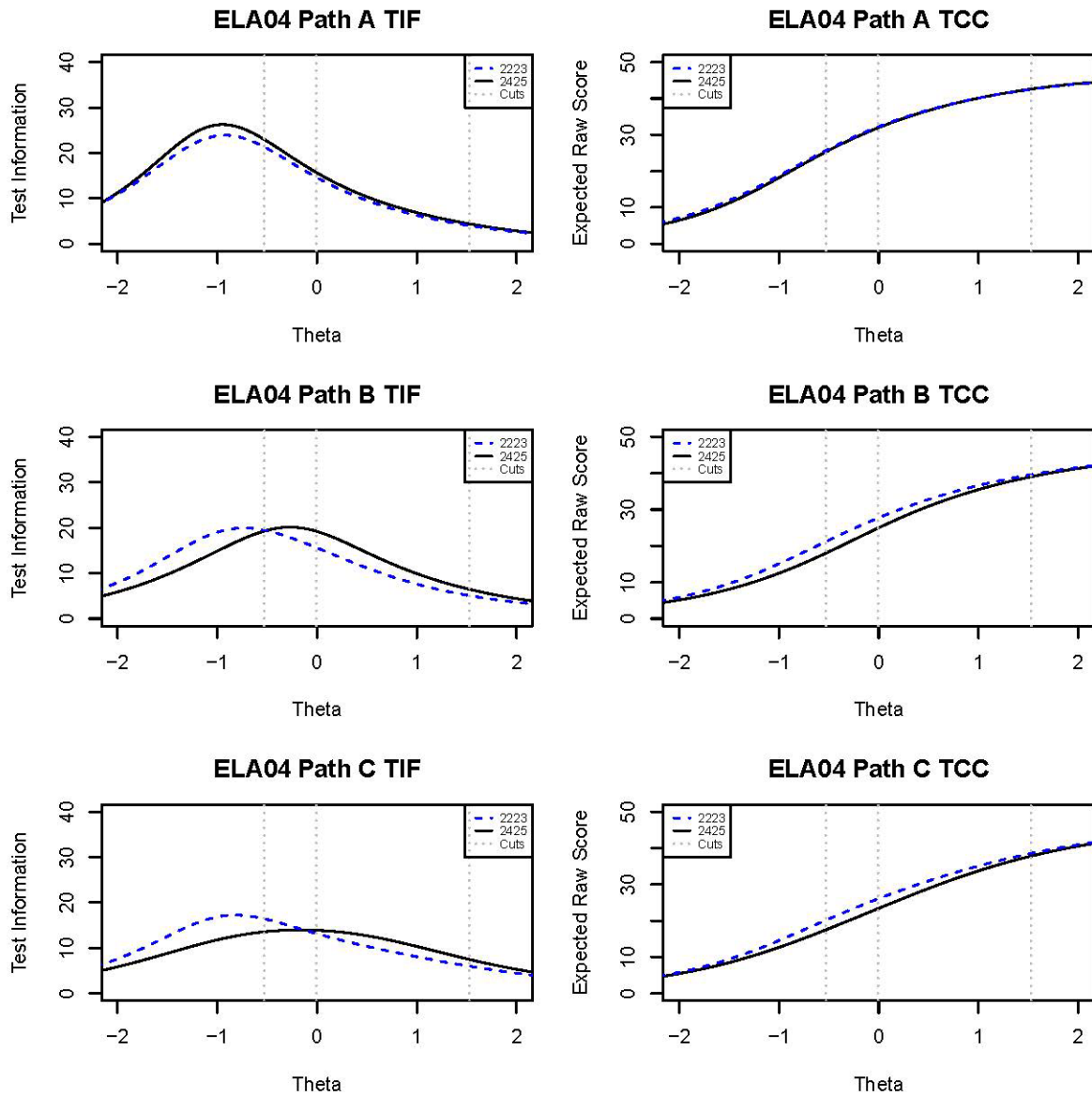
IREF	a	SE (a)	b	SE (b)
613828	1.06	0.10	-0.93	0.05
613850	1.01	0.06	-0.77	0.04
775250	0.77	0.14	0.35	0.12
657306	0.82	0.14	-0.24	0.06
657312	0.90	0.06	-0.67	0.04
657320	0.77	0.14	0.23	0.10
775312	0.49	0.08	0.34	0.09
613838	2.75	0.20	0.24	0.02
613943	0.74	0.14	-0.17	0.12
657266	1.20	0.13	-0.11	0.05
657278	2.69	0.19	0.26	0.02
657284	1.04	0.08	-0.13	0.04
657290	0.69	0.05	0.02	0.04
657292	0.68	0.05	0.40	0.05
657318	1.14	0.13	-0.66	0.05
657330	0.55	0.08	0.41	0.08
822894	0.98	0.16	0.53	0.08
822908	0.83	0.17	0.10	0.11
822914	1.01	0.11	-0.92	0.08
822924	1.02	0.10	0.02	0.05
775324	0.49	0.04	0.14	0.06
775246	0.75	0.09	0.09	0.07
775254	1.25	0.17	-0.65	0.05
775260	0.97	0.06	-0.15	0.03
822892	0.97	0.16	0.46	0.07
822872	1.08	0.10	-0.31	0.05
822896	1.00	0.16	-0.57	0.10
822902	0.87	0.10	-0.79	0.08
822916	1.24	0.11	-0.11	0.04
822864	0.81	0.09	0.29	0.07
822882	0.77	0.15	0.42	0.10
822898	1.51	0.23	0.33	0.05
822928	1.19	0.12	-0.79	0.06
822942	1.28	0.11	0.07	0.04

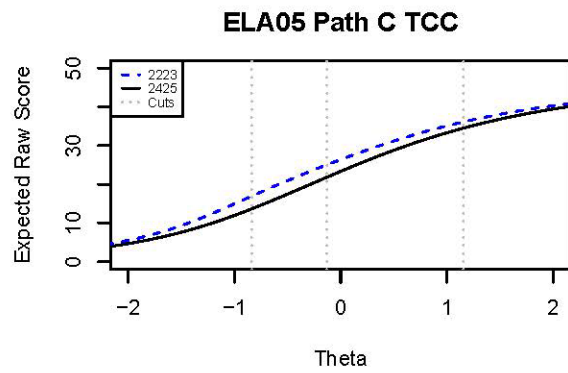
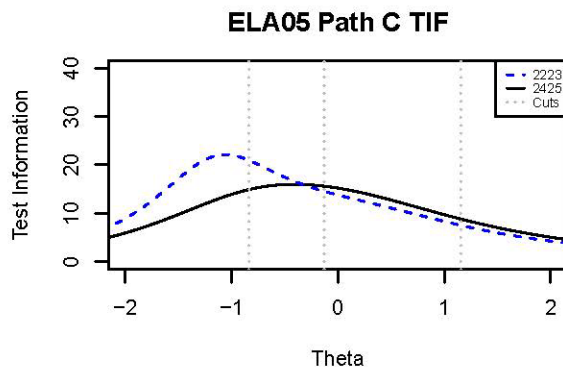
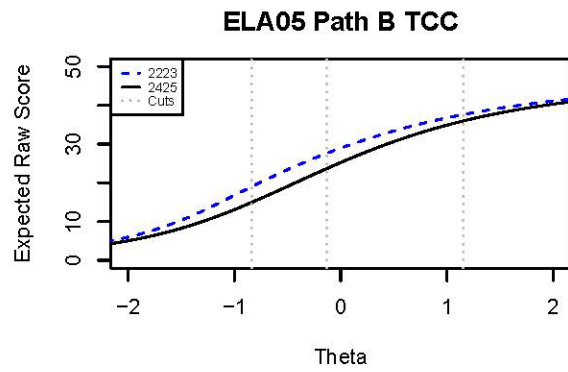
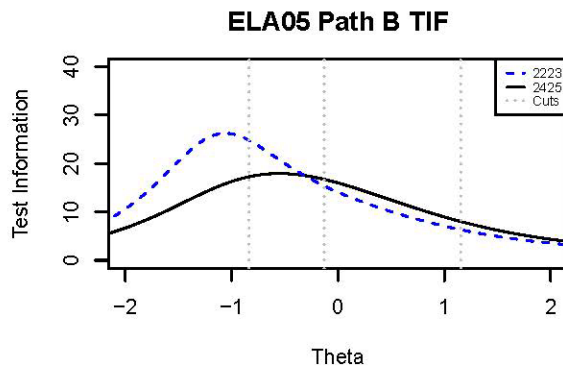
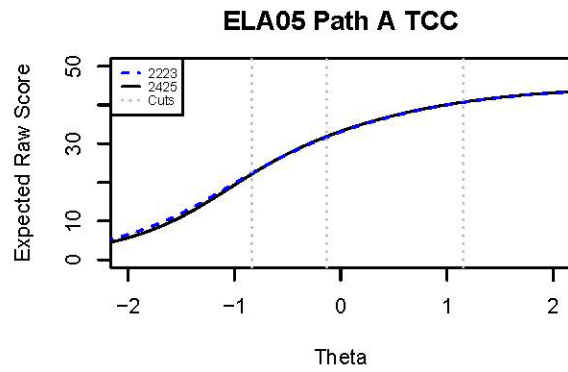
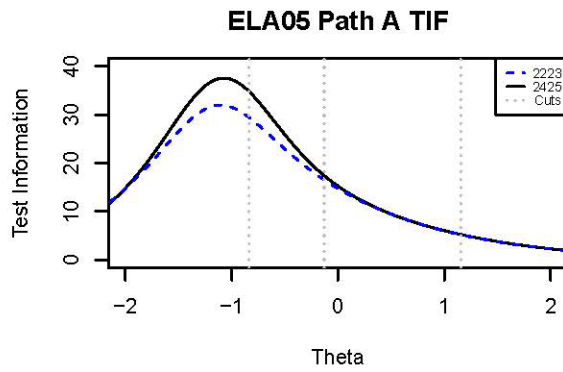
Table I-15. IRT Polytomous Item Parameters for ELA Writing Prompts

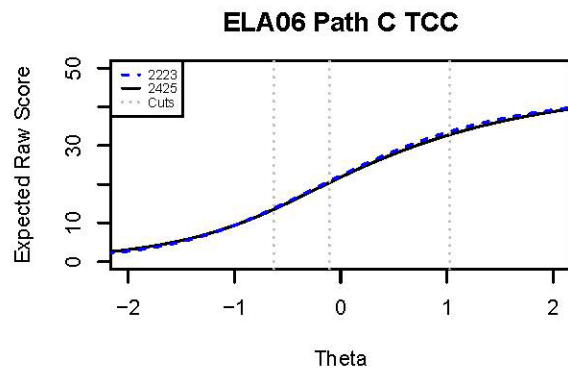
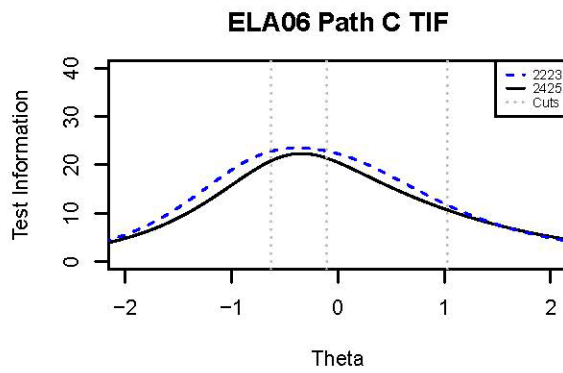
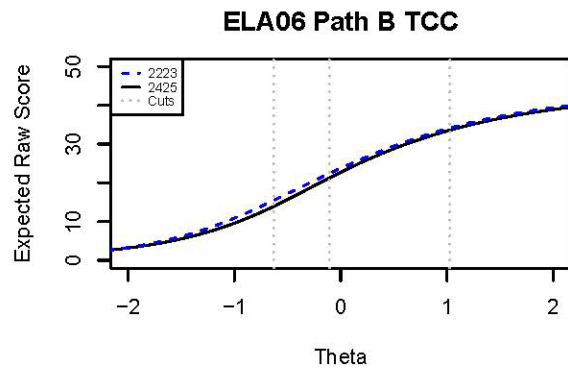
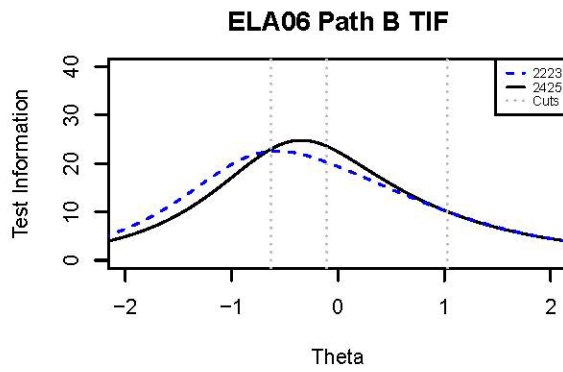
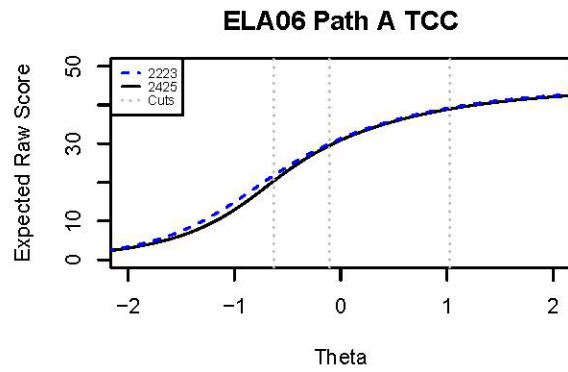
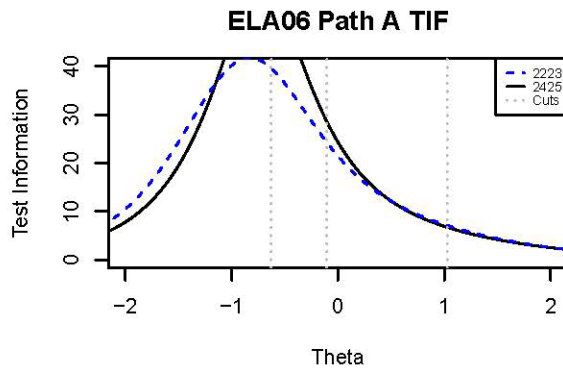
Grade	ID	a	b	d0	d1	b-d0	b-d1
3	125971C	0.6840	0.5211	1.0771	-1.0771	-0.5561	1.5982
3	125971I	0.7126	1.8053	1.0042	-1.0042	0.8011	2.8095
3	125971O	0.7834	1.5921	1.6280	-1.6280	-0.0360	3.2201
3	464856C	0.7418	-0.0199	0.8126	-0.8126	-0.8325	0.7926
3	464856I	0.7059	0.1523	0.6606	-0.6606	-0.5083	0.8129
3	464856O	0.6394	0.1897	1.3501	-1.3501	-1.1604	1.5397
4	126163C	0.7045	0.5089	1.0087	-1.0087	-0.4997	1.5176
4	126163I	0.7447	1.7838	0.4723	-0.4723	1.3115	2.2561
4	126163O	0.7374	2.0440	1.3395	-1.3395	0.7045	3.3835
4	469093C	0.8509	-0.1571	0.7724	-0.7724	-0.9295	0.6153
4	469093I	0.7407	0.3089	0.5270	-0.5270	-0.2181	0.8359
4	469093O	0.7416	0.3488	1.0937	-1.0937	-0.7448	1.4425
5	126972C	0.8106	0.3210	0.8910	-0.8910	-0.5700	1.2120
5	126972I	0.8932	1.6482	0.8692	-0.8692	0.7790	2.5174
5	126972O	0.9110	1.3556	0.7726	-0.7726	0.5830	2.1283
5	471924C	0.8713	-0.3914	0.6892	-0.6892	-1.0806	0.2978
5	471924I	0.8418	0.3436	0.9387	-0.9387	-0.5951	1.2823
5	471924O	0.8010	-0.0543	0.9881	-0.9881	-1.0425	0.9338
6	127286C	0.8353	0.4033	0.7377	-0.7377	-0.3344	1.1409
6	127286I	0.7846	0.8985	0.9704	-0.9704	-0.0719	1.8689
6	127286O	0.8372	1.5394	1.3975	-1.3975	0.1419	2.9369
6	471934C	0.8835	-0.0342	0.6401	-0.6401	-0.6743	0.6059
6	471934I	0.9852	0.0006	0.4699	-0.4699	-0.4693	0.4705
6	471934O	0.7926	0.3549	1.2038	-1.2038	-0.8489	1.5587
7	127658C	0.8392	0.2944	0.8990	-0.8990	-0.6046	1.1934
7	127658I	0.8735	1.3833	1.0378	-1.0378	0.3455	2.4211
7	127658O	1.0541	1.3370	1.6467	-1.6467	-0.3097	2.9836
7	471948C	0.6427	0.1409	0.7419	-0.7419	-0.6010	0.8828
7	471948I	0.8558	-0.1643	0.5405	-0.5405	-0.7049	0.3762
7	471948O	0.8977	0.2850	1.1199	-1.1199	-0.8349	1.4049
8	127794C	0.9048	-0.1088	0.8878	-0.8878	-0.9966	0.7790
8	127794I	1.2151	0.6522	1.1176	-1.1176	-0.4654	1.7697
8	127794O	1.2725	0.6348	1.2348	-1.2348	-0.6000	1.8695
8	471958C	0.8053	-0.2635	0.5480	-0.5480	-0.8115	0.2845
8	471958I	0.7503	0.4558	0.7508	-0.7508	-0.2950	1.2065
8	471958O	0.6934	0.9647	1.3724	-1.3724	-0.4077	2.3371
HS	126858C	0.8108	0.1786	0.9382	-0.9382	-0.7596	1.1169
HS	126858I	0.6797	1.6778	1.2315	-1.2315	0.4462	2.9093
HS	126858O	0.8016	0.9870	1.6437	-1.6437	-0.6567	2.6307
HS	471963C	0.8138	-0.3419	0.5427	-0.5427	-0.8846	0.2008
HS	471963I	0.7729	0.5071	0.9274	-0.9274	-0.4203	1.4345
HS	471963O	0.7428	0.3829	1.1897	-1.1897	-0.8069	1.5726

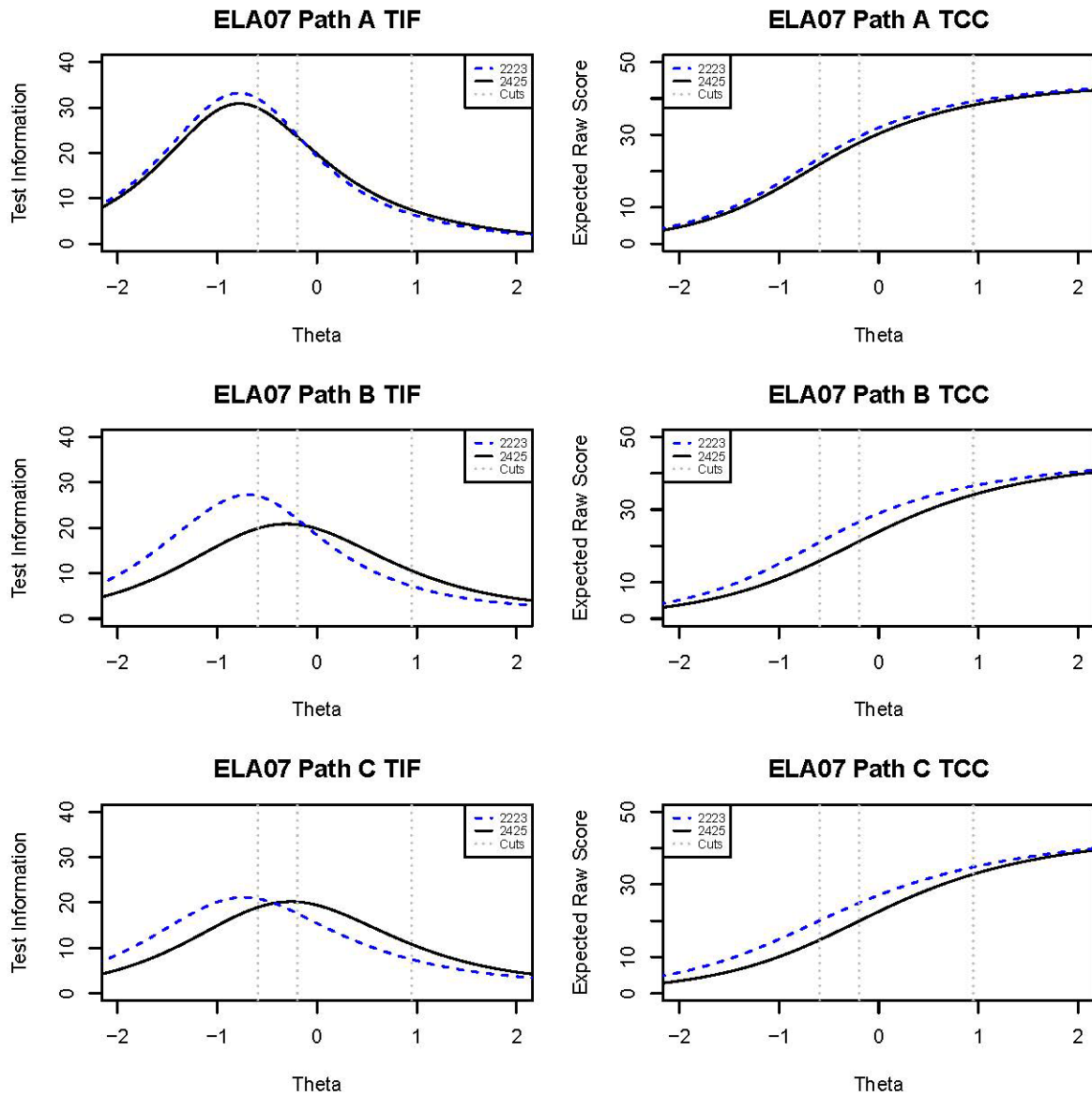
APPENDIX J
TEST CHARACTERISTIC CURVES &
TEST INFORMATION FUNCTIONS

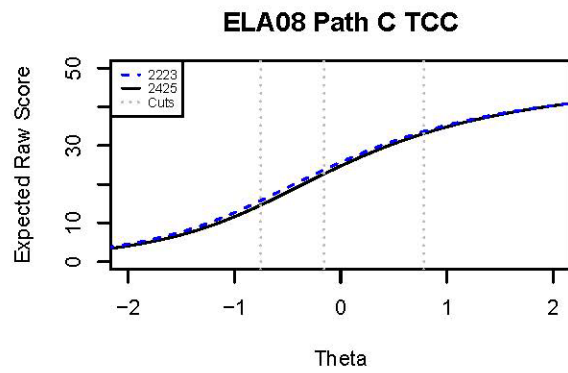
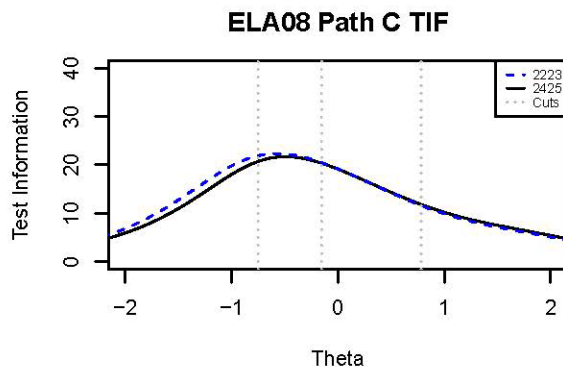
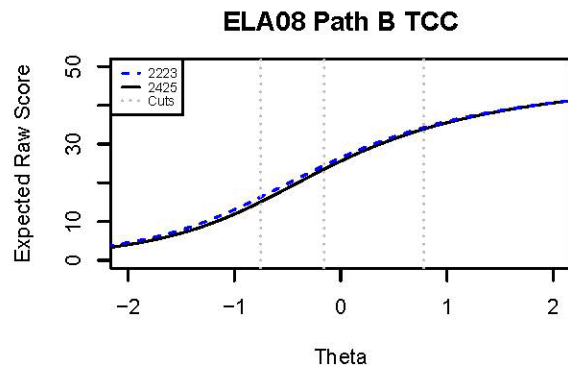
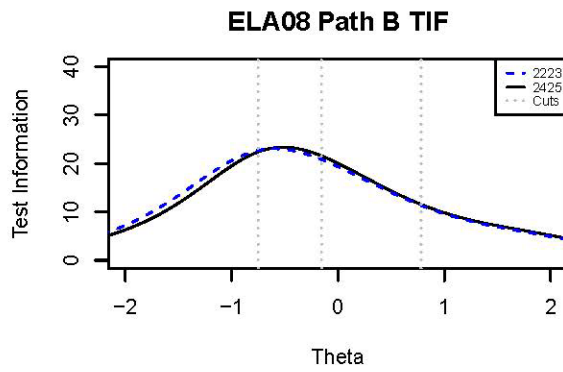
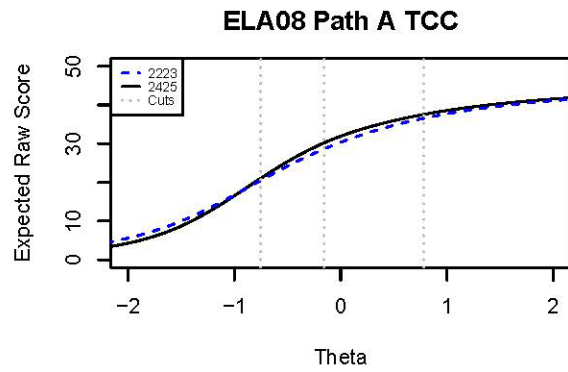
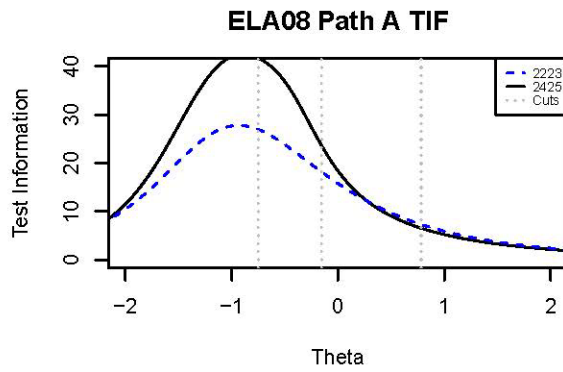


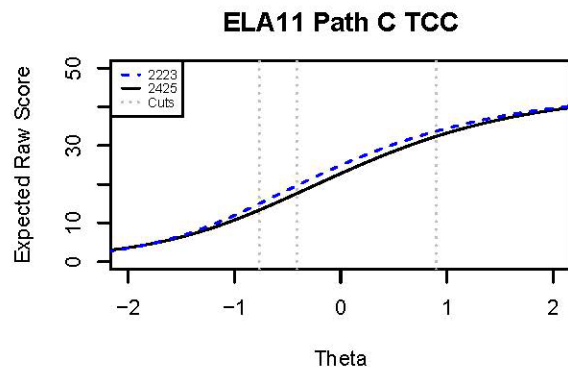
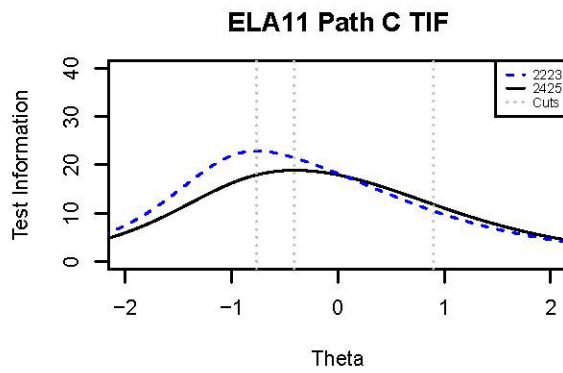
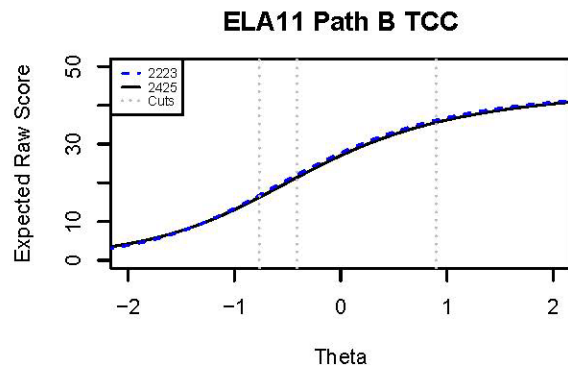
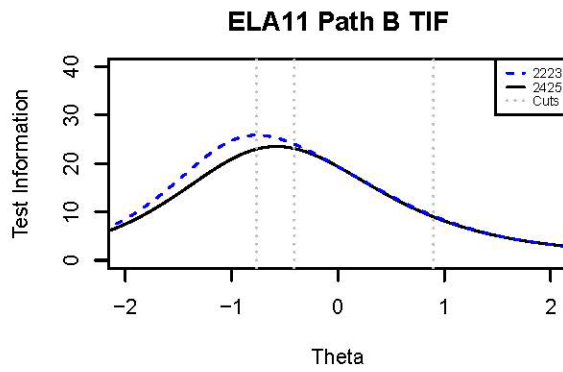
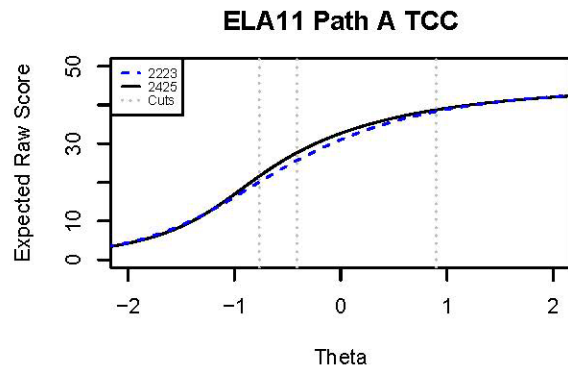
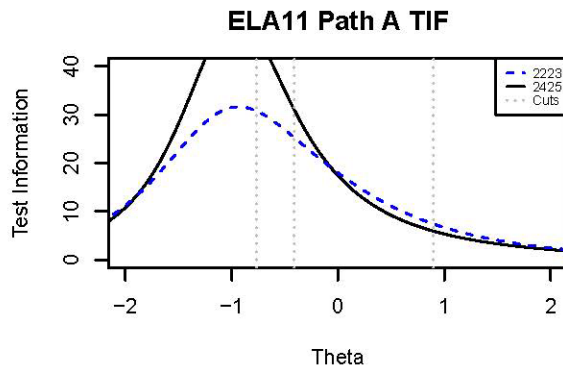


