Multi-State Alternate Assessment

2017–18 Technical Report

100 EDUCATION WAY, DOVER, NH 03820 (800) 431-8901
WWW.MEASUREDPROGRESS.ORG

TABLE OF CONTENTS

CHAPTER 1 OVERVIEW OF MSAA AND 2018 UPDATES 1
1.1 PURPOSES AND USES OF THE MSAA 1
1.2 INTENDED MSAA SCORE INTERPRETATIONS AND USES 1
1.3 VALIDITY ARGUMENTS FOR THE MSAA 2
1.4 UPDATES FOR THE 2018 ADMINISTRATION 3
CHAPTER 2 OVERVIEW OF THE MSAA 5
2.1 HISTORY OF THE MSAA 5
2.1.1 Core Beliefs 6
2.1.2 Stakeholders 7
2.2 MSAA PARTICIPATION 8
CHAPTER 3 TEST CONTENT 9
3.1 HISTORY OF ALTERNATE ACADEMIC ACHIEVEMENT STANDARDS AND CORE CONTENT CONNECTORS 9
3.1.1 The Learning Progression Frameworks 9
3.1.2 Core Content Connectors 10
3.2 ALIGNMENT AND LINKAGES 12
3.3 2018 MSAA ASSESSMENT DESIGN 13
3.3.1 Operational Design 13
3.3.2 Operational Items and Embedded Field-Test Items 16
3.3.3 Item Design and Administration 17
3.3.4 Item Components 20
3.3.4.1 Selected-Response: Reading, Writing (stand-alone items and multiple-part selected response writing prompt), Mathematics 20
3.3.4.2 Constructed-Response: Mathematics 20
3.3.4.3 Open-Response: Writing Prompt 20
3.4 CONTENT AND BLUEPRINTS 21
3.4.1 Mathematics 21
3.4.2 English Language Arts 22
CHAPTER 4 TEST DEVELOPMENT 24
4.1 GENERAL PHILOSOPHY AND ROLE OF ITEM DEVELOPMENT SUBCOMMITTEE IN TEST DEVELOPMENT 24
CHAPTER 5 TRAINING AND ADMINISTRATION 27
5.1 TEST ADMINISTRATOR AND TEST COORDINATOR TRAINING 27
5.2 TEST ADMINISTRATOR TRAINING MODULES 27
5.3 TEST COORDINATOR TRAINING MODULES 28
5.4 BEST PRACTICE VIDEOS 29
5.5 TEST ADMINISTRATION MANUAL 29
5.6 DIRECTIONS FOR TEST ADMINISTRATION (DTA) 30
5.7 TEST COORDINATOR AND TEST ADMINISTRATOR USER GUIDES 31

5.8 OPERATIONAL ADMINISTRATION ... 31
5.8.1 MSAA Service Center ... 31
5.8.2 Additional Supports ... 32
5.8.3 Monitoring and Quality Control ... 32
5.8.4 Operational Test Survey Results ... 33
CHAPTER 6 SCORING ... 36
6.1 SELECTED-RESPONSE AND CONSTRUCTED-RESPONSE ITEM SCORING PROCESSES ... 36
6.1.1 Overview of Scoring Process within the System and Test Administrator/Scorer Training ... 36
6.2 OPEN-RESPONSE WRITING PROMPTS SCORING PROCESSES ... 37
6.2.1 Overview of Open-Response Writing Entry Process within the Assessment System and Test Administrator Training ... 37
6.2.2 Benchmarking and Identification of Scoring ... 38
6.2.3 Scorer Recruitment and Qualifications ... 38
6.2.4 Measured Progress Staff and Scoring Leadership ... 39
6.2.5 Training ... 39
6.2.6 Qualification ... 41
6.2.7 Methodology for Scoring Operational Open-Response Writing Prompts ... 42
6.2.8 Monitoring of Scoring Quality Control ... 45
6.2.8.1 Calibration Sets ... 46
6.2.8.2 Read-behind Scoring ... 46
6.2.8.3 Double-blind Scoring ... 47
6.2.9 Quality and Production Management Reports ... 48
6.2.10 Interrater Agreement ... 49
CHAPTER 7 REPORTING ... 51
7.1 DEVELOPMENT AND APPROVAL OF REPORT SPECIFIC DOCUMENTS ... 51
7.2 SPECIFIC PRIMARY REPORTS GENERATED FOR SCHOOLS, DISTRICTS, AND STATES ... 52
7.2.1 Student Report ... 53
7.2.2 School Roster Report ... 54
7.2.3 Summary Reports ... 55
7.2.4 Quality Assurance ... 55
CHAPTER 8 CLASSICAL ITEM ANALYSIS ... 57
8.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES ... 57
8.2 DIFFERENTIAL ITEM FUNCTIONING ... 62
8.3 DIMENSIONALITY ANALYSIS ... 64
CHAPTER 9 ITEM RESPONSE THEORY SCALING AND EQUATING ... 71
9.1 ITEM RESPONSE THEORY ... 71
9.2 CALIBRATION PROCEDURE ... 73
9.3 ITEM RESPONSE THEORY RESULTS ... 75
9.4 EQUATING ... 82
9.5 MSAA PERFORMANCE LEVELS, CUT SCORES, AND STANDARDS VALIDATION ... 84
9.6 REPORTED SCALE SCORES ... 88
CHAPTER 10 RELIABILITY ... 90
10.1 IRT MARGINAL RELIABILITY ... 90

10.2 SUBGROUP RELIABILITY 92
10.3 RELIABILITY OF PERFORMANCE-LEVEL CATEGORIZATION 93
10.3.1 Accuracy and Consistency 94
CHAPTER 11 VALIDITY 96
11.1 PRIMARY INTENDED SCORE INTERPRETATION 96
11.1.1 Primary Intended Score Use 1 103
11.1.2 Primary Intended Score Use 2 104
11.1.3 Primary Intended Score Use 3 105
REFERENCES 107
APPENDICES 111
APPENDIX A LIST OF ACRONYMNS
APPENDIX B ACCOMMODATION FREQUENCIES
APPENDIX C PARTICIPATION RATES
APPENDIX D TEST BLUEPRINTS
APPENDIX E ITEM REVIEW, BIAS/SENSITIVITY REVIEW, AND TECHNICAL ADVISORY COMMITTEE MEMBERS
APPENDIX F IT-REPORTING PROCESSING AND REPORTING BUSINESS REQUIREMENTS DOCUMENTS
APPENDIX G MSAA 2018 GUIDE FOR SCORE REPORT INTERPRETATION
APPENDIX H MSAA QUALIFICATION RATES
APPENDIX I ITEM-LEVEL CLASSICAL STATISTICS
APPENDIX J DIFFERENTIAL ITEM FUNCTIONING RESULTS
APPENDIX K ITEM RESPONSE THEORY PARAMETERS
APPENDIX L TEST CHARACTERISTIC CURVES & TEST INFORMATION FUNCTIONS
APPENDIX M STANDARDS VALIDATION REPORT
APPENDIX N RAW TO SCALED SCORE LOOKUP TABLES
APPENDIX O SCORE DISTRIBUTIONS
APPENDIX P CLASSICAL RELIABILITY
APPENDIX Q DECISION ACCURACY AND CONSISTENCY RESULTS

CHAPTER 1 OVERVIEW OF MSAA AND 2018 UPDATES

1.1 PURPOSES AND USES OF THE MSAA

The Multi-State Alternate Assessment (the MSAA)[1] is a comprehensive, summative assessment system designed to promote increasingly higher academic outcomes for students with the most significant disabilities in preparation for a broader array of post-secondary outcomes. The MSAA is designed to measure grade level academic content that is aligned with, and derived from, MSAA Partner States' content standards. This test 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 they are able. The MSAA is administered in the areas of ELA and mathematics in grades 3–8 and 11.

The MSAA was developed to ensure that all students with the most significant cognitive disabilities are able to participate in a summative assessment that is a measure of what they know and can do in relation to grade-level state content standards. The MSAA is a component of a system of curriculum, instruction, and professional development that allows students with the most significant cognitive disabilities to access grade-level content aligned with grade-level state content standards.

The MSAA Partner States' long-term goal is to ensure that students with the most significant cognitive disabilities achieve increasingly higher academic outcomes and leave high school with the potential to pursue productive post-secondary options. A well-designed summative assessment alone is insufficient to achieve this goal.

The MSAA is an alternate assessment aligned with alternate academic achievement standards (AA-AAAS) as described in the Elementary and Secondary Education Act (ESEA). This law mandates that all students participate in assessments that measure student achievement of grade-level content standards. In order to ensure that MSAA measures student achievement of the Core Content Connectors, this technical report provides the standard psychometric data and descriptions of technical procedures found in all state assessment technical reports. In addition, this MSAA technical report identifies four primary intended interpretations and uses of MSAA scores and cites the assumptions, elements, and evidence that are relevant to those interpretations and uses.

1.2 INTENDED MSAA SCORE INTERPRETATIONS AND USES

MSAA is designed, developed, and implemented to support four primary intended score interpretations and uses.

[1] See Appendix A for a list of acronyms.

Primary Intended MSAA Score *Interpretation*

MSAA scores provide trustworthy information about important knowledge and skills in grade-level numeracy and literacy that students with the most significant cognitive disabilities are attaining.

Primary Intended MSAA Score *Uses*

1. The MSAA and its results can be used to help schools and districts (a) monitor trends in student performance, and (b) design professional development for teachers.
2. Help teachers to better integrate assessment with their instructional planning.
3. Give parents information about (a) what their child knows and can do, and (b) their child's progress from year to year.

The intended score interpretation and uses stated here align with the original statements of intended score interpretations and uses in the 2015 NCSC technical report (see the "claim" and "uses" statements on page 8).

Chapter 11, *Validity Arguments*, contains assumptions, elements, and evidence that support the intended interpretations and uses of MSAA scores.

1.3 VALIDITY ARGUMENTS FOR THE MSAA

The 2018 technical report describes several psychometric and procedural aspects of the MSAA, contributing to the accumulation of validity evidence to support MSAA score interpretations and uses. Because the interpretations and uses of test scores, rather than the test itself, are evaluated for validity, this report presents documentation to substantiate intended interpretations and uses (AERA, APA, & NCME 2014). Each section in this report contributes important information about MSAA: test development, test alignment, test administration, scoring, reliability, performance levels, and reporting. The information to support validity arguments for intended MSAA score interpretations and uses, summarized in chapter 11, are formed as assumptions that underlie each interpretation and use elements of each of those assumptions, and specific evidence that supports each element of the assumptions.

The phrase "**intended score interpretations for uses**" appears several times in the *Standards for Educational Psychological Testing* and is the core of the field's views on validity and validation. For the MSAA and other assessment programs, the phrase refers broadly to **test scores** (e.g., total 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* (AERA et al., 2014) also gives a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five general areas: test content, response processes, internal structure, relationship to

other variables, and consequences of testing. Although each of these sources may address different aspects of validity, they are not distinct types of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations and uses. In addition, they 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 overall validity arguments for intended interpretations and uses of MSAA scores and other information.

1.4 UPDATES FOR THE 2018 ADMINISTRATION

For the 2018 assessment, the MSAA Partner States comprised Arizona, Arkansas, Maine, Maryland, Montana, the Pacific Assessment Consortium (PAC-6: Guam and the Commonwealth of Northern Mariana Islands [CNMI]), South Dakota, Tennessee, the U.S. Virgin Islands (USVI), and Washington, D.C. USVI received a federal Hardship Waiver for the 2017–18 assessment; therefore no USVI students took the assessment and are not included in any of the analyses contained in this report. USVI continued as a member state in MSAA.

The 2017–18 MSAA utilized a two-stage adaptive design, meaning that student performance in the first stage of the assessment determined the choice of the second stage to be administered to the student. This was the second year of the stage adaptive design administration.

An open-response writing prompt was administered operationally for the first time in 2018. Student scores from the writing prompt were linked to the existing ELA scale to expand its interpretation. In conjunction with this expansion, the ELA cut scores, which were established in 2015, were articulated vertically to enhance interpretation of scores using the expanded PLDs. (The ELA PLDs were expanded to account for inclusion of scores from the open response writing prompt.)

The mathematics cut scores, which were established in 2015, also were articulated vertically to enhance interpretation of scores using the existing PLDs. (In contrast to ELA, the mathematics PLDs were not expanded or otherwise edited.) Details on the procedures and results from the ELA and mathematics standards articulation and validation processes appear in section 9.5.

Reading foundational items, as field-test, and Tier 1 writing (multiple-part selected-response items), as operational, were reintroduced to the MSAA after additional research and/or development of these item types last year. The grade 3 and 4 English language arts (ELA) foundational reading items were revised to better align to the standards. In the fall of 2016, the Item Development Subcommittee reevaluated the assessment of the reading foundational content category. The subcommittee decided to include this content category in the summative assessment, but to assess the content category using the anchor standard of fluency (RF 4) rather than phonics and word analysis (RF 3). Based on these decisions, item specifications were developed and approved by the Item Development Subcommittee. Using these item specifications, new items were developed and reviewed by the Item Development Subcommittee members as well as the Item Review Subcommittee.

The scoring model for the Tier 1 writing prompts (multiple-part selected-response items) was reviewed and it was determined that each item would be worth 0/1 point. Additionally, as indicated above, Tier 2 and Tier 3 open-response writing prompts were administered operationally for the first time in 2017–18. The balance of the MSAA operational items remained fundamentally the same as the 2016–17 operational assessment.