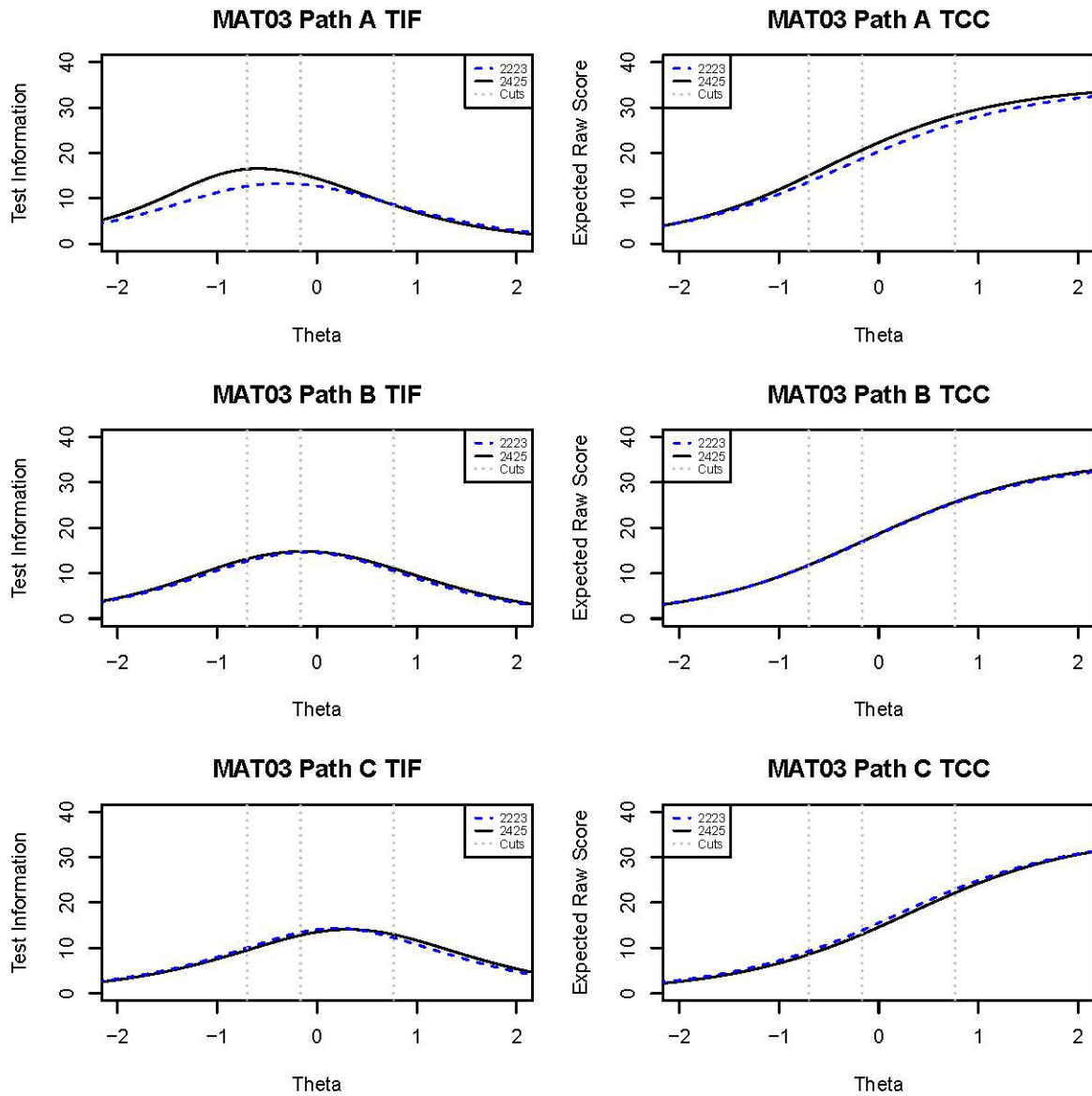


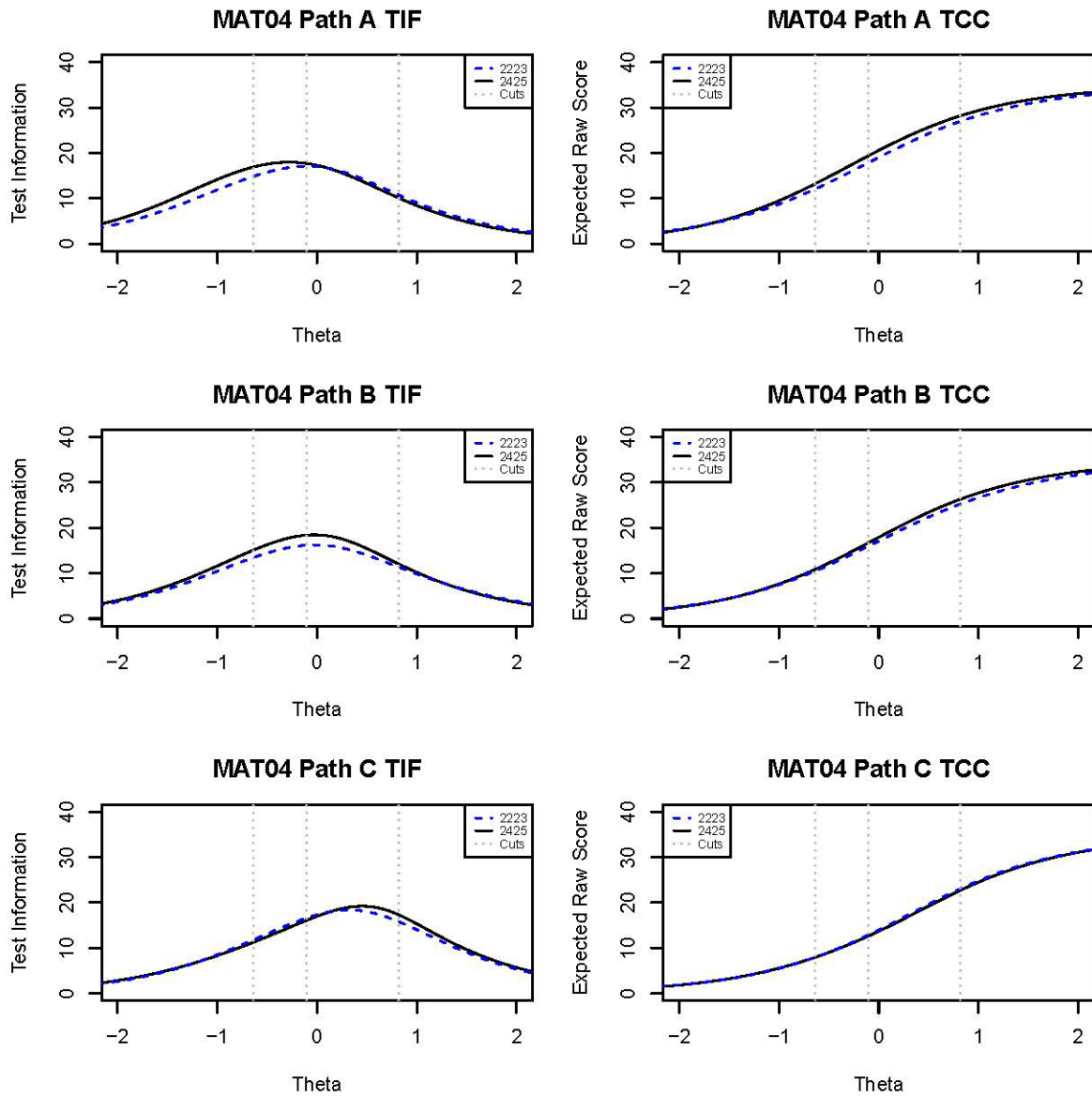


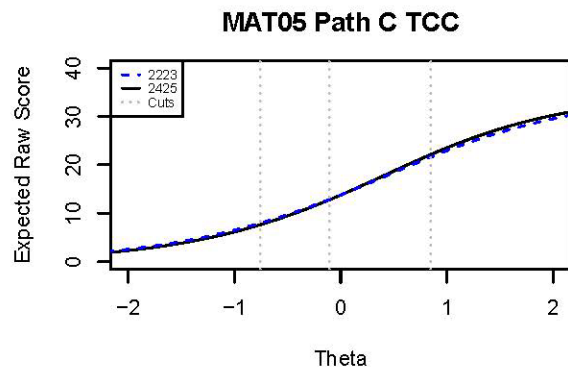
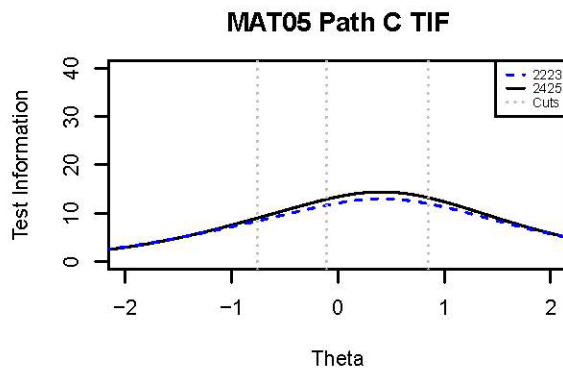
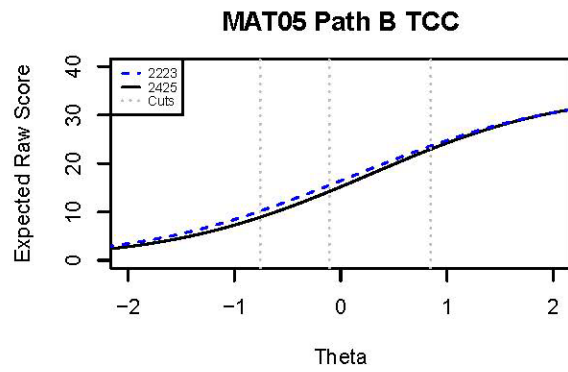
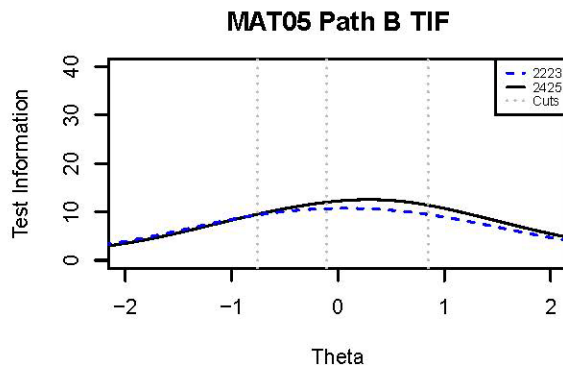
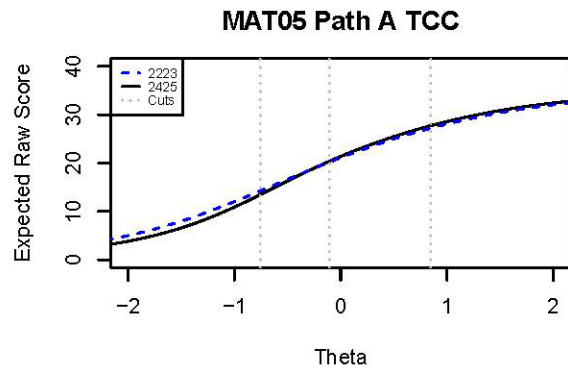
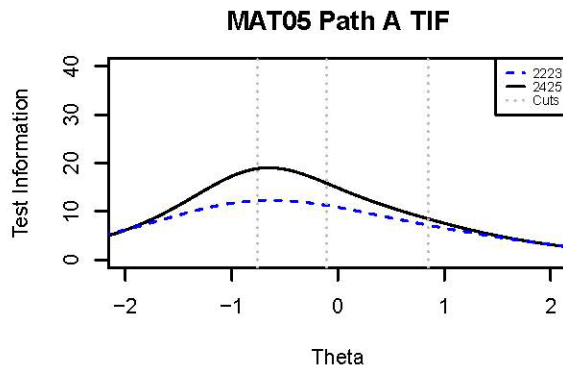


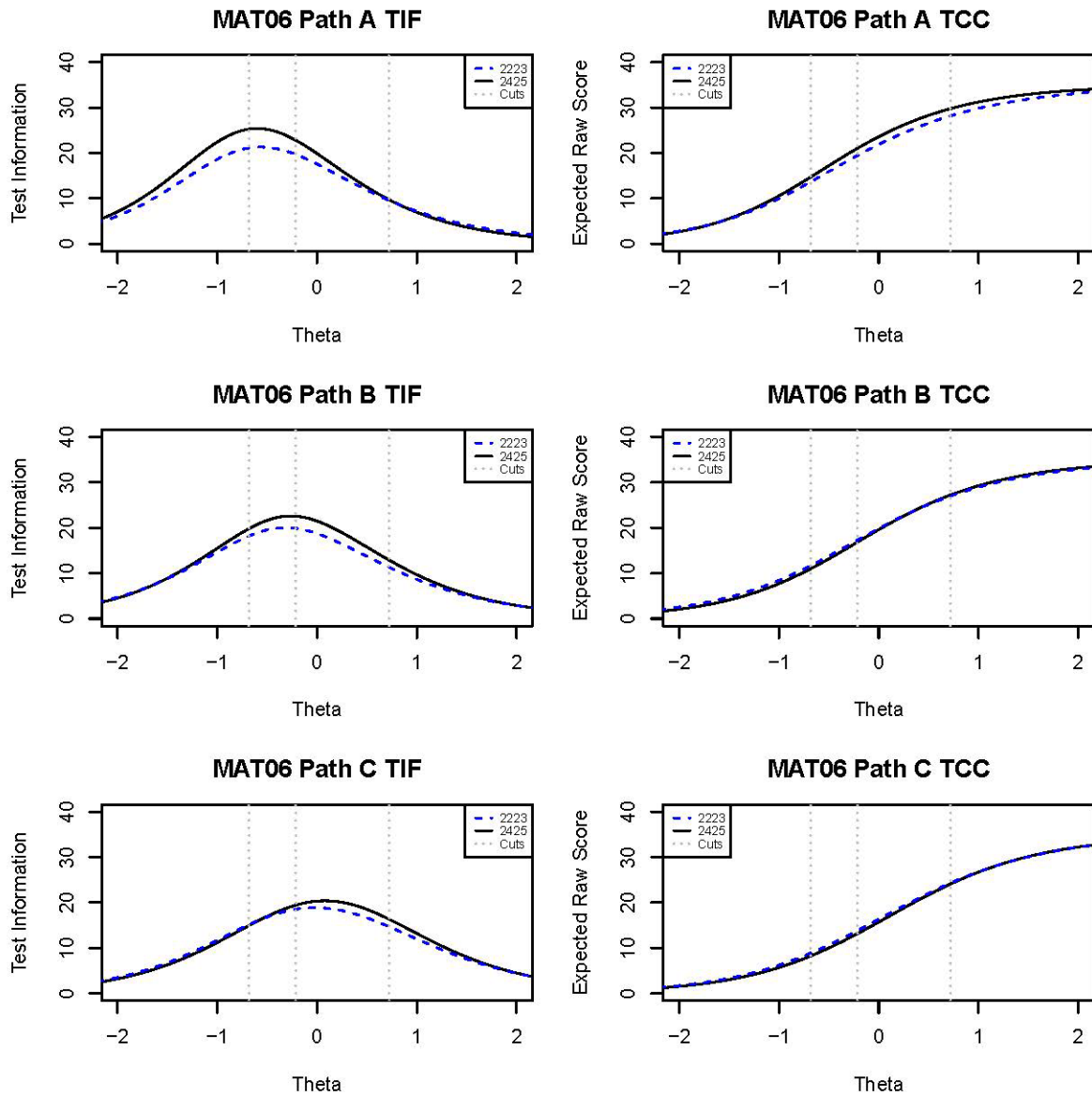


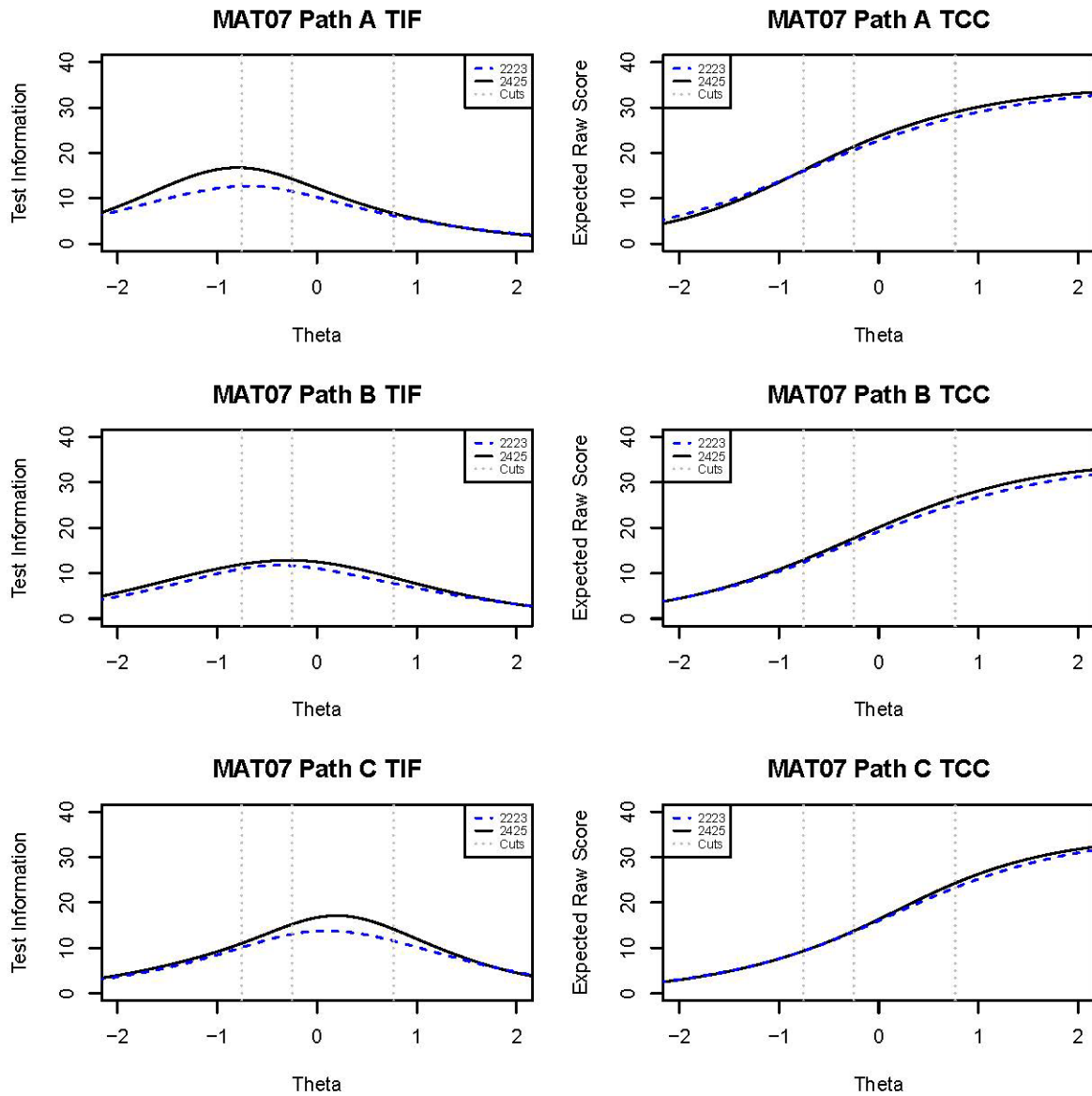


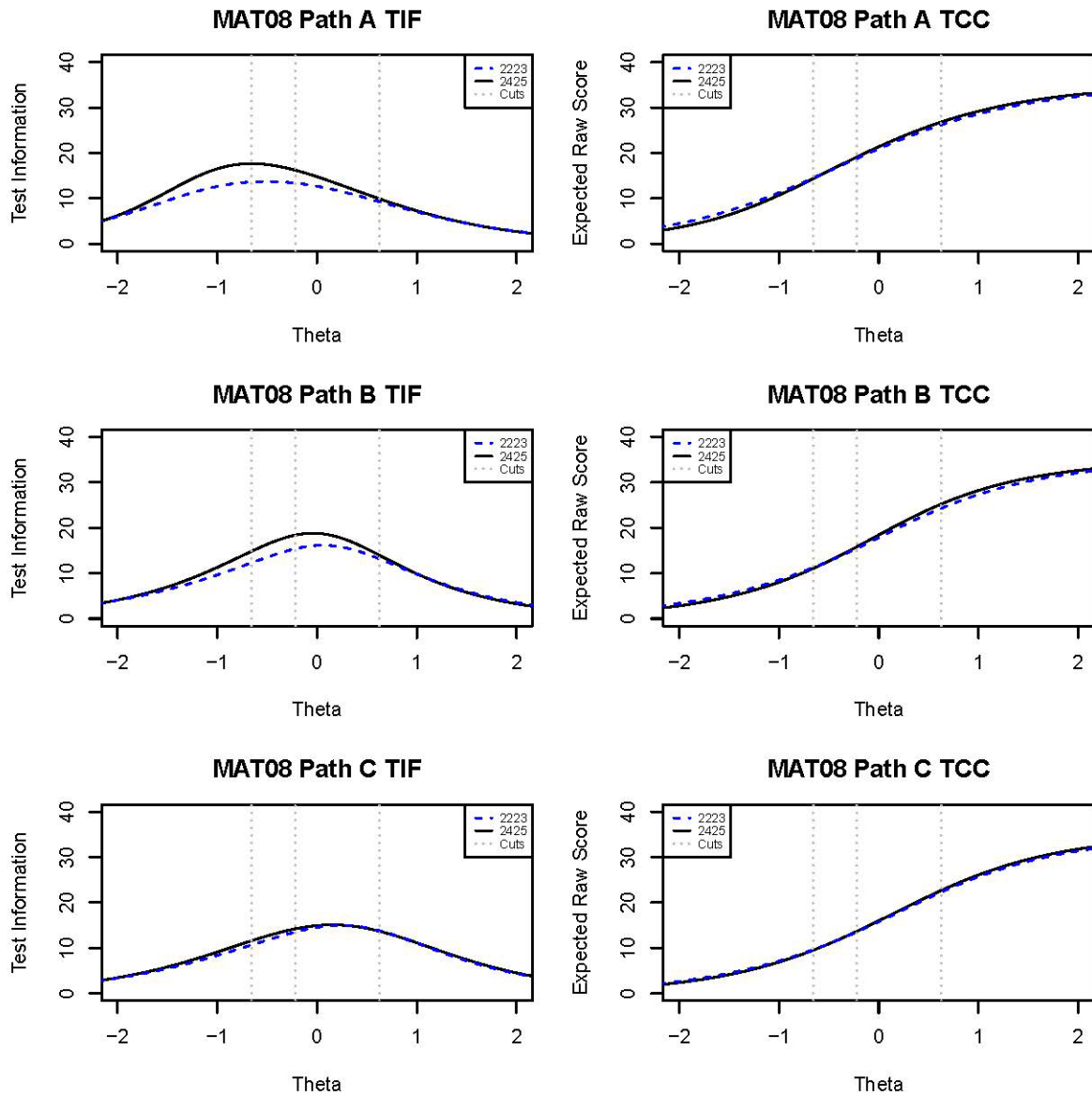


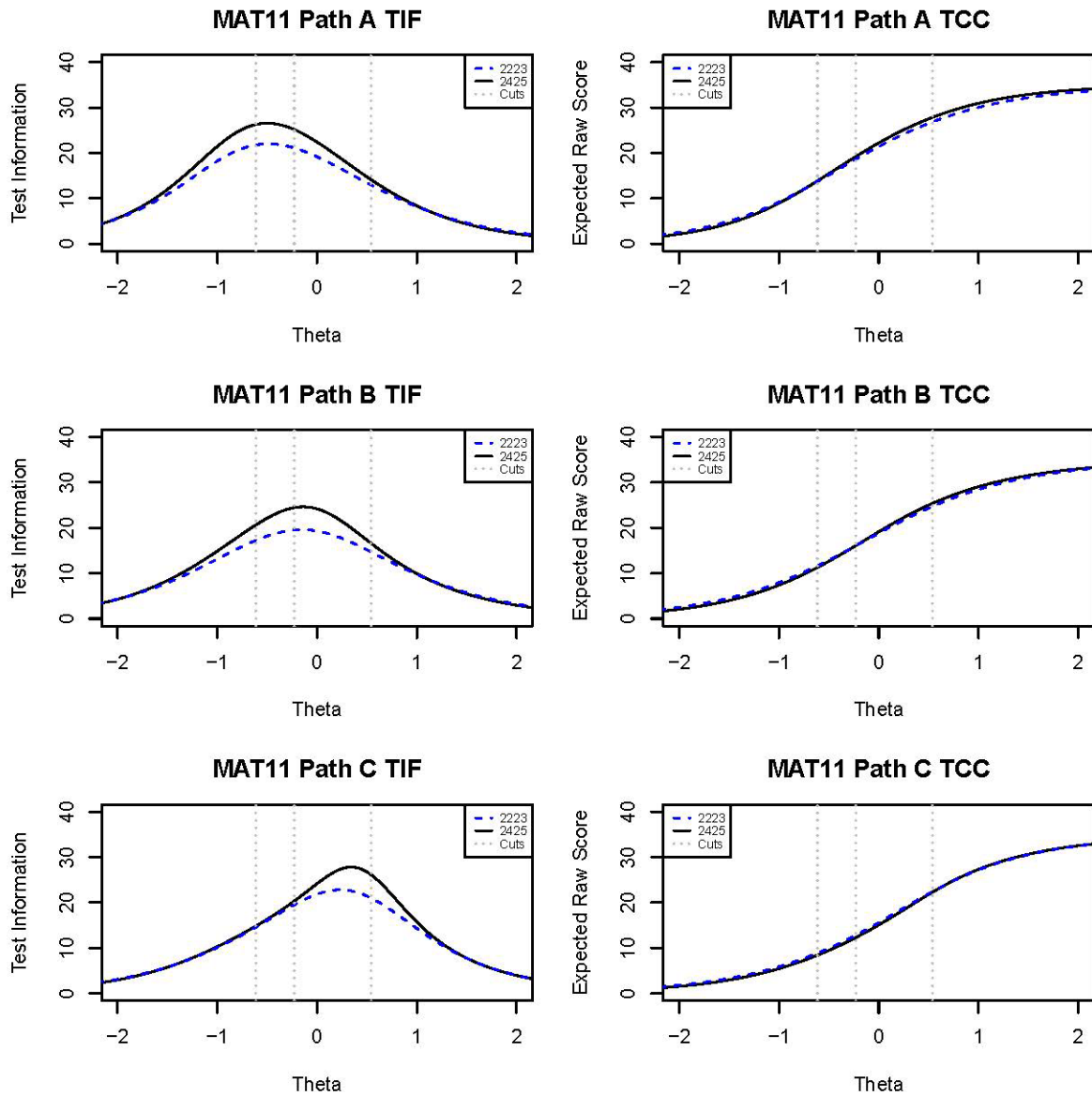












APPENDIX K
RAW TO SCALED SCORE LOOK-UP TABLES

Table K-1. Raw to Scaled Score Look-up Table—ELA Grade 3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	16.92	1	1200	19.71	1
A	1	1200	10.01	1	1201	9.82	1
A	2	1206	6.64	1	1208	6.30	1
A	3	1211	5.25	1	1212	4.95	1
A	4	1213	4.47	1	1215	4.22	1
A	5	1216	3.96	1	1217	3.74	1
A	6	1217	3.59	1	1219	3.41	1
A	7	1219	3.32	1	1220	3.16	1
A	8	1220	3.11	1	1221	2.97	1
A	9	1222	2.95	1	1222	2.82	1
A	10	1223	2.81	1	1224	2.70	1
A	11	1224	2.70	1	1225	2.60	1
A	12	1225	2.61	1	1225	2.52	1
A	13	1226	2.54	1	1226	2.45	1
A	14	1227	2.48	1	1227	2.40	1
A	15	1227	2.43	1	1228	2.36	1
A	16	1228	2.39	1	1229	2.33	1
A	17	1229	2.36	1	1230	2.31	1
A	18	1230	2.33	1	1230	2.29	1
A	19	1231	2.31	1	1231	2.28	1
A	20	1231	2.30	1	1232	2.28	1
A	21	1232	2.30	1	1233	2.28	1
A	22	1233	2.30	1	1233	2.29	1
A	23	1233	2.30	1	1234	2.31	2
A	24	1234	2.31	2	1235	2.33	2
A	25	1235	2.33	2	1236	2.36	2
A	26	1236	2.35	2	1236	2.40	2
A	27	1237	2.37	2	1237	2.44	2
A	28	1238	2.41	2	1238	2.49	2
A	29	1238	2.45	2	1239	2.55	2
A	30	1239	2.50	2	1240	2.62	3
A	31	1240	2.55	3	1241	2.70	3
A	32	1241	2.62	3	1242	2.79	3
A	33	1242	2.71	3	1243	2.90	3
A	34	1243	2.80	3	1244	3.02	3
A	35	1244	2.92	3	1245	3.15	3
A	36	1245	3.05	3	1246	3.31	3
A	37	1247	3.22	3	1248	3.49	3
A	38	1248	3.41	3	1249	3.70	3
A	39	1249	3.65	3	1251	3.94	3
A	40	1251	3.96	3	1253	4.22	3
A	41	1253	4.34	3	1255	4.56	4
A	42	1256	4.85	4	1257	4.97	4
A	43	1258	5.55	4	1260	5.51	4
A	44	1262	6.58	4	1264	6.30	4
A	45	1267	8.29	4	1269	7.66	4
A	46	1276	12.11	4	1277	10.97	4
A	47	1290	20.55	4	1290	19.29	4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	0	1200	19.01	1	1200	23.05	1
B	1	1200	11.36	1	1203	10.44	1
B	2	1208	7.68	1	1210	6.76	1
B	3	1212	6.12	1	1215	5.36	1
B	4	1216	5.23	1	1218	4.58	1
B	5	1218	4.64	1	1220	4.07	1
B	6	1221	4.22	1	1222	3.72	1
B	7	1222	3.90	1	1223	3.45	1
B	8	1224	3.65	1	1225	3.24	1
B	9	1226	3.46	1	1226	3.08	1
B	10	1227	3.30	1	1227	2.95	1
B	11	1228	3.16	1	1228	2.84	1
B	12	1229	3.05	1	1229	2.75	1
B	13	1230	2.96	1	1230	2.67	1
B	14	1232	2.89	1	1231	2.61	1
B	15	1233	2.82	1	1232	2.56	1
B	16	1233	2.77	1	1233	2.52	1
B	17	1234	2.73	2	1234	2.49	2
B	18	1235	2.70	2	1235	2.47	2
B	19	1236	2.68	2	1236	2.45	2
B	20	1237	2.67	2	1237	2.45	2
B	21	1238	2.66	2	1237	2.44	2
B	22	1239	2.67	2	1238	2.45	2
B	23	1240	2.69	3	1239	2.46	2
B	24	1241	2.71	3	1240	2.48	3
B	25	1242	2.74	3	1241	2.51	3
B	26	1243	2.79	3	1242	2.54	3
B	27	1244	2.84	3	1243	2.59	3
B	28	1245	2.90	3	1243	2.64	3
B	29	1246	2.96	3	1244	2.70	3
B	30	1247	3.04	3	1245	2.78	3
B	31	1248	3.13	3	1247	2.87	3
B	32	1249	3.23	3	1248	2.98	3
B	33	1250	3.34	3	1249	3.10	3
B	34	1252	3.47	3	1250	3.24	3
B	35	1253	3.61	3	1251	3.40	3
B	36	1254	3.77	4	1253	3.58	3
B	37	1256	3.96	4	1254	3.78	4
B	38	1258	4.17	4	1256	4.01	4
B	39	1260	4.42	4	1258	4.28	4
B	40	1262	4.72	4	1260	4.59	4
B	41	1264	5.09	4	1263	4.97	4
B	42	1267	5.56	4	1266	5.45	4
B	43	1271	6.17	4	1269	6.07	4
B	44	1275	7.00	4	1273	6.93	4
B	45	1280	8.29	4	1279	8.24	4
B	46	1289	11.06	4	1288	11.07	4
B	47	1290	11.06	4	1290	11.50	4
B	0	1200	17.97	1	1200	24.31	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	1	1200	11.96	1	1203	10.78	1
B	2	1205	8.00	1	1211	6.97	1
B	3	1210	6.33	1	1215	5.52	1
B	4	1214	5.37	1	1218	4.73	1
B	5	1217	4.75	1	1221	4.22	1
B	6	1219	4.30	1	1223	3.86	1
B	7	1221	3.97	1	1225	3.60	1
B	8	1222	3.70	1	1226	3.39	1
B	9	1224	3.50	1	1228	3.23	1
B	10	1225	3.33	1	1229	3.11	1
B	11	1227	3.19	1	1230	3.00	1
B	12	1228	3.07	1	1231	2.92	1
B	13	1229	2.97	1	1232	2.85	1
B	14	1230	2.89	1	1233	2.80	1
B	15	1231	2.82	1	1234	2.75	2
B	16	1232	2.76	1	1235	2.71	2
B	17	1233	2.71	1	1236	2.69	2
B	18	1233	2.67	1	1237	2.66	2
B	19	1235	2.64	2	1238	2.65	2
B	20	1235	2.62	2	1239	2.63	2
B	21	1236	2.61	2	1240	2.63	3
C	22	1237	2.60	2	1241	2.63	3
C	23	1238	2.61	2	1242	2.64	3
C	24	1239	2.62	2	1243	2.65	3
C	25	1240	2.64	3	1244	2.67	3
C	26	1241	2.67	3	1245	2.70	3
C	27	1242	2.71	3	1246	2.73	3
C	28	1242	2.76	3	1247	2.78	3
C	29	1243	2.83	3	1248	2.83	3
C	30	1244	2.90	3	1249	2.89	3
C	31	1246	2.99	3	1250	2.96	3
C	32	1247	3.10	3	1251	3.04	3
C	33	1248	3.23	3	1252	3.13	3
C	34	1249	3.38	3	1253	3.23	3
C	35	1250	3.55	3	1255	3.35	4
C	36	1252	3.76	3	1256	3.49	4
C	37	1254	4.02	4	1257	3.64	4
C	38	1255	4.32	4	1259	3.83	4
C	39	1257	4.69	4	1261	4.05	4
C	40	1260	5.13	4	1263	4.31	4
C	41	1263	5.67	4	1265	4.65	4
C	42	1266	6.30	4	1268	5.09	4
C	43	1270	7.03	4	1271	5.69	4
C	44	1275	7.89	4	1275	6.53	4
C	45	1281	9.13	4	1280	7.86	4
C	46	1289	11.47	4	1288	10.75	4
C	47	1290	11.47	4	1290	10.97	4

Table K-2. Raw to Scaled Score Look-up Table—ELA Grade 4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	18.51	1	1200	20.83	1
A	1	1200	11.66	1	1200	9.76	1
A	2	1203	7.49	1	1207	6.35	1
A	3	1207	5.79	1	1211	4.98	1
A	4	1211	4.83	1	1213	4.21	1
A	5	1213	4.21	1	1215	3.71	1
A	6	1215	3.77	1	1217	3.36	1
A	7	1217	3.45	1	1219	3.10	1
A	8	1218	3.20	1	1220	2.90	1
A	9	1220	3.00	1	1221	2.75	1
A	10	1221	2.85	1	1222	2.62	1
A	11	1222	2.72	1	1223	2.52	1
A	12	1223	2.62	1	1224	2.43	1
A	13	1224	2.54	1	1225	2.37	1
A	14	1225	2.48	1	1226	2.31	1
A	15	1225	2.43	1	1227	2.27	1
A	16	1226	2.40	1	1227	2.24	1
A	17	1227	2.37	1	1228	2.21	1
A	18	1228	2.36	1	1229	2.20	1
A	19	1229	2.35	1	1230	2.19	1
A	20	1229	2.36	1	1230	2.18	1
A	21	1230	2.37	1	1231	2.19	1
A	22	1231	2.39	1	1232	2.20	1
A	23	1232	2.42	1	1232	2.22	1
A	24	1233	2.46	1	1233	2.24	1
A	25	1233	2.50	1	1234	2.27	2
A	26	1234	2.55	2	1235	2.31	2
A	27	1235	2.61	2	1236	2.36	2
A	28	1236	2.68	2	1236	2.41	2
A	29	1237	2.76	2	1237	2.48	2
A	30	1238	2.85	2	1238	2.55	2
A	31	1239	2.94	2	1239	2.64	2
A	32	1240	3.05	3	1239	2.73	2
A	33	1241	3.17	3	1241	2.84	3
A	34	1242	3.30	3	1242	2.97	3
A	35	1244	3.45	3	1243	3.11	3
A	36	1245	3.62	3	1245	3.28	3
A	37	1247	3.81	3	1246	3.46	3
A	38	1248	4.02	3	1247	3.68	3
A	39	1250	4.27	3	1249	3.92	3
A	40	1252	4.57	3	1251	4.22	3
A	41	1254	4.93	3	1253	4.57	3
A	42	1257	5.40	3	1256	5.02	3
A	43	1260	6.04	4	1258	5.62	3
A	44	1264	7.00	4	1262	6.51	4
A	45	1269	8.68	4	1267	8.03	4
A	46	1278	12.76	4	1276	11.69	4
A	47	1290	19.25	4	1290	20.57	4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	0	1200	20.93	1	1200	24.07	1
B	1	1200	13.51	1	1201	10.94	1
B	2	1203	8.75	1	1209	7.25	1
B	3	1209	6.82	1	1214	5.74	1
B	4	1213	5.74	1	1217	4.89	1
B	5	1216	5.04	1	1219	4.33	1
B	6	1218	4.54	1	1221	3.94	1
B	7	1220	4.17	1	1223	3.65	1
B	8	1222	3.89	1	1225	3.42	1
B	9	1223	3.66	1	1226	3.24	1
B	10	1225	3.47	1	1227	3.09	1
B	11	1226	3.31	1	1229	2.97	1
B	12	1227	3.18	1	1230	2.87	1
B	13	1229	3.07	1	1231	2.79	1
B	14	1230	2.98	1	1232	2.73	1
B	15	1231	2.90	1	1233	2.67	1
B	16	1232	2.84	1	1234	2.63	2
B	17	1233	2.79	1	1235	2.61	2
B	18	1233	2.75	1	1236	2.59	2
B	19	1235	2.72	2	1237	2.58	2
B	20	1236	2.70	2	1237	2.58	2
B	21	1236	2.69	2	1238	2.58	2
B	22	1237	2.69	2	1239	2.60	2
B	23	1238	2.70	2	1240	2.62	3
B	24	1239	2.72	2	1241	2.65	3
B	25	1240	2.75	3	1242	2.68	3
B	26	1241	2.79	3	1243	2.73	3
B	27	1242	2.85	3	1244	2.78	3
B	28	1243	2.91	3	1245	2.83	3
B	29	1244	2.99	3	1246	2.89	3
B	30	1245	3.07	3	1247	2.96	3
B	31	1246	3.17	3	1248	3.03	3
B	32	1247	3.29	3	1249	3.11	3
B	33	1249	3.42	3	1250	3.21	3
B	34	1250	3.57	3	1252	3.31	3
B	35	1251	3.74	3	1253	3.43	3
B	36	1253	3.93	3	1254	3.57	3
B	37	1255	4.15	3	1256	3.74	3
B	38	1256	4.42	3	1257	3.94	3
B	39	1258	4.73	3	1259	4.18	4
B	40	1261	5.10	4	1261	4.48	4
B	41	1263	5.57	4	1263	4.87	4
B	42	1266	6.15	4	1266	5.37	4
B	43	1270	6.90	4	1269	6.05	4
B	44	1275	7.93	4	1273	7.02	4
B	45	1281	9.52	4	1279	8.55	4
B	46	1288	12.03	4	1288	11.75	4
B	47	1290	12.03	4	1290	11.99	4
C	0	1200	19.43	1	1200	23.74	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	1	1200	12.91	1	1201	11.16	1
C	2	1203	8.63	1	1209	7.43	1
C	3	1208	6.86	1	1213	5.90	1
C	4	1212	5.86	1	1217	5.03	1
C	5	1215	5.21	1	1219	4.47	1
C	6	1217	4.75	1	1221	4.07	1
C	7	1219	4.42	1	1223	3.77	1
C	8	1221	4.16	1	1225	3.54	1
C	9	1223	3.96	1	1226	3.35	1
C	10	1224	3.80	1	1228	3.20	1
C	11	1226	3.67	1	1229	3.08	1
C	12	1227	3.57	1	1230	2.98	1
C	13	1228	3.49	1	1231	2.90	1
C	14	1230	3.42	1	1232	2.83	1
C	15	1231	3.36	1	1233	2.78	1
C	16	1232	3.32	1	1234	2.73	2
C	17	1233	3.29	1	1235	2.70	2
C	18	1234	3.26	2	1236	2.68	2
C	19	1235	3.24	2	1237	2.67	2
C	20	1236	3.23	2	1238	2.66	2
C	21	1237	3.23	2	1239	2.66	2
C	22	1239	3.23	2	1239	2.67	2
C	23	1239	3.23	2	1241	2.69	3
C	24	1241	3.24	3	1242	2.72	3
C	25	1242	3.26	3	1243	2.75	3
C	26	1243	3.28	3	1244	2.79	3
C	27	1244	3.31	3	1245	2.83	3
C	28	1245	3.34	3	1246	2.88	3
C	29	1246	3.39	3	1247	2.94	3
C	30	1247	3.44	3	1248	3.00	3
C	31	1249	3.51	3	1249	3.07	3
C	32	1250	3.58	3	1250	3.15	3
C	33	1251	3.67	3	1251	3.24	3
C	34	1253	3.78	3	1252	3.35	3
C	35	1254	3.91	3	1254	3.47	3
C	36	1255	4.06	3	1255	3.61	3
C	37	1257	4.24	3	1257	3.77	3
C	38	1259	4.45	4	1258	3.97	3
C	39	1261	4.72	4	1260	4.21	4
C	40	1263	5.04	4	1262	4.51	4
C	41	1265	5.44	4	1264	4.89	4
C	42	1268	5.95	4	1267	5.39	4
C	43	1272	6.62	4	1270	6.05	4
C	44	1276	7.57	4	1274	7.00	4
C	45	1282	9.08	4	1279	8.51	4
C	46	1288	11.28	4	1288	11.70	4
C	47	1290	11.28	4	1290	11.70	4

Table K-3. Raw to Scaled Score Look-up Table—ELA Grade 5

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	21.09	1	1200	20.31	1
A	1	1200	9.42	1	1201	9.41	1
A	2	1207	6.07	1	1208	6.23	1
A	3	1211	4.73	1	1211	4.87	1
A	4	1214	3.97	1	1214	4.08	1
A	5	1216	3.48	1	1216	3.57	1
A	6	1217	3.14	1	1218	3.20	1
A	7	1219	2.88	1	1219	2.93	1
A	8	1220	2.68	1	1220	2.72	1
A	9	1221	2.53	1	1222	2.56	1
A	10	1222	2.40	1	1223	2.43	1
A	11	1223	2.30	1	1223	2.32	1
A	12	1224	2.22	1	1224	2.24	1
A	13	1225	2.16	1	1225	2.17	1
A	14	1225	2.11	1	1226	2.12	1
A	15	1226	2.07	1	1227	2.08	1
A	16	1227	2.05	1	1227	2.05	1
A	17	1228	2.03	1	1228	2.04	1
A	18	1228	2.03	1	1229	2.03	1
A	19	1229	2.03	1	1230	2.03	1
A	20	1230	2.05	1	1230	2.04	1
A	21	1230	2.07	1	1231	2.06	1
A	22	1231	2.10	1	1232	2.09	2
A	23	1232	2.14	2	1232	2.13	2
A	24	1233	2.19	2	1233	2.17	2
A	25	1233	2.26	2	1234	2.23	2
A	26	1234	2.33	2	1235	2.30	2
A	27	1235	2.41	2	1236	2.38	2
A	28	1236	2.50	2	1236	2.47	2
A	29	1237	2.61	2	1237	2.57	2
A	30	1238	2.73	2	1238	2.68	2
A	31	1239	2.86	2	1239	2.81	2
A	32	1240	3.00	3	1240	2.96	3
A	33	1241	3.15	3	1242	3.12	3
A	34	1243	3.32	3	1243	3.31	3
A	35	1244	3.51	3	1244	3.51	3
A	36	1246	3.73	3	1246	3.74	3
A	37	1247	3.97	3	1248	4.01	3
A	38	1249	4.25	3	1250	4.32	3
A	39	1251	4.59	3	1252	4.69	3
A	40	1254	5.02	3	1254	5.14	3
A	41	1257	5.59	4	1257	5.73	4
A	42	1260	6.42	4	1261	6.56	4
A	43	1265	7.87	4	1266	7.96	4
A	44	1274	11.41	4	1275	11.24	4
A	45	1290	25.47	4	1290	23.67	4
B	0	1200	21.75	1	1200	21.99	1
B	1	1200	12.93	1	1200	10.92	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	2	1205	8.40	1	1208	7.42	1
B	3	1210	6.56	1	1213	5.91	1
B	4	1214	5.51	1	1216	5.03	1
B	5	1217	4.83	1	1219	4.44	1
B	6	1219	4.35	1	1221	4.03	1
B	7	1221	4.00	1	1223	3.72	1
B	8	1223	3.73	1	1224	3.48	1
B	9	1224	3.53	1	1226	3.29	1
B	10	1225	3.37	1	1227	3.15	1
B	11	1227	3.25	1	1228	3.03	1
B	12	1228	3.16	1	1229	2.94	1
B	13	1229	3.08	1	1230	2.86	1
B	14	1230	3.03	1	1232	2.80	2
B	15	1232	2.99	2	1232	2.75	2
B	16	1232	2.96	2	1233	2.72	2
B	17	1233	2.94	2	1234	2.69	2
B	18	1234	2.94	2	1235	2.66	2
B	19	1235	2.94	2	1236	2.65	2
B	20	1236	2.94	2	1237	2.64	2
B	21	1237	2.96	2	1238	2.63	2
B	22	1238	2.98	2	1239	2.64	2
B	23	1239	3.01	2	1239	2.64	2
B	24	1240	3.05	3	1241	2.66	3
B	25	1241	3.10	3	1242	2.68	3
B	26	1242	3.15	3	1242	2.71	3
B	27	1244	3.21	3	1243	2.76	3
B	28	1245	3.28	3	1244	2.81	3
B	29	1246	3.36	3	1245	2.88	3
B	30	1247	3.45	3	1246	2.96	3
B	31	1248	3.56	3	1248	3.06	3
B	32	1250	3.68	3	1249	3.17	3
B	33	1251	3.82	3	1250	3.31	3
B	34	1253	3.98	3	1251	3.48	3
B	35	1254	4.17	3	1253	3.67	3
B	36	1256	4.39	4	1254	3.90	3
B	37	1258	4.65	4	1256	4.16	4
B	38	1260	4.96	4	1258	4.48	4
B	39	1263	5.34	4	1260	4.86	4
B	40	1265	5.79	4	1263	5.33	4
B	41	1269	6.36	4	1266	5.90	4
B	42	1273	7.18	4	1270	6.65	4
B	43	1279	8.61	4	1275	7.82	4
B	44	1288	12.48	4	1283	10.59	4
B	45	1290	14.46	4	1290	15.29	4
C	0	1200	22.31	1	1200	22.87	1
C	1	1200	12.47	1	1202	10.60	1
C	2	1206	8.23	1	1210	7.29	1
C	3	1211	6.52	1	1214	5.87	1
C	4	1215	5.55	1	1217	5.05	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	5	1218	4.91	1	1220	4.51	1
C	6	1220	4.46	1	1222	4.13	1
C	7	1222	4.13	1	1224	3.85	1
C	8	1224	3.88	1	1226	3.63	1
C	9	1225	3.69	1	1227	3.46	1
C	10	1227	3.54	1	1228	3.33	1
C	11	1228	3.42	1	1230	3.22	1
C	12	1229	3.33	1	1231	3.14	1
C	13	1230	3.25	1	1232	3.07	2
C	14	1232	3.20	2	1233	3.01	2
C	15	1233	3.16	2	1234	2.96	2
C	16	1234	3.13	2	1235	2.93	2
C	17	1235	3.12	2	1236	2.89	2
C	18	1236	3.11	2	1237	2.87	2
C	19	1237	3.11	2	1238	2.85	2
C	20	1238	3.11	2	1239	2.84	2
C	21	1239	3.13	2	1240	2.84	3
C	22	1240	3.15	3	1241	2.85	3
C	23	1241	3.18	3	1242	2.86	3
C	24	1242	3.21	3	1243	2.89	3
C	25	1243	3.25	3	1244	2.92	3
C	26	1245	3.30	3	1245	2.97	3
C	27	1246	3.36	3	1247	3.03	3
C	28	1247	3.42	3	1248	3.10	3
C	29	1248	3.50	3	1249	3.19	3
C	30	1249	3.59	3	1250	3.29	3
C	31	1251	3.69	3	1251	3.40	3
C	32	1252	3.81	3	1253	3.53	3
C	33	1254	3.94	3	1254	3.67	3
C	34	1255	4.10	3	1255	3.83	3
C	35	1257	4.28	4	1257	4.01	4
C	36	1259	4.49	4	1259	4.22	4
C	37	1261	4.73	4	1261	4.45	4
C	38	1263	5.02	4	1263	4.72	4
C	39	1265	5.36	4	1265	5.04	4
C	40	1268	5.77	4	1268	5.44	4
C	41	1271	6.33	4	1271	5.96	4
C	42	1276	7.15	4	1275	6.73	4
C	43	1281	8.67	4	1280	8.06	4
C	44	1290	12.73	4	1288	11.36	4
C	45	1290	13.11	4	1290	12.85	4

Table K-4. Raw to Scaled Score Look-up Table—ELA Grade 6

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	27.50	1	1200	31.88	1
A	1	1203	8.74	1	1205	10.12	1
A	2	1209	5.73	1	1212	6.70	1
A	3	1213	4.45	1	1216	5.27	1
A	4	1215	3.71	1	1219	4.45	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	5	1217	3.23	1	1221	3.91	1
A	6	1219	2.88	1	1223	3.53	1
A	7	1220	2.61	1	1225	3.25	1
A	8	1221	2.40	1	1226	3.03	1
A	9	1222	2.24	1	1227	2.85	1
A	10	1223	2.10	1	1228	2.71	1
A	11	1224	1.99	1	1229	2.60	1
A	12	1225	1.89	1	1231	2.51	2
A	13	1226	1.82	1	1231	2.44	2
A	14	1226	1.75	1	1232	2.38	2
A	15	1227	1.71	1	1233	2.33	2
A	16	1227	1.67	1	1234	2.30	2
A	17	1228	1.65	1	1235	2.28	2
A	18	1229	1.64	1	1235	2.26	2
A	19	1229	1.64	1	1236	2.26	2
A	20	1230	1.65	1	1237	2.26	3
A	21	1231	1.67	2	1238	2.27	3
A	22	1231	1.70	2	1238	2.29	3
A	23	1232	1.74	2	1239	2.31	3
A	24	1232	1.80	2	1240	2.35	3
A	25	1233	1.86	2	1241	2.39	3
A	26	1234	1.94	2	1242	2.44	3
A	27	1234	2.03	2	1243	2.50	3
A	28	1235	2.13	2	1243	2.57	3
A	29	1236	2.24	2	1244	2.65	3
A	30	1237	2.37	3	1245	2.74	3
A	31	1238	2.52	3	1246	2.85	3
A	32	1239	2.69	3	1247	2.98	3
A	33	1240	2.88	3	1249	3.13	3
A	34	1241	3.10	3	1250	3.30	3
A	35	1243	3.34	3	1251	3.51	4
A	36	1244	3.62	3	1253	3.76	4
A	37	1246	3.95	3	1255	4.06	4
A	38	1248	4.33	3	1257	4.44	4
A	39	1251	4.79	4	1259	4.94	4
A	40	1254	5.39	4	1262	5.60	4
A	41	1257	6.23	4	1266	6.53	4
A	42	1262	7.62	4	1272	7.95	4
A	43	1270	10.80	4	1280	10.86	4
A	44	1290	23.68	4	1290	14.45	4
B	0	1200	26.35	1	1200	31.93	1
B	1	1200	10.79	1	1205	7.86	1
B	2	1208	7.24	1	1211	5.12	1
B	3	1213	5.75	1	1214	4.03	1
B	4	1216	4.88	1	1216	3.43	1
B	5	1218	4.31	1	1218	3.05	1
B	6	1220	3.89	1	1219	2.79	1
B	7	1222	3.59	1	1221	2.59	1
B	8	1224	3.35	1	1222	2.45	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	9	1225	3.16	1	1223	2.33	1
B	10	1226	3.02	1	1224	2.25	1
B	11	1227	2.90	1	1224	2.18	1
B	12	1228	2.81	1	1225	2.12	1
B	13	1230	2.74	1	1226	2.08	1
B	14	1231	2.68	2	1227	2.05	1
B	15	1231	2.65	2	1227	2.03	1
B	16	1232	2.62	2	1228	2.02	1
B	17	1233	2.61	2	1229	2.02	1
B	18	1234	2.62	2	1230	2.02	1
B	19	1235	2.63	2	1231	2.02	2
B	20	1236	2.66	2	1231	2.03	2
B	21	1237	2.69	3	1232	2.05	2
B	22	1238	2.74	3	1232	2.07	2
B	23	1239	2.79	3	1233	2.10	2
B	24	1240	2.85	3	1234	2.13	2
B	25	1241	2.92	3	1235	2.16	2
B	26	1242	2.99	3	1235	2.20	2
B	27	1243	3.08	3	1236	2.25	2
B	28	1244	3.17	3	1237	2.31	3
B	29	1245	3.27	3	1238	2.37	3
B	30	1247	3.38	3	1239	2.44	3
B	31	1248	3.51	3	1240	2.52	3
B	32	1249	3.65	3	1241	2.62	3
B	33	1251	3.81	4	1242	2.74	3
B	34	1252	3.99	4	1243	2.88	3
B	35	1254	4.20	4	1244	3.05	3
B	36	1256	4.44	4	1246	3.25	3
B	37	1258	4.74	4	1247	3.50	3
B	38	1261	5.09	4	1249	3.82	3
B	39	1263	5.54	4	1251	4.23	4
B	40	1267	6.13	4	1254	4.80	4
B	41	1271	6.96	4	1257	5.62	4
B	42	1276	8.29	4	1261	6.99	4
B	43	1285	11.35	4	1269	10.09	4
B	44	1290	12.56	4	1290	25.00	4
C	0	1200	27.71	1	1200	31.33	1
C	1	1200	11.34	1	1205	9.88	1
C	2	1208	7.38	1	1212	6.69	1
C	3	1212	5.77	1	1216	5.35	1
C	4	1216	4.86	1	1219	4.58	1
C	5	1218	4.26	1	1221	4.08	1
C	6	1220	3.83	1	1223	3.72	1
C	7	1222	3.51	1	1225	3.45	1
C	8	1223	3.27	1	1226	3.24	1
C	9	1225	3.08	1	1228	3.07	1
C	10	1226	2.93	1	1229	2.94	1
C	11	1227	2.81	1	1230	2.82	1
C	12	1228	2.71	1	1231	2.73	2
C	13	1229	2.63	1	1232	2.66	2
C	14	1231	2.57	2	1233	2.59	2
C	15	1231	2.53	2	1234	2.54	2

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	16	1232	2.50	2	1235	2.50	2
C	17	1233	2.48	2	1235	2.47	2
C	18	1234	2.48	2	1236	2.45	2
C	19	1235	2.49	2	1237	2.44	3
C	20	1235	2.51	2	1238	2.44	3
C	21	1236	2.54	2	1239	2.44	3
C	22	1237	2.57	3	1240	2.45	3
C	23	1238	2.62	3	1240	2.47	3
C	24	1239	2.68	3	1241	2.50	3
C	25	1240	2.75	3	1242	2.53	3
C	26	1241	2.83	3	1243	2.57	3
C	27	1242	2.92	3	1244	2.62	3
C	28	1243	3.02	3	1245	2.68	3
C	29	1244	3.14	3	1246	2.75	3
C	30	1245	3.27	3	1247	2.83	3
C	31	1247	3.41	3	1248	2.92	3
C	32	1248	3.58	3	1249	3.03	3
C	33	1250	3.77	3	1250	3.16	3
C	34	1251	3.99	4	1251	3.31	4
C	35	1253	4.24	4	1253	3.49	4
C	36	1255	4.54	4	1254	3.71	4
C	37	1257	4.90	4	1256	3.98	4
C	38	1260	5.34	4	1258	4.32	4
C	39	1263	5.88	4	1260	4.77	4
C	40	1266	6.59	4	1263	5.38	4
C	41	1271	7.56	4	1267	6.26	4
C	42	1277	9.12	4	1272	7.65	4
C	43	1287	12.98	4	1280	10.53	4
C	44	1290	13.01	4	1290	14.18	4

Table K-5. Raw to Scaled Score Look-up Table—ELA Grade 7

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	21.93	1	1200	22.06	1
A	1	1203	9.30	1	1203	9.36	1
A	2	1210	6.21	1	1210	6.19	1
A	3	1214	4.93	1	1214	4.86	1
A	4	1216	4.19	1	1217	4.10	1
A	5	1219	3.71	1	1219	3.61	1
A	6	1220	3.37	1	1220	3.27	1
A	7	1222	3.11	1	1222	3.01	1
A	8	1223	2.91	1	1223	2.82	1
A	9	1224	2.75	1	1224	2.66	1
A	10	1225	2.63	1	1225	2.54	1
A	11	1226	2.52	1	1226	2.44	1
A	12	1227	2.44	1	1227	2.36	1
A	13	1228	2.37	1	1228	2.30	1
A	14	1229	2.32	1	1229	2.25	1
A	15	1230	2.28	1	1229	2.22	1
A	16	1231	2.25	1	1230	2.19	1
A	17	1231	2.23	1	1231	2.17	1
A	18	1232	2.22	1	1232	2.16	1
A	19	1233	2.21	1	1232	2.16	1
A	20	1234	2.22	1	1233	2.16	1
A	21	1234	2.23	1	1234	2.18	1
A	22	1236	2.25	2	1234	2.20	1
A	23	1236	2.28	2	1236	2.23	2

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	24	1237	2.32	2	1236	2.26	2
A	25	1238	2.36	2	1237	2.31	2
A	26	1238	2.41	2	1238	2.37	2
A	27	1239	2.48	2	1238	2.44	2
A	28	1240	2.55	3	1239	2.52	2
A	29	1241	2.63	3	1240	2.61	3
A	30	1242	2.73	3	1241	2.72	3
A	31	1243	2.85	3	1242	2.85	3
A	32	1244	2.98	3	1243	2.99	3
A	33	1246	3.13	3	1245	3.16	3
A	34	1247	3.31	3	1246	3.36	3
A	35	1248	3.52	3	1247	3.60	3
A	36	1250	3.77	3	1249	3.87	3
A	37	1252	4.07	3	1251	4.19	3
A	38	1254	4.43	3	1253	4.58	3
A	39	1256	4.88	4	1255	5.05	4
A	40	1259	5.46	4	1258	5.66	4
A	41	1262	6.30	4	1262	6.50	4
A	42	1267	7.68	4	1267	7.87	4
A	43	1275	10.92	4	1275	11.07	4
A	44	1290	23.40	4	1290	23.65	4
B	0	1200	24.76	1	1200	23.91	1
B	1	1204	10.77	1	1204	10.60	1
B	2	1212	7.29	1	1211	7.20	1
B	3	1216	5.82	1	1216	5.75	1
B	4	1219	4.98	1	1219	4.92	1
B	5	1222	4.42	1	1221	4.38	1
B	6	1224	4.03	1	1223	3.99	1
B	7	1226	3.74	1	1225	3.71	1
B	8	1227	3.51	1	1227	3.50	1
B	9	1229	3.34	1	1228	3.34	1
B	10	1230	3.19	1	1229	3.21	1
B	11	1231	3.08	1	1231	3.11	1
B	12	1232	2.99	1	1232	3.04	1
B	13	1233	2.92	1	1233	2.99	1
B	14	1234	2.86	1	1234	2.95	1
B	15	1236	2.81	2	1235	2.92	1
B	16	1236	2.78	2	1236	2.91	2
B	17	1237	2.75	2	1237	2.91	2
B	18	1238	2.74	2	1238	2.91	2
B	19	1239	2.74	2	1239	2.92	2
B	20	1240	2.74	3	1239	2.94	2
B	21	1241	2.75	3	1241	2.97	3
B	22	1242	2.77	3	1242	3.00	3
B	23	1243	2.80	3	1243	3.03	3
B	24	1244	2.84	3	1244	3.08	3
B	25	1245	2.89	3	1245	3.12	3
B	26	1246	2.95	3	1246	3.18	3
B	27	1247	3.02	3	1247	3.24	3
B	28	1248	3.11	3	1248	3.31	3
B	29	1249	3.20	3	1250	3.39	3
B	30	1250	3.32	3	1251	3.48	3
B	31	1252	3.45	3	1252	3.59	3
B	32	1253	3.60	3	1254	3.71	3
B	33	1255	3.77	4	1255	3.85	4
B	34	1256	3.98	4	1257	4.01	4
B	35	1258	4.23	4	1258	4.20	4
B	36	1260	4.52	4	1260	4.44	4
B	37	1262	4.86	4	1262	4.72	4
B	38	1265	5.27	4	1265	5.06	4
B	39	1268	5.75	4	1268	5.46	4
B	40	1271	6.29	4	1271	5.95	4
B	41	1276	6.94	4	1275	6.57	4
B	42	1281	8.00	4	1280	7.54	4
B	43	1290	11.08	4	1288	10.03	4
B	44	1290	11.90	4	1290	11.87	4
C	0	1200	23.21	1	1200	24.81	1
C	1	1203	10.45	1	1205	10.30	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	2	1211	7.16	1	1212	6.87	1
C	3	1215	5.76	1	1216	5.42	1
C	4	1219	4.95	1	1219	4.59	1
C	5	1221	4.41	1	1222	4.05	1
C	6	1223	4.02	1	1224	3.67	1
C	7	1225	3.73	1	1225	3.38	1
C	8	1226	3.51	1	1227	3.17	1
C	9	1228	3.33	1	1228	3.00	1
C	10	1229	3.19	1	1229	2.86	1
C	11	1230	3.07	1	1230	2.76	1
C	12	1231	2.98	1	1231	2.67	1
C	13	1232	2.90	1	1232	2.60	1
C	14	1233	2.84	1	1233	2.55	1
C	15	1234	2.79	1	1234	2.50	1
C	16	1236	2.75	2	1235	2.47	1
C	17	1236	2.73	2	1236	2.45	2
C	18	1237	2.71	2	1236	2.43	2
C	19	1238	2.70	2	1237	2.43	2
C	20	1239	2.69	2	1238	2.43	2
C	21	1239	2.70	2	1239	2.44	2
C	22	1241	2.71	3	1239	2.46	2
C	23	1242	2.74	3	1240	2.48	3
C	24	1242	2.77	3	1241	2.52	3
C	25	1243	2.81	3	1242	2.56	3
C	26	1244	2.86	3	1243	2.61	3
C	27	1245	2.92	3	1244	2.68	3
C	28	1246	2.99	3	1245	2.76	3
C	29	1247	3.07	3	1246	2.85	3
C	30	1249	3.17	3	1247	2.96	3
C	31	1250	3.29	3	1248	3.09	3
C	32	1251	3.42	3	1249	3.24	3
C	33	1252	3.58	3	1251	3.42	3
C	34	1254	3.76	3	1252	3.63	3
C	35	1256	3.98	4	1254	3.89	3
C	36	1257	4.25	4	1256	4.20	4
C	37	1259	4.57	4	1258	4.57	4
C	38	1262	4.96	4	1260	5.02	4
C	39	1265	5.44	4	1263	5.57	4
C	40	1268	6.01	4	1267	6.21	4
C	41	1272	6.68	4	1271	6.92	4
C	42	1278	7.59	4	1277	7.79	4
C	43	1286	9.76	4	1285	9.87	4
C	44	1290	12.70	4	1290	13.06	4

Table K-6. Raw to Scaled Score Look-up Table—ELA Grade 8

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	24.61	1	1200	32.36	1
A	1	1200	10.62	1	1203	10.92	1
A	2	1207	6.51	1	1211	6.74	1
A	3	1211	4.82	1	1215	5.16	1
A	4	1214	3.91	1	1218	4.31	1
A	5	1216	3.35	1	1220	3.78	1
A	6	1217	2.97	1	1222	3.42	1
A	7	1219	2.70	1	1224	3.15	1
A	8	1220	2.51	1	1225	2.96	1
A	9	1221	2.36	1	1226	2.80	1
A	10	1222	2.25	1	1227	2.68	1
A	11	1223	2.16	1	1228	2.59	1
A	12	1223	2.09	1	1229	2.51	1
A	13	1224	2.04	1	1230	2.44	2
A	14	1225	2.00	1	1231	2.39	2

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	15	1226	1.97	1	1232	2.35	2
A	16	1226	1.95	1	1233	2.32	2
A	17	1227	1.94	1	1234	2.29	2
A	18	1228	1.94	1	1234	2.27	2
A	19	1229	1.94	1	1235	2.26	2
A	20	1229	1.94	1	1236	2.26	2
A	21	1229	1.95	1	1237	2.27	2
A	22	1231	1.97	2	1237	2.28	2
A	23	1231	1.99	2	1238	2.30	3
A	24	1232	2.02	2	1239	2.34	3
A	25	1233	2.06	2	1240	2.38	3
A	26	1234	2.11	2	1241	2.43	3
A	27	1234	2.19	2	1242	2.49	3
A	28	1235	2.28	2	1243	2.56	3
A	29	1236	2.40	2	1243	2.65	3
A	30	1237	2.56	2	1245	2.75	3
A	31	1238	2.74	3	1246	2.87	3
A	32	1240	2.96	3	1247	3.01	3
A	33	1241	3.22	3	1248	3.17	3
A	34	1242	3.51	3	1250	3.35	4
A	35	1244	3.84	3	1251	3.56	4
A	36	1246	4.22	3	1253	3.80	4
A	37	1248	4.65	3	1255	4.07	4
A	38	1251	5.18	4	1257	4.37	4
A	39	1253	5.82	4	1259	4.71	4
A	40	1257	6.64	4	1262	5.14	4
A	41	1262	7.76	4	1266	5.80	4
A	42	1268	9.49	4	1271	7.17	4
A	43	1278	13.10	4	1279	11.43	4
A	44	1290	20.01	4	1290	19.92	4
A	0	1200	23.40	1	1200	31.48	1
B	1	1200	11.89	1	1204	8.82	1
B	2	1206	7.90	1	1210	5.39	1
B	3	1211	6.21	1	1213	4.10	1
B	4	1214	5.23	1	1216	3.41	1
B	5	1217	4.59	1	1217	2.99	1
B	6	1219	4.13	1	1219	2.71	1
B	7	1221	3.79	1	1220	2.50	1
B	8	1222	3.52	1	1221	2.35	1
B	9	1224	3.32	1	1222	2.24	1
B	10	1225	3.16	1	1223	2.15	1
B	11	1226	3.03	1	1224	2.09	1
B	12	1227	2.93	1	1225	2.04	1
B	13	1228	2.86	1	1225	2.01	1
B	14	1229	2.80	1	1226	1.98	1
B	15	1230	2.76	2	1227	1.97	1
B	16	1231	2.73	2	1228	1.97	1
B	17	1232	2.71	2	1228	1.97	1
B	18	1233	2.71	2	1229	1.98	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	19	1234	2.71	2	1229	2.00	1
B	20	1235	2.72	2	1230	2.03	2
B	21	1236	2.74	2	1231	2.06	2
B	22	1237	2.77	2	1232	2.10	2
B	23	1238	2.81	3	1233	2.15	2
B	24	1239	2.85	3	1233	2.20	2
B	25	1240	2.90	3	1234	2.27	2
B	26	1241	2.96	3	1235	2.33	2
B	27	1242	3.02	3	1236	2.41	2
B	28	1243	3.10	3	1237	2.49	2
B	29	1244	3.18	3	1238	2.59	3
B	30	1245	3.28	3	1239	2.69	3
B	31	1246	3.39	3	1240	2.81	3
B	32	1248	3.51	3	1241	2.94	3
B	33	1249	3.65	3	1242	3.10	3
B	34	1251	3.81	4	1244	3.28	3
B	35	1252	3.98	4	1245	3.50	3
B	36	1254	4.17	4	1247	3.77	3
B	37	1256	4.37	4	1249	4.11	3
B	38	1258	4.60	4	1251	4.54	4
B	39	1261	4.87	4	1254	5.12	4
B	40	1264	5.26	4	1257	5.94	4
B	41	1267	5.93	4	1261	7.15	4
B	42	1272	7.30	4	1268	9.13	4
B	43	1280	11.20	4	1278	13.39	4
B	44	1290	18.14	4	1290	20.74	4
C	0	1200	23.54	1	1200	33.92	1
C	1	1200	11.43	1	1205	10.14	1
C	2	1206	7.60	1	1212	6.46	1
C	3	1211	5.98	1	1216	5.03	1
C	4	1214	5.04	1	1219	4.25	1
C	5	1217	4.43	1	1221	3.77	1
C	6	1219	3.99	1	1223	3.44	1
C	7	1221	3.67	1	1225	3.21	1
C	8	1222	3.42	1	1226	3.03	1
C	9	1223	3.22	1	1227	2.90	1
C	10	1225	3.07	1	1228	2.80	1
C	11	1226	2.94	1	1229	2.72	1
C	12	1227	2.85	1	1230	2.66	2
C	13	1228	2.77	1	1231	2.62	2
C	14	1229	2.71	1	1232	2.58	2
C	15	1229	2.67	1	1233	2.56	2
C	16	1231	2.64	2	1234	2.54	2
C	17	1232	2.62	2	1235	2.53	2
C	18	1233	2.61	2	1236	2.53	2
C	19	1233	2.61	2	1237	2.53	2
C	20	1234	2.62	2	1238	2.54	3
C	21	1235	2.64	2	1239	2.55	3
C	22	1236	2.66	2	1239	2.57	3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	23	1237	2.69	2	1240	2.60	3
C	24	1238	2.74	3	1241	2.63	3
C	25	1239	2.78	3	1242	2.67	3
C	26	1240	2.84	3	1243	2.72	3
C	27	1241	2.91	3	1244	2.78	3
C	28	1242	2.99	3	1245	2.85	3
C	29	1243	3.07	3	1246	2.92	3
C	30	1244	3.17	3	1247	3.01	3
C	31	1245	3.29	3	1248	3.11	3
C	32	1247	3.42	3	1250	3.23	4
C	33	1248	3.57	3	1251	3.36	4
C	34	1250	3.73	4	1252	3.51	4
C	35	1251	3.92	4	1254	3.68	4
C	36	1253	4.12	4	1256	3.87	4
C	37	1255	4.35	4	1258	4.09	4
C	38	1257	4.59	4	1260	4.33	4
C	39	1260	4.86	4	1262	4.64	4
C	40	1263	5.23	4	1265	5.08	4
C	41	1266	5.83	4	1268	5.81	4
C	42	1271	7.07	4	1273	7.26	4
C	43	1279	10.68	4	1281	11.33	4
C	44	1290	19.10	4	1290	17.35	4

Table K-7. Raw to Scaled Score Look-up Table—ELA Grade HS

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	22.42	1	1200	24.55	1
A	1	1207	9.90	1	1210	8.31	1
A	2	1214	6.10	1	1216	5.08	1
A	3	1218	4.54	1	1219	3.80	1
A	4	1221	3.68	1	1222	3.11	1
A	5	1223	3.13	1	1223	2.67	1
A	6	1224	2.75	1	1224	2.38	1
A	7	1225	2.48	1	1226	2.17	1
A	8	1226	2.27	1	1226	2.01	1
A	9	1227	2.11	1	1227	1.89	1
A	10	1228	1.98	1	1228	1.80	1
A	11	1229	1.88	1	1229	1.72	1
A	12	1230	1.80	1	1229	1.67	1
A	13	1230	1.74	1	1230	1.62	1
A	14	1231	1.70	1	1231	1.59	1
A	15	1232	1.66	1	1231	1.57	1
A	16	1232	1.64	1	1232	1.56	1
A	17	1233	1.63	1	1232	1.55	1
A	18	1233	1.62	1	1233	1.55	1
A	19	1234	1.63	1	1233	1.56	1
A	20	1234	1.64	1	1234	1.58	1
A	21	1235	1.67	1	1234	1.60	1
A	22	1236	1.70	2	1235	1.63	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	23	1236	1.74	2	1236	1.67	2
A	24	1237	1.79	2	1236	1.72	2
A	25	1238	1.85	2	1237	1.77	2
A	26	1238	1.92	2	1238	1.84	2
A	27	1239	2.00	2	1238	1.92	2
A	28	1240	2.10	3	1239	2.00	2
A	29	1241	2.21	3	1240	2.10	3
A	30	1241	2.33	3	1241	2.22	3
A	31	1242	2.47	3	1242	2.35	3
A	32	1243	2.64	3	1243	2.49	3
A	33	1245	2.82	3	1244	2.66	3
A	34	1246	3.04	3	1245	2.86	3
A	35	1247	3.28	3	1246	3.09	3
A	36	1249	3.57	3	1248	3.35	3
A	37	1251	3.92	3	1249	3.68	3
A	38	1253	4.34	3	1251	4.07	3
A	39	1255	4.86	4	1254	4.57	3
A	40	1258	5.53	4	1256	5.22	4
A	41	1262	6.43	4	1260	6.11	4
A	42	1267	7.83	4	1265	7.46	4
A	43	1275	10.90	4	1273	10.33	4
A	44	1290	21.48	4	1290	23.46	4
B	0	1200	21.35	1	1200	24.03	1
B	1	1206	10.00	1	1210	9.17	1
B	2	1214	6.56	1	1216	6.00	1
B	3	1218	5.14	1	1220	4.67	1
B	4	1221	4.35	1	1223	3.92	1
B	5	1223	3.84	1	1225	3.44	1
B	6	1225	3.49	1	1226	3.10	1
B	7	1226	3.23	1	1228	2.85	1
B	8	1227	3.03	1	1229	2.66	1
B	9	1229	2.87	1	1230	2.51	1
B	10	1230	2.75	1	1231	2.39	1
B	11	1231	2.66	1	1232	2.30	1
B	12	1232	2.58	1	1233	2.23	1
B	13	1233	2.52	1	1234	2.17	1
B	14	1234	2.47	1	1234	2.13	1
B	15	1234	2.43	1	1235	2.10	1
B	16	1235	2.40	1	1236	2.07	2
B	17	1236	2.38	2	1237	2.05	2
B	18	1237	2.37	2	1237	2.04	2
B	19	1238	2.37	2	1238	2.04	2
B	20	1238	2.37	2	1239	2.04	2
B	21	1239	2.39	2	1239	2.05	2
B	22	1240	2.40	3	1240	2.06	3
B	23	1241	2.43	3	1241	2.08	3
B	24	1242	2.46	3	1241	2.11	3
B	25	1242	2.50	3	1242	2.14	3
B	26	1243	2.55	3	1243	2.19	3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	27	1244	2.61	3	1244	2.24	3
B	28	1245	2.68	3	1245	2.30	3
B	29	1246	2.76	3	1245	2.38	3
B	30	1247	2.86	3	1246	2.47	3
B	31	1248	2.97	3	1247	2.58	3
B	32	1249	3.11	3	1248	2.71	3
B	33	1251	3.27	3	1249	2.86	3
B	34	1252	3.45	3	1251	3.05	3
B	35	1254	3.68	3	1252	3.28	3
B	36	1255	3.96	4	1254	3.57	3
B	37	1257	4.30	4	1256	3.93	4
B	38	1260	4.73	4	1258	4.39	4
B	39	1262	5.29	4	1260	4.99	4
B	40	1266	6.01	4	1264	5.78	4
B	41	1270	6.96	4	1268	6.84	4
B	42	1276	8.31	4	1274	8.29	4
B	43	1285	11.04	4	1283	10.95	4
B	44	1290	13.56	4	1290	14.22	4
C	0	1200	23.22	1	1200	25.14	1
C	1	1208	10.24	1	1211	9.30	1
C	2	1215	6.68	1	1217	6.13	1
C	3	1219	5.23	1	1221	4.80	1
C	4	1222	4.44	1	1224	4.05	1
C	5	1225	3.93	1	1226	3.56	1
C	6	1226	3.59	1	1228	3.22	1
C	7	1228	3.34	1	1229	2.98	1
C	8	1229	3.16	1	1230	2.80	1
C	9	1231	3.02	1	1232	2.67	1
C	10	1232	2.92	1	1233	2.58	1
C	11	1233	2.83	1	1234	2.50	1
C	12	1234	2.77	1	1235	2.45	1
C	13	1235	2.73	1	1236	2.42	2
C	14	1236	2.69	2	1237	2.39	2
C	15	1237	2.67	2	1237	2.38	2
C	16	1238	2.65	2	1238	2.37	2
C	17	1239	2.65	2	1239	2.36	2
C	18	1240	2.64	3	1240	2.36	3
C	19	1241	2.65	3	1241	2.36	3
C	20	1242	2.66	3	1242	2.37	3
C	21	1243	2.67	3	1243	2.38	3
C	22	1244	2.70	3	1243	2.39	3
C	23	1244	2.72	3	1244	2.41	3
C	24	1245	2.75	3	1245	2.43	3
C	25	1246	2.79	3	1246	2.46	3
C	26	1247	2.83	3	1247	2.49	3
C	27	1248	2.89	3	1248	2.53	3
C	28	1249	2.95	3	1249	2.58	3
C	29	1251	3.01	3	1250	2.64	3
C	30	1252	3.10	3	1251	2.71	3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	31	1253	3.19	3	1252	2.80	3
C	32	1254	3.30	3	1253	2.90	3
C	33	1255	3.43	4	1254	3.02	3
C	34	1257	3.58	4	1255	3.16	4
C	35	1258	3.76	4	1257	3.34	4
C	36	1260	3.98	4	1258	3.55	4
C	37	1262	4.25	4	1260	3.81	4
C	38	1264	4.58	4	1262	4.14	4
C	39	1267	5.02	4	1264	4.58	4
C	40	1270	5.60	4	1267	5.17	4
C	41	1274	6.42	4	1271	6.03	4
C	42	1279	7.72	4	1276	7.39	4
C	43	1287	10.58	4	1284	10.22	4
C	44	1290	12.10	4	1290	13.20	4