The 2017–18 assessment included field-test items in both mathematics and ELA (reading and writing) with differing levels of complexity. There were three forms per grade for each content area. Detailed information is covered in Chapter 3.

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 2018 Guide for Score Report Interpretation Guide*. The TAM, *MSAA Online Assessment System User Guide for Test Administrators*, *MSAA Online Assessment System User Guide for Test Coordinators*, DTA, and online training modules were revised in order to streamline information and provide more clarity to test administrators (TAs) and test coordinators (TCs). Additional detailed information is available in Chapter 5.

CHAPTER 2 OVERVIEW OF THE MSAA

The MSAA assesses ELA and mathematics at grades 3–8 and 11 and is aligned with the States' 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 distinct levels of complexity, representing different levels of skill and knowledge acquisition by students.

Students with the most significant cognitive disabilities often need materials and instructional strategies that are substantially adapted and scaffolded, which have built-in supports 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 test items are developed with many scaffolds and supports embedded within the items. Students are provided additional support based on their individual requirements, including other allowable ways for test administrators (TAs) to present each item.

This assessment is designed to be administered one-on-one, delivered in an online format or via a paper-pencil/hybrid format as an accommodation if appropriate. The needs of the student are also addressed through other supports, such as assessment features built into the platform, accommodations such as using assistive technology, a scribe, and/or sign language. Appendix B contains the 2017–18 summary of accommodation usage frequencies for the MSAA. 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 32–35 operational items, primarily selected-response with some constructed-response items. The operational writing items of the ELA test consist of selected-response items, along with a multiple-part selected-response writing prompt and an open-response writing prompt at each grade level. Each content area assessment is accomplished in two "test sessions." There are also embedded field-test items in Session 1 for each grade and content area.

2.1 HISTORY OF THE MSAA

Work leading up to the MSAA began in late 2010, when the National Center and State Collaborative (NCSC) began development of the NCSC Alternate Assessment, which was designed to meet the requirements of the Elementary and Secondary Education Act and the Individuals with Disabilities Education Act and is based on Alternate Achievement Standards (AA-AAS) for students with the most significant cognitive disabilities. This work culminated in the operationalized NCSC assessment in spring 2015. The work of NCSC ended following the spring 2015 administration. For additional information about the NCSC

assessment, please refer to the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* (see References for URL) or contact the MSAA Partner States at MSAA@azed.gov. The MSAA Partner States continued the work of NCSC following many of the same principles, purposes, and core beliefs. The first administration of MSAA was in the spring of 2016.

2.1.1 Core Beliefs

The core beliefs about the MSAA began with NCSC and were laid out in the prior planning and development of the AA-AAS. As recorded in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual*, as states and organizational partners implemented the NCSC development plan, they found they had to come to a consensus on topics that were a mix of practice and theory in the comprehensive context of teaching and learning for the students. 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.

The MSAA Partner States, as their NCSC counterparts before, believe that accessibility is central to the validity argument of the assessment, and that accessibility to the academic content based on college- and career-ready academic standards begins with rigorous curriculum and instruction resources and training for teachers. 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 on communicative competence to ensure all students have a way first to learn the concepts, and then to show what they know and can do on the assessment. The NCSC Theory of Action and Validity Approach, available at ncscpartners.org/Media/Default/PDFs/Resources/NCSCBrief9.pdf, 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 community.

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) or by contacting the MSAA Partner States at MSAA@azed.gov.

2.1.2 Stakeholders

Many stakeholders are involved in the development of the 2018 MSAA. MSAA State Leads are key representatives from each Partner State who together compose the decision-making body for the MSAA. Members of this body participate in various subcommittees that focus on specific aspects of the assessment and have decision-making authority on behalf of the MSAA Partner States for each subcommittee's focal area.

The MSAA Manuals, User Guides, and Training Subcommittee that oversees development of the *Test Administration Manual* (TAM), *MSAA Online Assessment System User Guide for Test Administrators, MSAA Online Assessment System User Guide for Coordinators*, online training modules, and best practice videos consists of representatives from Arizona, Maryland, Montana, PAC-6, South Dakota, Tennessee, and Washington, D.C.

The Psychometric and Test Construction Subcommittee assists in planning Technical Advisory Committee (TAC) meetings, contributes to psychometric decisions, reviews item performance statistics for each field-tested item, approves the test design, provides decisions and approvals related to the core item constructed sets, makes decisions throughout the standards validation process, provides the content of the End-of-Test Survey, determines relevant policies, receives the survey results after administration, and advises on the structure of the technical report. This subcommittee is composed of members from Arizona, Maine, Maryland, Montana, and South Dakota.

The Item Development Subcommittee is composed of representatives from Arizona, Maine, Maryland, Montana, PAC-6, South Dakota, and Tennessee. This subcommittee provides overall input and direction related to development of field-test items; reviews all item development tasks; participates in development planning, item and passage reviews, and post-IRC edits; reviews alternative text; participates in Accessible Portable Item Protocol (APIP) reviews, plus the computer-based and paper-based materials review; and provides direction on updates to the graphics and editorial style guides, teacher directions, and the front matter for the *Directions for Test Administration* (DTA). In addition to the Item Development Subcommittee, stakeholders from schools and districts across the MSAA states participate in the field-test item development process during the passage content and bias meeting and the item content and bias meeting. Additional detailed information is available in Chapter 4.

Members of the Benchmarking and Scoring Subcommittee, which includes representatives from Arizona, Washington, D.C, PAC-6, and South Dakota, review the scoring specifications and participate during the scoring process. The Platform Subcommittee determines the platform development priorities and reviews all changes. This subcommittee is composed of representatives from Arizona, Maine, Maryland, and Tennessee. The Communicative Competence Subcommittee, composed of representatives from Arizona, Maryland, Montana, and PAC-6, provides direction related to student communicative competence and the provision of resources around the application of the early stopping rule and the student response check.

Finally, report revisions and decisions are the responsibility of the Reports Subcommittee, with representation from Arizona, Maryland, South Dakota, and Tennessee.

2.2 MSAA PARTICIPATION

The criteria for student participation in the 2018 MSAA reflect the pervasive nature of a significant cognitive disability. Both the mathematics and ELA content areas are considered when determining who should participate in this assessment. Table 2-1 below shows the participation criteria and the descriptors used to determine eligibility for each student.

Table 2-1. 2017–18 MSAA: Participation 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.*
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 the 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 C shows the 2017–18 summary of participation rates by demographic category for the MSAA for both mathematics and ELA.

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 an appropriate mode of communication are identified during the assessment process. Post-assessment, teachers have the opportunity to use the Communication Tool Kit developed by NCSC to help these students develop an appropriate mode of communication. The Tool Kit can be found here: https://wiki.ncscpartners.org/index.php/Communication_Tool_Kit.

CHAPTER 3 TEST CONTENT

3.1 History of Alternate Academic Achievement Standards and Core Content Connectors

Designed specifically for students with the most significant cognitive disabilities, the NCSC AA-AAS 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-AAS 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 development of Core Content Connectors (CCCs) to connect the Learning Progression Frameworks (LPFs) to the CCSS took place.

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 in an educationally sound way 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.

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 with their peers without disabilities from grade level to grade level. They are the starting point for instruction, not necessarily 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. Table 3-1 shows a series of CCCs within one big idea across multiple grades for the mathematics strand of geometry to demonstrate the content sequence maintained by the CCCs.

Table 3-1. 2017–18 MSAA: Example of the Core Content Connectors, Developed by NCSC, Across Grades—Mathematics Strand: Geometry Big Idea: Shapes and Figures—Their Attributes, Properties, and Corresponding Parts

Grades	*Geometry Core Content Connectors*			
Gr 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, pentagons, hexagons, ovals, equilateral, isosceles, and scalene triangles
Gr 3-4	3.GM.1h1 Identify shared attributes of shapes	4.GM.1h2 Classify two-dimensional shapes based on attributes (number of angles)		
Gr 5-6	5.GM.1a1 Recognize properties of simple plane figures	5.GM.1b1 Distinguish plane figures by their properties		
Gr 7-8	7.GM.1e Construct or draw plane figures using properties	8.GM.1g1 Recognize congruent and similar figures		
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		

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).

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 is about coherent connections across various aspects within and across a system (Forte, 2013a, 2013b). Traditional alignment procedures describe the degree of intersection, overlap, or relationship among content embedded in state content standards, assessment, and instruction (Webb, 1997).

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. 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. The following table lists the studies, when each was conducted, and the alignment question being addressed.

Table 3-2. 2017–18 MSAA: 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-AAS 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.

3.3 2018 MSAA Assessment Design

3.3.1 Operational Design

The operational MSAA is designed to produce valid and reliable mathematics and ELA scores. The mathematics and reading portions of the test are composed primarily of selected-response items. In mathematics, all grade levels, except for grades 6 and 7, also include constructed-response items. Writing is composed of selected-response stand-alone items, a multiple-part selected-response writing prompt, and an open-response writing prompt. For the 2017–18 assessment, grades 3 and 4 include foundational reading items for field testing.

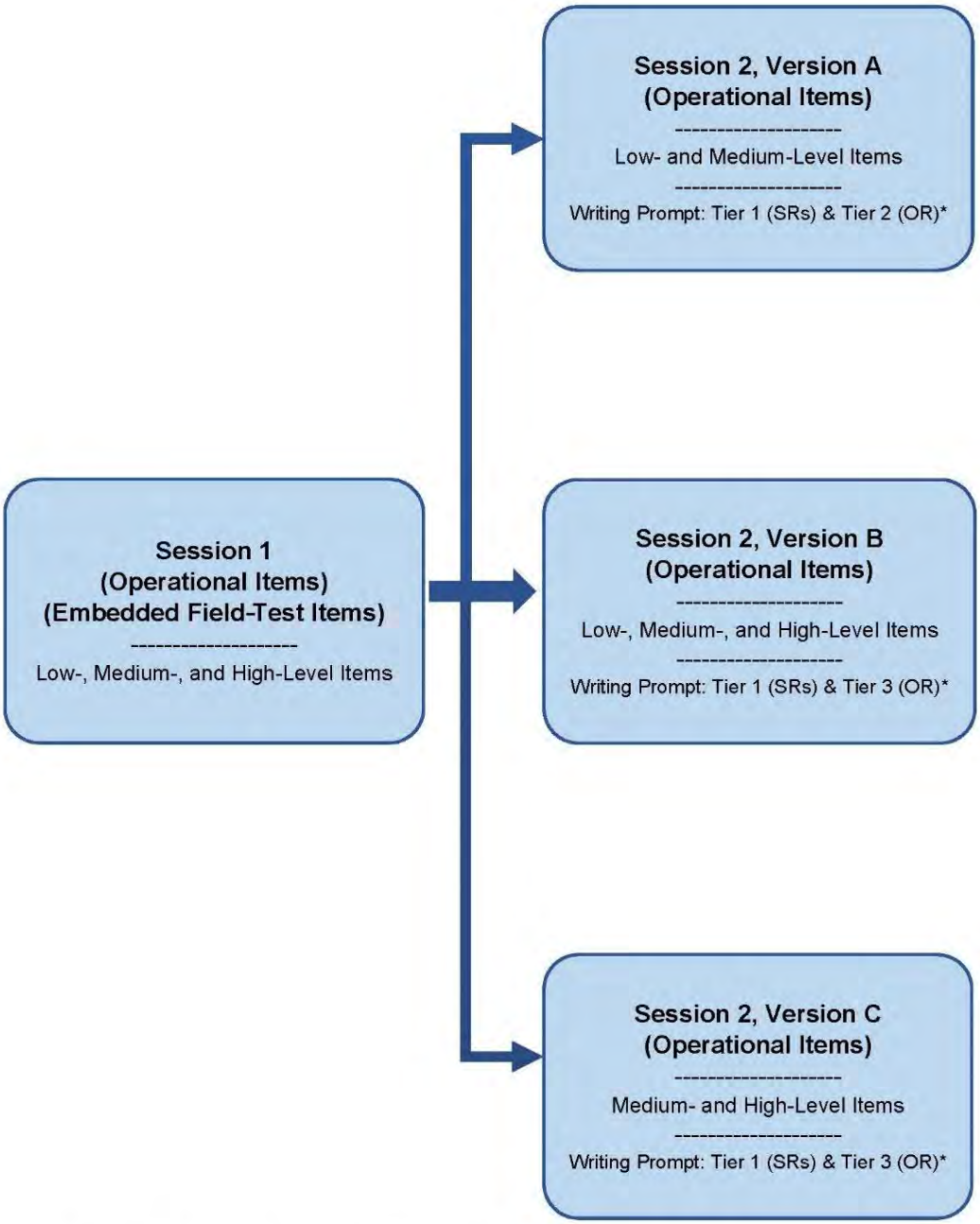
The items varied in complexity. Tier level designations are used during development. Items are built as item families where each tier within the family addresses both the content complexity and the degree of scaffolding and support provided with the items. Each item family provides four decreasingly complex versions (items) of the task referred to as Tier 4 (most complex), Tier 3 (less complex), Tier 2 (less complex than Tier 3), and Tier 1 (least complex). For the writing prompts there are three tiers of items. Tier 1 is a multiple-part selected-response item series, where all items build on each other toward the creation of a final product. Tier 2 and Tier 3 are open-response writing prompts that vary in complexity due to the amount of support provided at each tier. Additional detailed information about item development is provided in section 3.3.3.