Table K-8. Raw to Scaled Score Look-up Table—Mathematics Grade 3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	21.39	1	1200	19.92	1
A	1	1200	12.24	1	1200	12.97	1
A	2	1206	8.40	1	1204	8.96	1
A	3	1212	6.71	1	1209	7.26	1
A	4	1215	5.71	1	1213	6.27	1
A	5	1218	5.04	1	1216	5.62	1
A	6	1221	4.56	1	1219	5.15	1
A	7	1223	4.20	1	1221	4.80	1
A	8	1225	3.93	1	1223	4.53	1
A	9	1226	3.72	1	1225	4.32	1
A	10	1228	3.56	1	1227	4.14	1
A	11	1229	3.44	1	1229	4.00	1
A	12	1231	3.35	1	1230	3.88	1
A	13	1232	3.29	1	1232	3.79	1
A	14	1233	3.24	1	1233	3.71	1
A	15	1234	3.22	1	1234	3.65	1
A	16	1236	3.21	2	1236	3.60	2
A	17	1237	3.21	2	1237	3.57	2
A	18	1238	3.23	2	1238	3.55	2
A	19	1239	3.26	2	1240	3.54	2
A	20	1241	3.30	2	1241	3.55	2
A	21	1242	3.35	3	1242	3.58	3
A	22	1243	3.42	3	1244	3.62	3
A	23	1245	3.51	3	1245	3.68	3
A	24	1246	3.61	3	1247	3.76	3
A	25	1248	3.74	3	1248	3.87	3
A	26	1249	3.91	3	1250	4.02	3
A	27	1251	4.11	3	1252	4.21	3
A	28	1253	4.37	3	1253	4.46	3
A	29	1255	4.70	4	1256	4.79	4
A	30	1258	5.14	4	1258	5.26	4
A	31	1261	5.75	4	1261	5.95	4
A	32	1264	6.66	4	1265	7.03	4
A	33	1270	8.23	4	1271	8.97	4
A	34	1278	11.82	4	1281	13.58	4
A	35	1290	24.68	4	1290	23.71	4
B	0	1200	29.01	1	1200	29.91	1
B	1	1205	12.47	1	1205	12.68	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	2	1214	8.65	1	1214	8.66	1
B	3	1219	7.01	1	1219	6.96	1
B	4	1223	6.05	1	1223	5.98	1
B	5	1227	5.42	1	1226	5.33	1
B	6	1229	4.97	1	1229	4.87	1
B	7	1231	4.62	1	1231	4.52	1
B	8	1233	4.36	1	1233	4.25	1
B	9	1235	4.15	2	1235	4.03	2
B	10	1237	3.98	2	1237	3.86	2
B	11	1239	3.84	2	1238	3.72	2
B	12	1240	3.74	2	1240	3.60	2
B	13	1241	3.65	2	1241	3.51	2
B	14	1243	3.58	3	1242	3.44	3
B	15	1244	3.53	3	1244	3.38	3
B	16	1246	3.50	3	1245	3.34	3
B	17	1247	3.48	3	1246	3.32	3
B	18	1248	3.48	3	1247	3.31	3
B	19	1249	3.49	3	1249	3.31	3
B	20	1251	3.52	3	1250	3.34	3
B	21	1252	3.56	3	1251	3.38	3
B	22	1253	3.62	3	1253	3.44	3
B	23	1255	3.69	4	1254	3.53	4
B	24	1256	3.79	4	1255	3.65	4
B	25	1258	3.92	4	1257	3.81	4
B	26	1260	4.08	4	1259	4.00	4
B	27	1261	4.27	4	1260	4.25	4
B	28	1263	4.52	4	1262	4.57	4
B	29	1266	4.85	4	1265	4.97	4
B	30	1268	5.28	4	1267	5.51	4
B	31	1271	5.88	4	1271	6.24	4
B	32	1275	6.78	4	1275	7.32	4
B	33	1280	8.33	4	1281	9.15	4
B	34	1289	11.95	4	1290	13.36	4
B	35	1290	16.00	4	1290	16.39	4
C	0	1200	24.16	1	1200	24.11	1
C	1	1201	12.06	1	1200	12.62	1
C	2	1210	8.38	1	1209	8.69	1
C	3	1215	6.78	1	1214	7.02	1
C	4	1219	5.83	1	1218	6.07	1
C	5	1222	5.20	1	1221	5.43	1
C	6	1224	4.74	1	1224	4.98	1
C	7	1227	4.40	1	1226	4.63	1
C	8	1228	4.14	1	1228	4.36	1
C	9	1230	3.94	1	1230	4.14	1
C	10	1232	3.79	1	1232	3.95	1
C	11	1233	3.67	1	1233	3.79	1
C	12	1235	3.58	2	1235	3.65	2
C	13	1236	3.51	2	1236	3.53	2
C	14	1238	3.46	2	1238	3.42	2
C	15	1239	3.43	2	1239	3.33	2
C	16	1240	3.40	2	1240	3.24	2
C	17	1242	3.39	3	1242	3.17	3
C	18	1243	3.40	3	1243	3.12	3
C	19	1244	3.41	3	1244	3.08	3
C	20	1245	3.44	3	1245	3.05	3
C	21	1247	3.48	3	1246	3.05	3
C	22	1248	3.54	3	1248	3.06	3
C	23	1250	3.61	3	1249	3.09	3
C	24	1251	3.71	3	1250	3.15	3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	25	1253	3.83	3	1252	3.25	3
C	26	1254	3.97	4	1253	3.38	3
C	27	1256	4.15	4	1255	3.57	4
C	28	1258	4.38	4	1256	3.83	4
C	29	1260	4.68	4	1258	4.20	4
C	30	1262	5.07	4	1261	4.71	4
C	31	1265	5.62	4	1264	5.47	4
C	32	1269	6.46	4	1267	6.62	4
C	33	1274	7.92	4	1273	8.60	4
C	34	1282	11.31	4	1283	13.08	4
C	35	1290	20.84	4	1290	22.15	4

Table K-9. Raw to Scaled Score Look-up Table—Mathematics Grade 4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	30.19	1	1200	27.76	1
A	1	1203	10.67	1	1203	9.53	1
A	2	1211	7.35	1	1210	6.67	1
A	3	1216	5.95	1	1214	5.46	1
A	4	1219	5.16	1	1217	4.76	1
A	5	1222	4.63	1	1219	4.30	1
A	6	1224	4.26	1	1221	3.98	1
A	7	1226	3.99	1	1223	3.73	1
A	8	1227	3.77	1	1225	3.54	1
A	9	1229	3.60	1	1226	3.40	1
A	10	1230	3.47	1	1227	3.28	1
A	11	1232	3.36	2	1229	3.18	1
A	12	1233	3.27	2	1230	3.11	1
A	13	1234	3.19	2	1231	3.05	1
A	14	1235	3.14	2	1232	3.00	2
A	15	1237	3.09	2	1233	2.97	2
A	16	1238	3.07	2	1234	2.95	2
A	17	1239	3.05	3	1235	2.94	2
A	18	1240	3.05	3	1236	2.94	2
A	19	1241	3.06	3	1238	2.95	2
A	20	1242	3.09	3	1239	2.98	3
A	21	1243	3.14	3	1240	3.01	3
A	22	1245	3.21	3	1241	3.06	3
A	23	1246	3.29	3	1242	3.13	3
A	24	1247	3.41	3	1243	3.21	3
A	25	1249	3.54	3	1245	3.32	3
A	26	1250	3.72	3	1246	3.45	3
A	27	1252	3.93	4	1247	3.62	3
A	28	1254	4.20	4	1249	3.84	3
A	29	1256	4.54	4	1251	4.12	4
A	30	1258	4.99	4	1253	4.51	4
A	31	1261	5.60	4	1256	5.05	4
A	32	1265	6.50	4	1259	5.86	4
A	33	1270	8.02	4	1263	7.27	4
A	34	1278	11.50	4	1271	10.53	4
A	35	1290	21.19	4	1290	28.70	4
B	0	1200	27.72	1	1200	31.00	1
B	1	1202	10.18	1	1204	10.52	1
B	2	1209	7.04	1	1212	7.26	1
B	3	1214	5.71	1	1216	5.89	1
B	4	1217	4.95	1	1220	5.10	1
B	5	1219	4.46	1	1222	4.58	1
B	6	1221	4.11	1	1224	4.22	1
B	7	1223	3.85	1	1226	3.95	1
B	8	1225	3.65	1	1228	3.74	1
B	9	1226	3.49	1	1229	3.58	1
B	10	1228	3.37	1	1231	3.44	1
B	11	1229	3.27	1	1232	3.33	2
B	12	1230	3.20	1	1234	3.25	2

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	13	1231	3.14	1	1235	3.17	2
B	14	1233	3.09	2	1236	3.11	2
B	15	1234	3.06	2	1237	3.06	2
B	16	1235	3.04	2	1238	3.03	2
B	17	1236	3.04	2	1239	3.01	3
B	18	1237	3.04	2	1240	2.99	3
B	19	1238	3.06	2	1242	3.00	3
B	20	1239	3.08	3	1243	3.01	3
B	21	1240	3.13	3	1244	3.05	3
B	22	1242	3.18	3	1245	3.10	3
B	23	1243	3.26	3	1246	3.17	3
B	24	1244	3.35	3	1247	3.26	3
B	25	1245	3.47	3	1249	3.38	3
B	26	1247	3.62	3	1250	3.54	3
B	27	1249	3.81	3	1252	3.74	4
B	28	1250	4.05	3	1254	3.99	4
B	29	1252	4.37	4	1255	4.32	4
B	30	1254	4.79	4	1258	4.77	4
B	31	1257	5.38	4	1261	5.39	4
B	32	1261	6.27	4	1264	6.31	4
B	33	1266	7.82	4	1269	7.91	4
B	34	1274	11.39	4	1277	11.59	4
B	35	1290	25.51	4	1290	22.11	4
C	0	1200	35.42	1	1200	35.65	1
C	1	1207	10.77	1	1207	10.75	1
C	2	1215	7.44	1	1215	7.42	1
C	3	1219	6.05	1	1220	6.03	1
C	4	1223	5.27	1	1223	5.24	1
C	5	1225	4.76	1	1226	4.72	1
C	6	1228	4.39	1	1228	4.35	1
C	7	1230	4.11	1	1230	4.06	1
C	8	1232	3.89	2	1232	3.82	2
C	9	1233	3.71	2	1234	3.62	2
C	10	1235	3.56	2	1235	3.45	2
C	11	1236	3.43	2	1237	3.30	2
C	12	1238	3.32	2	1238	3.17	2
C	13	1239	3.23	3	1239	3.05	3
C	14	1240	3.16	3	1241	2.96	3
C	15	1241	3.09	3	1242	2.88	3
C	16	1243	3.05	3	1243	2.82	3
C	17	1244	3.01	3	1244	2.78	3
C	18	1245	2.99	3	1245	2.75	3
C	19	1246	2.99	3	1246	2.75	3
C	20	1247	3.00	3	1247	2.77	3
C	21	1248	3.04	3	1248	2.81	3
C	22	1250	3.09	3	1249	2.87	3
C	23	1251	3.18	4	1251	2.96	4
C	24	1252	3.28	4	1252	3.07	4
C	25	1254	3.42	4	1253	3.21	4
C	26	1255	3.59	4	1255	3.38	4
C	27	1257	3.81	4	1256	3.59	4
C	28	1259	4.08	4	1258	3.85	4
C	29	1261	4.41	4	1260	4.18	4
C	30	1263	4.85	4	1262	4.61	4
C	31	1266	5.44	4	1265	5.20	4
C	32	1269	6.32	4	1268	6.10	4
C	33	1274	7.83	4	1273	7.66	4
C	34	1283	11.33	4	1282	11.34	4
C	35	1290	17.36	4	1290	18.47	4

Table K-10. Raw to Scaled Score Look-up Table—Mathematics Grade 5

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	24.06	1	1200	23.88	1
A	1	1200	11.56	1	1200	12.18	1
A	2	1207	7.76	1	1206	8.05	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	3	1212	6.11	1	1211	6.35	1
A	4	1216	5.16	1	1215	5.40	1
A	5	1218	4.54	1	1217	4.80	1
A	6	1221	4.10	1	1220	4.39	1
A	7	1223	3.79	1	1222	4.09	1
A	8	1224	3.55	1	1224	3.87	1
A	9	1226	3.37	1	1225	3.71	1
A	10	1227	3.24	1	1227	3.58	1
A	11	1228	3.14	1	1228	3.50	1
A	12	1230	3.07	1	1230	3.43	1
A	13	1231	3.02	1	1231	3.39	1
A	14	1232	3.00	2	1232	3.37	2
A	15	1233	3.00	2	1234	3.36	2
A	16	1235	3.02	2	1235	3.36	2
A	17	1236	3.05	2	1236	3.38	2
A	18	1237	3.10	2	1238	3.40	2
A	19	1238	3.17	2	1239	3.43	2
A	20	1239	3.25	2	1240	3.46	3
A	21	1241	3.35	3	1242	3.49	3
A	22	1242	3.47	3	1243	3.54	3
A	23	1244	3.60	3	1245	3.59	3
A	24	1245	3.74	3	1246	3.65	3
A	25	1247	3.90	3	1248	3.73	3
A	26	1249	4.09	3	1249	3.85	3
A	27	1251	4.31	3	1251	4.00	3
A	28	1253	4.57	4	1253	4.22	4
A	29	1256	4.91	4	1255	4.51	4
A	30	1258	5.35	4	1258	4.91	4
A	31	1262	5.96	4	1261	5.49	4
A	32	1266	6.86	4	1264	6.40	4
A	33	1271	8.38	4	1270	8.07	4
A	34	1280	11.91	4	1279	12.25	4
A	35	1290	21.01	4	1290	23.06	4
B	0	1200	27.87	1	1200	29.74	1
B	1	1202	12.02	1	1203	12.11	1
B	2	1211	8.32	1	1212	8.18	1
B	3	1216	6.75	1	1217	6.55	1
B	4	1220	5.86	1	1221	5.64	1
B	5	1223	5.28	1	1224	5.05	1
B	6	1226	4.89	1	1226	4.64	1
B	7	1228	4.60	1	1228	4.33	1
B	8	1230	4.39	1	1230	4.10	1
B	9	1232	4.22	2	1232	3.92	2
B	10	1233	4.09	2	1234	3.77	2
B	11	1235	3.99	2	1235	3.65	2
B	12	1237	3.91	2	1236	3.55	2
B	13	1238	3.84	2	1238	3.48	2
B	14	1239	3.79	2	1239	3.41	2
B	15	1241	3.74	3	1241	3.37	3
B	16	1243	3.72	3	1242	3.33	3
B	17	1244	3.70	3	1243	3.32	3
B	18	1245	3.69	3	1244	3.31	3
B	19	1247	3.70	3	1246	3.32	3
B	20	1248	3.72	3	1247	3.34	3
B	21	1250	3.76	3	1248	3.38	3
B	22	1251	3.81	3	1249	3.43	3
B	23	1253	3.88	4	1251	3.50	3
B	24	1254	3.97	4	1252	3.60	3
B	25	1256	4.08	4	1254	3.71	4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	26	1257	4.23	4	1255	3.85	4
B	27	1259	4.40	4	1257	4.03	4
B	28	1261	4.63	4	1259	4.25	4
B	29	1264	4.93	4	1261	4.54	4
B	30	1266	5.33	4	1263	4.93	4
B	31	1269	5.90	4	1266	5.48	4
B	32	1273	6.77	4	1270	6.33	4
B	33	1278	8.30	4	1275	7.84	4
B	34	1287	11.95	4	1283	11.51	4
B	35	1290	15.94	4	1290	18.59	4
C	0	1200	30.65	1	1200	32.04	1
C	1	1205	11.79	1	1205	11.53	1
C	2	1213	8.18	1	1214	7.85	1
C	3	1219	6.65	1	1219	6.34	1
C	4	1222	5.77	1	1222	5.50	1
C	5	1225	5.20	1	1225	4.95	1
C	6	1228	4.80	1	1228	4.56	1
C	7	1230	4.51	1	1230	4.27	1
C	8	1232	4.28	2	1232	4.04	2
C	9	1234	4.09	2	1234	3.85	2
C	10	1236	3.94	2	1235	3.70	2
C	11	1237	3.81	2	1237	3.57	2
C	12	1239	3.71	2	1238	3.47	2
C	13	1240	3.62	3	1239	3.39	2
C	14	1242	3.56	3	1241	3.32	3
C	15	1243	3.50	3	1242	3.27	3
C	16	1244	3.47	3	1243	3.24	3
C	17	1246	3.45	3	1245	3.23	3
C	18	1247	3.44	3	1246	3.23	3
C	19	1248	3.46	3	1247	3.24	3
C	20	1250	3.48	3	1248	3.27	3
C	21	1251	3.52	3	1249	3.31	3
C	22	1252	3.59	3	1251	3.37	3
C	23	1254	3.67	4	1252	3.45	3
C	24	1255	3.77	4	1254	3.54	4
C	25	1257	3.90	4	1255	3.67	4
C	26	1258	4.06	4	1257	3.81	4
C	27	1260	4.26	4	1258	4.00	4
C	28	1262	4.51	4	1260	4.23	4
C	29	1264	4.84	4	1262	4.54	4
C	30	1267	5.27	4	1265	4.94	4
C	31	1270	5.88	4	1267	5.50	4
C	32	1274	6.80	4	1271	6.35	4
C	33	1279	8.39	4	1276	7.81	4
C	34	1288	12.08	4	1284	11.25	4
C	35	1290	15.39	4	1290	17.19	4

Table K-11. Raw to Scaled Score Look-up Table—Mathematics Grade 6

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	33.52	1	1200	26.44	1
A	1	1209	8.85	1	1205	8.87	1
A	2	1215	6.14	1	1211	6.10	1
A	3	1219	5.00	1	1215	4.92	1
A	4	1222	4.35	1	1218	4.24	1
A	5	1224	3.92	1	1220	3.80	1
A	6	1226	3.61	1	1222	3.49	1
A	7	1227	3.39	1	1223	3.26	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	8	1229	3.21	1	1225	3.08	1
A	9	1230	3.07	1	1226	2.95	1
A	10	1231	2.96	1	1227	2.84	1
A	11	1233	2.88	2	1228	2.75	1
A	12	1234	2.81	2	1229	2.69	1
A	13	1235	2.76	2	1230	2.64	1
A	14	1236	2.73	2	1231	2.60	1
A	15	1237	2.71	2	1232	2.58	1
A	16	1238	2.70	2	1233	2.57	2
A	17	1239	2.70	3	1234	2.58	2
A	18	1240	2.72	3	1235	2.59	2
A	19	1241	2.74	3	1236	2.62	2
A	20	1242	2.78	3	1237	2.67	2
A	21	1243	2.84	3	1238	2.72	2
A	22	1244	2.90	3	1239	2.80	3
A	23	1245	2.98	3	1240	2.88	3
A	24	1246	3.08	3	1241	2.99	3
A	25	1247	3.20	3	1242	3.11	3
A	26	1249	3.34	3	1244	3.26	3
A	27	1250	3.52	3	1245	3.45	3
A	28	1252	3.75	4	1247	3.67	3
A	29	1254	4.04	4	1249	3.95	3
A	30	1256	4.42	4	1251	4.31	4
A	31	1258	4.96	4	1253	4.82	4
A	32	1261	5.76	4	1256	5.58	4
A	33	1266	7.16	4	1261	6.90	4
A	34	1274	10.41	4	1268	9.95	4
A	35	1290	25.84	4	1290	33.37	4
B	0	1200	30.30	1	1200	37.18	1
B	1	1207	8.03	1	1211	8.70	1
B	2	1213	5.58	1	1217	6.10	1
B	3	1217	4.55	1	1221	5.01	1
B	4	1219	3.96	1	1224	4.38	1
B	5	1221	3.58	1	1226	3.97	1
B	6	1223	3.31	1	1228	3.67	1
B	7	1224	3.11	1	1230	3.45	1
B	8	1225	2.96	1	1231	3.29	1
B	9	1227	2.84	1	1233	3.15	2
B	10	1228	2.75	1	1234	3.05	2
B	11	1229	2.68	1	1235	2.96	2
B	12	1230	2.63	1	1236	2.89	2
B	13	1231	2.59	1	1237	2.84	2
B	14	1232	2.56	1	1239	2.79	3
B	15	1233	2.55	2	1240	2.76	3
B	16	1234	2.54	2	1241	2.74	3
B	17	1235	2.55	2	1242	2.73	3
B	18	1236	2.57	2	1243	2.73	3
B	19	1237	2.60	2	1244	2.73	3
B	20	1237	2.63	2	1245	2.75	3
B	21	1238	2.68	2	1246	2.78	3
B	22	1240	2.75	3	1247	2.82	3
B	23	1241	2.82	3	1248	2.88	3
B	24	1242	2.92	3	1249	2.95	3
B	25	1243	3.03	3	1250	3.05	3
B	26	1244	3.17	3	1252	3.17	4
B	27	1246	3.34	3	1253	3.32	4
B	28	1247	3.56	3	1255	3.52	4
B	29	1249	3.84	3	1256	3.77	4
B	30	1251	4.21	4	1258	4.12	4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	31	1253	4.72	4	1261	4.60	4
B	32	1257	5.50	4	1264	5.33	4
B	33	1261	6.86	4	1268	6.60	4
B	34	1268	10.05	4	1275	9.57	4
B	35	1290	32.95	4	1290	24.28	4
C	0	1200	38.69	1	1200	32.33	1
C	1	1211	9.01	1	1208	8.87	1
C	2	1218	6.28	1	1215	6.17	1
C	3	1222	5.13	1	1219	5.02	1
C	4	1225	4.48	1	1221	4.35	1
C	5	1227	4.05	1	1224	3.91	1
C	6	1229	3.75	1	1225	3.60	1
C	7	1231	3.52	1	1227	3.36	1
C	8	1232	3.35	1	1228	3.17	1
C	9	1234	3.22	2	1230	3.02	1
C	10	1235	3.11	2	1231	2.91	1
C	11	1236	3.03	2	1232	2.81	1
C	12	1237	2.96	2	1233	2.74	2
C	13	1238	2.91	2	1234	2.69	2
C	14	1239	2.88	3	1235	2.65	2
C	15	1240	2.85	3	1236	2.63	2
C	16	1242	2.84	3	1237	2.62	2
C	17	1243	2.84	3	1238	2.62	2
C	18	1244	2.85	3	1239	2.64	3
C	19	1245	2.87	3	1240	2.67	3
C	20	1246	2.90	3	1241	2.71	3
C	21	1247	2.94	3	1242	2.76	3
C	22	1248	3.00	3	1243	2.83	3
C	23	1249	3.07	3	1244	2.91	3
C	24	1250	3.16	3	1245	3.00	3
C	25	1252	3.27	4	1247	3.12	3
C	26	1253	3.41	4	1248	3.25	3
C	27	1255	3.58	4	1249	3.42	3
C	28	1256	3.80	4	1251	3.63	4
C	29	1258	4.09	4	1253	3.90	4
C	30	1260	4.47	4	1255	4.25	4
C	31	1263	5.00	4	1257	4.75	4
C	32	1266	5.79	4	1260	5.50	4
C	33	1271	7.17	4	1264	6.79	4
C	34	1278	10.35	4	1271	9.82	4
C	35	1290	20.72	4	1290	28.60	4

Table K-12. Raw to Scaled Score Look-up Table—Mathematics Grade 7

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	19.84	1	1200	19.97	1
A	1	1200	10.98	1	1200	11.04	1
A	2	1206	7.52	1	1206	7.49	1
A	3	1211	6.04	1	1211	6.01	1
A	4	1214	5.19	1	1214	5.16	1
A	5	1217	4.64	1	1217	4.62	1
A	6	1219	4.25	1	1219	4.23	1
A	7	1221	3.96	1	1221	3.95	1
A	8	1223	3.74	1	1222	3.74	1
A	9	1224	3.58	1	1224	3.57	1
A	10	1226	3.45	1	1225	3.43	1
A	11	1227	3.35	1	1227	3.32	1
A	12	1228	3.27	1	1228	3.23	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	13	1230	3.21	1	1229	3.15	1
A	14	1231	3.18	1	1231	3.09	1
A	15	1232	3.16	1	1232	3.03	1
A	16	1233	3.15	1	1233	2.99	1
A	17	1234	3.16	2	1234	2.96	2
A	18	1236	3.19	2	1235	2.95	2
A	19	1237	3.24	2	1236	2.96	2
A	20	1238	3.30	2	1238	2.99	2
A	21	1239	3.38	2	1239	3.06	2
A	22	1241	3.48	3	1240	3.15	3
A	23	1242	3.59	3	1241	3.27	3
A	24	1244	3.74	3	1243	3.43	3
A	25	1245	3.90	3	1244	3.63	3
A	26	1247	4.10	3	1246	3.86	3
A	27	1249	4.34	3	1248	4.14	3
A	28	1251	4.62	3	1250	4.47	3
A	29	1254	4.98	4	1252	4.87	3
A	30	1256	5.44	4	1255	5.37	4
A	31	1259	6.07	4	1258	6.06	4
A	32	1263	7.01	4	1262	7.06	4
A	33	1268	8.59	4	1267	8.76	4
A	34	1277	12.23	4	1277	12.67	4
A	35	1290	25.06	4	1290	25.82	4
B	0	1200	21.08	1	1200	24.89	1
B	1	1200	11.14	1	1202	10.80	1
B	2	1208	7.85	1	1210	7.23	1
B	3	1213	6.45	1	1215	5.80	1
B	4	1216	5.65	1	1218	5.01	1
B	5	1219	5.13	1	1221	4.50	1
B	6	1221	4.76	1	1223	4.15	1
B	7	1224	4.49	1	1225	3.90	1
B	8	1226	4.28	1	1226	3.70	1
B	9	1227	4.12	1	1228	3.54	1
B	10	1229	3.99	1	1230	3.41	1
B	11	1231	3.88	1	1231	3.30	1
B	12	1232	3.80	1	1232	3.21	1
B	13	1234	3.73	2	1234	3.13	2
B	14	1235	3.68	2	1235	3.07	2
B	15	1236	3.64	2	1236	3.02	2
B	16	1238	3.62	2	1237	2.99	2
B	17	1239	3.60	2	1239	2.97	2
B	18	1240	3.61	3	1239	2.96	2
B	19	1242	3.62	3	1241	2.97	3
B	20	1243	3.65	3	1242	3.00	3
B	21	1244	3.69	3	1243	3.05	3
B	22	1246	3.75	3	1245	3.11	3
B	23	1247	3.83	3	1246	3.19	3
B	24	1249	3.92	3	1247	3.30	3
B	25	1250	4.04	3	1249	3.44	3
B	26	1252	4.19	3	1250	3.63	3
B	27	1254	4.38	4	1252	3.87	3
B	28	1256	4.62	4	1254	4.18	4
B	29	1258	4.92	4	1256	4.60	4
B	30	1261	5.33	4	1259	5.16	4
B	31	1264	5.89	4	1262	5.94	4
B	32	1267	6.74	4	1266	7.11	4
B	33	1272	8.20	4	1272	9.06	4
B	34	1281	11.58	4	1282	13.47	4

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
B	35	1290	21.33	4	1290	22.14	4
C	0	1200	25.93	1	1200	28.96	1
C	1	1204	11.20	1	1206	10.50	1
C	2	1212	7.87	1	1214	7.23	1
C	3	1217	6.46	1	1219	5.92	1
C	4	1221	5.65	1	1222	5.19	1
C	5	1224	5.12	1	1225	4.72	1
C	6	1227	4.72	1	1227	4.37	1
C	7	1229	4.42	1	1229	4.10	1
C	8	1231	4.16	1	1231	3.89	1
C	9	1233	3.95	1	1233	3.70	1
C	10	1235	3.77	2	1234	3.55	2
C	11	1236	3.61	2	1236	3.41	2
C	12	1238	3.48	2	1237	3.30	2
C	13	1239	3.38	2	1239	3.21	2
C	14	1240	3.29	3	1239	3.13	2
C	15	1242	3.22	3	1241	3.07	3
C	16	1243	3.17	3	1242	3.03	3
C	17	1244	3.14	3	1244	3.01	3
C	18	1245	3.12	3	1245	3.00	3
C	19	1246	3.12	3	1246	3.01	3
C	20	1248	3.14	3	1247	3.03	3
C	21	1249	3.18	3	1248	3.08	3
C	22	1250	3.23	3	1250	3.14	3
C	23	1251	3.30	3	1251	3.23	3
C	24	1253	3.40	3	1252	3.34	3
C	25	1254	3.52	4	1254	3.49	4
C	26	1256	3.68	4	1255	3.67	4
C	27	1257	3.87	4	1257	3.90	4
C	28	1259	4.11	4	1259	4.18	4
C	29	1261	4.42	4	1261	4.55	4
C	30	1264	4.84	4	1263	5.03	4
C	31	1266	5.42	4	1266	5.70	4
C	32	1270	6.31	4	1270	6.69	4
C	33	1275	7.88	4	1276	8.39	4
C	34	1284	11.58	4	1285	12.35	4
C	35	1290	18.94	4	1290	18.51	4

Table K-13. Raw to Scaled Score Look-up Table—Mathematics Grade 8

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	25.07	1	1200	21.96	1
A	1	1201	11.18	1	1200	12.14	1
A	2	1209	7.45	1	1206	8.12	1
A	3	1214	5.85	1	1211	6.38	1
A	4	1217	4.94	1	1215	5.39	1
A	5	1220	4.36	1	1218	4.74	1
A	6	1222	3.97	1	1220	4.29	1
A	7	1224	3.69	1	1222	3.97	1
A	8	1226	3.50	1	1224	3.74	1
A	9	1227	3.36	1	1225	3.57	1
A	10	1228	3.26	1	1227	3.44	1
A	11	1230	3.19	1	1228	3.36	1
A	12	1231	3.14	1	1230	3.31	1
A	13	1232	3.11	1	1231	3.28	1
A	14	1233	3.10	1	1232	3.26	1
A	15	1235	3.10	2	1233	3.26	1
A	16	1236	3.12	2	1235	3.27	2

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	17	1237	3.14	2	1236	3.29	2
A	18	1238	3.18	2	1237	3.32	2
A	19	1239	3.23	2	1239	3.35	2
A	20	1241	3.29	3	1240	3.38	3
A	21	1242	3.36	3	1241	3.43	3
A	22	1243	3.44	3	1243	3.48	3
A	23	1245	3.54	3	1244	3.54	3
A	24	1246	3.66	3	1245	3.62	3
A	25	1247	3.80	3	1247	3.72	3
A	26	1249	3.97	3	1249	3.85	3
A	27	1251	4.18	4	1250	4.02	3
A	28	1253	4.43	4	1252	4.25	4
A	29	1255	4.75	4	1254	4.55	4
A	30	1257	5.18	4	1257	4.97	4
A	31	1260	5.76	4	1260	5.56	4
A	32	1264	6.64	4	1263	6.45	4
A	33	1269	8.17	4	1268	7.99	4
A	34	1278	11.74	4	1277	11.57	4
A	35	1290	24.06	4	1290	24.83	4
B	0	1200	30.01	1	1200	27.00	1
B	1	1206	10.83	1	1202	11.70	1
B	2	1214	7.62	1	1211	7.97	1
B	3	1219	6.26	1	1216	6.40	1
B	4	1222	5.47	1	1220	5.52	1
B	5	1225	4.95	1	1222	4.94	1
B	6	1227	4.57	1	1225	4.53	1
B	7	1229	4.29	1	1227	4.23	1
B	8	1231	4.07	1	1229	4.00	1
B	9	1233	3.90	1	1230	3.81	1
B	10	1235	3.76	2	1232	3.67	1
B	11	1236	3.65	2	1233	3.55	1
B	12	1237	3.56	2	1235	3.45	2
B	13	1239	3.49	2	1236	3.37	2
B	14	1240	3.44	3	1237	3.30	2
B	15	1241	3.40	3	1238	3.25	2
B	16	1242	3.37	3	1240	3.22	3
B	17	1244	3.35	3	1241	3.19	3
B	18	1245	3.35	3	1242	3.18	3
B	19	1246	3.36	3	1243	3.18	3
B	20	1247	3.39	3	1244	3.20	3
B	21	1248	3.42	3	1246	3.23	3
B	22	1250	3.48	3	1247	3.28	3
B	23	1251	3.54	4	1248	3.34	3
B	24	1252	3.63	4	1249	3.44	3
B	25	1254	3.75	4	1251	3.56	4
B	26	1255	3.89	4	1252	3.71	4
B	27	1257	4.06	4	1254	3.91	4
B	28	1259	4.29	4	1256	4.16	4
B	29	1261	4.58	4	1258	4.49	4
B	30	1263	4.97	4	1260	4.92	4
B	31	1266	5.51	4	1263	5.51	4
B	32	1269	6.34	4	1266	6.40	4
B	33	1274	7.75	4	1271	7.92	4
B	34	1282	11.05	4	1280	11.43	4
B	35	1290	19.39	4	1290	21.88	4
C	0	1200	27.37	1	1200	29.34	1
C	1	1204	10.95	1	1204	11.85	1
C	2	1212	7.67	1	1213	8.09	1
C	3	1217	6.24	1	1218	6.51	1
C	4	1220	5.41	1	1222	5.62	1

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	5	1223	4.85	1	1225	5.04	1
C	6	1226	4.44	1	1227	4.63	1
C	7	1228	4.13	1	1229	4.32	1
C	8	1229	3.88	1	1231	4.09	1
C	9	1231	3.68	1	1233	3.91	1
C	10	1232	3.52	1	1235	3.76	2
C	11	1233	3.39	1	1236	3.64	2
C	12	1235	3.28	2	1237	3.55	2
C	13	1236	3.19	2	1239	3.47	2
C	14	1237	3.12	2	1240	3.41	3
C	15	1239	3.07	2	1241	3.36	3
C	16	1240	3.03	3	1243	3.32	3
C	17	1241	3.01	3	1244	3.30	3
C	18	1242	3.00	3	1245	3.29	3
C	19	1243	3.01	3	1246	3.30	3
C	20	1244	3.04	3	1248	3.32	3
C	21	1245	3.08	3	1249	3.36	3
C	22	1246	3.14	3	1250	3.42	3
C	23	1248	3.22	3	1251	3.49	4
C	24	1249	3.32	3	1253	3.59	4
C	25	1250	3.45	3	1254	3.71	4
C	26	1252	3.62	4	1256	3.86	4
C	27	1253	3.82	4	1257	4.04	4
C	28	1255	4.08	4	1259	4.28	4
C	29	1257	4.41	4	1261	4.58	4
C	30	1259	4.85	4	1264	4.97	4
C	31	1262	5.45	4	1267	5.52	4
C	32	1266	6.36	4	1270	6.35	4
C	33	1271	7.91	4	1275	7.77	4
C	34	1279	11.56	4	1283	11.10	4
C	35	1290	22.74	4	1290	18.65	4