For the 2017–18 assessment, three two-stage adaptive forms were developed for both ELA and mathematics to accommodate the inclusion of field-test items within Session 1. The forms follow guidelines informed by the respective content-area test blueprints (test blueprints are discussed in section 3.3.4). Each form contains 9–10 field-test items. The operational items are presented in two sessions. The Session 1 operational items are the same across all forms. Session 1 is considered Stage 1. Session 1 is taken by all students, while Session 2, which is considered Stage 2, is assigned to students based on how they perform on Session 1. There are three versions of Session 2, of varying difficulty, that may be assigned. Version C is intended to be slightly more complex and difficult than Version B, and Version B is intended to be slightly more complex and difficult than Version A. There are, thus, three possible paths for a student to take through the multistage test. All students take Stage 1, and, depending on how they perform on Stage 1, are assigned either 2A, 2B, or 2C.

In 2017–18 there is a high level of overlap between the items in each version, but enough variation to ensure varying degrees of the desired separation of the test information functions (TIFs) across the paths. For more information on TIF, see Chapter 9, and for more information on the measurement reliability stemming from these TIF values, see Chapter 10.

A Tier 1 writing prompt is included for Session 2A, 2B, and 2C. A Tier 2 writing prompt is included in Session 2A, and a Tier 3 writing prompt is included in Sessions 2B and 2C. Figure 3-1 illustrates the two-stage adaptive design with field-test items indicated in Session 1, along with the levels of items that were used. The three paths (Session 1 plus Session 2A, Session 1 plus Session 2B, and Session 1 plus Session 2C) for the operational assessment exist for each of the three field-test forms.

Figure 3-1. 2017—18 MSAA: Two-Stage Adaptive Design



*ELA Only, SR=Selected Response, OR=Open Response

3.3.2 Operational Items and Embedded Field-Test Items

As shown in Table 3-3, the mathematics tests consist of 35 operational items across the testing sessions per grade, consisting of 15 Session 1 items and 20 Session 2 items. Each field-test form has 10 different field-test items for a total of 30 field-test items across the three field-test forms.

Table 3-3. 2017–18 MSAA: Mathematics Items

Grade	*Total Operational Items*	*Field-Test Items*
	Administered to each student	*Total across three field-test forms*
3	35	30
4	35	30
5	35	30
6	35	30
7	35	30
8	35	30
11	35	30

As discussed earlier, there are three versions of Session 2. There are a number of items that are the same across Session 2A, 2B, and 2C in Tier 2 and Tier 3.

As shown in Table 3-4, the ELA tests administer 38-39 operational items and two writing prompts, consisting of 18 Session 1 items and 20-21 Session 2 items. Each field-test form has 8-9 field-test items for a total of 24 to 27 field-test items across the three field-test forms.

Table 3-4. 2017–18 MSAA: ELA Items

Grade	*Total Operational Items*	*Writing Prompt Operational Items*	*Selected-Response Field-Test Items*
	Administered to each student	*Total across three forms*	
3	39	2	27
4	38	2	27
5	38	2	27
6	39	2	24
7	39	2	24
8	39	2	24
11	39	2	27

As discussed earlier, there are three versions of Session 2. There are 2-3 passage sets and 4-7 writing items, for a total of 9-15 items, that overlap between Session 2A, 2B, and 2C.

The field-test items are selected from items developed in 2016–17. During the item development process, these items follow a typical development cycle, including reviews by MSAA State Leads and by

Item Content and Bias and Sensitivity panelists. The 2017–18 field-test items are selected based on the following criteria:

- mathematics and ELA items represent a variety of tier levels (including the writing stand-alone component);
- 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 Psychometrics for any validity and reliability concerns. All constructed tests, as well as the field-test items, are posted on a secure FTP site for the 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.

3.3.3 Item Design and Administration

The MSAA item design and administration intends 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 Partner States 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 States' Content Standards, and the heterogeneous characteristics of the target student population. The assessment items vary systematically in complexity yet remain aligned with the focal knowledge, skill, and ability (FKSA) 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. As discussed in 3.3.1 (Operational Design), items are built as item families where each tier within the family address both the content complexity and the degree of scaffolding and support provided with the items.

Overall Item Structure

Multiple item families are developed for each CCC, as introduced in 3.3.1. An item family is a cluster of items specific to one CCC. Each tier provides variable features and supports that offer multiple entry points for a variety of students to demonstrate their knowledge and skill. All items in an item family assess grade-

level academic concepts defined by either the FKSAs (for most of the tiers) or the Essential Understandings (EUs, at the lowest tier).

The tier guidelines and item specifications used in MSAA item development were developed and implemented in the initial design phase of NCSC. As outlined in the tier 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 tier address the EU which 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 tier or item type, all items include an introductory sentence and 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 Tier 1); the answer is worth 0 or 1 point. ELA multiple-part selected-response items 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 Tier 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 items. A typical example of a multiple-part selected-response item 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 tier 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 constructed-response items require the student to interact in some way with a set of materials to provide a response. These items are scored 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. 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 (as field-test) 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 Tier 1 writing prompt is a multiple-part selected-response item where the items build on each other toward the creation of a final product. For each item in the series, the student selects a response from two options, with the answer worth 0 or 1 point. There are four- to six-part items depending on the grade.

Unlike Tier 1, the Tier 2 and Tier 3 writing prompts are open-response writing prompts that vary in complexity with the amount of support provided at each tier. The Tier 2 writing prompt provides a graphic organizer and a template with sentence starters that a student utilizes to create his or her product based on information he or she included in the graphic organizer. The Tier 3 writing prompt provides a graphic organizer and a template which 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 Tiers 2 and 3, the student response is evaluated against a grade- and tier-specific rubric. (To reference the specific rubrics, see them embedded as an appendix in the *MSAA 2018 Guide for Score Report Interpretation*) Open-response writing prompt items were developed for Tiers 2 and 3 only. As outlined in Chapter 1, the writing prompt items are operational in each grade for the 2018 MSAA.

Administration

For every grade level, the ELA and mathematics test required two test sessions. TAs begin with Session 1 of either the mathematics test or the ELA test. Descriptions of the test sessions are shown in Tables 3-5 and 3-6.

Table 3-5. 2017–18 MSAA: ELA Test Sessions

Session 1: ELA	*Session 2: ELA, includes Writing Prompts*
Literary and informational reading passages and associated selected-response reading items	Literary and informational reading passages and associated selected-response reading items
Selected-response writing stand-alone items	One multiple-part selected-response writing prompt
Reading Foundational items field-test (grades 3 and 4 only)	One open-response writing prompt

Table 3-6. 2017–18 MSAA: Mathematics Test Sessions

Session 1: Mathematics	*Session 2: Mathematics*
Selected-response mathematics items	Selected-response mathematics items
Constructed-response mathematics items in selected grades	Constructed-response mathematics items in selected grades

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*

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 performance using the required Mathematics Scoring Rubrics. The Mathematics Scoring Rubrics provide scoring standards that must be used in evaluating student responses. 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 it can be understood by a scorer.

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 best practices in educational measurement. Tables 3-7 and 3-8 show the broad targets developed to guide the item development process and to inform test construction. They provide general guidance for identifying areas of emphasis in the development of the mathematics and ELA tests. The test blueprints in Appendix D 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. In an effort 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 Mathematics

Mathematics items are aligned with prioritized CCCs, which are in turn connected to the CCSS and States' Content Standards, as well as the LPFs. Mathematical knowledge across the CCCs is assessed through selected-response items and constructed-response items. (Constructed-response items are present at grades 3, 4, 5, 8, and 11 only.) The need for constructed-response items is determined by the FKSA associated with a given CCC.

Table 3-7. 2017–18 MSAA: Guidelines for Distribution of Mathematics Content by Grade Level

Mathematics Content Category	*Gr 3*	*Gr 4*	*Gr 5*	*Gr 6*	*Gr 7*	*Gr 8*	*Gr 11*
Operations and Algebraic Thinking	10-11	10-11	3-4				
Number and Operations Base Ten	7	3-4	14				
Number and Operations Fractions	7	10-11	7				
Measurement and Data	7	7	7				
Geometry	3-4	3-4	3-4	3-4	7	10-11	3-4
Ratio and Proportions				10-11	14		
Expressions and Equations				7	3-4	7	
The Number System				10-11	7	3-4	
Statistics and Probability				3-4	3-4	7	7
Functions						7	
Algebra and Functions							17-18
Number and Quantity							7
Total	**35**	**35**	**35**	**35**	**35**	**35**	**35**

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 D identify which CCCs require two items.

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

3.4.2 English Language Arts

For the 2018 MSAA, the ELA items in reading and writing are aligned with prioritized CCCs, which are in turn connected to the CCSS and States' Content Standards, as well as 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 States' Content Standards.

For the 2018 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 11, 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 foundational content category addresses the anchor standard of fluency. For 2017–18 the foundational reading items are being field-tested.

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 the use of standard English conventions (for the open-response writing prompt only).

Table 3-8. 2017–18 MSAA: Guidelines for Distribution of ELA Content by Grade Level

ELA Content Category	*Gr 3*	*Gr 4*	*Gr 5*	*Gr 6*	*Gr 7*	*Gr 8*	*Gr 11*
Reading Literary	38-40%	38-40%	38-40%	38-40%	38-40%	38-40%	38-40%
Reading Informational	38-40%	38-40%	38-40%	38-40%	38-40%	38-40%	38-40%
Reading Vocabulary	10-14%	10-14%	10-14%	10-14%	10-14%	10-14%	10-14%
Reading Foundation	0	0	0	0	0	0	0
Writing	10%	10%	10%	10%	10%	10%	10%

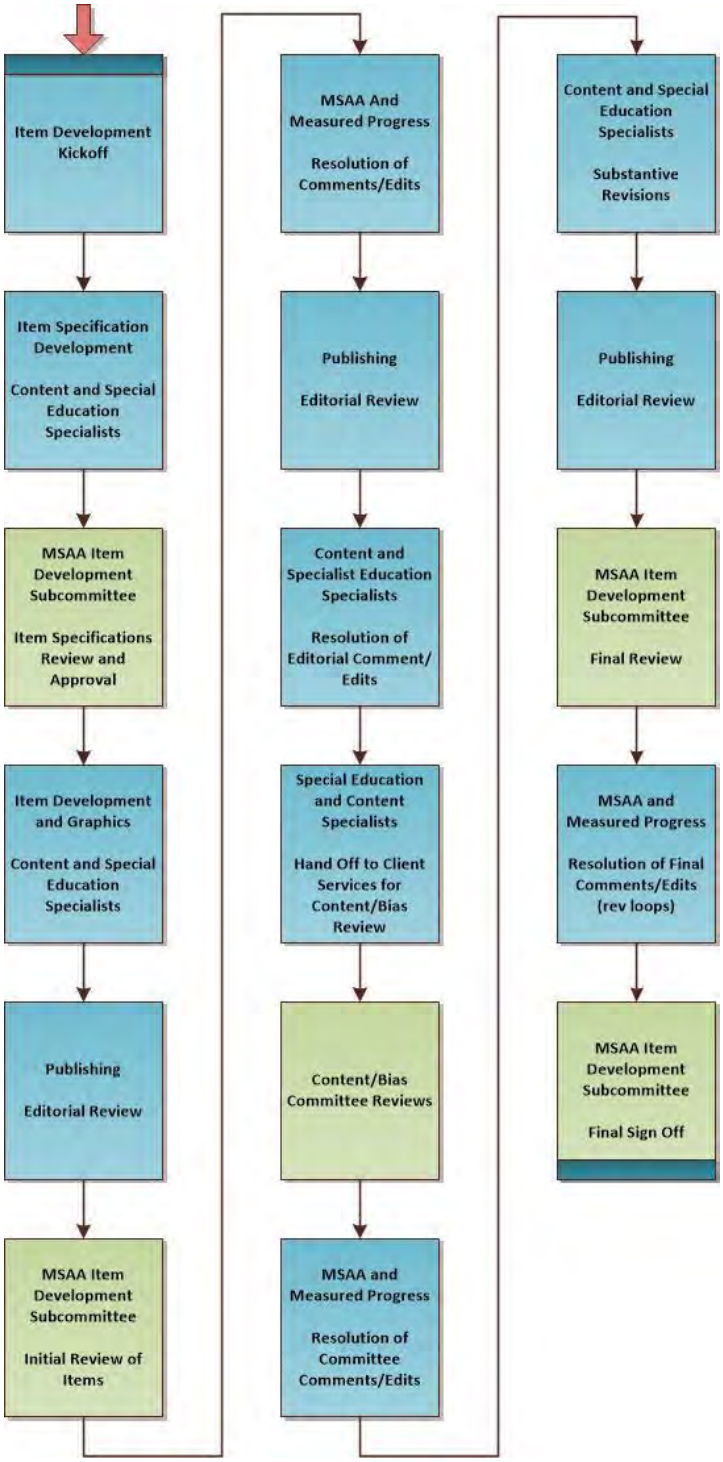
CHAPTER 4 TEST DEVELOPMENT

4.1 General Philosophy and Role of Item Development Subcommittee 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. The MSAA is designed to assess the academic content of the CCCs through an assessment design that consists of items written at various levels of complexity and provides built-in supports to meet the individual needs of the students. The two-stage adaptive assessment allows students to demonstrate what they know and can do. Given the wide diversity of the student population, great emphasis is placed on ensuring the MSAA is appropriate and accessible to all students.

The MSAA items on the 2017–18 administration are from the previous NCSC 2014–15 administration, the 2015–16 MSAA administration, and/or the 2016–17 MSAA administration. As described in Chapter 3, the items selected as field-test items are developed by MSAA. The item development process is an iterative one, which allows for multiple opportunities for review of the items by various stakeholders including MSAA State Leads, content experts and Partner State representative reviewers that are selected by MSAA State Leads, and external passage and item content and bias review participants. Figure 4-1 provides a flowchart outlining the item-development process.

Figure 4-1. 2017—18 MSAA: Item Development Process



The external passage content and bias group, comprising of general and special education teachers, review passages prior to the start of item development for the ELA assessment. External item content and bias groups, comprising of general and special education teachers, convene in the summer to review newly developed items for content or bias and sensitivity. Each ELA and mathematics content group reviews items for content related considerations, such as alignment to the FKSA or EU, ratings of Depth of Knowledge, clarity of the item content, and consistency of teacher directions. Separate bias and sensitivity groups review the ELA and mathematics items for bias and sensitivity considerations, as well as accessibility considerations. The list of participants from the item content and bias review is included in Appendix E. The Item Development Subcommittee, which is made up of the MSAA State Leads, provides overall direction and guidance regarding field-test item development. This multistage development and review process provides ample opportunity to evaluate items for their accessibility, appropriateness, and adherence to the principles of Universal Design. In this way, accessibility serves as a primary area of consideration throughout the item development process. This 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.

CHAPTER 5 TRAINING AND ADMINISTRATION

5.1 TEST ADMINISTRATOR AND TEST COORDINATOR TRAINING

The MSAA Partner States 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 TAs are utilizing the same well-defined administration procedures and the provided training, manuals, and supporting documents, administration is optimally 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 at their role and responsibility ensures they are appropriately prepared to support the TAs.

As the 2018 administration is a computer-administered test, the administration procedures are consistent with the hardware and software requirements of the test specifications. MSAA required 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) *MSAA Online Assessment System User Guide for Test Administrators*; (d) *MSAA Online Assessment System User Guide for Test Coordinators*, and (e) grade-, content-, and form-specific DTA. TCs and TAs receive both the online training and the supporting documents to ensure fidelity of implementation and the validity of the assessment results as well as to help MSAA prevent, detect, and respond to irregularities in academic testing and maintain testing integrity practices for technology-based assessments.

5.2 TEST ADMINISTRATOR TRAINING MODULES

The online training modules for TAs are available prior to the beginning of the testing window and throughout the testing window. 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 are updated for the 2018 administration in correspondence with the updates to the required documents. There are 6 modules (see Table 5-1). Each module takes approximately 25 minutes.

Table 5-1. 2017–18 MSAA: Training Modules for Test Administrators

Module 1: MSAA Overview
Module 2: Test Design and Experience
Module 3: Navigating the MSAA Online Assessment System
Module 4: Completing the Student Information
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; one module must be viewed before the link for the subsequent module to become accessible.

There are questions pertaining to information from the module at the end of 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. If TAs do not fulfill this certification requirement, they are not allowed access to the secure test materials. The TAs are notified within the MSAA system whether they pass the final quiz. TAs are allowed multiple attempts to obtain a score of 80% or higher on the final quiz.

In addition to the module trainings, TAs are instructed to become familiar with the online system by accessing sample items. MSAA utilizes the same sample items as the 2017 administration, which were developed by content and measurement experts for teachers, administrators, and policymakers for the NCSC assessment. The sample items do not address all assessed content at each grade level and are not representative of every item type. Rather, the sample items provide a preview of the array of items and illustrate multiple item features supporting ways in which students with a wide range of learner characteristics can interact with the assessment process.

5.3 TEST COORDINATOR TRAINING MODULES

In addition to the training modules for TAs described above, online modules specific to the role of TCs are made available both before and during the testing window. These training modules are customized to address the specific responsibilities of the TCs and to provide important information from the documents TCs are required to use: the (1) TAM and (2) *MSAA Online Assessment System User Guide for Test Coordinators*. Like the TA training modules, the TC training modules are updated based on the revisions made to the required documents. There are 6 modules; each of which run 20–25 minutes (see Table 5-2).

Table 5-2. 2017–18 MSAA: Training Modules for Test Coordinators

Module 1: MSAA Overview
Module 2: Test Design and Experience
Module 3: Navigating the MSAA Online Assessment System
Module 4: Completing the Student Information
Module 5: Create Users and Orgs
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; one module must be viewed before the link to the subsequent module becomes accessible. There are quiz 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

New to 2018 are four "Best Practice" videos. The videos provide TAs with targeted information about the MSAA. Video 1 focuses on reviewing assessment features that are available within the MSAA online system, how to go to full screen mode and zooming within the browser, and 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 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. The videos are accessed through the MSAA system.

Table 5-3. 2017–18 MSAA: 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 MSAA. It is organized by:

- 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 and 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 where TAs & TCs can locate the 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 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 following are provided as part of each DTA (as applicable for a content area):

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

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* provides 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 screen shots that demonstrate the functionality of the MSAA system. The user guides also contain appendices that describe accessibility features, assistive technology compatibility, and provide 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 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 is March 19-May 4, 2018. Both the ELA and mathematics are completed within the same administration window. Regardless of administration format (i.e., online or paper), the student assessments are submitted by the TA on or before May 4, 2018. 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 ongoing monitoring and quality control processes are part of the MSAA. Support is provided to TCs and TAs through the MSAA service center, additional supports built into the MSAA system functionality, and by the MSAA State Partners. TA feedback is gathered through an end of administration test survey. The intent is that reviewing the service center logs and analyzing the test survey results will better inform MSAA State Partners on areas where clarification and further support is needed.

5.8.1 MSAA Service Center

To provide support to schools before, during, and after testing, Measured Progress operates and provides tiered technical support through the MSAA Service Center. The MSAA Service Center is available for extended hours throughout registration and the testing window (from 7:00 a.m. to 8:00 p.m. EST, Monday through Friday) to accommodate the multiple time zones in which the test is administered.

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 and e-mail address are promoted to the field through the MSAA system and related communications.

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

All activity is tracked in the new MSAA Service Center ticketing system, ServiceNow, and is included in weekly status reports that are provided to MSAA State Leads. 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 Measured Progress program management team periodically provides direct phone and e-mail support where logistical or procedural support is needed by MSAA State Leads. In cases with policy or consortium-wide implications, however, program management refers the State Lead to the MSAA Partner States and related policy documentation.

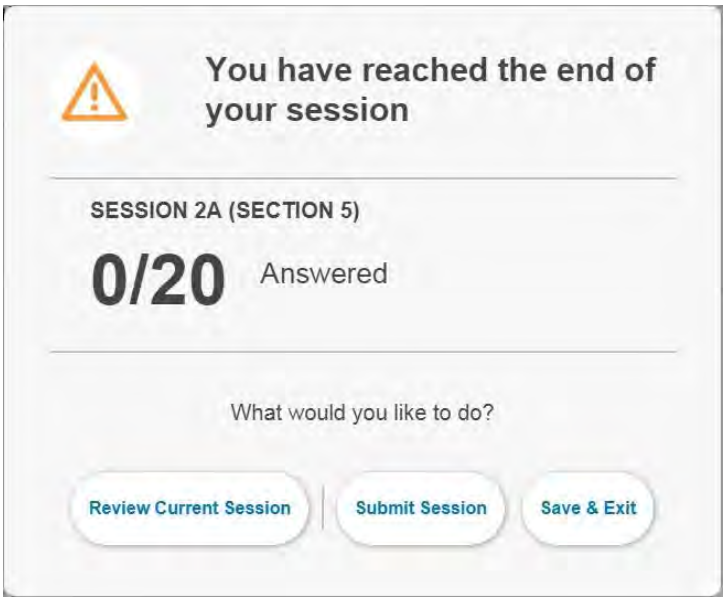
Furthermore, a banner messaging system in the MSAA system is implemented, as needed, to notify users of important information during the administration window. When activated, upon logging into the system a banner message appears at the top of the screen to notify users of system information and upcoming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



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 Measured Progress monitors activity and provides weekly updates to State Leads on the test statuses across MSAA Partner States and trends identified in support calls. This provides 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 MSAA.

5.8.4 Operational Test Survey Results

An End-of-Test Survey (EOTS) allows MSAA to gain knowledge from the experience of each TA administering the test. TAs are instructed to complete one EOTS after completing test administration for all of their students. The survey questions focus on several themes:

- technology use in the classroom,
- student behaviors and engagement, and
- instructional time and resources spent teaching the state content standards.

The results of the EOTS highlight several areas of concern that the MSAA Partners had identified prior to reviewing this data. The data support continued work in the following areas:

- increasing student engagement
- monitoring the available technology in classrooms to ensure the platform is up to date for compatibility
- providing professional development to support effective instructional strategies.

The survey data also identify the effectiveness of several improvements implemented in the 2018 MSAA to correct issues identified in the 2017 administration. These include:

- improving the online messages for submission of tests,
- providing clarification on exit procedures to reduce locked tests,
- creating best practice videos to address common administration questions, and
- reducing the need to scroll through items on screen.

One issue raised by the teachers in the EOTS data is a lack of familiarity and relatability with the contexts and scenarios used in some of the writing prompts and other items. The MSAA Partners' focus on developing test items and writing prompts that contain contexts and scenarios that are more relatable to students in this population is a sustained goal that requires several testing cycles.

Several questions on the survey address teachers' viewpoints and philosophies regarding teaching students with the most significant cognitive disabilities. The results again indicate the need for professional development that builds awareness and use of the available instructional and curricular materials, which illustrate various ways that students in this population can learn rigorous academic content.

At the same time, there continues to be a perception that the test is too difficult for some of the targeted population. The 2017 administration had introduced a stage adaptive design to address this issue. The MSAA Partners anticipated that the stage adaptive design would help to alleviate the concerns of many teachers by directing students to an appropriate level of difficulty within the test for each content area. The MSAA Partner States continue to work to ensure that future administrations' multistage tests have higher differentiation while still maintaining the required match to the blueprint.

Furthermore, responses from TAs regarding the high level of difficulty of the test reveal that many students are not fully engaging with the assessment. Individual comments regarding engagement suggest the need for implementation of best practices in administering the test to maximize student engagement.

The EOTS data also show that some students are using a variety of Augmentative and Alternative Communication (AAC) devices to access the test. In addition, most of the responses indicate that students use desktop computers, laptops, and tablets in the classroom with and without AAC devices; however, some of responses indicate that the students in these classrooms have no access to electronic devices outside of testing. This will be valuable information when evaluating compatibility with commonly used devices. Likewise, this information can be used to gauge the impact of limited prior exposure to computers on engagement with the online test.

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 which requires human scoring. The selected-response and constructed-response item types are described in detail in Chapter 3. The student may provide their responses to the items within the MSAA system. The system also allows for teacher entry of student responses for paper-based test delivery. The selected-response items are scored according to the answer keys provided in each test package. The mathematics constructed-response items are scored as a correct or incorrect student response which is then entered by the TA. At the completion of the operational test, all test data is extracted from the system and is 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 Measured Progress Information Technology Reporting (IT-Reporting) Department. The exported items go through a key verification check to confirm the selected-response and constructed-response item keys were entered correctly. A key verification check is conducted by the data analyst. Any items that may be flagged are provided to the content specialists to conduct a blind key check. The content specialists reviews the actual item and marks 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 of them 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 the comprehensive data analysis requirements are provided in the *IT Reporting Processing and Reporting Business Requirements* document, which can be reviewed in Appendix F.

Test Administrator/Scorer Training and Support

All TAs are required to 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.

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

If a TA needs scoring support related to the administration, scoring, entry of student responses and submission of student responses during the administration window, TAs are able to 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 11 are being operationally administered in the 2017–18 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 are able to 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 of the item. The item responses (no matter how they are entered) are then extracted from the online system and provided to Measured Progress 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 reviewing the parameters for the administration of the open-response writing prompt, as well as how to enter the student responses into the MSAA system. In addition, the best practice videos provide a student-TA representation that provides TAs with 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 includes the teacher scripting and directions related to any item setup and administration specifics and the materials for the open-response writing prompt.

Further direction is provided on the entering the open-response writing prompt student responses through the *MSAA System User Guide for Test Administrators*. Additionally, support for TAs is provided through the MSAA Service Center.

6.2.2 Benchmarking and Identification of Scoring

The open-response writing prompts were benchmarked during the 2015, 2016, and 2017 field tests. Measured Progress scoring experts (Scoring Supervisors and Scoring Team Leaders [STLs], defined below) worked collaboratively with NCSC representatives in 2015 and with MSAA representatives from the Benchmarking and Scoring Subcommittee in 2016 and 2017 to review student responses, assign a score based on the MSAA grade and tier 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; representatives from NCSC (2015) and the MSAA Benchmarking and Scoring Subcommittee (2016 and 2017) acknowledged their consensus on the signoff document for each prompt. Also, beginning in 2017 a scoring decisions document was developed and reviewed by the MSAA Benchmarking and Scoring Subcommittee which provided rationale and decision points made by the subcommittee to use 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. They are primarily obtained through Kelly Services, a temporary employment agency. All selected scorers hold the minimum of a four-year college degree which included ELA or writing coursework. 90% of the leadership and scorer group assigned to the MSAA have previous experience in scoring alternate assessments, and 40% have scored previous MSAA administrations. All scorers sign a nondisclosure/confidentiality agreement.