Table K-14. Raw to Scaled Score Look-up Table—Mathematics Grade HS

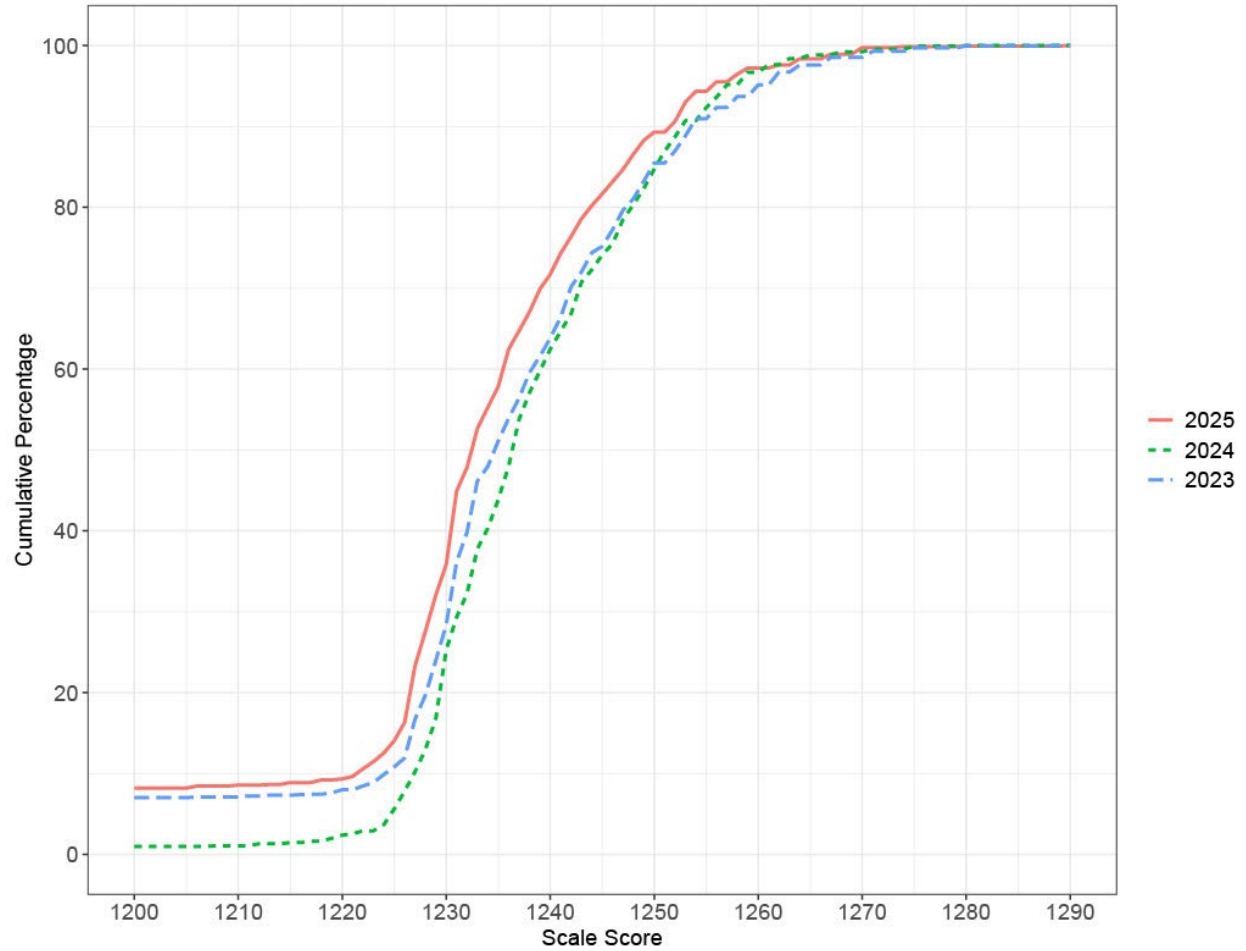
Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	0	1200	33.41	1	1200	30.75	1
A	1	1210	8.18	1	1208	8.44	1
A	2	1216	5.68	1	1214	5.87	1
A	3	1219	4.60	1	1218	4.79	1
A	4	1222	3.99	1	1220	4.16	1
A	5	1224	3.58	1	1222	3.75	1
A	6	1226	3.29	1	1224	3.46	1
A	7	1227	3.08	1	1226	3.25	1
A	8	1228	2.92	1	1227	3.08	1
A	9	1230	2.80	1	1228	2.96	1
A	10	1231	2.71	1	1230	2.86	1
A	11	1232	2.64	1	1231	2.78	1
A	12	1233	2.59	1	1232	2.72	1
A	13	1234	2.55	1	1233	2.68	1
A	14	1235	2.53	2	1234	2.65	1
A	15	1236	2.52	2	1235	2.63	2
A	16	1236	2.52	2	1236	2.62	2
A	17	1237	2.53	2	1237	2.63	2
A	18	1238	2.55	2	1237	2.64	2
A	19	1239	2.58	2	1238	2.66	2
A	20	1240	2.62	3	1239	2.69	2
A	21	1241	2.67	3	1240	2.74	3
A	22	1242	2.73	3	1241	2.79	3
A	23	1243	2.80	3	1243	2.86	3
A	24	1244	2.89	3	1244	2.94	3
A	25	1246	3.00	3	1245	3.04	3
A	26	1247	3.13	3	1246	3.17	3

Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
A	27	1248	3.29	3	1247	3.32	3
A	28	1250	3.49	4	1249	3.52	3
A	29	1251	3.74	4	1251	3.77	4
A	30	1253	4.07	4	1253	4.11	4
A	31	1256	4.53	4	1255	4.59	4
A	32	1259	5.22	4	1258	5.31	4
A	33	1263	6.39	4	1262	6.57	4
A	34	1269	9.11	4	1269	9.59	4
A	35	1290	35.79	4	1290	35.71	4
B	0	1200	39.36	1	1200	39.12	1
B	1	1212	9.55	1	1212	9.77	1
B	2	1219	6.60	1	1219	6.66	1
B	3	1223	5.37	1	1223	5.37	1
B	4	1226	4.67	1	1226	4.64	1
B	5	1228	4.21	1	1228	4.16	1
B	6	1231	3.89	1	1230	3.81	1
B	7	1232	3.64	1	1232	3.54	1
B	8	1234	3.45	1	1234	3.33	1
B	9	1235	3.28	2	1235	3.15	2
B	10	1237	3.14	2	1236	2.99	2
B	11	1238	3.02	2	1238	2.85	2
B	12	1239	2.91	2	1239	2.72	2
B	13	1240	2.82	3	1240	2.61	3
B	14	1241	2.73	3	1241	2.51	3
B	15	1242	2.65	3	1242	2.42	3
B	16	1243	2.58	3	1243	2.35	3
B	17	1244	2.53	3	1244	2.30	3
B	18	1245	2.49	3	1244	2.26	3
B	19	1246	2.47	3	1245	2.25	3
B	20	1247	2.46	3	1246	2.26	3
B	21	1248	2.48	3	1247	2.30	3
B	22	1249	2.52	3	1248	2.36	3
B	23	1250	2.59	4	1249	2.45	3
B	24	1251	2.68	4	1250	2.57	4
B	25	1252	2.81	4	1251	2.72	4
B	26	1254	2.98	4	1252	2.90	4
B	27	1255	3.20	4	1254	3.12	4
B	28	1257	3.47	4	1255	3.39	4
B	29	1258	3.82	4	1257	3.73	4
B	30	1260	4.28	4	1259	4.16	4
B	31	1263	4.90	4	1262	4.76	4
B	32	1266	5.82	4	1265	5.66	4
B	33	1271	7.38	4	1269	7.23	4
B	34	1279	10.96	4	1278	10.95	4
B	35	1290	22.49	4	1290	24.02	4
C	0	1200	33.66	1	1200	35.65	1
C	1	1209	9.27	1	1210	9.00	1
C	2	1216	6.39	1	1217	6.25	1
C	3	1220	5.17	1	1221	5.09	1
C	4	1223	4.47	1	1224	4.42	1
C	5	1225	4.02	1	1226	3.97	1
C	6	1227	3.69	1	1228	3.65	1
C	7	1229	3.45	1	1229	3.41	1
C	8	1230	3.26	1	1231	3.22	1
C	9	1232	3.11	1	1232	3.07	1
C	10	1233	2.98	1	1233	2.95	1
C	11	1234	2.88	1	1235	2.85	2
C	12	1235	2.80	2	1236	2.77	2
C	13	1236	2.74	2	1237	2.70	2
C	14	1237	2.69	2	1238	2.65	2
C	15	1238	2.65	2	1239	2.61	2
C	16	1239	2.63	2	1240	2.59	3
C	17	1240	2.62	3	1241	2.57	3
C	18	1241	2.62	3	1242	2.57	3
C	19	1242	2.64	3	1242	2.58	3
C	20	1243	2.67	3	1243	2.60	3
C	21	1244	2.72	3	1244	2.64	3

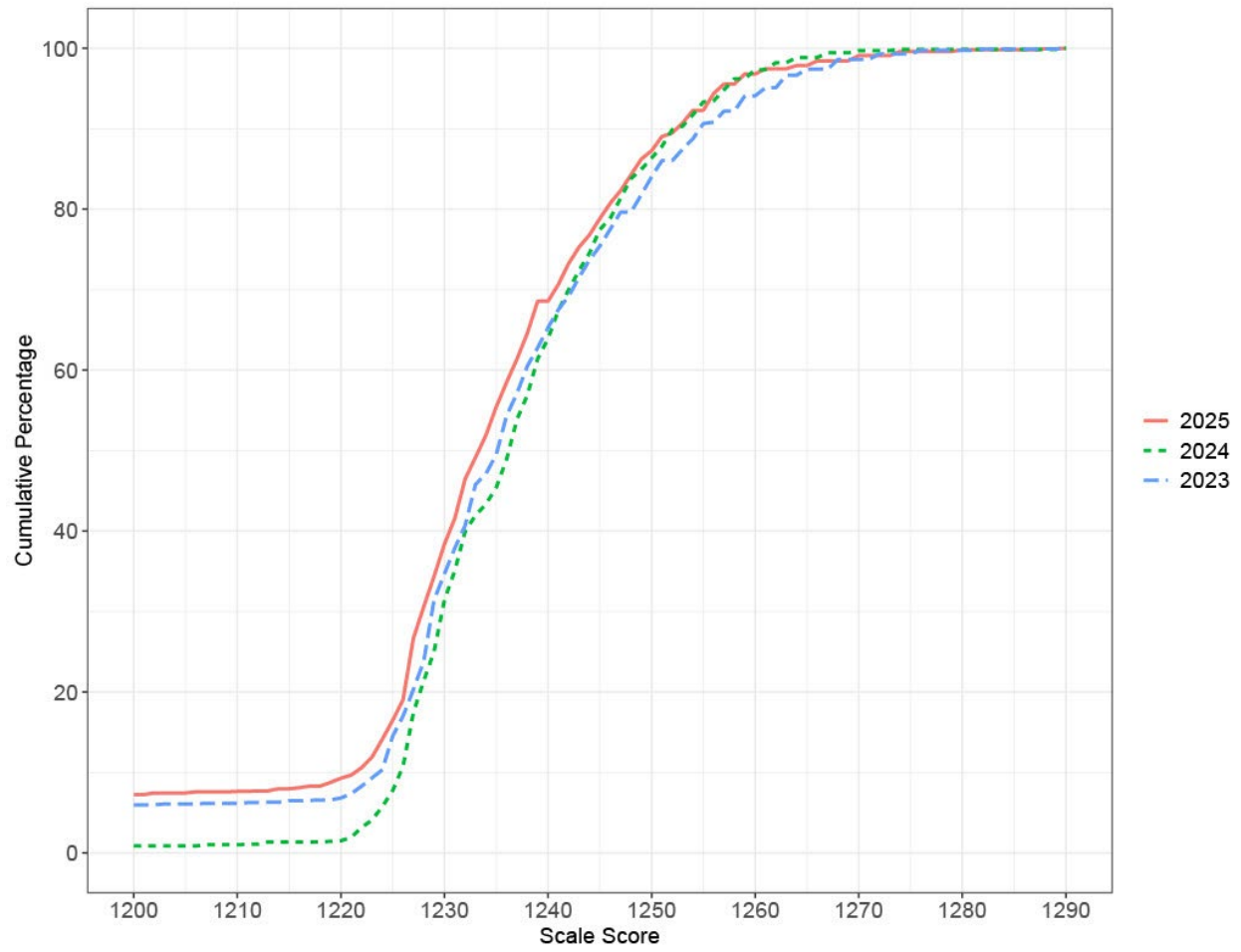
Path	Raw Score	2023 Scaled Score	2023 Standard Error	2023 Performance Level	2024 / 2025 Scaled Score	2024 / 2025 Standard Error	2024 / 2025 Performance Level
C	22	1245	2.79	3	1245	2.69	3
C	23	1247	2.88	3	1246	2.76	3
C	24	1248	2.99	3	1247	2.84	3
C	25	1249	3.13	3	1249	2.95	3
C	26	1250	3.30	4	1250	3.09	4
C	27	1252	3.51	4	1251	3.26	4
C	28	1254	3.78	4	1253	3.48	4
C	29	1255	4.11	4	1254	3.76	4
C	30	1258	4.55	4	1256	4.14	4
C	31	1260	5.16	4	1259	4.65	4
C	32	1264	6.06	4	1262	5.43	4
C	33	1269	7.60	4	1266	6.76	4
C	34	1277	11.20	4	1273	9.88	4
C	35	1290	24.64	4	1290	28.95	4

APPENDIX L
SCORE DISTRIBUTIONS

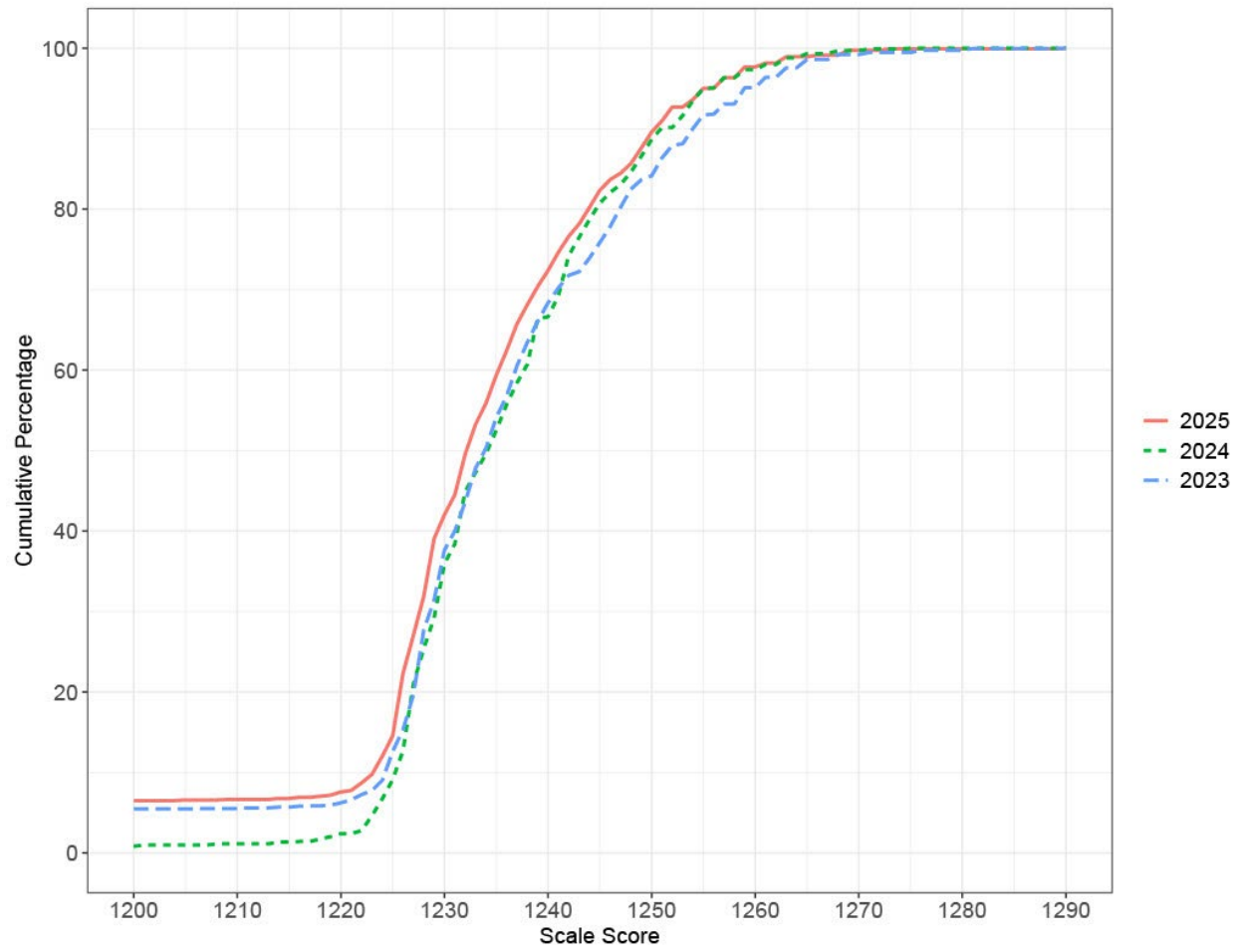
Cumulative Scale Score Distributions: ELA03



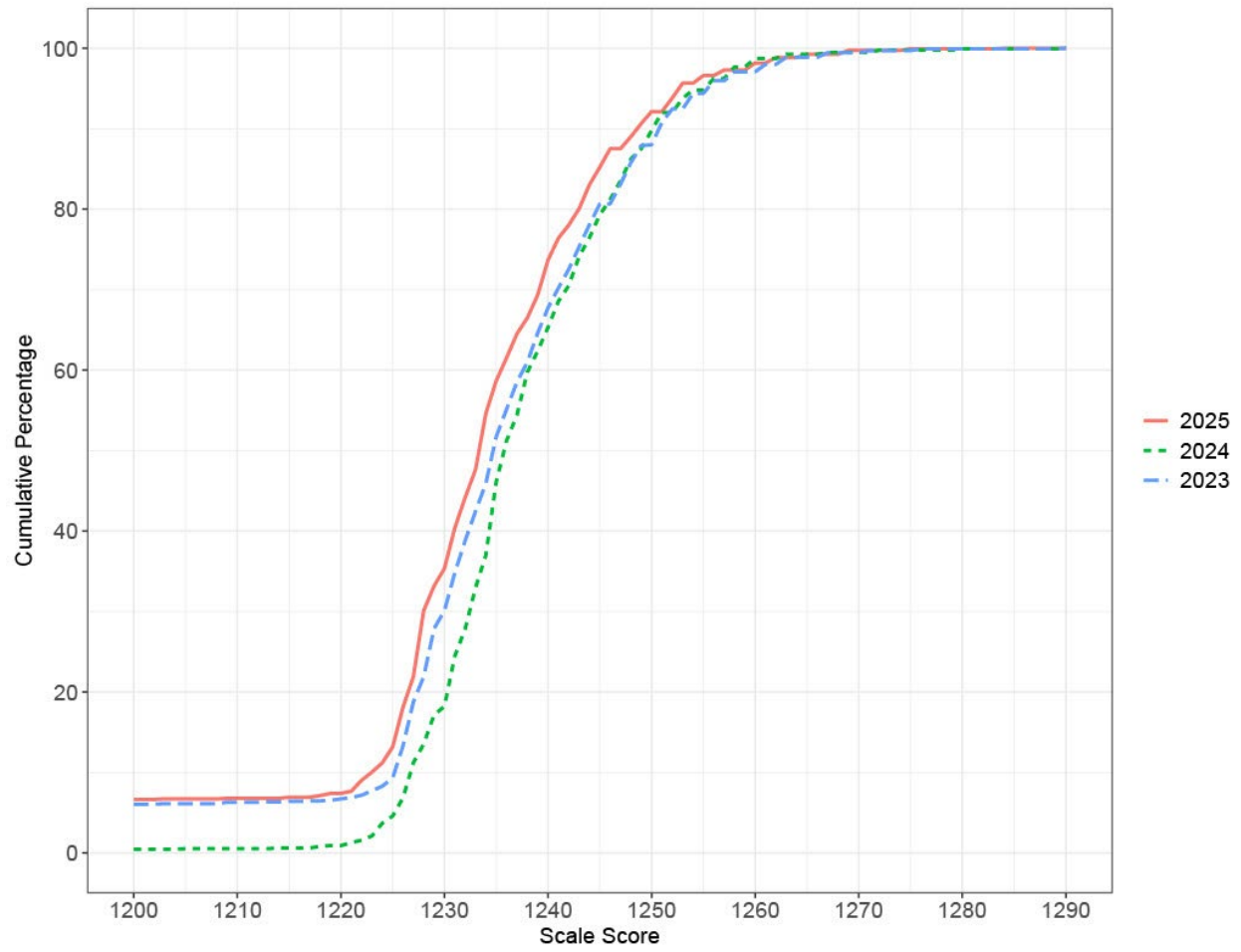
Cumulative Scale Score Distributions: ELA04



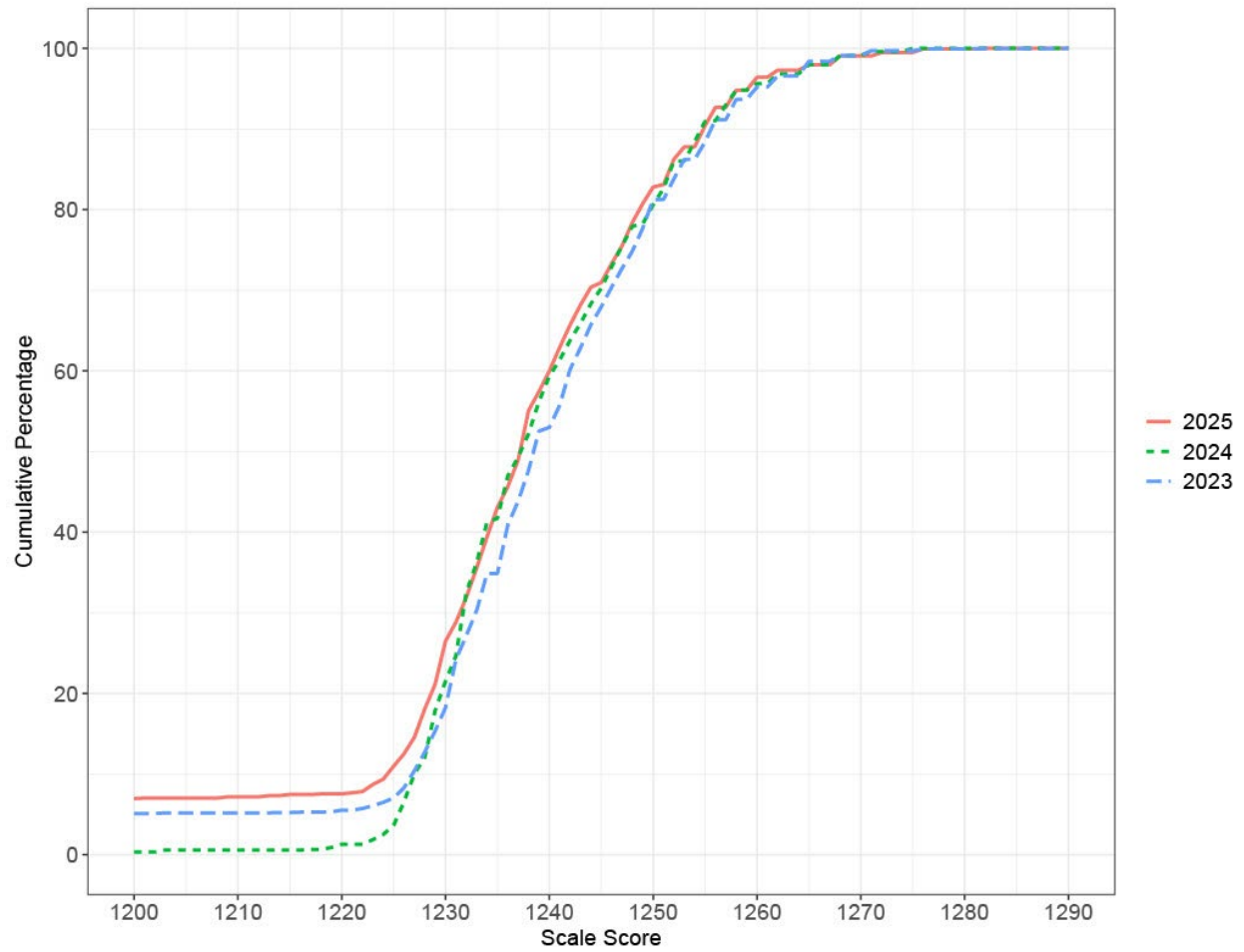
Cumulative Scale Score Distributions: ELA05



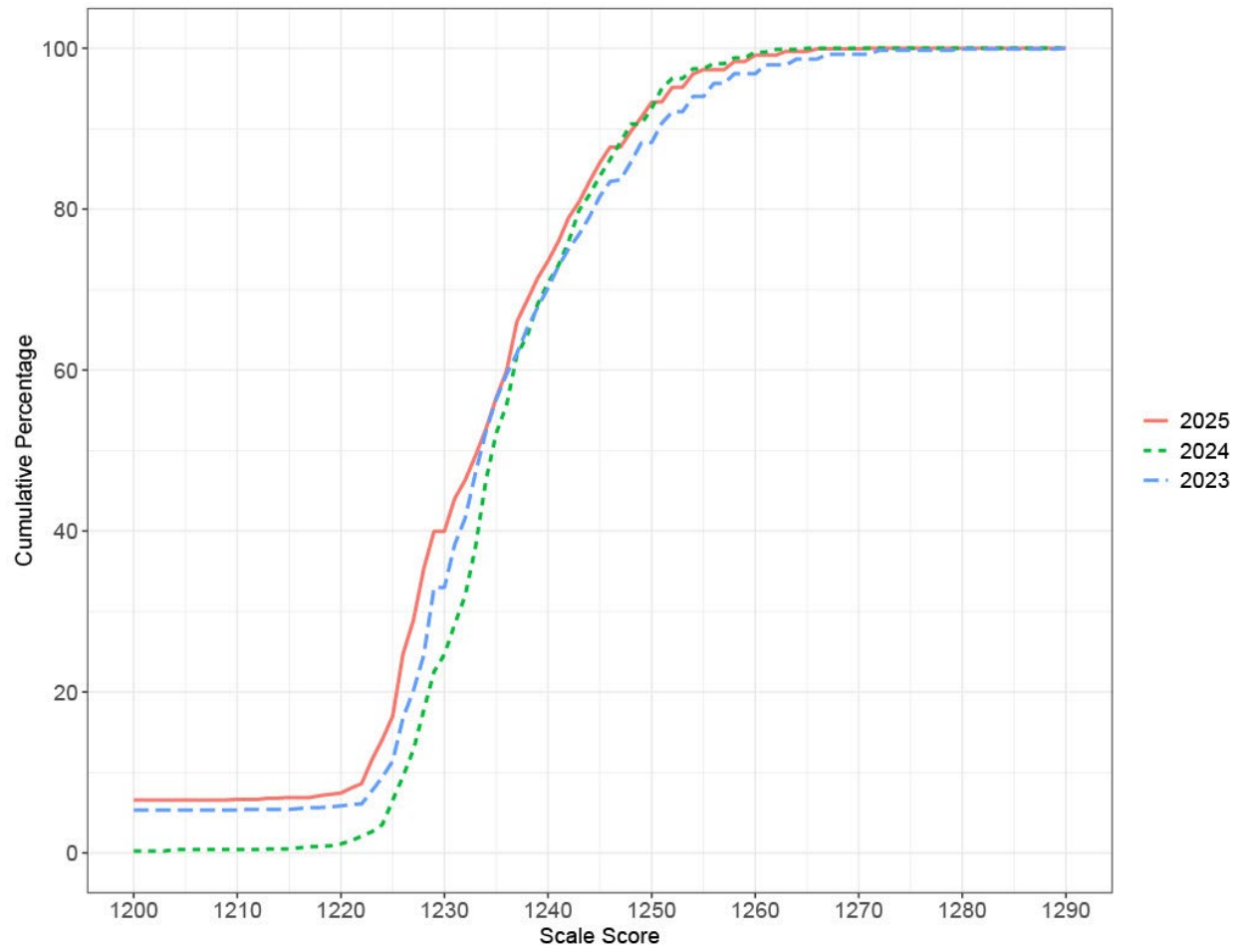
Cumulative Scale Score Distributions: ELA06



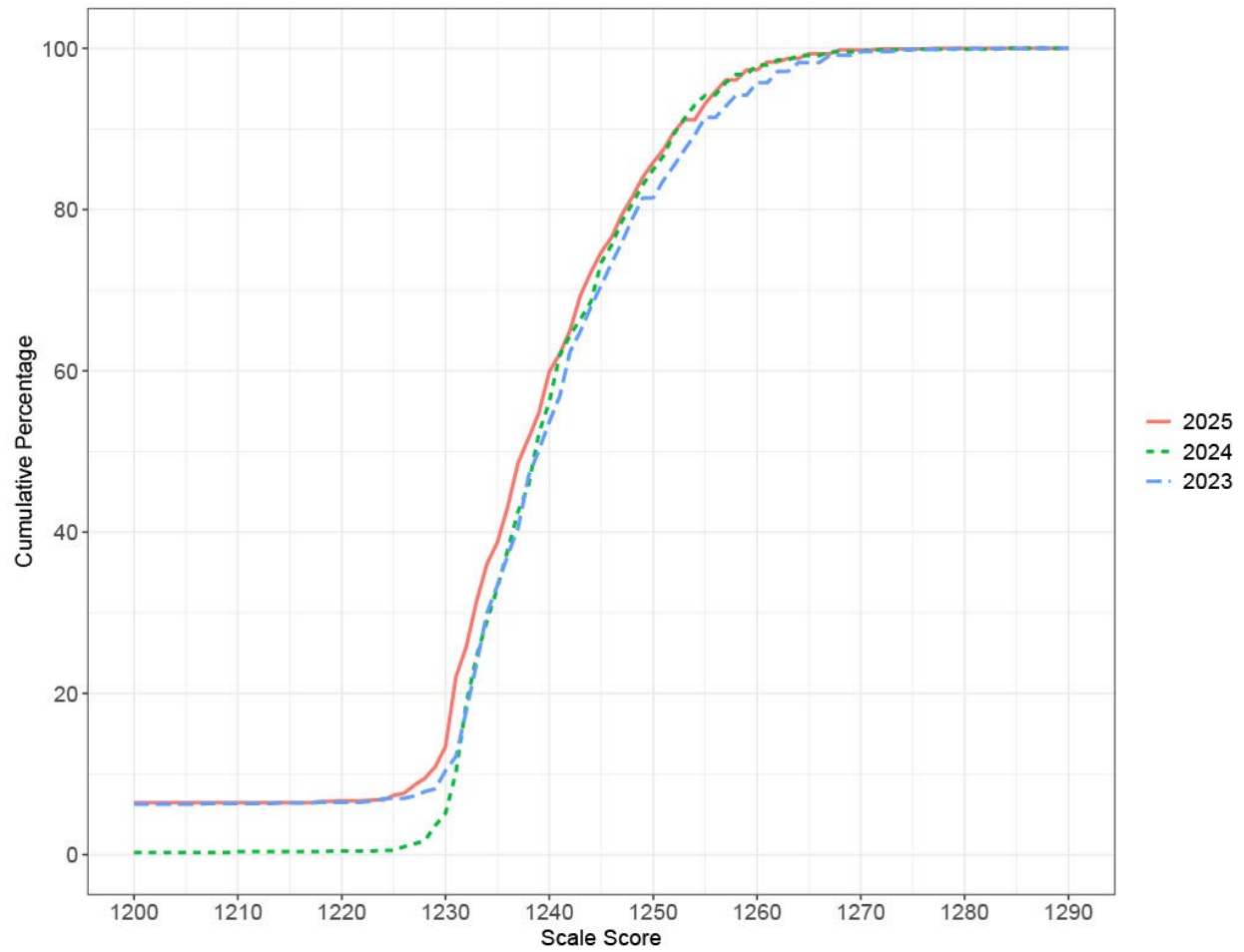
Cumulative Scale Score Distributions: ELA07



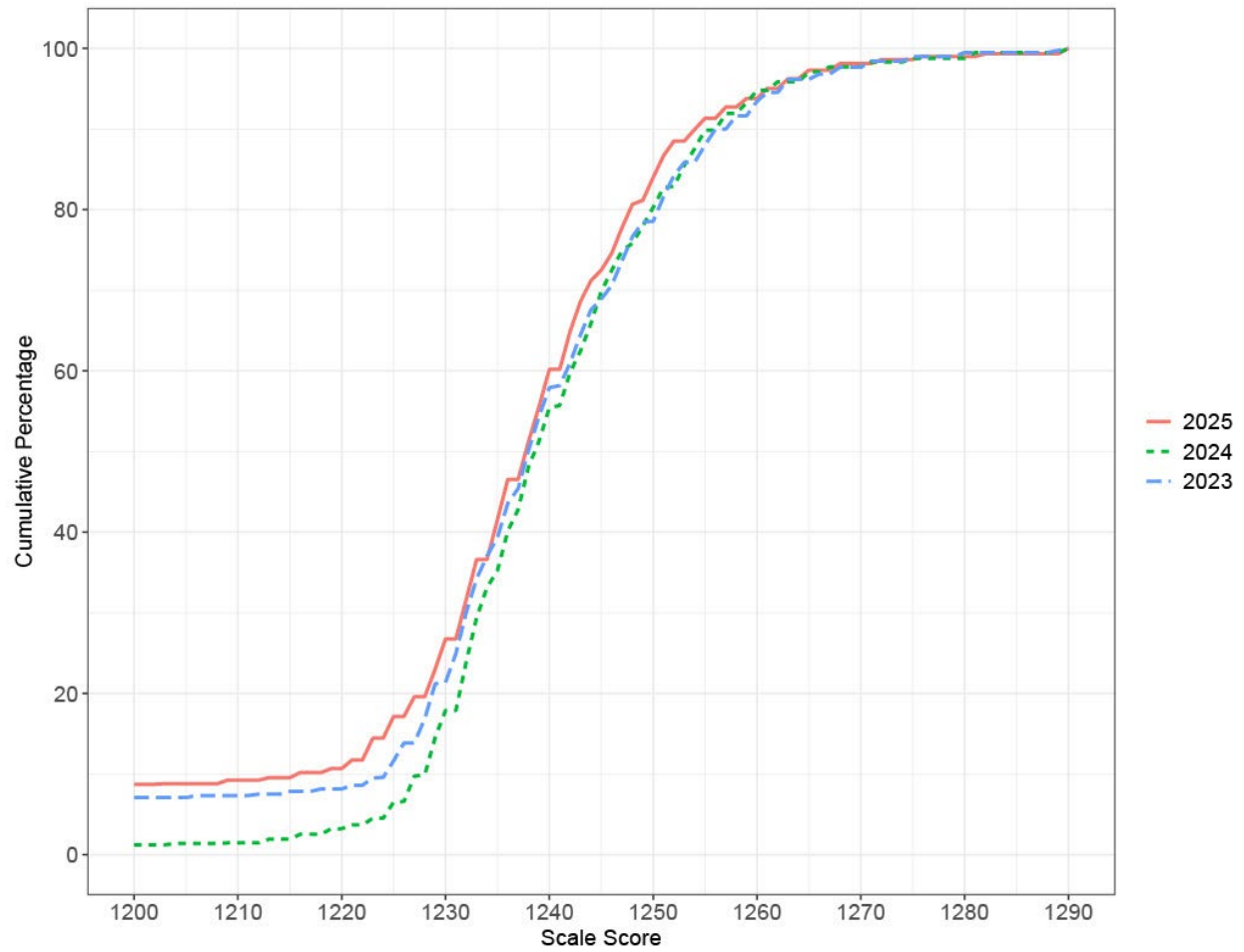
Cumulative Scale Score Distributions: ELA08



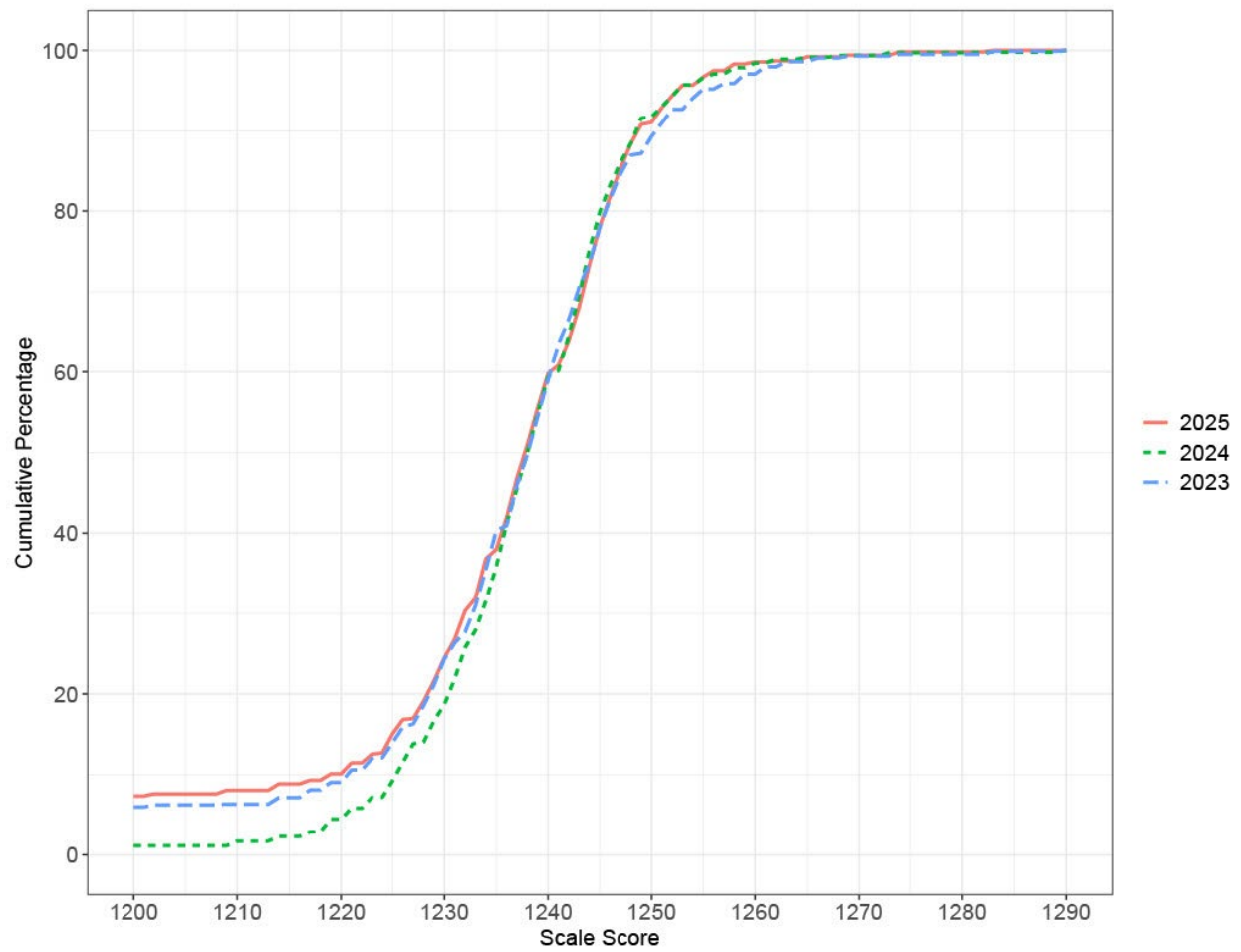
Cumulative Scale Score Distributions: ELA11



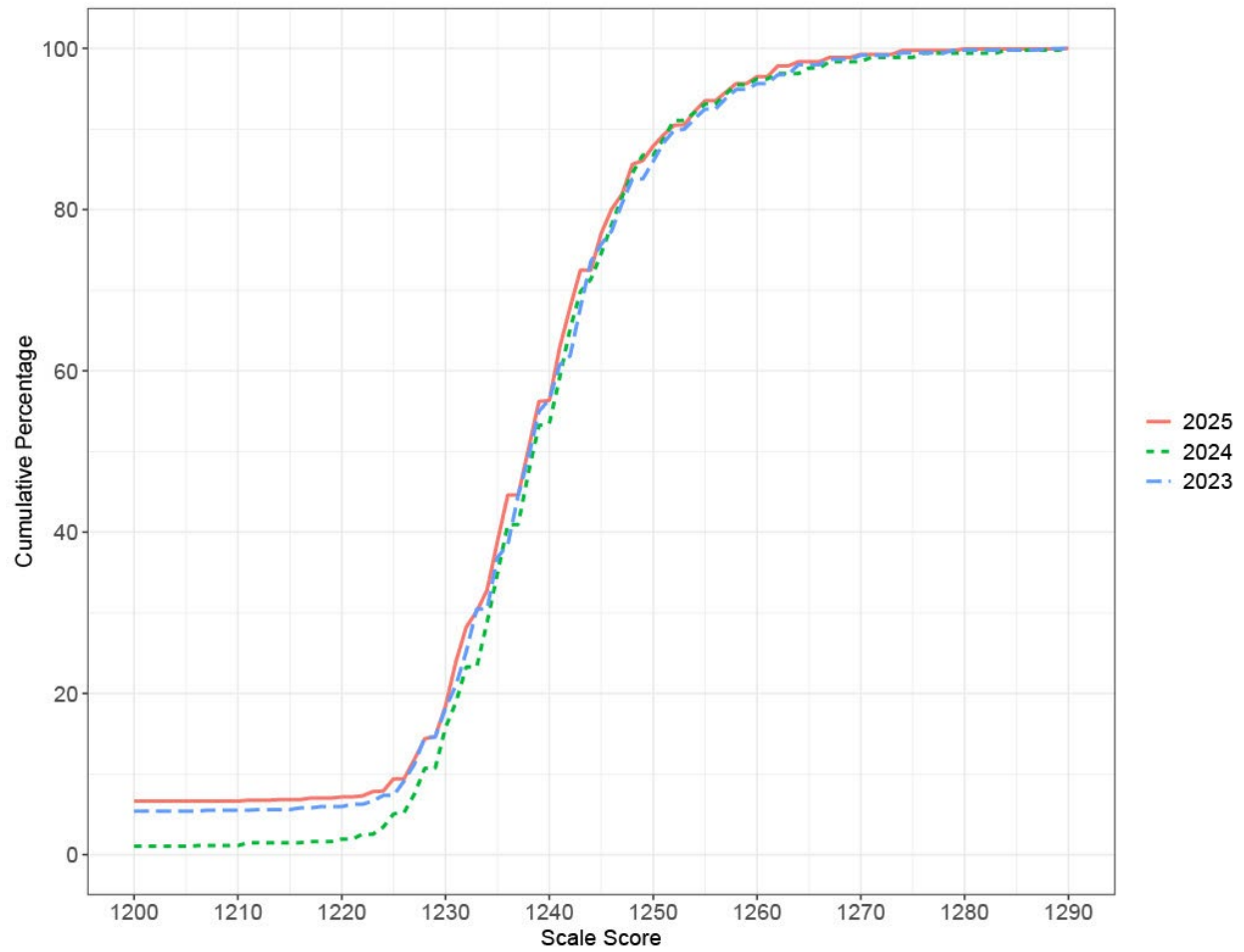
Cumulative Scale Score Distributions: MAT03



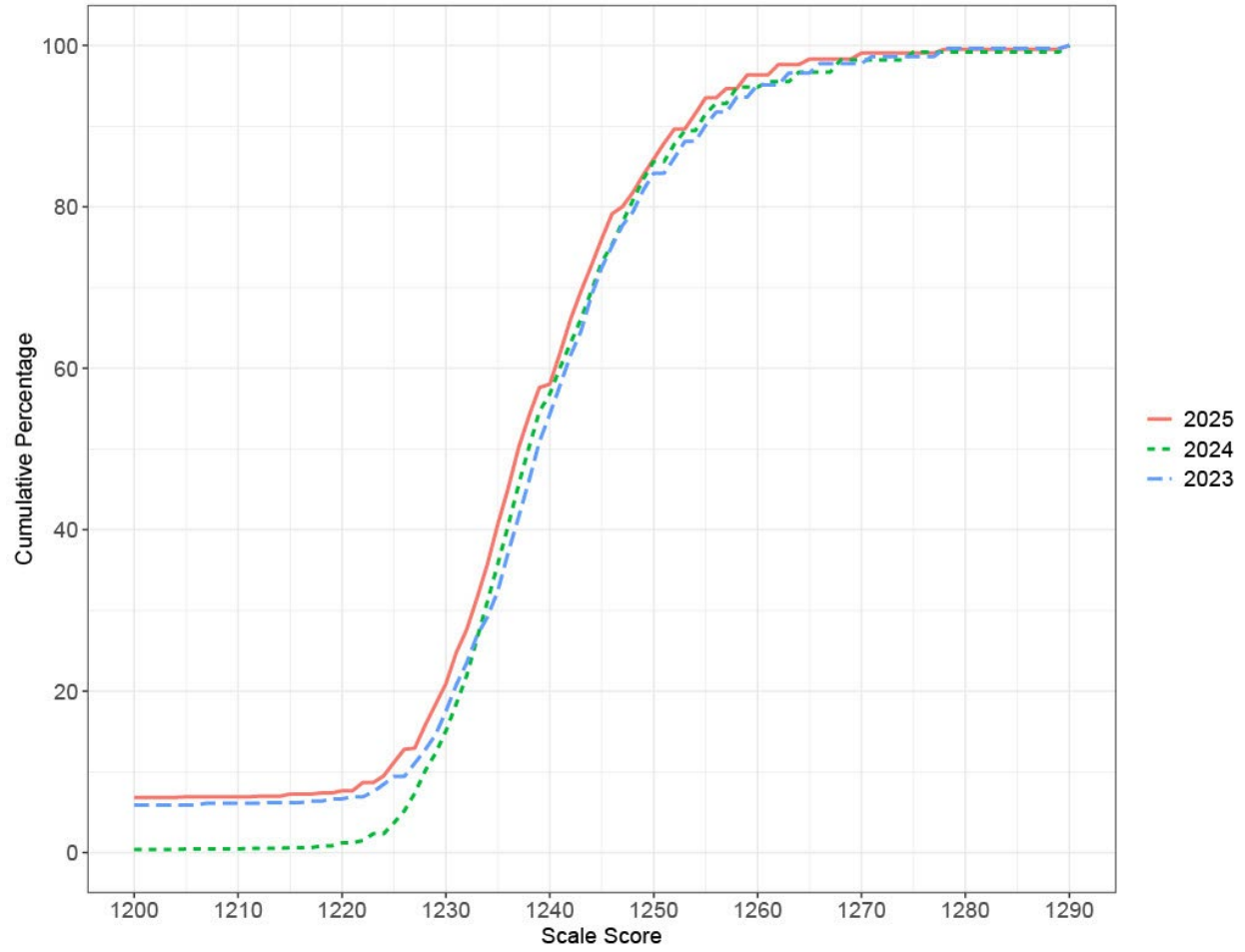
Cumulative Scale Score Distributions: MAT04



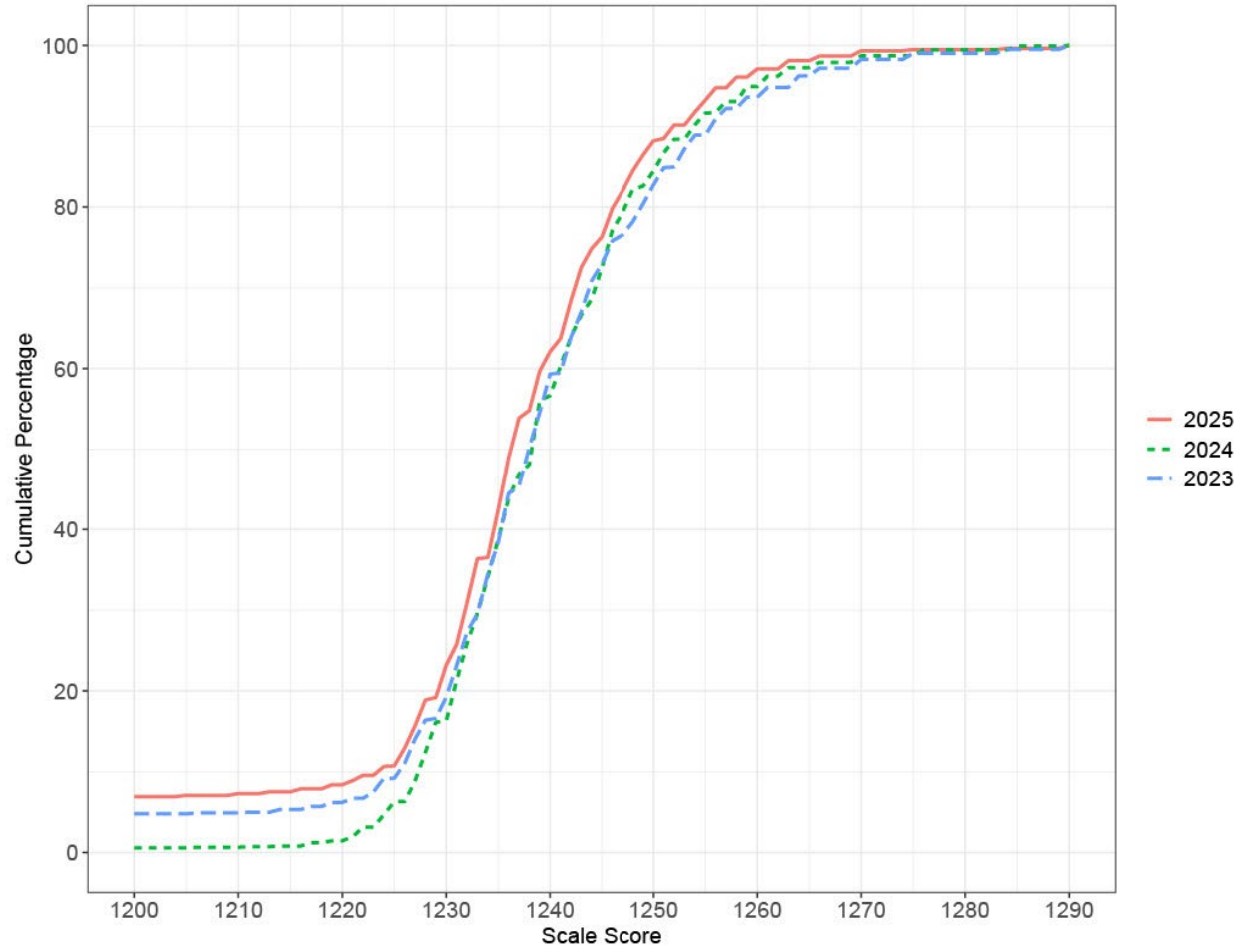
Cumulative Scale Score Distributions: MAT05



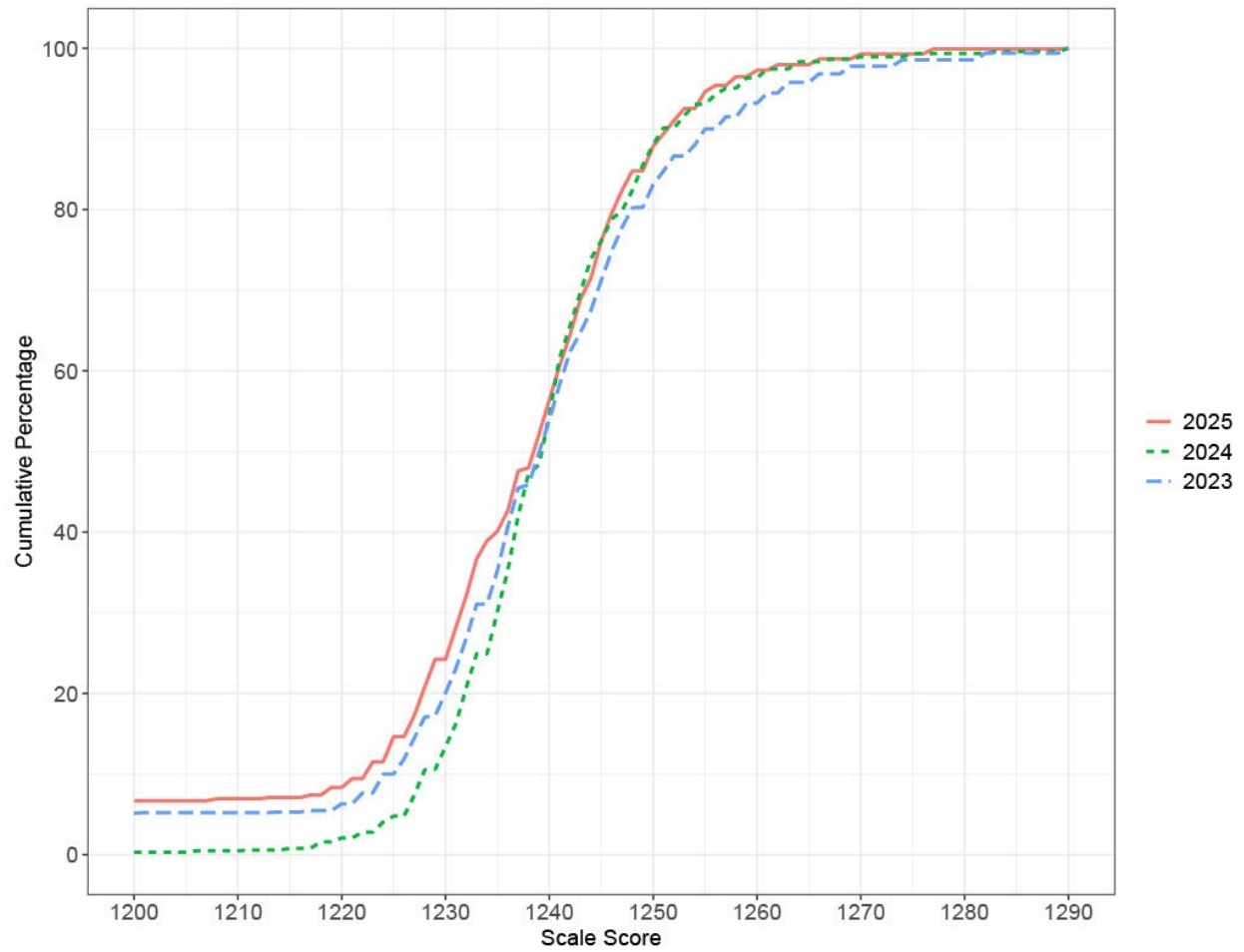
Cumulative Scale Score Distributions: MAT06



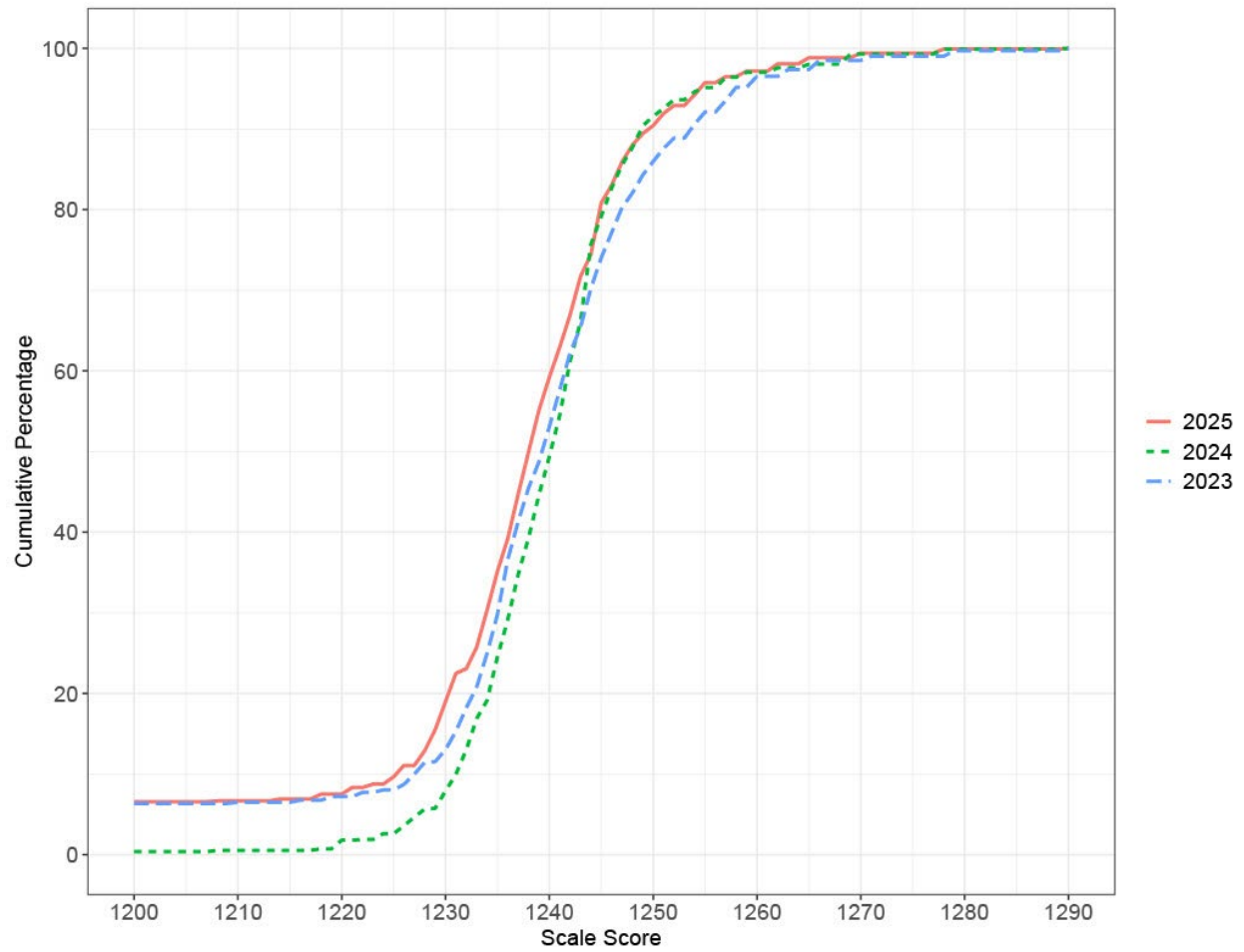
Cumulative Scale Score Distributions: MAT07



Cumulative Scale Score Distributions: MAT08



Cumulative Scale Score Distributions: MAT11



APPENDIX M

IRT SUBGROUP RELIABILITY

Note: Values are calculated only for subgroups with 100 or more students.

Table M-1. IRT Subgroup Reliability: ELA Grade 3

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,474	1200	1290	1236.64	10.83	0.93	2.84
Female	448	1200	1290	1236.73	11.02	0.93	2.85
Male	1,026	1200	1280	1236.61	10.75	0.93	2.84
Gender Undefined	474	1200	1270	1235.76	10.29	0.92	2.81
Hispanic or Latino	81	1220	1270	1236.14	10.27	NA	NA
American Indian or Alaska Native	45	1200	1258	1232.33	9.89	NA	NA
Asian	139	1210	1290	1236.99	11.66	0.93	2.89
Black or African American	17	1200	1253	1231.59	10.76	NA	NA
Native Hawaiian or Pacific Islander	533	1200	1274	1237.10	10.71	0.93	2.83
White (non-Hispanic)	78	1223	1270	1236.23	10.07	NA	NA
Two or More Races (non-Hispanic)	474	1200	1270	1235.76	10.29	0.92	2.81
No Primary race/Ethnicity Undefined	268	1200	1270	1234.65	10.07	0.91	2.80
Currently receiving LEP services	212	1200	1290	1236.27	11.46	0.93	2.88
Not receiving LEP services	282	1210	1290	1238.23	11.75	0.93	2.90
LEP: All Other Students	69	1200	1264	1233.19	10.84	NA	NA
Economically Disadvantaged Students	11	1224	1256	1237.00	9.91	NA	NA
Non-economically Disadvantaged Students	221	1200	1290	1236.40	11.55	0.93	2.88
SES: All Other Students	517	1200	1254	1231.73	6.92	0.85	2.62
Migrant	952	1200	1290	1239.26	11.58	0.93	2.95
Non-migrant	34	1210	1254	1231.21	8.34	NA	NA
Undefined Migrant Status	1,440	1200	1290	1236.77	10.85	0.93	2.84
Augmentative Communication	62	1200	1290	1233.40	11.63	NA	NA
No Augmentative Communication	1,410	1200	1280	1236.80	10.78	0.93	2.84
Undefined Augmentative Communications	172	1200	1253	1229.41	7.47	0.85	2.73
Hearing Loss	1,302	1200	1290	1237.60	10.85	0.93	2.85
Within Normal Limits	153	1206	1256	1230.37	7.11	0.86	2.62
Visual Impairment	881	1200	1270	1235.24	9.82	0.91	2.77
Within Normal Limits	233	1206	1280	1240.18	11.15	0.93	2.95
Undefined Visual Impairment	118	1223	1290	1244.29	12.67	0.93	3.23
Sensory Stimuli Response	89	1200	1270	1241.92	12.00	NA	NA
Follow Directions	153	1200	1253	1228.20	8.37	0.86	2.87
Special School	461	1200	1267	1233.17	8.22	0.89	2.66
Regular School Self-contained	860	1206	1290	1240.01	11.05	0.92	2.93
Regular School Primarily Self-contained	1,474	1200	1290	1236.64	10.83	0.93	2.84
Regular School Resource Room	448	1200	1290	1236.73	11.02	0.93	2.85
Regular School General Education	1,026	1200	1280	1236.61	10.75	0.93	2.84
Communicates Primarily Through Cries	474	1200	1270	1235.76	10.29	0.92	2.81
Uses Intentional Communication	81	1220	1270	1236.14	10.27	NA	NA
Uses Symbolic Language	45	1200	1258	1232.33	9.89	NA	NA

Table M-2. IRT Subgroup Reliability: ELA Grade 4

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,428	1200	1290	1237.21	11.60	0.93	3.00
Female	455	1200	1290	1237.70	12.30	0.93	3.08
Male	973	1202	1288	1236.98	11.26	0.93	2.96
Gender Undefined	476	1210	1274	1237.58	10.74	0.92	2.95
Hispanic or Latino	93	1221	1266	1237.43	9.83	NA	NA
American Indian or Alaska Native	41	1206	1259	1232.85	9.20	NA	NA
Asian	147	1219	1288	1237.18	11.51	0.93	2.98
Black or African American	18	1221	1250	1230.28	7.54	NA	NA
Native Hawaiian or Pacific Islander	494	1200	1288	1237.10	12.26	0.93	3.05
White (non-Hispanic)	77	1212	1270	1234.42	10.31	NA	NA
Two or More Races (non-Hispanic)	476	1210	1274	1237.58	10.74	0.92	2.95
No Primary race/Ethnicity Undefined	246	1202	1274	1236.17	10.62	0.92	2.91
Currently receiving LEP services	240	1202	1288	1236.53	12.35	0.93	3.04
Not receiving LEP services	301	1212	1288	1238.11	11.76	0.93	3.03
LEP: All Other Students	86	1202	1274	1234.83	13.16	NA	NA
Economically Disadvantaged Students	243	1202	1288	1237.45	12.52	0.93	3.07
Non-economically Disadvantaged Students	423	1202	1266	1230.73	7.20	0.85	2.70
SES: All Other Students	997	1200	1290	1239.98	12.04	0.93	3.13
Non-migrant	8	1228	1244	1234.75	5.65	NA	NA
Undefined Migrant Status	43	1202	1251	1233.63	10.43	NA	NA
Augmentative Communication	1,385	1200	1290	1237.32	11.62	0.93	3.00
No Augmentative Communication	46	1200	1254	1231.11	11.31	NA	NA
Hearing Loss	1,376	1202	1290	1237.43	11.57	0.93	3.00
Within Normal Limits	6	1225	1244	1232.67	7.55	NA	NA
Visual Impairment	123	1200	1256	1228.20	8.27	0.86	2.85
Within Normal Limits	1,305	1202	1290	1238.06	11.51	0.93	3.01
Undefined Visual Impairment	153	1202	1280	1230.78	8.83	0.89	2.78
Sensory Stimuli Response	790	1202	1280	1235.89	10.52	0.92	2.90
Follow Directions	244	1200	1290	1240.39	12.71	0.93	3.20
Special School	156	1221	1288	1241.94	12.52	0.92	3.25
Regular School Self-contained	85	1224	1274	1243.16	12.21	NA	NA
Regular School Primarily Self-contained	104	1200	1251	1227.69	8.55	0.87	2.90
Regular School Resource Room	408	1202	1274	1232.59	9.39	0.91	2.77
Regular School General Education	916	1206	1290	1240.35	11.52	0.92	3.11
Communicates Primarily Through Cries	1,428	1200	1290	1237.21	11.60	0.93	3.00
Uses Intentional Communication	455	1200	1290	1237.70	12.30	0.93	3.08
Uses Symbolic Language	973	1202	1288	1236.98	11.26	0.93	2.96

Table M-3. IRT Subgroup Reliability: ELA Grade 5

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,415	1200	1290	1236.07	10.64	0.92	2.83
Female	474	1205	1273	1236.95	10.62	0.92	2.85
Male	941	1200	1290	1235.63	10.63	0.92	2.82
Gender Undefined	445	1200	1269	1236.06	10.26	0.92	2.80
Hispanic or Latino	110	1200	1290	1235.56	11.14	0.92	2.88
American Indian or Alaska Native	50	1220	1269	1235.04	10.42	NA	NA
Asian	141	1218	1273	1237.51	11.70	0.93	2.90
Black or African American	15	1200	1237	1228.00	8.55	NA	NA
Native Hawaiian or Pacific Islander	518	1200	1273	1235.83	10.47	0.92	2.82
White (non-Hispanic)	66	1223	1269	1234.80	11.02	NA	NA
Two or More Races (non-Hispanic)	445	1200	1269	1236.06	10.26	0.92	2.80
No Primary race/Ethnicity Undefined	211	1222	1269	1236.10	9.88	0.92	2.80
Currently receiving LEP services	227	1200	1273	1235.84	11.69	0.93	2.88
Not receiving LEP services	289	1200	1290	1236.93	11.47	0.93	2.93
LEP: All Other Students	81	1218	1269	1236.63	12.39	NA	NA
Economically Disadvantaged Students	7	1227	1255	1235.43	9.76	NA	NA
Non-economically Disadvantaged Students	220	1205	1273	1236.71	12.05	0.94	2.91
SES: All Other Students	398	1200	1269	1230.63	6.96	0.86	2.56
Non-migrant	1,011	1200	1290	1238.26	11.07	0.92	2.93
Undefined Migrant Status	6	1224	1239	1229.00	5.33	NA	NA
Augmentative Communication	34	1221	1257	1232.18	9.02	NA	NA
No Augmentative Communication	1,381	1200	1290	1236.17	10.66	0.92	2.83
Undefined Augmentative Communications	50	1200	1257	1231.60	9.33	NA	NA
Hearing Loss	1,363	1200	1290	1236.23	10.65	0.92	2.83
Within Normal Limits	127	1200	1252	1228.23	6.04	0.81	2.52
Visual Impairment	1,288	1200	1290	1236.85	10.69	0.92	2.86
Within Normal Limits	154	1200	1257	1230.17	8.76	0.90	2.66
Sensory Stimuli Response	787	1200	1273	1234.97	9.79	0.92	2.75
Follow Directions	230	1200	1290	1238.31	11.36	0.92	2.98
Special School	167	1223	1269	1240.25	10.51	0.92	2.96
Regular School Self-contained	77	1225	1273	1243.39	11.87	NA	NA
Regular School Primarily Self-contained	116	1200	1255	1229.37	6.93	0.85	2.55
Regular School Resource Room	405	1200	1269	1232.41	8.71	0.90	2.64
Regular School General Education	894	1200	1290	1238.61	10.98	0.92	2.95
Communicates Primarily Through Cries	1,415	1200	1290	1236.07	10.64	0.92	2.83
Uses Intentional Communication	474	1205	1273	1236.95	10.62	0.92	2.85
Uses Symbolic Language	941	1200	1290	1235.63	10.63	0.92	2.82

Table M-4. IRT Subgroup Reliability: ELA Grade 6

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,380	1203	1284	1236.25	9.62	0.92	2.62
Female	457	1209	1275	1237.06	10.27	0.92	2.69
Male	923	1203	1284	1235.85	9.26	0.92	2.58
Gender Undefined	461	1218	1275	1235.88	9.61	0.92	2.60
Hispanic or Latino	97	1221	1265	1235.82	8.84	NA	NA
American Indian or Alaska Native	42	1222	1253	1233.93	7.12	NA	NA
Asian	154	1209	1284	1235.99	10.81	0.93	2.69
Black or African American	15	1227	1244	1233.00	6.01	NA	NA
Native Hawaiian or Pacific Islander	455	1203	1269	1236.55	9.72	0.92	2.63
White (non-Hispanic)	81	1221	1253	1234.64	7.68	NA	NA
Two or More Races (non-Hispanic)	461	1218	1275	1235.88	9.61	0.92	2.60
No Primary race/Ethnicity Undefined	162	1222	1269	1234.78	7.59	0.89	2.46
Currently receiving LEP services	204	1203	1257	1234.24	8.54	0.91	2.49
Not receiving LEP services	284	1215	1265	1235.04	8.70	0.91	2.51
LEP: All Other Students	69	1203	1257	1232.99	9.31	NA	NA
Economically Disadvantaged Students	6	1227	1249	1234.67	8.59	NA	NA
Non-economically Disadvantaged Students	213	1203	1257	1234.02	8.43	0.91	2.49
SES: All Other Students	349	1203	1262	1231.63	7.00	0.88	2.32
Migrant	1,023	1218	1284	1237.86	9.89	0.92	2.72
Non-migrant	8	1224	1245	1231.88	6.81	NA	NA
Undefined Migrant Status	31	1215	1257	1231.48	8.96	NA	NA
Augmentative Communication	1,349	1203	1284	1236.36	9.61	0.92	2.62
No Augmentative Communication	44	1221	1262	1233.16	8.60	NA	NA
Undefined Augmentative Communications	1,331	1203	1284	1236.37	9.64	0.92	2.62
Hearing Loss	5	1222	1243	1232.20	8.11	NA	NA
Within Normal Limits	94	1209	1253	1228.59	5.89	NA	NA
Visual Impairment	1,286	1203	1284	1236.81	9.60	0.92	2.64
Within Normal Limits	145	1203	1262	1230.97	7.80	0.90	2.37
Sensory Stimuli Response	832	1215	1284	1235.59	9.10	0.91	2.55
Follow Directions	202	1222	1269	1238.47	10.03	0.92	2.75
Special School	134	1223	1275	1240.49	10.37	0.91	2.91
Regular School Self-contained	67	1224	1269	1240.76	9.87	NA	NA
Regular School Primarily Self-contained	89	1215	1253	1230.39	7.17	NA	NA
Regular School Resource Room	392	1203	1269	1232.20	7.54	0.90	2.36
Regular School General Education	899	1218	1284	1238.60	9.81	0.91	2.76
Communicates Primarily Through Cries	1,380	1203	1284	1236.25	9.62	0.92	2.62
Uses Intentional Communication	457	1209	1275	1237.06	10.27	0.92	2.69
Uses Symbolic Language	923	1203	1284	1235.85	9.26	0.92	2.58