Table 6-1 summarizes the qualifications of the 2018 MSAA scoring leadership and scorers.

Table 6-1. 2017–18 MSAA: Qualifications of Scoring Leadership and Scorers

Scoring Responsibility	*Educational Credentials*			
	Doctorate	*Master's*	*Bachelor's*	*Number*
Scoring Leadership[1]	6.67%	13.33%	80.00%	15
Scorers	7.32%	24.39%	68.29%	41

[1] Scoring Leadership = Scoring Supervisors and Scoring Team Leaders

6.2.4 Measured Progress Staff and Scoring Leadership

The MSAA operational open-response writing prompts were scored in Dover, New Hampshire, between June 4 and June 15, 2018. The following staff members are listed below.

- Assistant Director (AD), Scoring Operations: Primarily responsible for coordinating scheduling, budgeting, and logistics of all Scoring Centers. In addition, the AD for Scoring Operations coordinates the scoring of special education contracts, has overall responsibility for MSAA scoring-related activities, and serves as the Scoring Services Project Manager for MSAA.
- ELA Group Manager for Scoring: Responsible for managing scoring-related activities and monitoring reports, as well as leadership and scorer training to ensure overall consistency of scoring.
- Scoring Content Specialist: Responsible for overseeing scoring activities across grades and monitoring accuracy and productivity across groups.
- Special Education Specialist: Responsible for overseeing scoring activities and acting as special education lead in coordination with the Measured Progress scoring staff.
- *iScore* Operations Manager: Responsible for setup and maintenance of *iScore* system for scoring and coordinating technical communication.
- Scoring Supervisor: Responsible for selecting calibration responses, training STLs and scorers, resolving arbitrations, and monitoring the consistency of scoring for items in assigned grades. Scoring Supervisors may also participate in benchmarking and identifying qualification sets prior to the onset of scoring.
- Scoring Team Leader (STL): Responsible for performing quality-control measures, resolving arbitrations, and monitoring the accuracy of a small group, usually consisting of not more than six scorers. STLs may also participate in benchmarking and identifying qualification sets prior to the onset of scoring.

6.2.5 Training

Scoring Content Specialists and Scoring Supervisors assigned to train the STLs and scorers thoroughly review the decisions and materials that resulted from the benchmarking meetings in preparation for training. One Scoring Supervisor is assigned to each tier's writing prompts across grades. The Scoring Content Specialists and Scoring Supervisors were responsible for creating pre-recorded training modules to

use in this year's training. Leadership training took place on May 30 and June 1. STLs are required to meet or exceed the accuracy standard of 80% exact agreement on all items and at least 90% exact/adjacent[2] agreement on each trait. This process is applied to each of the three writing traits[3] 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 and Scoring Supervisors conduct training on each open-response writing prompt before scorers are allowed access to student responses. Scorers are divided into two groups. One group focuses on Level 2 items and the other on Level 3 items. Training sessions for scorers are facilitated by the Scoring Content Specialists and a Scoring Supervisor 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 addressing 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 three traits of the MSAA analytic rubrics for writing and how the scoring for each trait are applied to student work. (See Writing Scoring Rubrics embedded as an appendix in the *MSAA 2018 Guide for Score Report Interpretation* provided in Appendix G.)
    - Pertinent information on the testing instructions and item stimuli.
    - 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 module announces the anchor response score and explains the scoring rationale, allowing scorers to internalize typical characteristics of each score point.
- Scorers are instructed to refer back to the anchor set frequently during scoring.

[2] "Adjacent agreement" means that the two scores only differed by one score point.
3 The three writing traits are organization, idea development, and conventions. See rubrics embedded in Appendix G.

- After completing the module, training continues with the Scoring Content Specialist and/or the Scoring Supervisor presenting the supplementary training materials 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, contain both very low and very high attributes). None of the practice papers contain responses which would require identification as nonscorable responses.
- During the review of practice responses, the trainer(s) often focus 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 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 for scoring each item. If scorers attain a score match of at least 80% exact and 90% exact/adjacent agreement on all traits for the first qualification set, they are considered a "qualified scorer" and permitted to score live student responses. For the other scorers, the Scoring Supervisor conducts a 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 receives the retraining referenced above. However, they would only be required to qualify on trait 2 in qualification set 2. When the data indicates that a qualified scorer has demonstrated a weakness in a particular trait, that qualified scorer receives 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.

Once the first open-response writing prompt for a grade and tier is completely scored, the training process is repeated for the next prompt. This continues until all 14 open-response writing prompts are scored. Qualification statistics are located in Appendix H.

6.2.7 Methodology 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 Measured Progress *iScore* 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* provides anonymity to individual students and ensures 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 into the MSAA system. For Tier 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, then 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 Tier 3 prompts, the computer-generated text and the uploaded evidence serve to provide a holistic demonstration of student ability and are considered in combination with one another when both are available. When only one portion is available, the prompt is scored like a Tier 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* linked student responses; no student names are visible to scorers unless it appears on material uploaded by the TA.
- Measured Progress 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 only those who are scoring or working for Measured Progress in a scoring management capacity.
- During scoring, *iScore* enables a constant measuring and monitoring of scorers for scoring accuracy and consistency. Each scorer's reading rate and total number of scored responses are also monitored.
- Scorers are required to maintain an acceptable scoring accuracy rate (80% exact/90% exact/adjacent agreement) on a daily basis as measured across read-behinds, double-blinds, and daily calibration sets. (These are described below.)
- Scorers who repeatedly fell below standard are retrained or dismissed from scoring that item.

- 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 Tables 6-6 through 6-8.)

Table 6-3 represents the total number of student responses scored by writing prompt in each grade.

Table 6-3. 2017–18 MSAA: Student Responses per Grade

Grade	*Number of Student Responses*		
	WRCC002	*WRCC003*	*Total*
3	1335	1818	3153
4	1513	1951	3464
5	1555	2024	3579
6	1398	2777	4175
7	1238	2412	3650
8	1445	2259	3704
11	1230	2205	3435

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

Scoring Rules

All open-response writing prompts are scored against a three-trait rubric (see rubrics embedded in Appendix G). The scoring scale options of 0, 1, 2, and 3 were applied to each trait. When a response does not conform to score point parameters, scorers can designate the response as one of the following:

- **Blank:** There is no attempt to respond to the item; no uploaded material is provided, and no response has been typed.
- **Unreadable:** The text on the scorer's computer screen is indecipherable or too faint to read accurately.
- **Non-English:** The response is written in a language other than English.
- **Repeats the Prompt:** The response is a direct copy of the prompt without any original text.
- **No Score:** The response requires clarification or adjudication by Scoring Leadership and scorers can only assign this designation with approval from Scoring Leadership.

Table 6-4 displays the resolution process for each of the responses described above.

Table 6-4. 2017–18 MSAA: 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 the Scoring Leadership.
Unreadable	Those responses judged unreadable are forwarded to a special queue within *iScore* 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.”)
Non-English	Responses written in a language other than English are marked Non-English and are converted to zeros (“0”) for reporting purposes.
No Score	Responses that require additional clarification or adjudication are escalated to Scoring Leadership for response appraisal and scoring. This includes responses where it appears that more than one student’s work has been uploaded to the response. 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 are scored based on its merits and then forwarded for review. If further attention is warranted, the Client Services team notify the appropriate MSAA partner state. Responses that are determined to be nonscorable are resolved by the Measured Progress leadership team and in consultation with the MSAA Benchmarking and Scoring Subcommittee, if necessary.

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

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

Scoring flagged a total of twenty-three responses as "Alert" during the scoring process and were forwarded to the appropriate Partner State representatives. See Table 6-5.

Table 6-5. 2017–18 MSAA: Responses Flagged with "Alert"

MSAA Partner State	*Number of "Alert" Responses Flagged*
Arkansas	2
Arizona	4
Guam	1
Maryland	7
Montana	2
Tennessee	7

Note: No responses were flagged "Alert" from Maine, South Dakota or Washington, D.C.

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 Benchmarking and Scoring Subcommittee representatives, along with the Measured Progress Special Education (SPED) and Scoring teams, monitor reports daily. Read-behind and double-blind statistics are reviewed daily. Calibration sets are administered and reviewed repeatedly during the course of the project. Scoring Leadership and Content Specialists from the Scoring Services and SPED departments at

Measured Progress pay close attention to the disaggregated read-behind, double-blind, and calibration statistics.

Scorers in need of additional 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 remove 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 MSAA scoring, the work of two scorers and one STL was voided (grade 3, Tier 2).

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 consiting 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 these scorers continue to be closely monitored. Over the course of scoring the MSAA, only 10 scorers (across all seven grades and 14 items), required retraining. In most cases, scorers who received retraining successfully returned to scoring and, as mentioned previously, only three scorers had work voided during the course of scoring.

6.2.8.2 *Read-behind Scoring*

Read-behinds provide a crucial tool in verifying scorer accuracy. The STLs complete read-behinds on individual scorers on a daily basis. An 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 varies depending on the accuracy of each scorer. 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 80% exact and 90% exact/adjacent threshold, Scoring Leadership have the prerogative to void their scores for the day and/or stop them from scoring that item. Scoring Leadership monitors scoring

accuracy and consistency by reviewing the read-behinds performed by the STLs as well as completing read-behinds on the STLs whenever possible.

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 that was 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 inter-rater agreement rates, double-blind scoring also 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, evaluate the response, with their 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 of retraining.

Final Score Resolution

Scoring Leadership provides resolution scores for responses that do not have exact agreement on all traits after read-behind or double-blind scoring. Tables 6-6, 6-7, and 6-8 provide examples of how the final score of record may be determined through resolutions.

Table 6-6. 2017–18 MSAA: Examples of Scoring Resolutions: Read-Behind Scoring [1] (Trait 1-Trait 2-Trait 3)

Scorer Score	*Leadership Score*	*Score*
3-3-3	3-3-3	3-3-3
3-2-2	3-3-3	3-3-3
3-3-3	2-2-2	2-2-2

[1] In these cases, the leadership score overrides the scorer score.

Table 6-7. 2017–18 MSAA: Examples of Scoring Resolutions: Double-Blind Scoring[1] (Trait 1-Trait 2-Trait 3)

Scorer #1	*Scorer #2*	*Leadership Resolution*	*Final*
3-3-3	3-3-2	3-3-3	3-3-3
3-2-2	1-1-1	3-3-2	3-3-2
2-1-1	1-1-1	2-2-1	2-2-1
1-1-1	3-3-3	2-2-2	2-2-2

[1] All adjacent or discrepant scores are resolved in arbitration; in these cases the leadership score becomes the final score of record.

Table 6-8. 2017–18 MSAA: Examples of Scoring Resolutions: Edit Scoring[1] (Trait 1-Trait 2-Trait 3)

Scorer #1	*Scorer #2*	*STL #1 RB*	*STL #2 RB*	*Scoring Supervisor Resolution*	*Final*
3-2-2	3-2-2	-	-	-	3-2-2
2-2-2	3-2-2	2-2-2	2-2-2	-	2-2-2
0-1-1	1-2-1	1-2-1	1-2-1	-	1-2-1
3-2-2	2-1-1	3-2-2	3-1-2	3-2-2	3-2-2
1-0-1	1-1-2	1-1-1	1-1-2	1-1-2	1-1-2

[1] If a response receives more than one read-behind and the scores supplied by the STLs do not agree, a resolution score is needed. In these cases, the Scoring Supervisor provides a final score.

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 Measured Progress Scoring team and the MSAA Benchmarking and 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 following reports, listed in Table 6-9, provide the comprehensive tools and statistical information needed to execute quality control and manage production.

Table 6-9. 2017–18 MSAA: 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 note 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 note 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 scores across each trait.
Summary Report	The Summary Report lists the total number of student responses loaded into *iScore*. This report includes how many reads have been completed to date and how many reads 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 how much agreement is actually present compared to how much agreement would be expected to be present 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 6-10 shows that agreement between raters achieved the Substantial Agreement or Almost Perfect Agreement ranges for most of the open-response writing prompts across grades. In four cases, the kappa agreement rate is at the high end of the Moderate Agreement range. (See Table 6-15: grade 3 Organization and Idea Development traits for one prompt; grades 5, 7, and 8 Idea Development category for one prompt in each grade).

Table 6-10. 2017–18 MSAA: Kappa Agreement—Operational Open-Response Writing Prompts

Grade	*Item*	*Organization* *Trait 1*	*Idea Development* *Trait 2*	*Conventions* *Trait 3*
Grade 3	WRCC002	0.67	0.65	0.84
	WRCC003	0.54	0.54	0.82
Grade 4	WRCC002	0.63	0.63	0.79
	WRCC003	0.78	0.72	0.80
Grade 5	WRCC002	0.61	0.57	0.82
	WRCC003	0.73	0.74	0.81
Grade 6	WRCC002	0.74	0.75	0.80
	WRCC003	0.66	0.63	0.81
Grade 7	WRCC002	0.72	0.74	0.78
	WRCC003	0.55	0.49	0.77
Grade 8	WRCC002	0.77	0.74	0.84
	WRCC003	0.64	0.59	0.83
Grade 11	WRCC002	0.75	0.67	0.80
	WRCC003	0.65	0.63	0.77

Note: For identification purposes in *iScore,* Tier 2 prompts are designated as WRCC002 across all grades and Tier 3 prompts are designated 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

Decision Rules Document

To ensure that reported results for MSAA are accurate relative to collected data, a decision rules document delineating processing rules is prepared, edited in collaboration with the MSAA Reports Subcommittee, and then approved by all participating MSAA Partner States prior to processing of the results. The decision rules 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, and then approved by all participating MSAA Partner States prior to reporting.