Table M-5. IRT Subgroup Reliability: ELA Grade 7

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,282	1201	1282	1240.28	11.01	0.92	3.04
Female	432	1201	1276	1241.15	11.05	0.92	3.09
Male	850	1209	1282	1239.83	10.97	0.92	3.02
Gender Undefined	401	1213	1276	1239.71	10.54	0.91	2.98
Hispanic or Latino	85	1226	1272	1243.39	10.62	NA	NA
American Indian or Alaska Native	33	1223	1260	1235.18	7.51	NA	NA
Asian	146	1215	1268	1239.63	9.60	0.90	2.95
Black or African American	18	1228	1252	1238.78	7.03	NA	NA
White (non-Hispanic)	478	1201	1276	1240.53	11.86	0.92	3.11
Two or More Races (non-Hispanic)	63	1213	1265	1239.65	10.78	NA	NA
No Primary race/Ethnicity Undefined	401	1213	1276	1239.71	10.54	0.91	2.98
Currently receiving LEP services	110	1223	1262	1239.29	8.38	0.88	2.87
Not receiving LEP services	252	1209	1272	1239.23	10.68	0.92	2.98
LEP: All Other Students	291	1209	1276	1241.05	11.35	0.92	3.09
Economically Disadvantaged Students	96	1209	1268	1238.51	9.76	NA	NA
Non-economically Disadvantaged Students	7	1218	1242	1231.14	8.38	NA	NA
SES: All Other Students	247	1209	1272	1238.87	10.51	0.92	2.96
Migrant	304	1201	1262	1233.05	7.19	0.85	2.68
Non-migrant	966	1213	1282	1242.57	11.03	0.91	3.16
Undefined Migrant Status	12	1228	1268	1238.75	10.91	NA	NA
Augmentative Communication	28	1223	1260	1236.07	10.49	NA	NA
No Augmentative Communication	1,254	1201	1282	1240.37	11.01	0.92	3.05
Undefined Augmentative Communications	57	1201	1268	1236.46	10.82	NA	NA
Hearing Loss	1,220	1209	1282	1240.48	10.98	0.92	3.05
Within Normal Limits	5	1225	1256	1234.40	12.62	NA	NA
Visual Impairment	90	1201	1250	1231.70	6.61	NA	NA
Within Normal Limits	1,192	1209	1282	1240.93	11.01	0.92	3.07
Undefined Visual Impairment	134	1201	1265	1234.77	8.78	0.89	2.80
Sensory Stimuli Response	765	1213	1276	1239.46	10.35	0.91	2.97
Follow Directions	179	1209	1268	1239.86	10.53	0.91	2.98
Special School	144	1223	1282	1247.12	11.83	0.90	3.51
Regular School Self-contained	60	1226	1276	1247.85	12.14	NA	NA
Regular School Primarily Self-contained	69	1201	1268	1232.26	9.20	NA	NA
Regular School Resource Room	321	1213	1276	1234.92	8.82	0.90	2.72
Regular School General Education	892	1209	1282	1242.83	10.87	0.91	3.17
Uses Intentional Communication	1,282	1201	1282	1240.28	11.01	0.92	3.04
Uses Symbolic Language	432	1201	1276	1241.15	11.05	0.92	3.09
Communicates Primarily Through Cries	850	1209	1282	1239.83	10.97	0.92	3.02

Table M-6. IRT Subgroup Reliability: ELA Grade 8

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,199	1210	1271	1235.52	9.49	0.92	2.57
Female	413	1218	1266	1235.46	9.24	0.92	2.56
Male	786	1210	1271	1235.55	9.63	0.92	2.58
Gender Undefined	403	1210	1266	1235.25	9.44	0.92	2.56
Hispanic or Latino	92	1213	1260	1237.47	9.56	NA	NA
American Indian or Alaska Native	37	1223	1254	1232.57	6.66	NA	NA
Asian	108	1216	1254	1234.46	8.28	0.91	2.49
Black or African American	13	1223	1250	1233.62	8.29	NA	NA
Native Hawaiian or Pacific Islander	425	1213	1271	1235.26	9.77	0.93	2.56
White (non-Hispanic)	67	1221	1263	1236.91	10.20	NA	NA
Two or More Races (non-Hispanic)	403	1210	1266	1235.25	9.44	0.92	2.56
No Primary race/Ethnicity Undefined	113	1218	1260	1234.60	8.98	0.92	2.52
Currently receiving LEP services	181	1213	1271	1234.87	9.97	0.93	2.57
Not receiving LEP services	222	1213	1271	1236.12	9.84	0.92	2.63
LEP: All Other Students	65	1219	1258	1233.14	9.17	NA	NA
Economically Disadvantaged Students	12	1218	1252	1233.92	9.86	NA	NA
Non-economically Disadvantaged Students	180	1213	1271	1234.88	10.19	0.93	2.56
SES: All Other Students	284	1210	1258	1229.78	6.63	0.88	2.27
Migrant	905	1213	1271	1237.29	9.58	0.92	2.67
Non-migrant	10	1228	1248	1237.90	6.15	NA	NA
Undefined Migrant Status	37	1220	1254	1231.81	8.51	NA	NA
Augmentative Communication	1,162	1210	1271	1235.64	9.50	0.92	2.58
No Augmentative Communication	53	1215	1255	1230.38	8.26	NA	NA
Undefined Augmentative Communications	1,143	1210	1271	1235.74	9.49	0.92	2.58
Hearing Loss	67	1215	1250	1227.99	6.58	NA	NA
Within Normal Limits	1,132	1210	1271	1235.97	9.45	0.92	2.59
Visual Impairment	141	1213	1263	1229.73	8.59	0.92	2.34
Within Normal Limits	718	1210	1266	1234.97	8.80	0.91	2.52
Sensory Stimuli Response	180	1219	1271	1237.22	9.86	0.92	2.68
Follow Directions	101	1223	1266	1240.42	9.26	0.90	2.84
Special School	59	1223	1263	1242.47	9.72	NA	NA
Regular School Self-contained	73	1210	1250	1228.16	6.33	NA	NA
Regular School Primarily Self-contained	276	1215	1260	1231.41	7.85	0.91	2.34
Regular School Resource Room	850	1213	1271	1237.49	9.48	0.92	2.67
Regular School General Education	1,199	1210	1271	1235.52	9.49	0.92	2.57
Communicates Primarily Through Cries	413	1218	1266	1235.46	9.24	0.92	2.56
Uses Intentional Communication	786	1210	1271	1235.55	9.63	0.92	2.58
Uses Symbolic Language	403	1210	1266	1235.25	9.44	0.92	2.56

Table M-7. IRT Subgroup Reliability: ELA Grade HS

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,180	1218	1277	1240.51	9.06	0.93	2.31
Female	436	1218	1277	1240.54	9.19	0.93	2.32
Male	744	1218	1268	1240.50	8.99	0.93	2.31
Gender Undefined	383	1218	1268	1239.87	8.58	0.93	2.27
Hispanic or Latino	98	1226	1261	1239.47	7.08	NA	NA
American Indian or Alaska Native	44	1226	1263	1237.48	7.09	NA	NA
Asian	94	1225	1268	1240.19	8.96	NA	NA
Black or African American	11	1227	1255	1235.09	8.31	NA	NA
Native Hawaiian or Pacific Islander	462	1218	1277	1241.10	9.28	0.93	2.35
White (non-Hispanic)	48	1227	1272	1241.75	10.12	NA	NA
Two or More Races (non-Hispanic)	383	1218	1268	1239.87	8.58	0.93	2.27
No Primary race/Ethnicity Undefined	75	1226	1263	1238.35	7.51	NA	NA
Currently receiving LEP services	193	1220	1268	1239.50	9.80	0.94	2.31
Not receiving LEP services	206	1220	1268	1240.04	9.26	0.93	2.30
LEP: All Other Students	90	1225	1261	1240.50	9.28	NA	NA
Economically Disadvantaged Students	190	1220	1268	1240.18	10.01	0.94	2.36
Non-economically Disadvantaged Students	226	1220	1256	1235.32	5.83	0.88	2.00
SES: All Other Students	946	1218	1277	1241.72	9.24	0.93	2.39
Non-migrant	8	1231	1259	1244.25	10.66	NA	NA
Undefined Migrant Status	38	1218	1257	1235.71	8.19	NA	NA
Augmentative Communication	1,142	1218	1277	1240.67	9.05	0.93	2.32
No Augmentative Communication	70	1218	1261	1235.90	8.34	NA	NA
Undefined Augmentative Communications	1,105	1220	1277	1240.77	9.02	0.93	2.32
Hearing Loss	5	1238	1261	1248.00	8.37	NA	NA
Within Normal Limits	72	1218	1253	1232.68	5.55	NA	NA
Visual Impairment	1,108	1218	1277	1241.02	9.01	0.93	2.34
Within Normal Limits	156	1218	1263	1235.08	6.95	0.91	2.06
Sensory Stimuli Response	675	1218	1272	1239.83	8.76	0.93	2.27
Follow Directions	242	1225	1268	1243.74	8.83	0.92	2.49
Special School	80	1231	1261	1244.88	8.52	NA	NA
Regular School Self-contained	27	1229	1277	1247.26	10.55	NA	NA
Regular School Primarily Self-contained	65	1218	1244	1231.80	4.93	NA	NA
Regular School Resource Room	249	1220	1265	1236.49	7.33	0.92	2.08
Regular School General Education	866	1218	1277	1242.32	9.00	0.92	2.41
Communicates Primarily Through Cries	1,180	1218	1277	1240.51	9.06	0.93	2.31
Uses Intentional Communication	436	1218	1277	1240.54	9.19	0.93	2.32
Uses Symbolic Language	744	1218	1268	1240.50	8.99	0.93	2.31

Table M-8. IRT Subgroup Reliability: Mathematics Grade 3

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,460	1200	1290	1240.12	11.18	0.86	3.92
Female	440	1200	1290	1239.64	10.89	0.84	3.94
Male	1,020	1200	1290	1240.32	11.30	0.86	3.92
Gender Undefined	469	1200	1290	1239.51	11.23	0.85	3.96
Hispanic or Latino	81	1213	1274	1240.30	10.21	NA	NA
American Indian or Alaska Native	43	1213	1262	1238.98	10.95	NA	NA
Asian	135	1209	1285	1239.56	10.78	0.86	3.86
Black or African American	15	1221	1257	1239.00	9.45	NA	NA
Native Hawaiian or Pacific Islander	532	1200	1290	1239.50	10.71	0.85	3.89
White (non-Hispanic)	77	1216	1285	1239.62	10.26	NA	NA
Two or More Races (non-Hispanic)	469	1200	1290	1239.51	11.23	0.85	3.96
No Primary race/Ethnicity Undefined	262	1200	1290	1238.19	11.25	0.85	4.03
Currently receiving LEP services	207	1200	1285	1238.43	11.66	0.86	4.03
Not receiving LEP services	282	1200	1290	1240.29	10.95	0.86	3.88
LEP: All Other Students	65	1200	1265	1234.80	13.08	NA	NA
Economically Disadvantaged Students	10	1229	1262	1245.80	9.90	NA	NA
Non-economically Disadvantaged Students	216	1200	1290	1238.33	12.08	0.85	4.12
SES: All Other Students	509	1200	1274	1236.01	9.43	0.82	3.92
Migrant	947	1200	1290	1242.28	11.43	0.86	3.93
Non-migrant	4	1246	1257	1250.75	4.86	NA	NA
Undefined Migrant Status	33	1216	1252	1235.64	9.76	NA	NA
Augmentative Communication	1,427	1200	1290	1240.22	11.19	0.86	3.92
No Augmentative Communication	60	1200	1285	1234.82	13.80	NA	NA
Undefined Augmentative Communications	1,398	1200	1290	1240.34	11.01	0.86	3.90
Hearing Loss	167	1200	1262	1232.61	10.62	0.81	4.26
Within Normal Limits	1,293	1200	1290	1241.09	10.88	0.85	3.88
Visual Impairment	153	1200	1259	1235.10	9.92	0.82	4.02
Within Normal Limits	868	1200	1290	1238.94	10.30	0.85	3.86
Undefined Visual Impairment	232	1200	1290	1243.84	11.76	0.86	4.00
Sensory Stimuli Response	118	1216	1290	1245.44	11.60	0.85	3.97
Follow Directions	89	1200	1285	1243.44	13.09	NA	NA
Special School	149	1200	1265	1232.23	10.94	0.82	4.33
Regular School Self-contained	456	1200	1285	1237.54	10.21	0.84	3.93
Regular School Primarily Self-contained	855	1200	1290	1242.87	10.73	0.85	3.85
Regular School Resource Room	1,460	1200	1290	1240.12	11.18	0.86	3.92
Regular School General Education	440	1200	1290	1239.64	10.89	0.84	3.94
Communicates Primarily Through Cries	1,020	1200	1290	1240.32	11.30	0.86	3.92
Uses Intentional Communication	469	1200	1290	1239.51	11.23	0.85	3.96
Uses Symbolic Language	81	1213	1274	1240.30	10.21	NA	NA

Table M-9. IRT Subgroup Reliability: Mathematics Grade 4

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,427	1203	1283	1238.31	9.49	0.87	3.33
Female	452	1203	1274	1237.61	9.06	0.86	3.30
Male	975	1203	1283	1238.64	9.67	0.87	3.35
Gender Undefined	476	1203	1283	1238.59	9.47	0.87	3.32
Hispanic or Latino	91	1203	1274	1237.55	9.68	NA	NA
American Indian or Alaska Native	42	1215	1256	1234.88	9.82	NA	NA
Asian	148	1215	1283	1238.54	9.41	0.87	3.34
Black or African American	18	1211	1251	1232.11	8.18	NA	NA
Native Hawaiian or Pacific Islander	490	1203	1283	1238.47	9.13	0.86	3.29
White (non-Hispanic)	77	1211	1255	1236.74	8.12	NA	NA
Two or More Races (non-Hispanic)	476	1203	1283	1238.59	9.47	0.87	3.32
No Primary race/Ethnicity Undefined	246	1203	1283	1237.70	9.42	0.86	3.35
Currently receiving LEP services	239	1203	1283	1238.41	9.54	0.87	3.34
Not receiving LEP services	298	1203	1283	1239.32	9.98	0.88	3.36
LEP: All Other Students	87	1223	1283	1238.07	9.38	NA	NA
Economically Disadvantaged Students	7	1226	1236	1232.86	3.72	NA	NA
Non-economically Disadvantaged Students	242	1203	1283	1239.02	9.76	0.87	3.35
SES: All Other Students	423	1203	1265	1234.73	7.61	0.79	3.39
Migrant	997	1203	1283	1239.83	9.80	0.88	3.31
Non-migrant	7	1221	1249	1238.14	10.07	NA	NA
Undefined Migrant Status	44	1203	1258	1235.68	10.21	NA	NA
Augmentative Communication	1,383	1203	1283	1238.40	9.46	0.87	3.33
No Augmentative Communication	42	1203	1253	1235.21	10.18	NA	NA
Undefined Augmentative Communications	1,379	1203	1283	1238.40	9.47	0.87	3.32
Hearing Loss	6	1233	1244	1238.83	4.17	NA	NA
Within Normal Limits	118	1203	1252	1232.88	8.35	0.79	3.59
Visual Impairment	1,309	1203	1283	1238.80	9.44	0.87	3.31
Within Normal Limits	156	1215	1260	1235.29	7.55	0.80	3.34
Undefined Visual Impairment	789	1203	1283	1237.02	9.14	0.86	3.36
Sensory Stimuli Response	242	1203	1274	1240.26	9.40	0.87	3.26
Follow Directions	155	1211	1283	1242.08	10.60	0.88	3.38
Special School	85	1224	1269	1243.45	9.21	NA	NA
Regular School Self-contained	103	1203	1247	1231.93	8.49	0.79	3.68
Regular School Primarily Self-contained	407	1203	1283	1235.27	8.90	0.84	3.44
Regular School Resource Room	917	1211	1283	1240.38	9.17	0.87	3.25
Regular School General Education	1,427	1203	1283	1238.31	9.49	0.87	3.33
Communicates Primarily Through Cries	452	1203	1274	1237.61	9.06	0.86	3.30
Uses Intentional Communication	975	1203	1283	1238.64	9.67	0.87	3.35
Uses Symbolic Language	476	1203	1283	1238.59	9.47	0.87	3.32

Table M-10. IRT Subgroup Reliability: Mathematics Grade 5

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,411	1200	1290	1240.20	9.79	0.85	3.76
Female	474	1211	1274	1240.03	8.87	0.83	3.69
Male	937	1200	1290	1240.29	10.23	0.85	3.79
Gender Undefined	445	1217	1280	1240.63	9.72	0.85	3.73
Hispanic or Latino	109	1222	1290	1240.94	10.10	0.84	3.84
American Indian or Alaska Native	50	1223	1280	1240.80	12.29	NA	NA
Asian	141	1200	1274	1239.67	9.86	0.85	3.75
Black or African American	14	1225	1248	1234.71	7.94	NA	NA
Native Hawaiian or Pacific Islander	515	1200	1274	1239.37	9.31	0.83	3.76
White (non-Hispanic)	67	1214	1267	1239.27	9.50	NA	NA
Two or More Races (non-Hispanic)	445	1217	1280	1240.63	9.72	0.85	3.73
No Primary race/Ethnicity Undefined	212	1222	1280	1240.29	9.81	0.85	3.72
Currently receiving LEP services	225	1200	1274	1239.42	10.66	0.85	3.90
Not receiving LEP services	289	1217	1290	1239.63	9.96	0.85	3.75
LEP: All Other Students	80	1200	1280	1240.41	11.49	NA	NA
Economically Disadvantaged Students	7	1230	1254	1241.71	8.14	NA	NA
Non-economically Disadvantaged Students	219	1200	1280	1240.07	10.83	0.85	3.93
SES: All Other Students	396	1200	1280	1236.54	8.26	0.79	3.72
Non-migrant	1,008	1200	1290	1241.68	9.96	0.85	3.77
Undefined Migrant Status	7	1214	1246	1234.57	11.00	NA	NA
Augmentative Communication	34	1217	1255	1235.21	8.43	NA	NA
No Augmentative Communication	1,377	1200	1290	1240.33	9.79	0.85	3.76
Undefined Augmentative Communications	49	1217	1274	1238.12	10.07	NA	NA
Hearing Loss	1,360	1200	1290	1240.28	9.78	0.85	3.76
Within Normal Limits	124	1222	1260	1235.59	6.92	0.71	3.68
Visual Impairment	1,287	1200	1290	1240.65	9.92	0.85	3.76
Within Normal Limits	156	1200	1274	1236.58	10.94	0.84	3.99
Sensory Stimuli Response	782	1217	1280	1239.35	8.66	0.82	3.69
Follow Directions	228	1222	1290	1241.61	9.99	0.85	3.79
Special School	168	1225	1274	1242.79	10.49	0.87	3.75
Regular School Self-contained	77	1227	1280	1246.42	11.46	NA	NA
Regular School Primarily Self-contained	113	1217	1262	1236.10	7.22	0.74	3.68
Regular School Resource Room	405	1200	1280	1238.04	9.98	0.85	3.76
Regular School General Education	893	1200	1290	1241.70	9.67	0.84	3.76
Communicates Primarily Through Cries	1,411	1200	1290	1240.20	9.79	0.85	3.76
Uses Intentional Communication	474	1211	1274	1240.03	8.87	0.83	3.69
Uses Symbolic Language	937	1200	1290	1240.29	10.23	0.85	3.79

Table M-11. IRT Subgroup Reliability: Mathematics Grade 6

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,376	1205	1290	1240.23	10.60	0.89	3.17
Female	457	1212	1290	1240.50	10.27	0.89	3.14
Male	919	1205	1290	1240.10	10.77	0.89	3.18
Gender Undefined	458	1215	1290	1239.43	9.96	0.89	3.08
Hispanic or Latino	97	1222	1290	1242.42	12.15	NA	NA
American Indian or Alaska Native	41	1224	1259	1239.34	9.45	NA	NA
Asian	152	1205	1290	1239.18	10.15	0.88	3.14
Black or African American	15	1231	1259	1241.20	8.25	NA	NA
Native Hawaiian or Pacific Islander	457	1215	1290	1239.96	10.71	0.89	3.17
White (non-Hispanic)	81	1220	1255	1238.85	8.46	NA	NA
Two or More Races (non-Hispanic)	458	1215	1290	1239.43	9.96	0.89	3.08
No Primary race/Ethnicity Undefined	161	1224	1262	1239.61	8.70	0.88	2.95
Currently receiving LEP services	203	1205	1278	1238.01	9.57	0.89	3.04
Not receiving LEP services	284	1215	1290	1238.95	10.15	0.88	3.11
LEP: All Other Students	68	1205	1257	1236.87	9.82	NA	NA
Economically Disadvantaged Students	6	1233	1251	1242.50	6.86	NA	NA
Non-economically Disadvantaged Students	211	1205	1278	1237.61	9.23	0.89	3.01
SES: All Other Students	347	1205	1270	1235.40	8.79	0.88	2.98
Migrant	1,021	1212	1290	1241.90	10.67	0.88	3.23
Non-migrant	8	1229	1259	1237.50	9.78	NA	NA
Undefined Migrant Status	32	1220	1252	1235.31	8.09	NA	NA
Augmentative Communication	1,344	1205	1290	1240.35	10.63	0.89	3.17
No Augmentative Communication	44	1220	1259	1234.98	9.05	NA	NA
Undefined Augmentative Communications	1,327	1205	1290	1240.42	10.62	0.89	3.17
Hearing Loss	5	1228	1248	1237.60	7.44	NA	NA
Within Normal Limits	91	1212	1252	1231.69	7.26	NA	NA
Visual Impairment	1,285	1205	1290	1240.84	10.54	0.89	3.18
Within Normal Limits	142	1205	1265	1233.36	8.47	0.86	3.03
Sensory Stimuli Response	831	1215	1290	1239.61	9.83	0.89	3.07
Follow Directions	201	1222	1290	1242.80	12.53	0.87	3.54
Special School	135	1224	1290	1244.49	9.64	0.86	3.22
Regular School Self-contained	67	1225	1278	1246.19	10.52	NA	NA
Regular School Primarily Self-contained	88	1212	1270	1233.74	9.80	NA	NA
Regular School Resource Room	388	1205	1290	1236.73	9.08	0.88	3.00
Regular School General Education	900	1220	1290	1242.38	10.64	0.88	3.24
Communicates Primarily Through Cries	1,376	1205	1290	1240.23	10.60	0.89	3.17
Uses Intentional Communication	457	1212	1290	1240.50	10.27	0.89	3.14
Uses Symbolic Language	919	1205	1290	1240.10	10.77	0.89	3.18

Table M-12. IRT Subgroup Reliability: Mathematics Grade 7

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,279	1200	1290	1239.39	10.70	0.87	3.60
Female	430	1200	1290	1239.10	10.88	0.88	3.60
Male	849	1205	1290	1239.55	10.62	0.87	3.59
Gender Undefined	400	1213	1290	1239.60	10.26	0.87	3.57
Hispanic or Latino	85	1222	1270	1241.49	10.08	NA	NA
American Indian or Alaska Native	33	1219	1263	1237.30	9.58	NA	NA
Asian	146	1200	1260	1238.45	9.89	0.86	3.53
Black or African American	18	1227	1290	1239.56	13.83	NA	NA
Native Hawaiian or Pacific Islander	477	1205	1290	1238.94	11.15	0.88	3.63
White (non-Hispanic)	62	1210	1266	1238.84	9.90	NA	NA
Two or More Races (non-Hispanic)	400	1213	1290	1239.60	10.26	0.87	3.57
No Primary race/Ethnicity Undefined	111	1219	1266	1238.02	9.22	0.86	3.43
Currently receiving LEP services	251	1200	1290	1238.00	10.24	0.86	3.57
Not receiving LEP services	291	1200	1284	1239.23	10.32	0.88	3.53
LEP: All Other Students	96	1210	1266	1236.78	8.60	NA	NA
Economically Disadvantaged Students	7	1210	1244	1227.29	10.40	NA	NA
Non-economically Disadvantaged Students	247	1200	1275	1237.41	9.90	0.87	3.53
SES: All Other Students	306	1200	1290	1234.10	8.53	0.81	3.56
Migrant	962	1200	1290	1241.04	10.69	0.88	3.60
Non-migrant	11	1226	1290	1242.45	17.37	NA	NA
Undefined Migrant Status	28	1226	1263	1237.36	9.57	NA	NA
Augmentative Communication	1,251	1200	1290	1239.44	10.73	0.87	3.60
No Augmentative Communication	56	1205	1258	1234.07	9.60	NA	NA
Undefined Augmentative Communications	1,218	1200	1290	1239.66	10.68	0.87	3.60
Hearing Loss	5	1219	1248	1234.60	12.78	NA	NA
Within Normal Limits	91	1205	1290	1230.57	9.17	NA	NA
Visual Impairment	1,188	1200	1290	1240.07	10.51	0.87	3.59
Within Normal Limits	135	1200	1275	1233.62	9.67	0.85	3.66
Undefined Visual Impairment	762	1200	1290	1238.58	9.66	0.86	3.49
Sensory Stimuli Response	176	1213	1270	1239.99	10.13	0.87	3.54
Follow Directions	146	1222	1290	1244.81	12.30	0.87	3.91
Special School	60	1226	1290	1247.85	12.58	NA	NA
Regular School Self-contained	69	1205	1290	1231.94	11.01	NA	NA
Regular School Primarily Self-contained	319	1200	1290	1234.85	9.20	0.83	3.57
Regular School Resource Room	891	1205	1290	1241.60	10.41	0.87	3.59
Regular School General Education	1,279	1200	1290	1239.39	10.70	0.87	3.60
Communicates Primarily Through Cries	430	1200	1290	1239.10	10.88	0.88	3.60
Uses Intentional Communication	849	1205	1290	1239.55	10.62	0.87	3.59
Uses Symbolic Language	400	1213	1290	1239.60	10.26	0.87	3.57

Table M-13. IRT Subgroup Reliability: Mathematics Grade 8

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,200	1200	1290	1239.89	10.82	0.88	3.54
Female	415	1200	1284	1239.60	9.67	0.87	3.42
Male	785	1200	1290	1240.04	11.38	0.88	3.61
Gender Undefined	403	1200	1284	1239.89	10.69	0.87	3.54
Hispanic or Latino	92	1219	1284	1241.20	10.49	NA	NA
American Indian or Alaska Native	37	1208	1264	1238.14	12.81	NA	NA
Asian	109	1219	1264	1237.97	9.61	0.87	3.38
Black or African American	13	1213	1290	1240.46	17.87	NA	NA
Native Hawaiian or Pacific Islander	426	1208	1284	1239.48	10.46	0.87	3.50
White (non-Hispanic)	66	1219	1284	1241.03	11.36	NA	NA
Two or More Races (non-Hispanic)	403	1200	1284	1239.89	10.69	0.87	3.54
No Primary race/Ethnicity Undefined	114	1200	1290	1240.54	13.14	0.85	4.01
Currently receiving LEP services	181	1208	1284	1238.81	10.82	0.89	3.51
Not receiving LEP services	222	1200	1284	1240.06	11.04	0.86	3.65
LEP: All Other Students	67	1219	1264	1236.87	9.24	NA	NA
Economically Disadvantaged Students	12	1200	1268	1238.33	19.04	NA	NA
Non-economically Disadvantaged Students	182	1200	1284	1238.59	11.25	0.86	3.65
SES: All Other Students	283	1208	1290	1234.33	9.06	0.84	3.41
Migrant	907	1200	1284	1241.58	10.74	0.87	3.58
Non-migrant	10	1225	1264	1243.90	11.46	NA	NA
Undefined Migrant Status	38	1217	1255	1236.18	8.57	NA	NA
Augmentative Communication	1,162	1200	1290	1240.01	10.86	0.88	3.55
No Augmentative Communication	52	1217	1253	1233.79	7.72	NA	NA
Undefined Augmentative Communications	1,145	1200	1290	1240.18	10.86	0.88	3.55
Hearing Loss	64	1200	1290	1232.69	11.76	NA	NA
Within Normal Limits	1,136	1200	1284	1240.29	10.62	0.88	3.51
Visual Impairment	136	1200	1284	1235.50	12.23	0.87	3.80
Within Normal Limits	724	1208	1290	1239.03	9.87	0.87	3.42
Undefined Visual Impairment	180	1213	1284	1241.91	11.00	0.88	3.60
Sensory Stimuli Response	101	1227	1284	1245.02	10.16	0.85	3.68
Follow Directions	59	1208	1284	1245.56	12.65	NA	NA
Special School	73	1200	1290	1232.85	11.18	NA	NA
Regular School Self-contained	275	1200	1284	1235.98	10.43	0.88	3.48
Regular School Primarily Self-contained	852	1208	1284	1241.75	10.32	0.87	3.53
Regular School Resource Room	1,200	1200	1290	1239.89	10.82	0.88	3.54
Regular School General Education	415	1200	1284	1239.60	9.67	0.87	3.42
Communicates Primarily Through Cries	785	1200	1290	1240.04	11.38	0.88	3.61
Uses Intentional Communication	403	1200	1284	1239.89	10.69	0.87	3.54
Uses Symbolic Language	92	1219	1284	1241.20	10.49	NA	NA

Table M-14. IRT Subgroup Reliability: Mathematics Grade HS

Description	Number of Students	Scale Score Minimum	Scale Score Maximum	Scale Score Mean	Scale Score Standard Deviation	IRT Marginal Reliability	Standard Error
All	1,171	1200	1290	1239.98	9.24	0.88	2.95
Female	433	1208	1278	1239.20	8.56	0.88	2.86
Male	738	1200	1290	1240.44	9.59	0.87	3.00
Gender Undefined	382	1214	1270	1239.51	8.21	0.87	2.84
Hispanic or Latino	97	1221	1262	1239.61	7.39	NA	NA
American Indian or Alaska Native	44	1221	1265	1237.73	9.07	NA	NA
Asian	93	1200	1270	1238.97	10.37	NA	NA
Black or African American	11	1228	1259	1237.00	9.77	NA	NA
Native Hawaiian or Pacific Islander	458	1208	1278	1240.12	9.32	0.89	2.92
White (non-Hispanic)	48	1218	1278	1242.10	10.35	NA	NA
Two or More Races (non-Hispanic)	382	1214	1270	1239.51	8.21	0.87	2.84
No Primary race/Ethnicity Undefined	74	1214	1257	1237.59	7.39	NA	NA
Currently receiving LEP services	186	1208	1278	1239.27	9.50	0.89	2.99
Not receiving LEP services	200	1208	1278	1239.42	9.02	0.88	2.93
LEP: All Other Students	88	1214	1262	1239.95	8.28	NA	NA
Economically Disadvantaged Students	183	1208	1278	1239.77	9.61	0.89	2.98
Non-economically Disadvantaged Students	225	1214	1265	1236.40	7.31	0.84	2.87
SES: All Other Students	938	1200	1290	1240.84	9.48	0.88	2.97
Migrant	8	1233	1247	1240.25	4.74	NA	NA
Non-migrant	38	1208	1262	1236.05	9.57	NA	NA
Undefined Migrant Status	1,133	1200	1290	1240.11	9.20	0.87	2.94
Augmentative Communication	70	1208	1254	1235.39	8.67	NA	NA
No Augmentative Communication	1,096	1200	1290	1240.22	9.17	0.87	2.94
Hearing Loss	5	1241	1265	1252.20	9.20	NA	NA
Within Normal Limits	72	1208	1248	1234.44	7.06	NA	NA
Visual Impairment	1,099	1200	1290	1240.34	9.25	0.88	2.94
Within Normal Limits	152	1200	1262	1235.52	7.64	0.75	3.09
Sensory Stimuli Response	675	1208	1278	1239.22	8.83	0.88	2.91
Follow Directions	239	1214	1290	1243.16	9.28	0.86	2.94
Special School	79	1218	1270	1243.82	10.04	NA	NA
Regular School Self-contained	26	1230	1278	1244.92	10.02	NA	NA
Regular School Primarily Self-contained	65	1208	1252	1234.31	7.26	NA	NA
Regular School Resource Room	247	1218	1270	1237.23	8.03	0.87	2.89
Regular School General Education	859	1200	1290	1241.20	9.38	0.87	2.96
Communicates Primarily Through Cries	1,171	1200	1290	1239.98	9.24	0.88	2.95
Uses Intentional Communication	433	1208	1278	1239.20	8.56	0.88	2.86
Uses Symbolic Language	738	1200	1290	1240.44	9.59	0.87	3.00

APPENDIX N
DECISION ACCURACY AND CONSISTENCY RESULTS

Table N-1. Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Performance Level

Content Area	Grade	Kappa	DA Overall	DA Level 1	DA Level 2	DA Level 3	DA Level 4	DC Overall	DC Level 1	DC Level 2	DC Level 3	DC Level 4
ELA	3	0.64	0.83	0.91	0.64	0.81	0.85	0.76	0.88	0.51	0.74	0.74
ELA	4	0.67	0.85	0.92	0.66	0.86	0.80	0.78	0.88	0.55	0.80	0.68
ELA	5	0.67	0.84	0.90	0.77	0.83	0.80	0.77	0.86	0.66	0.77	0.69
ELA	6	0.64	0.82	0.89	0.72	0.84	0.84	0.74	0.83	0.62	0.77	0.74
ELA	7	0.63	0.81	0.89	0.59	0.82	0.79	0.74	0.86	0.46	0.74	0.73
ELA	8	0.65	0.83	0.90	0.74	0.82	0.82	0.75	0.87	0.65	0.72	0.73
ELA	HS	0.66	0.83	0.92	0.60	0.86	0.77	0.77	0.86	0.48	0.81	0.72
Mathematics	3	0.55	0.77	0.86	0.60	0.78	0.83	0.69	0.80	0.45	0.74	0.70
Mathematics	4	0.52	0.75	0.78	0.65	0.83	0.79	0.66	0.68	0.57	0.72	0.70
Mathematics	5	0.50	0.74	0.75	0.66	0.77	0.87	0.64	0.64	0.56	0.69	0.76
Mathematics	6	0.59	0.78	0.85	0.63	0.83	0.86	0.70	0.76	0.54	0.73	0.79
Mathematics	7	0.54	0.76	0.81	0.61	0.81	0.82	0.67	0.75	0.49	0.72	0.76
Mathematics	8	0.58	0.79	0.87	0.57	0.79	0.85	0.71	0.83	0.41	0.75	0.72
Mathematics	HS	0.57	0.77	0.85	0.56	0.83	0.88	0.69	0.76	0.47	0.74	0.79

Table N-2. Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint

Content Area	Grade	Level 1/ Level 2 Decision Accuracy	Level 1/ Level 2 Decision Consistency	Level 1/ Level 2 False Positive	Level 1/ Level 2 False Negative	Level 1/ Level 2 Decision Accuracy	Level 2/ Level 3 Decision Consistency	Level 2/ Level 3 False Positive	Level 2/ Level 3 False Negative	Level 2/ Level 3 Decision Accuracy	Level 3/ Level 4 Decision Consistency	Level 3/ Level 4 False Positive	Level 3/ Level 4 False Negative
ELA	3	0.92	0.88	0.04	0.04	0.94	0.92	0.03	0.03	0.97	0.96	0.02	0.01
ELA	4	0.93	0.90	0.04	0.04	0.94	0.92	0.03	0.03	0.98	0.97	0.01	0.01
ELA	5	0.92	0.89	0.04	0.04	0.94	0.92	0.03	0.03	0.97	0.96	0.01	0.01
ELA	6	0.93	0.89	0.04	0.04	0.93	0.90	0.04	0.03	0.97	0.95	0.02	0.01
ELA	7	0.92	0.89	0.04	0.03	0.93	0.90	0.03	0.04	0.95	0.93	0.02	0.03
ELA	8	0.94	0.91	0.04	0.03	0.93	0.90	0.05	0.03	0.96	0.95	0.02	0.02
ELA	HS	0.93	0.90	0.03	0.04	0.93	0.90	0.03	0.04	0.96	0.95	0.01	0.02
Mathematics	3	0.92	0.88	0.04	0.04	0.89	0.84	0.04	0.07	0.96	0.95	0.02	0.01
Mathematics	4	0.89	0.85	0.05	0.06	0.90	0.86	0.06	0.04	0.96	0.94	0.02	0.02
Mathematics	5	0.89	0.85	0.05	0.06	0.88	0.83	0.06	0.06	0.96	0.95	0.02	0.01
Mathematics	6	0.92	0.89	0.03	0.04	0.91	0.87	0.06	0.03	0.96	0.94	0.02	0.02
Mathematics	7	0.89	0.84	0.06	0.05	0.90	0.86	0.05	0.05	0.96	0.95	0.02	0.02
Mathematics	8	0.93	0.90	0.04	0.03	0.90	0.86	0.03	0.07	0.95	0.93	0.03	0.02
Mathematics	HS	0.90	0.87	0.04	0.06	0.90	0.86	0.06	0.04	0.97	0.95	0.02	0.01

Note: Due to the small sample size, students in Levels 3 and 4 were collapsed for purposes of the decision accuracy and consistency analysis.