The 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 Partner States. The reporting requirements and corresponding report design templates were developed by Measured Progress 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, Measured Progress works with the MSAA Reports Subcommittee to identify modifications to the templates used last year that would ensure the data elements, layout, and report text were meaningful for reporting the spring 2018 MSAA results. Once finalized, the results of this collaborative process were presented to participating MSAA State Leads for final approval. The Student Report underwent a major redesign based on the recommendations of the Technical Advisory Committee (TAC). The TAC had recommended that the Student Reports be made more user-friendly and usable, to better engage parents/guardians and teachers, who are the primary users of this report. The biggest change made was to present the scores in the front page of the report and present the textual content in the back. Performance level descriptors (PLDs) were added along with the “What Next” text, which gives recommendations to guide the parents/guardians. For ELA, subscores were added for Reading and Writing. Another major change to the report was that color was added, whereas previous reports were black and white. Care was taken to ensure that the colors chosen were 509-compliant. Elements of the report were repositioned for greater readability. Some MSAA states engaged their district stakeholders to solicit feedback from focus groups. The final report design incorporated this feedback along with the feedback from the TAC.

MSAA 2018 Guide for Score Report Interpretation

Measured Progress uses an iterative process to annually update the Guide for Score Report Interpretation with the MSAA Reports Subcommittee. Updates are made to ensure the guide provides information that is most helpful to 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 MSAA, student participation criteria, score reporting overview, and samples of the various types of reports available to schools and districts. Guidelines are provided to 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 PLDs for ELA and mathematics, a sample individual student report, and the writing prompt scoring rubrics. The final, approved *2018 MSAA Guide for Score Report Interpretation* is delivered to the MSAA Partner States for state-specific revisions and distribution.

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

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

- student reports;
- school roster reports; and
- school, district, and state summary reports.

Reports are generated for each school, district, or state that has results, as defined by the MSAA decision rules 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 State Leads, only test coordinators (TCs) are granted access to the online reports. Access was controlled by user-permissioned accounts, as illustrated in Table 7-1:

Table 7-1. 2017–18 MSAA: Report Accessibility by Role

Reports	Test Coordinator		
	State	*District*	*School*
Student	Yes	Yes	Yes
School Roster	Yes	Yes	Yes
School Summary	Yes	Yes	Yes
District Summary	Yes	Yes	No
State Summary	Yes	No	No

For the purposes of the assessment system, MSAA State Leads are regarded as State TCs. As such, they are also able to add new district and school TCs to the online system and to any block users no longer in the TC role from accessing the system. For 2018, these reports were provided to schools, districts, and parents as soon as possible following the standards validation process and cut score approval process with each state's Board of Education/Superintendents, which occurred in the summer of 2018.

The primary results reported are the student's 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 (more information on these processes appears in chapter 9), 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 percent of students in each performance level are summarized by school, district, and state on both the roster and summary reports. This allows for the comparison of individual student performance in relation to the state, as well as comparison of school and district results with the overall state results.

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 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 permissioned users to download, print, and disseminate to parents and guardians. Each report contains the student name, test grade, and school on the front and back of the report. The back page also includes the state student ID for additional confirmation of the student's identification. Sample student reports are included in the *MSAA 2018 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 an easy-to-understand manner by providing the expected range of scores the student would likely earn if tested again.

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. This page also contains a short overview of the results included on the back page, as well as a link to where more information may be accessed online. 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 percent of students by performance level are summarized for the school, district, and state at the top of the roster. The decision rules and roster report requirements identify which of the 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 Partner States do not suppress information when the number of students participating was small. This practice places a burden on users to understand the data in the context of small numbers and to use all of the provided information to understand the results, as explained in the *MSAA 2018 Guide for Score Report Interpretation*.

Student results are listed below the summary section and are identified by name and state student identification number. It is intended that these data points be used in conjunction with the *MSAA 2018 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 decision rules 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 that did not test)
- Average Scale Score
- Performance Level (number and percentage at each performance level by grade in the state, district, school)

This summary provides a comparative snapshot of results and participation information at a high level and included both participation and performance summary information, allowing users to evaluate both aspects of their assessment results as guided by the *MSAA 2018 Guide for Score Report Interpretation*.

7.2.4 Quality Assurance

Proprietary quality-assurance measures at Measured Progress 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 data are handed off to different functions within the IT-Reporting Department, the sending function verifies that the data are accurate prior to handoff. Additionally, when a function receives a data set, the first step is to verify the data for accuracy.

A second level of quality-assurance measure 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 decision rules to independently program the reporting table. The production and quality-assurance tables are compared, and only after there is 100% agreement are the tables released for report generation.

The third aspect of quality control at Measured Progress involved the Software Quality Assurance (SQA) team who 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. Software Quality Assurance develops a test plan that includes previously agreed upon report designs and decision rule documents. Test cases housed in an internal test cases repository software are then executed including but not limited to the following:

1. Testing data counts of data imported
2. Testing data quality of individual fields for valid values, such as Gender, Ethnicity, etc.
3. Validate scripts developed by the software developers to ensure they match business requirements and technical specifications.

Included in this testing effort to ensure the quality of the data the SQA team uses a sample of schools and districts which is selected based on multiple criteria. A few are identified below:

4. Unique student testing records
5. Students complete testing
6. Students partial complete testing
7. Invalidated students

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

Included in the final execution the Software Quality Assurance team executes test cases validating student printed reports and student labels for accuracy in comparison to the previously agreed to report design specifications.

Once all the test cases have been passed the SQA team notifies the Measured Progress Client Services department for final sign off and communication.

CHAPTER 8 CLASSICAL ITEM ANALYSIS

As noted in Brown (1983), "A test is only as good as the items it contains." A complete evaluation of a test's quality must include an evaluation of each 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 assess only knowledge or skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. Items should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students, particularly racial, ethnic, or gender groups.

Both qualitative and quantitative analyses have been conducted to ensure that the 2017–18 MSAA ELA (reading and writing) and mathematics items met these standards. Qualitative analyses are described in earlier chapters of this report; this chapter focuses on quantitative evaluations. Statistical evaluations are presented in three parts: (1) difficulty indices, (2) item-test correlations, and (3) differential item functioning (DIF) statistics. The item analyses presented here are based on the administration of MSAA in spring 2018.

8.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

All items have been evaluated in terms of item difficulty according to standard classical test theory practices. Classical statistics provided in this chapter should be cautiously interpreted because some items are only administered to a subgroup of examinees, and each subgroup can be quite different in their underlying proficiencies. As described in Chapters 3 and 4, the 2017–18 MSAA was a stage adaptive test, consisting of three possible paths through the test. For simplicity, hereinafter, each path is referred to as:

- Path A: Stage 1 and Stage 2 Version A
- Path B: Stage 1 and Stage 2 Version B
- Path C: Stage 1 and Stage 2 Version C

As mentioned earlier, each version in Stage 2 was intended to be slightly more difficult, with C being the most complex and challenging (Note: As explained in section 5.6.17, stage correlates with session number.) The lowest-achieving examinees were routed to Stage 2 Version A, and so on. The examinees who were administered a particular path exhibited a much smaller range of achievement as compared with the entire population who took the assessment. This smaller range varied in the obvious way across the three paths. Because of this restriction of range and because of the differences across the three paths, the classical statistics are not comparable between items on different paths and are not comparable to statistics based on all the examinees (e.g., statistics for the Stage 1 items and statistics from past years).

Difficulty is defined as the average proportion of points achieved on an item and is measured by obtaining the average score on an item and dividing it by the maximum possible score for the item. Selected-response and 1-point open-response items are scored dichotomously (correct versus incorrect); for these items, the difficulty index is simply the proportion of students who correctly answered the item. An index of 0.0 indicates that all students received no credit for the item; an index of 1.0 indicates that all students received full credit for the item.

Discrimination is defined as the correlation between student performance on a single item and total test score on the particular path. Within classical test theory, the item-test correlation is referred to as the item's discrimination because it indicates the extent to which successful performance on an item discriminates between high and low scores on the particular path on which the item occurred. Because of the restriction of range complications mentioned above, the increase in the number of items with poor classical discrimination statistics (as compared to past years) was expected.

A summary of the item difficulty and item discrimination statistics for each content area and grade is presented in Table 8-1. The mean difficulty values shown in the table are within typically observed ranges and are similar to those for the assessment 2016–2017 MSAA reported in last year's technical report. The mean discrimination values are slightly, but consistently, larger than those reported last year, though still similar to typically observed ranges. A total of 24 out of 685 items displayed negative discrimination statistics. A closer examination revealed that 23 out of 24 items with negative discrimination statistics appeared in either Stage 2 Version A only or Stage 2 Versions A and B only. As mentioned above, the lower mean discrimination statistics and the increase in negative values is not surprising given the nature of the adaptive test, where the restriction of range occurs.

Table 8-1. 2017–18 MSAA: Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade

Content Area	*Grade*	*Number of Items*	p-*value*				*Discrimination*			
			Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
ELA	3	58	0.22	0.86	0.59	0.16	0.05	0.52	0.32	0.10
	4	58	0.20	0.93	0.63	0.17	0.05	0.46	0.30	0.09
	5	56	0.22	0.87	0.55	0.17	-0.03	0.47	0.29	0.12
	6	54	0.26	0.84	0.58	0.15	0.09	0.54	0.33	0.10
	7	60	0.26	0.83	0.58	0.14	0.05	0.50	0.30	0.12
	8	54	0.19	0.87	0.60	0.16	-0.03	0.48	0.32	0.12
	11	58	0.20	0.91	0.62	0.15	0.06	0.55	0.33	0.11
Mathematics	3	60	0.21	0.87	0.50	0.18	-0.07	0.46	0.25	0.11
	4	58	0.10	0.76	0.48	0.15	-0.03	0.42	0.24	0.10
	5	62	0.08	0.73	0.44	0.15	-0.21	0.46	0.23	0.15
	6	55	0.27	0.83	0.56	0.14	0.01	0.46	0.28	0.11
	7	57	0.26	0.83	0.52	0.14	-0.10	0.43	0.24	0.13
	8	57	0.25	0.77	0.49	0.14	-0.15	0.46	0.29	0.13
	11	56	0.27	0.72	0.48	0.12	-0.08	0.43	0.23	0.14

Note: *p*-values are dependent on the number of options within the items. All *p*-values are calculated on items having either two (chance correct is .50) or three (chance correct is .33) options.

The individual item statistics can be found in Appendix I. Note that the classical statistics should be interpreted with caution because the items are primarily two- or three-option selected-response items, and some items were only administered to a subset of examinees. Because the items were developed to correspond to different tiers, the item statistics have been summarized by tier (Tables 8-2 and 8-3). Also, the item statistics were summarized by path, representing the different paths in the stage adaptive design (Tables 8-4 and 8-5). The classical statistics are not comparable between items on different tiers and between items on different paths because of the restriction of range and the differences across paths mentioned above.