APPENDIX O
LIST OF ACRONYMS AND COMMONLY USED TERMS

Acronyms Used in the 2025 MSAA Technical Report

Acronym	Explanation
2PL	two-parameter logistic
AA-AAS	Alternate Assessment Aligned with Alternate Achievement Standards (utilized under ESEA until 2015)
AA-AAAS	Alternate Assessment Aligned with Alternate Academic Achievement Standards (current use under ESSA)
AAC	augmentative and alternative communication
AERA	American Educational Research Association
ANOVA	analysis of variance
APA	American Psychological Association
APIP	Accessible Portable Item Protocol
BIE	Bureau of Indian Education
CBT	Computer-based test
CCC	Core Content Connector
CCSS	Common Core State Standards
CSEM	conditional standard error of measurement
CTT	Classical Test Theory
DAC	decision accuracy and consistency
DETECT	Dimensionality Evaluation to Enumerate Contributing Traits
DIF	differential Item functioning
DIMTEST	computer program used by Cognia
DNU	do not use
DTA	Directions for Test Administration
ELA	English language arts
EOTS	end-of-test survey
ESEA	Elementary and Secondary Education Act
ESR	early stopping rule
ESSA	Every Student Succeeds Act
EU	essential understanding
FKSA	focal knowledge, skills, and ability
GM	geometry (in standards)
GRM	graded-response model
HOSS	highest obtainable scale score
ICC	item characteristic curve
ICCC	item category characteristic curve
ICTC	item category threshold curve
IEP	individualized education program
IIF	Item information function
IRC	Item Review Committee (includes Content Review Committee and Bias-Sensitivity Review Committee)
IRT	Item Response Theory

Acronym	Explanation
IT	information technology
KSA	knowledge, skills, and ability
LEP	limited English proficiency
LOSS	lowest obtainable scale score
LPF	Learning Progression Framework
MSAA	Multi-State Alternate Assessment
NCME	National Council on Measurement in Education
PARSCALE	Item response theory (IRT) software program that can perform item analysis and test scoring for dichotomous and polytomous IRT models
PBT	Paper-based test
PLAAFP	present level of academic achievement and functional performance?
PLD	performance level descriptor
R9-stringer	student who responds to nine (or more) consecutive multiple-choice items with the exact same option
SD	standard deviation
SEM	standard error of measurement
SIU	score interpretations and uses
SQA	Software Quality Assurance
SRC	student response check
STL	scoring team leader
TA	test administrators
TAC	Technical Advisory Committee
TAM	Test Administration Manual
TC	test coordinators
TCC	test characteristic curve
TIF	test information function
UWC	use with caution
WRCC	Writing Common—a code within iScore that is used to identify items for scoring
NCSC	National Center and State Collaborative

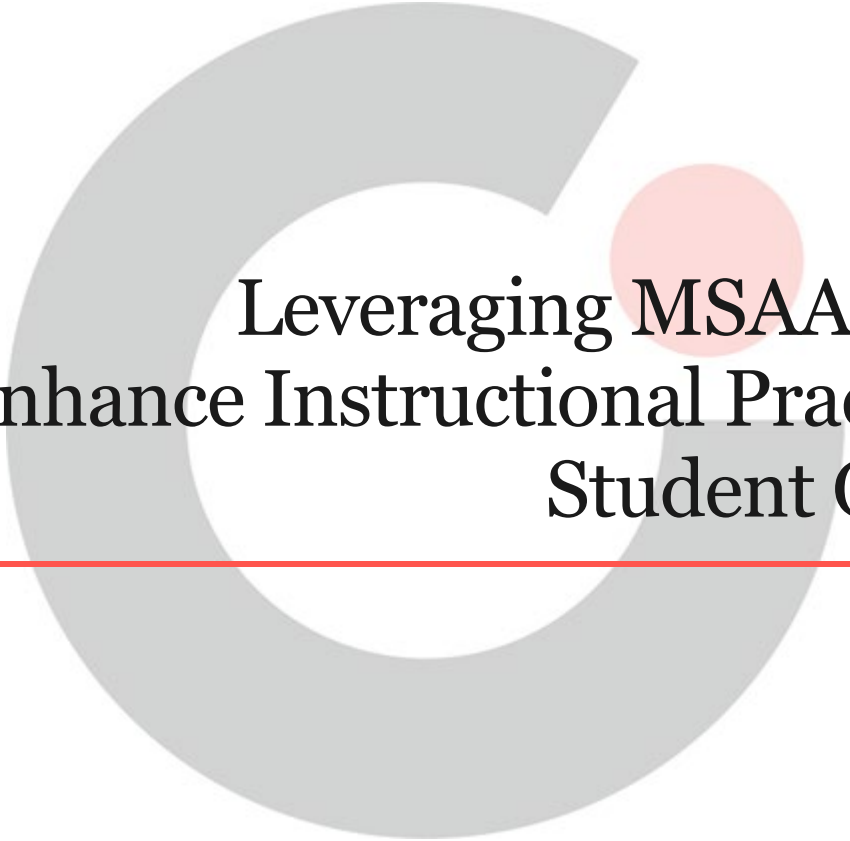
Commonly Used Terms

Term	Explanation
<i>a</i>	The item response theory index of item discrimination, analogous to the point-biserial and biserial correlations in classical test theory. It reflects the slope of the item response function. Often ranging from 0.1 to 2.0 in practice, a higher value indicates a better-performing item.
Accreditation	Accreditation by an outside agency affirms that an organization has met a certain level of standards. Certification testing programs may become accredited by meeting specified standards in test development, psychometrics, bylaws, management, etc.
<i>b</i>	The item response theory index of item difficulty or location, analogous to the P-value (P+) of classical test theory. Typically ranging from -3.0 to 3.0 in practice, a higher value indicates a more difficult item.
Biserial Correlation	A classical index of item discrimination, highly similar to the more commonly used point-biserial. The biserial correlation assumes that the item scores and test scores reflect an underlying normal distribution, which is not always the case.

Term	Explanation
Blueprint	A test blueprint, or test specification, details how an exam is to be constructed. It includes valuable information, such as the total number of items, the number of items in each content area or domain, the number of items that are recall vs. reasoning, and the item formats to be utilized.
c	The item response theory pseudo-guessing parameter, representing the lower asymptote of the item response function. It is theoretically near the value of $1/k$, where k is the number of alternatives. For example, with the typical four-option multiple choice item, a candidate has a base chance of 25% of guessing the correct answer.
Certification	A non-mandatory testing program which certifies that candidates have achieved a minimum standard or knowledge or performance.
Classical Test Theory	A psychometric analysis and test development paradigm based on correlations, proportions, and other statistics that are relatively simple compared to IRT. It is therefore more appropriate for smaller samples, especially for fewer than 100.
Classification	The use of tests for classifying candidates into categories, such as pass/fail, non-master/master, or basic/proficient/advanced.
Computerized Adaptive Testing	A dynamic method of test administration where items are selected one at a time to match item difficulty and candidate ability as closely as possible. This helps prevent candidates being presented with items that are too difficult or too easy for them, which has multiple benefits. Often, the test only takes half as many items to obtain a similar level of accuracy to form-based tests. This reduces the testing time per examinee and reduces the total number of times an item is exposed, as well as increasing security by the fact that nearly every candidate will receive a distinct set of items.
Cutscore	Also known as a passing score, the cutscore is the score that a candidate must achieve to obtain a certain classification, such as “pass” on a licensure or certification exam.
Criterion-Referenced	A test score (not a test) is criterion-referenced if it is interpreted regarding a specified criterion and not compared to scores of other candidates. For instance, providing the number-correct score does not relate any information regarding a candidate’s relative standing.
Distractors	Distractors are the incorrect options of a multiple-choice item. A distractor analysis is an important part of psychometric review, as it helps determine if one is acting as a keyed response.
Equating	The process of determining comparable scores on different forms of an examination. For example, if Form A is more difficult than Form B, it might be desirable to adjust scores on Form A upward for the purposes of comparing them to scores on Form B. Usually, this is done statistically based on items that are on both forms, which are called equater, anchor, or common items. Since the groups who took the two forms are different, this is called a common items non-equivalent groups design.
Form	A specific set of items that are administered together for a test. For example, if a test included a certain set of 100 items this year, and a distinct set of 100 items next year, these would be two distinct forms.
Item	The basic component of a test, often colloquially referred to as a “question,” but items are not necessarily phrased as a question. They can be as varied as true/false statements, rating scales, and performance task simulations, in addition to the ubiquitous multiple-choice item.
Item Bank	A repository of items for a testing program, including items at all stages, such as newly written, reviewed, pretested, active, and retired.
Item Difficulty	A statistical index of how easy/hard the item is with respect to the underlying ability/trait. That is, an item is difficult if not many people get it correct or respond in the keyed direction.
Item Discrimination	A statistical index of the quality of the item, assessing how well it differentiates examinees of high vs. low ability. Items with low discrimination are considered inadequate quality and are candidates to be revised or retired.

Term	Explanation
Item Response Theory (IRT)	A comprehensive approach to psychometric analysis and test development that utilizes complex mathematical models. This provides several benefits, including the ability to design CATs, but requires larger sample sizes. A common rule of thumb is 100 candidates for the one-parameter model and 500 for the three-parameter model.
Key	The key is the correct response to an item.
Knowledge, Skills, and Abilities (KSAs)	A critical step in testing for either employment or professional credentials is to determine the KSAs that are important in a job. This is often done via a job analysis study.
Licensure	A testing program mandated by a government body. The test must be passed to perform the task in question, whether it is to work in the profession or drive a car.
Norm-Referenced	A test score (not a test) is norm-referenced if it is interpreted regarding the performance of other candidates. Percentile rank is an example of this, because it does not provide any information regarding how many items the candidate got correct.
P-value	A classical index of item difficulty, presented as the proportion of candidates who correctly responded to the item. A value above 0.90 indicates an easy item, while a value below 0.50 indicates a relatively difficult item. Note that it is inverted; a higher value indicates less difficulty.
Point-Biserial Correlation	A classical index of item discrimination, calculated as the Pearson correlation between the item score and the total test score. If below 0.0, low-scoring candidates are doing better than high-scoring candidates, and the item should be revised or retired. Low positive values are marginal, higher positive values are ideal.
Field-test Item	An item that is administered to candidates simply for the purpose of obtaining data for future psychometric analysis. The results on this item are not included in the score. It is often prudent to include a small number of pretest items in a test.
Reliability	A measure of the repeatability or consistency of the measurement process. Often, this is indexed by a single number, most commonly the internal consistency index coefficient alpha or its dichotomous formulation, KR-20. Under most conditions, these range from 0.0 to 1.0, with 1.0 being a perfectly reliable measurement. However, just because a test is reliable does not mean that it is valid, i.e., measures what it is supposed to measure.
Scaling	A process of converting scores obtained on an exam to an arbitrary scale. This is done so that all the forms and exams used by a testing organization are on a common scale. For example, suppose an organization had two testing programs, one with 50 items and one with 150 items. All scores could be put on the same scale to standardize score reporting.
Standard-Setting	A formal study conducted by a testing organization to determine standards for a testing program, which are manifested as a cutscore. Common methods include the Angoff, Bookmark, Contrasting Groups, and Borderline Survey methods.
Subject Matter Expert	An extremely vital person in the test development process. SMEs are necessary to write items, review items, participate in standard-setting studies and job analyses, and oversee the testing program to ensure its fidelity to its true intent.
Validity	The concept that test scores can be interpreted as intended. For example, a test for certification in a profession should reflect basic knowledge of that profession, and not intelligence or other constructs, and scores can therefore be interpreted as evidencing professional competence. Validity must be formally established and maintained by empirical studies as well as sound psychometric and test development practices.

APPENDIX P
MSAA VALIDITY SURVEY



Leveraging MSAA Scores to Enhance Instructional Practices and Student Outcomes

Survey Results

8/7/2024

**Conducted by Cognia on Behalf of the
Multi-State Alternate Assessment (MSAA)**

Leveraging MSAA Scores to Enhance Instructional Practices and Student Outcomes

Executive Summary

Conducted by Cognia on behalf of MSAA, this survey aimed to evaluate MSAA score utilization and identify areas for improvement. From August to December 2023, 176 stakeholders in MSAA states/entities responded, providing insights into current practices and challenges. Key findings reveal a strong focus on using MSAA scores for individual instructional practices, but a gap exists in monitoring broader trends. Challenges include relevance concerns and limited utility perceptions. The survey underscores the importance of professional development in enhancing MSAA score interpretation. Recommendations include providing diverse resources and addressing specific needs.

Introduction

The MSAA (Multi-State Alternate Assessment) aims to empower teachers by providing them with scores to inform their instruction. To assess the effectiveness of score utilization, a survey was conducted targeting stakeholders in MSAA states/entities. The survey aimed to gauge the utilization of scores and whether teachers received training in interpreting and using these scores. This summary presents the methodology, results, and discussions derived from the survey.

Method

The survey was created and administered using Survey Monkey, with two rounds of revisions during the pilot and operational phases. The survey, conducted between August and December 2023, collected 176 responses from three states—Arizona, Montana, and South Dakota—and one entity, American Samoa. The survey sought feedback on various topics, including whether partner states have been presenting MSAA results, how they utilize these results, and whether they have received professional development on score interpretation.

Results

Demographics Findings

The survey's first section featured six demographic questions to grasp respondents' backgrounds. While the MSAA serves various states and entities, our respondents represent only a portion of this community. Despite limited demographic representation, their input provides valuable insights into MSAA utilization. Key findings from demographic questions include:

- Geographical settings predominantly comprise rural areas with small districts, many having fewer than 20 schools.
- Respondents possess extensive experience working with students with cognitive disabilities, especially in elementary and middle schools.

- Special education directors or coordinators were the largest respondent group represented.

The table below details the demographic questions and corresponding answers from the survey:

What is the size of your school district?	N	Percent
1-20 schools	143	83%
21-100 schools	21	12%
More than 100 schools	9	5%
What geographical setting is your school's district located in?	N	Percent
Urban	35	22%
Rural	86	53%
Suburban	24	15%
Charter	17	10%
How many students are eligible to participate in the MSAA?	N	Percent
Fewer than 10 students	106	62%
10-20 students	30	18%
More than 20 students	35	20%

How many years have you worked with students with significant cognitive disabilities?	N	Percent
1-2 years	16	9%
3-5 years	31	18%
6-10 years	30	17%
More than 10 years	98	56%
In what grades are the students you serve? (Select all that apply)	N	Percent
Grades 3-5	129	37%
Grades 6-8	124	36%
Grades 9-12	96	28%
What is your primary role in your district, school, or program?	N	Percent
Special education director or coordinator	76	46%
Special education teacher	56	34%
District testing coordinator	26	16%
School testing coordinator	7	4%

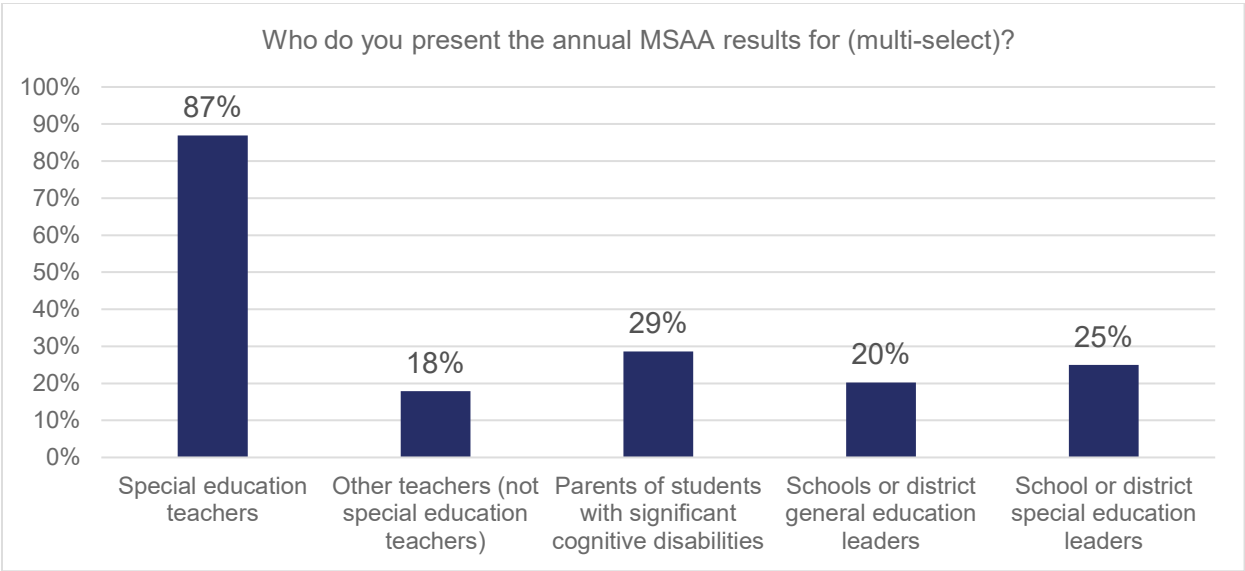
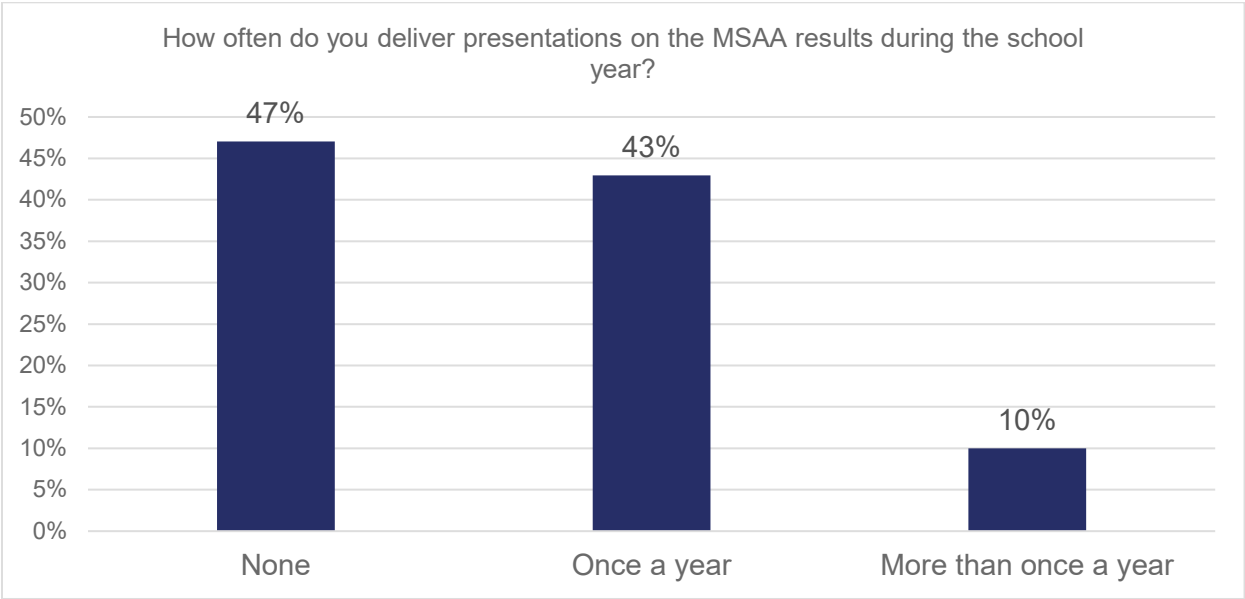
Presenting MSAA Results

The survey found that most respondents consistently present MSAA results at least once a year, predominantly focusing on special education teachers. These presentations are crucial for raising awareness and facilitating data-driven decision-making. Specifically, presentations serve several purposes, that include the following:

- Inform stakeholders, instruction planning, and provide professional development.
- Use during IEP meetings to discuss student progress and needs.
- Track student achievement, analyze data, and adjust instructional practices.
- Celebrate successes, share growth, and review guidelines and procedures.

It is our belief that addressing the infrequent or lack of effective presentations in schools and districts that need additional support, offers an opportunity to improve communication gaps between MSAA partner states/entities, and strengthen the collaboration and support within the MSAA community.

The charts below detail the questions regarding the presentation of MSAA results and corresponding answers from the survey. For open-ended questions, thematic analysis has been conducted to categorize each response into different themes, with the number of occurrences.



- What is the purpose of your MSAA results presentation (open-ended)?
 - Informing Stakeholders and Teachers (N=21)
 - Instructional Planning and Goal Setting (N=13)
 - Tracking Student Progress and Adjusting Instruction(N=13)
 - Celebrating Successes and Sharing Growth (N=4).
 - IEP meetings (N=3)

- Professional development to train teachers on assessment administration, interpretation, and instructional implications (N=3).

Using MSAA Results

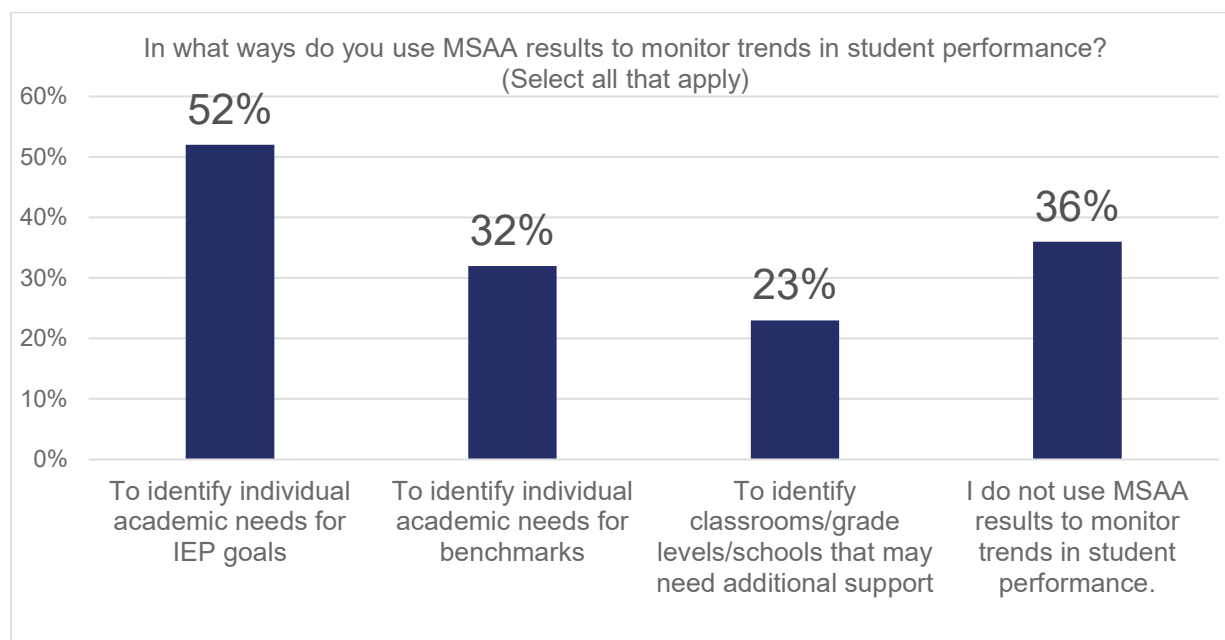
The survey findings indicate a strong utilization of MSAA results for informing instructional practices at the individual student level, however, there's a notable gap in utilizing these results to monitor broader trends in student group performance. Additionally, respondents highlighted various challenges and barriers to using MSAA results effectively, including:

- Concerns about validity and test administration issues.
- Perceived limited utility of results.
- Desire for alternative assessments and comprehensive data analysis.

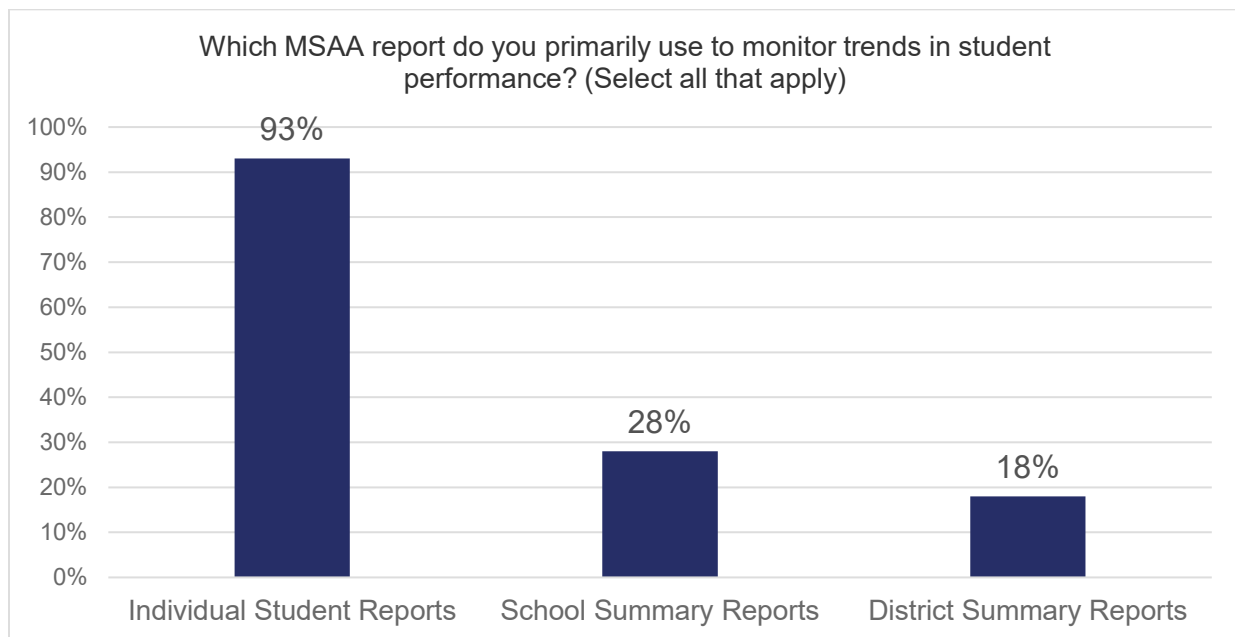
Respondents also emphasized the importance of various types of reports, such as Individual Student Reports, for tasks like IEP development, progress monitoring, and data-informed decision-making. However, some indicated limited benefits in using these reports. Suggestions for improving MSAA score utilization include:

- Enhancing test relevance.
- Providing more training and support for educators.
- Developing alternative assessments.
- Better data analysis tools and detailed reporting for trend monitoring and decision-making.

The charts below detail the questions regarding the utilization of MSAA results and corresponding answers from the survey. For open-ended questions, thematic analysis has been conducted to categorize each response into different themes, with the number of occurrences.



- What are the challenges and barriers to the use of MSAA results to monitor trends (open-ended)?
 - Enhance Validity and Relevance (N=28): Tailor questions to match the cognitive levels of students.
 - Improve Test Administration (N= 15): Streamline result access, shorter results turnaround time.
 - Limited Utility and Impact (N= 11): Increase support in utilizing MSAA effectively in goal-setting, progress monitoring, and instruction.
 - Explore Alternative Assessments (N= 9): Consider other assessments to better align with diverse student needs.
 - Data Analysis and Reporting Needs (N= 8): More detailed reports, growth data, and district-wide analyses to better understand student performance and trends.
- Please give us an example of how you use any of the three reports (open-ended).
 - IEP Development and Goal Setting (N=31)
 - Monitoring Progress and Growth (N= 30)
 - Comparing Results and Identifying Trends (N= 21)
 - Data-informed decision-making (N= 19)
 - Informing Parents and Stakeholders (N= 17)
 - Identifying Student Needs and Strengths (N= 15)

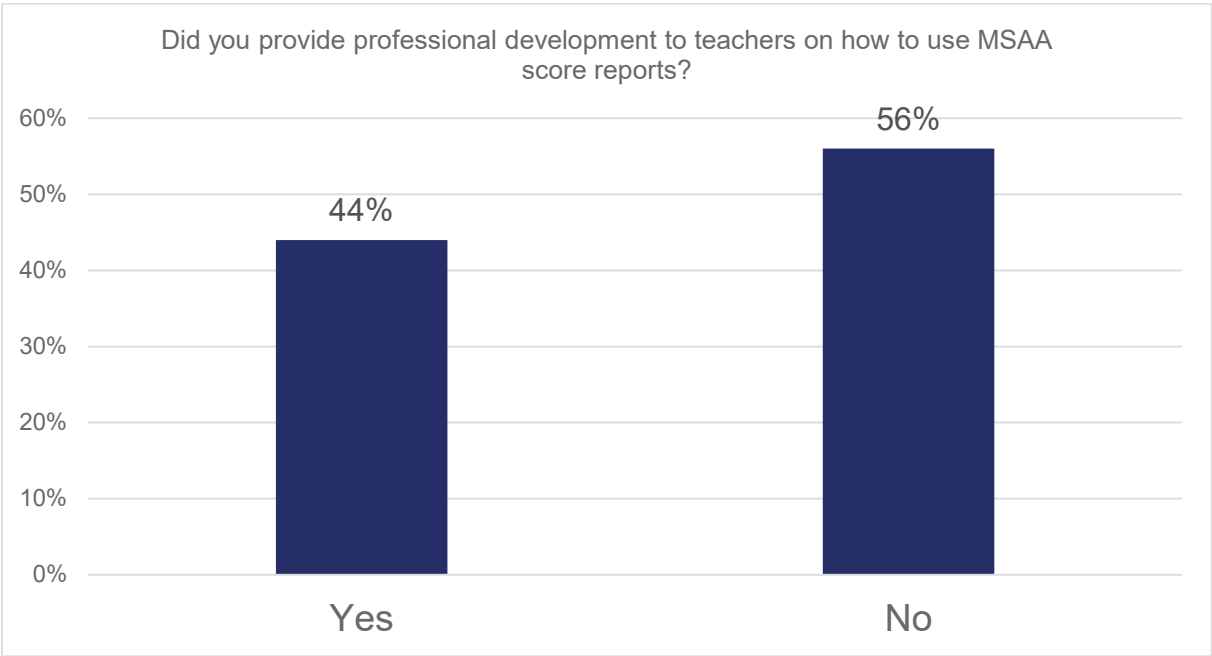


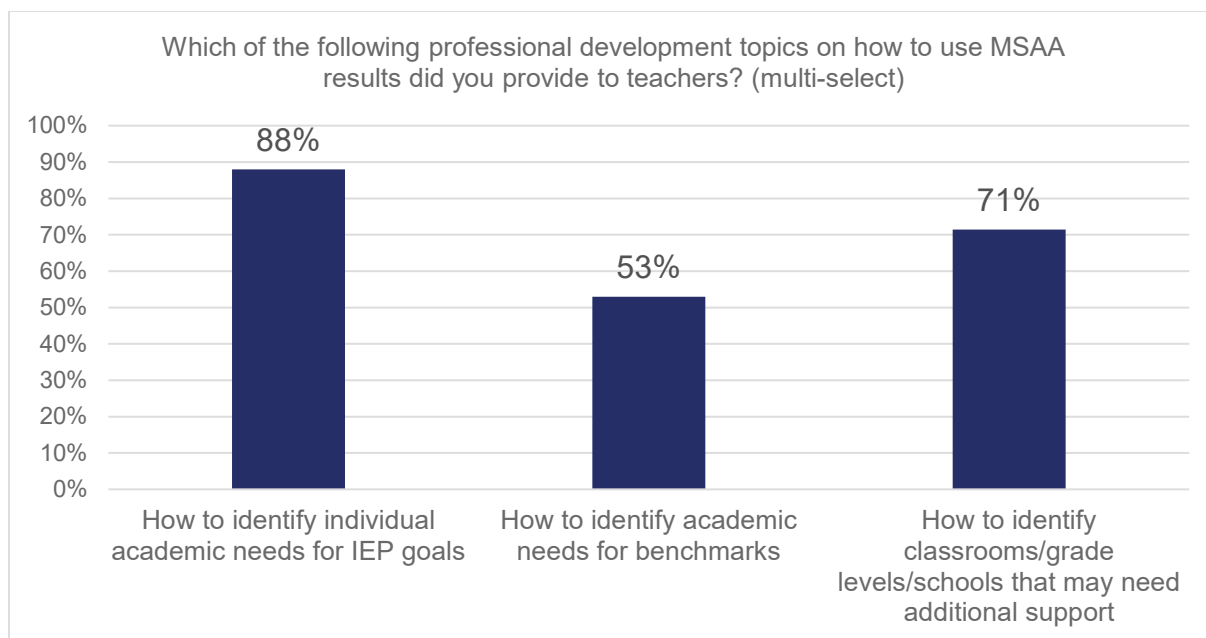
Providing Professional Development

The survey highlights the proactive efforts in providing professional development (PD) on MSAA score interpretation, with around 44% of respondents receiving such support. However, there's a suggestion to broaden the focus beyond individual academic needs for IEP goals. Respondents outlined several supports needed to enhance the accessibility of PD programs. These include:

- Requests for diverse PD resources like webinars and videos.
- Desires for easily accessible opportunities and guidance on MSAA result interpretation.
- Concerns about time constraints and resource availability.
- Support for individuals new to their roles.

The following charts and questions summarize the data collected from the survey, regarding Providing PDs on MSAA score interpretation. For open-ended questions, thematic analysis has been conducted to categorize each response into different themes, with the number of occurrences.





- What kind of support do you need to make designing PD more accessible (open-ended)?
 - Training Needs and Resources: Webinars, videos, pre-made presentations, and learning modules (N=36)
 - Accessibility and Availability: Online options, and asynchronous formats. (N= 14)
 - Interpretation and Implementation Guidance: Connections to instruction and using the data to support student needs. (N= 13)
 - Time and Resources (N= 7)
 - Support for New Roles: individuals new to their roles, such as directors or coordinators, who may be unfamiliar with MSAA. (N= 6)

Conclusion

In conclusion, the survey findings provide valuable insights into the utilization of MSAA scores, and the associated challenges faced by educators. The results underscore the importance of regular presentations to disseminate assessment data effectively, particularly among special education teachers who play a pivotal role in individualized instruction. While there is a strong indication in utilizing MSAA results at the individual level, there's room for improvement in extending this utilization to broader trends at student groups.

Challenges such as concerns about validity, issues with test administration, and perceptions of limited score report utility highlight the complexities in effectively using existing assessment data. Additionally, the provision of professional development on MSAA score interpretation reflects efforts to support educators, but there's a need to broaden the focus beyond the use of IEP identification. Furthermore, collaboration between stakeholders and state departments is crucial in addressing the identified barriers and enhancing the effectiveness of MSAA score utilization, to inform instructional practices and improve student outcomes.