Table 8-2. 2017–18 MSAA: Item-Level Classical Test Theory Statistics
Summary by Grade and Tier—ELA

Grade	*Tier*	*Number of Items*	p-*value*				*Discrimination*			
			Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
3	1	20	0.42	0.86	0.66	0.11	0.18	0.43	0.29	0.09
	2	14	0.22	0.83	0.56	0.23	0.05	0.45	0.31	0.10
	3	19	0.26	0.73	0.54	0.14	0.08	0.52	0.34	0.10
	4	5	0.32	0.75	0.61	0.17	0.26	0.40	0.32	0.05
4	1	20	0.46	0.84	0.71	0.10	0.17	0.46	0.31	0.09
	2	16	0.42	0.93	0.62	0.14	0.05	0.42	0.28	0.09
	3	14	0.24	0.72	0.55	0.15	0.10	0.44	0.33	0.09
	4	5	0.50	0.88	0.76	0.16	0.11	0.28	0.23	0.07
5	1	13	0.46	0.85	0.70	0.10	0.18	0.47	0.34	0.10
	2	24	0.22	0.86	0.46	0.15	0.07	0.47	0.25	0.13
	3	14	0.25	0.71	0.50	0.13	-0.03	0.44	0.30	0.13
	4	5	0.54	0.87	0.69	0.13	0.08	0.43	0.26	0.13
6	1	12	0.52	0.76	0.67	0.07	0.18	0.46	0.31	0.09
	2	21	0.26	0.84	0.59	0.17	0.12	0.49	0.33	0.09
	3	14	0.34	0.83	0.57	0.12	0.28	0.54	0.39	0.07
	4	7	0.30	0.57	0.44	0.09	0.09	0.45	0.22	0.12
7	1	17	0.48	0.83	0.65	0.09	0.16	0.40	0.30	0.06
	2	18	0.26	0.82	0.61	0.17	0.11	0.48	0.36	0.10
	3	18	0.30	0.66	0.49	0.10	0.11	0.50	0.28	0.15
	4	7	0.44	0.65	0.57	0.08	0.05	0.24	0.18	0.06
8	1	16	0.40	0.87	0.71	0.15	-0.03	0.43	0.30	0.13
	2	15	0.19	0.79	0.57	0.20	0.33	0.48	0.42	0.05
	3	18	0.37	0.65	0.51	0.09	0.04	0.46	0.29	0.13
	4	5	0.51	0.69	0.61	0.08	0.12	0.29	0.21	0.08
11	1	21	0.50	0.84	0.71	0.10	0.11	0.46	0.33	0.08
	2	19	0.20	0.73	0.56	0.15	0.06	0.55	0.34	0.14
	3	12	0.28	0.91	0.59	0.18	0.10	0.46	0.33	0.11
	4	6	0.34	0.70	0.59	0.14	0.18	0.39	0.28	0.07

Table 8-3. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Summary by Grade and Tier—Mathematics

Grade	Tier	Number of Items	p-value				Discrimination			
			Min	Max	Mean	SD	Min	Max	Mean	SD
3	1	10	0.36	0.74	0.61	0.14	-0.07	0.40	0.22	0.16
	2	24	0.21	0.87	0.49	0.16	-0.05	0.46	0.26	0.12
	3	22	0.22	0.77	0.43	0.17	0.10	0.45	0.25	0.09
	4	4	0.43	0.85	0.66	0.21	0.21	0.31	0.27	0.04
4	1	8	0.58	0.76	0.70	0.08	0.16	0.33	0.23	0.06
	2	23	0.10	0.66	0.42	0.11	-0.03	0.39	0.22	0.12
	3	24	0.27	0.75	0.45	0.14	0.05	0.40	0.25	0.10
	4	3	0.48	0.66	0.59	0.09	0.20	0.42	0.30	0.11
5	1	11	0.40	0.73	0.59	0.14	0.01	0.33	0.22	0.09
	2	25	0.23	0.72	0.41	0.13	-0.21	0.42	0.20	0.17
	3	23	0.08	0.65	0.38	0.13	-0.11	0.46	0.26	0.14
	4	3	0.45	0.69	0.56	0.12	0.23	0.35	0.29	0.06
6	1	10	0.48	0.83	0.68	0.11	0.01	0.35	0.24	0.12
	2	20	0.27	0.73	0.52	0.14	0.06	0.45	0.28	0.11
	3	19	0.35	0.73	0.51	0.12	0.11	0.46	0.30	0.10
	4	6	0.45	0.82	0.65	0.13	0.24	0.37	0.34	0.05
7	1	10	0.48	0.78	0.65	0.10	0.06	0.32	0.18	0.09
	2	21	0.34	0.83	0.51	0.14	-0.10	0.41	0.22	0.13
	3	20	0.26	0.62	0.44	0.10	-0.09	0.43	0.27	0.15
	4	6	0.36	0.66	0.55	0.13	0.21	0.43	0.33	0.09
8	1	10	0.41	0.74	0.59	0.10	-0.15	0.44	0.30	0.17
	2	20	0.25	0.71	0.47	0.14	-0.12	0.46	0.25	0.13
	3	21	0.27	0.77	0.46	0.14	0.12	0.46	0.31	0.11
	4	6	0.28	0.65	0.51	0.15	0.05	0.42	0.29	0.13
11	1	9	0.53	0.72	0.63	0.07	-0.03	0.33	0.18	0.12
	2	22	0.30	0.72	0.46	0.12	-0.08	0.43	0.24	0.15
	3	20	0.27	0.59	0.44	0.08	-0.03	0.41	0.22	0.14
	4	5	0.37	0.63	0.48	0.10	0.12	0.40	0.27	0.12

Table 8-4. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Summary by Grade and Path—ELA (All Item Types)

Grade	*Path*	*Number of Items*	p-*value*				*Discrimination*			
			Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
3	A	41	0.22	0.86	0.57	0.16	0.05	0.52	0.31	0.10
	B	41	0.26	0.86	0.59	0.15	0.05	0.52	0.33	0.10
	C	41	0.26	0.86	0.63	0.15	0.08	0.52	0.34	0.09
4	A	40	0.20	0.84	0.61	0.15	0.17	0.46	0.32	0.08
	B	40	0.24	0.84	0.62	0.14	0.10	0.46	0.33	0.08
	C	40	0.24	0.93	0.65	0.16	0.05	0.46	0.32	0.09
5	A	40	0.22	0.85	0.54	0.16	0.07	0.47	0.30	0.12
	B	40	0.25	0.86	0.57	0.16	-0.03	0.47	0.31	0.12
	C	40	0.25	0.87	0.58	0.16	-0.03	0.47	0.31	0.12
6	A	41	0.30	0.76	0.57	0.13	0.09	0.54	0.33	0.10
	B	41	0.26	0.83	0.58	0.14	0.09	0.54	0.33	0.10
	C	41	0.26	0.84	0.60	0.14	0.09	0.54	0.33	0.10
7	A	41	0.26	0.83	0.60	0.13	0.05	0.50	0.33	0.10
	B	41	0.30	0.83	0.59	0.14	0.05	0.50	0.31	0.13
	C	41	0.30	0.83	0.61	0.12	0.05	0.50	0.32	0.12
8	A	41	0.19	0.87	0.62	0.17	-0.03	0.48	0.35	0.11
	B	41	0.37	0.87	0.63	0.14	0.04	0.48	0.35	0.11
	C	41	0.41	0.87	0.64	0.13	0.12	0.48	0.35	0.10
11	A	41	0.20	0.84	0.64	0.15	0.11	0.55	0.36	0.10
	B	41	0.28	0.84	0.65	0.13	0.06	0.55	0.35	0.11
	C	41	0.28	0.91	0.65	0.14	0.10	0.55	0.35	0.10

Table 8-5. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Summary by Grade and Path—Mathematics (All Item Types)

Grade	Path	Number of Items	p-value				Discrimination			
			Min	Max	Mean	SD	Min	Max	Mean	SD
3	A	35	0.21	0.74	0.45	0.17	-0.07	0.46	0.24	0.13
	B	35	0.28	0.74	0.49	0.13	-0.05	0.46	0.28	0.11
	C	35	0.28	0.87	0.57	0.16	0.16	0.46	0.30	0.07
4	A	34	0.10	0.76	0.45	0.15	-0.03	0.40	0.24	0.11
	B	35	0.27	0.76	0.48	0.14	0.01	0.40	0.24	0.10
	C	35	0.27	0.76	0.53	0.14	0.07	0.42	0.26	0.08
5	A	35	0.08	0.73	0.42	0.17	-0.21	0.45	0.20	0.15
	B	35	0.23	0.73	0.43	0.15	-0.16	0.46	0.24	0.15
	C	35	0.25	0.73	0.50	0.14	0.09	0.46	0.32	0.08
6	A	35	0.27	0.83	0.53	0.15	0.01	0.43	0.24	0.10
	B	35	0.32	0.83	0.56	0.14	0.06	0.46	0.30	0.09
	C	35	0.35	0.83	0.62	0.12	0.16	0.46	0.34	0.07
7	A	35	0.26	0.78	0.49	0.14	-0.10	0.38	0.19	0.13
	B	35	0.33	0.78	0.51	0.11	-0.10	0.43	0.26	0.12
	C	35	0.36	0.83	0.56	0.12	0.16	0.43	0.33	0.07
8	A	35	0.25	0.74	0.46	0.15	-0.15	0.46	0.28	0.15
	B	35	0.28	0.77	0.49	0.13	-0.12	0.46	0.30	0.12
	C	35	0.40	0.77	0.56	0.10	0.13	0.46	0.35	0.08
11	A	35	0.27	0.72	0.46	0.12	-0.08	0.42	0.18	0.13
	B	35	0.31	0.72	0.46	0.11	-0.08	0.42	0.24	0.13
	C	35	0.37	0.72	0.52	0.10	0.07	0.43	0.30	0.09

8.2 DIFFERENTIAL ITEM FUNCTIONING

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 2017–18 MSAA, 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 at every total score. Then an overall average is calculated, weighting the total score distribution so that it is the same for the two groups.

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 be indicative of item bias. Course-taking patterns or differences in school curricula can lead to DIF but for construct-relevant reasons. 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 2017–18 MSAA, six subgroup comparisons were evaluated for DIF:

- Male compared with female
- White compared with Black
- White compared with Hispanic
- White compared with American Indian
- Not low socioeconomic status (SES) compared with low SES
- Not Limited English Proficiency (LEP) compared with LEP (including current, exited one year, and exited two years)

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 J 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 Tables J-1 and J-2). Dorans and Holland further stated 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 so content experts conducted review of items flagged for DIF .

The number of items with a "high" DIF index for each tier is shown in Tables 8-6 and 8-7. Since an item can exhibit DIF for multiple comparisons, the item was counted once if any of the comparisons showed "high" DIF. Tables 8-6 and 8-7 show that only a few items were classified as "high" DIF for each grade and each tier.

Table 8-6. 2017–18 MSAA: Number of Items with "High" DIF by Tier—ELA

Grade	*Tier 1*	*Tier 2*	*Tier 3*	*Tier 4*
3	0(20)	1(14)	2(19)	0(5)
4	1(20)	3(16)	4(14)	0(5)
5	2(13)	2(24)	0(14)	0(5)
6	1(12)	1(21)	2(14)	0(7)
7	1(17)	1(18)	4(18)	0(7)
8	0(16)	0(15)	3(18)	0(5)
11	0(21)	1(19)	0(12)	0(6)

Note: The numbers in the parentheses are the total number of items in each tier.

Table 8-7. 2017–18 MSAA: Number of Items with "High" DIF by Tier—Mathematics

Grade	*Tier 1*	*Tier 2*	*Tier 3*	*Tier 4*
3	1(10)	0(24)	0(22)	0(4)
4	2(8)	2(24)	2(22)	0(4)
5	1(11)	3(25)	0(23)	0(3)
6	1(10)	1(20)	0(19)	0(6)
7	2(10)	3(21)	1(20)	2(6)
8	0(10)	0(20)	1(21)	0(6)
11	0(9)	0(22)	0(20)	0(5)

Note: The numbers in the parentheses are the total number of items in each tier.

8.3 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 are used for calibrating, linking, scaling, and equating the 2017–18 MSAA operational tests.

The purpose of dimensionality analysis is to investigate whether violation of the assumption of test unidimensionality is statistically detectable and, if so, (a) the degree to which unidimensionality is violated and (b) the nature of the multidimensionality. Findings from dimensionality analyses performed on the 2017–18 MSAA operational items for ELA and mathematics 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, Froelich, & Gao, 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 within random noise of zero, indicating statistically independent item responses for examinees with equal expected total test scores. Nonzero conditional covariances are essentially violations of the principle of local independence, and local dependence implies multidimensionality. Thus, nonrandom patterns of positive and negative conditional covariances are indicative of multidimensionality.

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

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

The data used in this year's dimensionality analysis were modified in comparison to data used in past years. In previous years, the dimensionality analyses detected strong violations of local independence, and examinees referred to as "R9-stringers" were found to be the main source of the multidimensionality. Specifically, R9-stringers are students who respond to nine (or more) consecutive multiple-choice items with the exact same option. Data from examinees who were identified as R9-stringers were removed from the dataset used for calibrating the IRT model for the operational items. Because the goal of the dimensionality analysis is to evaluate the assumption of unidimensionality in the IRT model used for the calibration, the data used in the dimensionality analysis also had the R9-stringers' data removed.

DIMTEST and DETECT were separately applied to the three operational paths of each grade on the 2017–18 MSAA ELA and mathematics tests. The three paths resulted in three datasets to be analyzed for each ELA or mathematics grade-level test. Thus, a total of 42 analyses were conducted. 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 490 (grade 7 mathematics, Path A) to a high of 1530 (grade 8 mathematics, Path C). All but one sample size was larger than 600, and the sample sizes were over 1000 in 18 cases. A rough tabulation of the sample size distribution of the 42 datasets is given in Table 8-8.

Table 8-8. 2017–18 MSAA: Dataset Sample Sizes Used for Dimensionality Analyses.

Sample Size	Number of Datasets		
	Path A	Path B	Path C
< 600	1	0	0
600 to 800	4	6	0
800 to 1000	7	3	3
> 1000	2	5	11

DIMTEST was then applied to every dataset. Even though the sample sizes were not large for the MSAA 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 8-9 displays the multidimensional effect size estimates from DETECT.

Table 8-9. 2017–18 MSAA: Average Multidimensional Effect Sizes by Content Area and Grade[1]

Path	Content Area	Grade	Multidimensionality Effect Size	
			2016–17	2017–18
A	ELA	3	1.49	0.52
		4	1.80	0.51
		5	0.93	0.49
		6	1.73	0.40
		7	1.68	0.59
		8	1.71	0.50
		11	2.10	0.40
		Average	1.64	0.49
	Mathematics	3	1.81	0.55
		4	1.30	0.35
		5	1.37	0.56
		6	2.20	0.82
		7	1.95	0.89
		8	0.84	0.41
		11	1.57	0.57
		Average	1.58	0.59
B	ELA	3	0.87	0.37
		4	0.68	0.42
		5	1.49	0.47
		6	0.30	0.45
		7	1.23	0.50
		8	0.40	0.34
		11	0.83	0.39
		Average	0.83	0.42
	Mathematics	3	1.12	0.96
		4	1.12	0.66
		5	1.20	0.75
		6	1.02	0.44
		7	1.75	0.73
		8	1.08	0.68
		11	1.49	0.82
		Average	1.26	0.72

continued

Path	Content Area	Grade	Multidimensionality Effect Size	
			2016–17	2017–18
C	ELA	3	0.18	0.20
		4	0.14	0.22
		5	0.33	0.30
		6	-	0.27
		7	-	0.25
		8	-[2]	0.14
		11	0.20	0.13
		Average	0.21	0.22
	Mathematics	3	0.39	0.45
		4	0.54	0.50
		5	0.41	0.60
		6	0.30	0.24
		7	0.47	0.43
		8	1.00	0.47
		11	0.34	0.29
		Average	0.49	0.43

1.DETECT values for 2016–17 were computed with data that included data from R9-stringers; DETECT values for 2017-18 were computed after the removal of data from R9-stringers.

2. DETECT values not reported for 2016–17 grades 6,7, and 8 ELA Path C because of lack of rejection by DIMTEST hypothesis test.

We now review the results for 2017–18 displayed above in Table 8-9. First, the ELA tests tend to show lower DETECT indices than the mathematics tests. In fact, the ELA Path C tests were the only set of tests for a fixed combination of content area and path that consistently displayed low DETECT indices (weak or very weak multidimensionality) for every test. Even for mathematics, the Path C tests tended to be lower than for Paths A or B, although for this case the DETECT indices were generally moderate (5 moderate, 2 weak/very weak). Summarizing the remaining results in Table 8-9, the DETECT indices for the ELA Paths A and B tests were all at a moderate level; whereas for mathematics, the DETECT indices for Paths A and B were a mixture of moderate and strong values with Path A having more moderate than strong and Path B having more strong than moderate.

For comparison purposes, Table 8-9 also provides the results from last year, 2016–17. In considering the 2016–17 results, the most important factor to bear in mind is that the analyses were conducted on datasets that included the data from R9-stringers. Thus, nearly all the DETECT indices for Paths A and B were larger for 2016–17 (usually much larger) than for 2017–18. The average difference for the Path A tests was about 1.1, and the average difference for the Path B tests was about 0.5. The differences for Path C were negligible, which was understandable since higher performers were routed to take Path C, and it is not the path one

would expect to find students who displayed this repetitive response pattern when placed in this testing situation.

Next, an investigation was conducted to identify the possible source(s) of the violations of local independence as indicated in Table 8-9, especially regarding the tests with the moderate and strong DETECT indices. Hence, we investigated how DETECT divided the tests into clusters to see if there were any discernable patterns with respect to known substantive item characteristics.

In previous years when the data from R9-stringers were included in the data, we found a strong and consistent pattern related to the answer keys of the items: for tests administered on Paths A and B, the placement of the correct-response key option was a strong indicator of the cluster membership of nearly every multiple-choice item. In other words, nearly all the multiple-choice items fell into three clusters, where one cluster was dominated by items with a key of "A" (the first option), another was dominated by items with a key of "B" (the middle option, when it was available), and the third was dominated by items with a key of "C" (the last option). Note that the vast majority of multiple-choice items had only three answer-choice options, and the few items that did not have three options were items that had only two options, which we labeled "A" and "C" for cluster-membership labeling purposes. Because this key-clustering pattern had been so prominent in past years, we looked for it again this year and found that this same pattern again occurred to a significant degree, despite having removed the data from the R-9 stringers from the dataset. Specifically, for ELA nearly every grade-level test for Paths A and B had both an "A" cluster and a "C" cluster, and about half the tests had a "B" cluster. There was no significant difference between Path A and Path B for ELA. The results for the mathematics tests were similar to those for ELA: nearly every grade-level test for Paths A and B had both an "A" cluster and a "C" cluster. However, regarding the "B" cluster, while mathematics was similar to ELA for its Path A tests (about half had a "B" cluster), it differed from ELA for its Path B tests because every Path B mathematics test had a "B" cluster, while only about half the ELA Path B tests had a "B" cluster. This difference is supported by the DETECT indices for which the Path B mathematics tests had the largest values; and, thus, it is not surprising that these tests showed the strongest key-option clustering. For Path C, the results were similar to 2016–17 in that there was no consistent key-option clustering across either the mathematics or ELA grade-level tests.

In regard to the key-option clustering, in comparison to 2016–17, this year's results were weaker in three respects: (1) the DETECT indices were much lower, (2) the presence of "B" clusters was substantially reduced, and (3) the DETECT sign-pattern matrices indicated that the clusters were not as well defined as in 2016–17, with this year's results having much more statistical noise. The weaker strength of the key-option clustering was undoubtedly due to the removal of the R9-stringers' data, and an analysis was conducted with the R9-stringers' data included that verified this assertion.

For the ELA tests, the operational items included a writing prompt for the first time. Most of the writing prompts had been previously field-tested, and dimensionality analyses on the previous field-tests had indicated that we should expect to see the writing prompts contribute to a separate statistically detectable

dimension. Thus, it was not surprising to find that the DETECT clusters consistently indicated a cluster dominated by the writing prompt traits across all grade levels for all three paths. This multidimensionality led to the decision to calibrate the writing prompt traits after first calibrating and equating all the other operational ELA items. In this way the calibration essentially projected the IRT model for the writing prompt traits onto the existing scale.

As in the past years, these dimensionality analysis results for Paths A and B continue to indicate a violation of local independence having to do with how some student scores are related to the placement of the correct response options; however, the violations of local independence are greatly reduced from previous years because of the removal of the R9-stringers from the datasets. In general, it is important that violations of local independence be understood, monitored, and controlled on tests. The violations of local independence that are related to the ordering of the correct-response option in selected-response items are a phenomenon that will continue to require close study.

CHAPTER 9 ITEM RESPONSE THEORY SCALING AND EQUATING

This chapter describes the procedures used to calibrate, equate, and scale the 2017–18 MSAA. Throughout these psychometric analyses, a number of quality-control procedures and checks on the processes were implemented. These procedures included evaluation of item parameters and their standard errors for reasonableness, examining 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, 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 mathematical models to define a relationship between an unobserved measure of student performance, usually referred to as theta (θ), and the probability ($P(\theta)$) of 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 independent measures of the same construct (i.e., of the same θ). Another way to think of θ is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between θ and $P(\theta)$ (Hambleton & Swaminathan, 1985; Hambleton & van der Linden, 1997). The process of determining 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 considered to be an estimate of the student's true score or a general representation of student performance. It has characteristics that may be preferable to those of raw scores for equating purposes because it specifically models examinee responses at the item level, and also facilitates equating to an IRT-based item pool (Kolen & Brennan, 2014).

For the 2017–18 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 indexes 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 indexes 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 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:

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

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:

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

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:

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

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, $E\left(U_i\middle|\theta_j\right) = P_i(\theta_j)$. For more information about item calibration and determination, the reader is referred to Lord and Novick (1968), Hambleton and Swaminathan (1985), or Baker and Kim (2004).

9.2 Calibration Procedure

The 2016–17 MSAA was a pre-equated assessment, but the 2017–18 MSAA was post-equated because of a standards validation conducted after the operational administration. The 2017–18 test was still administered as a multistage test with the pre-equated model used to carry out the routing after the completion of Stage 1. However, the operational scale scores and performance-level classifications for the students were determined after all the items were post-equated and the standards validation was completed. In this section we describe the procedures that were used to conduct the calibrations.

As explained in section 8.3, in 2017–18 an additional culling 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. To repeat the definition, R9-stringers are students who respond to nine (or more) consecutive multiple-choice items with the exact same option. For 2017–18, the data from R9-

stringers were first identified and removed prior to conducting the operational and field-test calibrations. Table 9-1 summarizes the calibration sample sizes both prior to and after removing the data from the R9-stringers.

Table 9-1. 2017–18 MSAA: Summary of Testing Population for 2017–18

Subject	*Grade*	*Total before removing stringers*	*Total after removing stringers*	*Number of Stringers*	*Percent Stringers*
ELA	3	3,556	3,147	409	12
	4	3,866	3,352	514	13
	5	3,986	3,581	405	10
	6	4,080	3,623	457	11
	7	4,088	3,631	457	11
	8	4,144	3,669	475	11
	11	3,876	3,438	438	11
Mathematics	3	3,556	3,024	532	15
	4	3,866	3,242	624	16
	5	3,986	3,352	634	16
	6	4,080	3,631	449	11
	7	4,088	3,549	539	13
	8	4,144	3,611	533	13
	11	3,876	3,463	413	11

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 and every item was carefully examined for model fit. In particular, 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 0.25), 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 0.3.

The equating set (a subset of the operational items) was then carefully chosen to be representative of the test as a whole, and the equating items were evaluated to ensure only psychometrically stable items were used. For any equating design, it is critical that rigorous procedures are implemented to monitor the quality of the equating and check that the assumptions underlying the equating are not violated. Measured Progress psychometricians have conducted research studies (Hagge & Keller, 2009; Keller et al., 2008; Keller et al., 2007; Parker et al., 2009) in this regard and have developed tools to estimate equating error across years

under realistic violations of the equating assumptions. The Psychometrics and Research Department monitors well-known violations of IRT equating assumptions and uses the research to estimate their effects on the reliability and validity of the equating. Specifically, the equating data were analyzed in detail for scale drift through traditional delta analyses and *b*-*b* analyses. The delta analysis converts *p*-values to a type of *z*-score called delta scores using the inverse of the normal cumulative function, followed by a linear transformation to a metric with a mean of 13 and a standard deviation of 4 (Dorans & Holland, 1993; Michaelides, 2003).

The delta analysis then compared the old delta to the new delta using linear regression analysis. A standardized perpendicular difference from the regression line was calculated for each item; any item with a difference of a magnitude of 3 or greater was flagged for drift. The *b*-*b* analyses were similar in nature, with the main difference being that the IRT *b*-parameters are used rather than transformed *p*-values. Furthermore, special procedures were enacted during the calibration phase to check that the quality of the equating items was maintained consistently across years. Equating items that displayed lack of stability (e.g., standard error of the *b* parameters being large, inadequate model-data fit, etc.) were flagged and removed from equating usage. Using this equating set, the Stocking-Lord transformation constants were calculated to determine the relationship between the off-scale calibration and the base-year scale established in the first year of the program. The Stocking-Lord transformation was then applied to all the off-scale operational item parameters to bring them onto the base-year scale.

Next, the field-test items were calibrated. First, an off-scale calibration was conducted on all the operational and field-test items. Then the field-test items were evaluated for model-fit in the same way as described above for the operational items. Based on the model-fit evaluation, the field-test items were classified as either do-not-use (DNU) or use-with-caution (UWC) if any model-fit issues were identified. Items that were not classified as DNU were considered eligible. All items that were not classified as DNU were then brought onto the operational scale using the fixed-common-item-parameter (FCIP) calibration method. In this method, the operational items are first fixed to their on-scale values, and then the field-test items are brought onto the operational scale in a PARSCALE run. After the field-test items were brought onto scale, their model-fit was again evaluated as described above. All items not classified as DNU were then uploaded into the item bank.

9.3 Item Response Theory Results

The tables in Appendix K give the IRT item parameters for all the operational items on the 2017–18 MSAA tests by grade and content area. The statistics for the operational items are summarized in Tables 9-2 through 9-5. The mean item parameter estimates shown in the tables below are within generally acceptable and expected ranges. For easy reference, Table 9-2 displays the means and standard deviations averaged across all dichotomously scored operational items for each grade and content area.

Table 9-2. 2017–18 MSAA: IRT Summary Statistics for Dichotomously Scored Items

Content Area	*Grade*	*Number of Items*	*a*		*b*	
			mean	*SD*	*mean*	*SD*
ELA	3	52	0.71	0.27	-0.58	0.71
	4	52	0.76	0.31	-0.56	0.67
	5	50	0.65	0.32	-0.21	1.12
	6	48	0.88	0.36	-0.07	0.77
	7	54	0.80	0.39	-0.37	0.65
	8	48	0.83	0.37	-0.50	0.58
	11	52	0.91	0.36	-0.55	0.57
Mathematics	3	60	0.80	0.38	-0.07	0.78
	4	58	0.75	0.41	0.31	0.84
	5	62	0.63	0.28	0.40	1.02
	6	55	0.79	0.27	-0.11	0.64
	7	57	0.75	0.29	-0.02	0.74
	8	57	0.76	0.30	-0.09	0.57
	11	56	0.91	0.44	0.19	0.62

Because the items were developed to correspond to different tiers, the item statistics have also been summarized by tier for ELA (Table 9-3 for the dichotomous items and Table 9-4 for the writing prompt traits) and mathematics (Table 9-5).

Table 9-3. 2017–18 MSAA: IRT Summary Statistics by Grade and Tier—ELA Dichotomous Items

Grade	*Tier*	*Number of Items*	*a*		*b*	
			Mean	*SD*	*Mean*	*SD*
3	1	20	0.85	0.33	-1.08	0.47
	2	14	0.7	0.2	-0.18	0.81
	3	19	0.63	0.19	-0.09	0.73
	4	5	0.73	0.28	0.17	0.63
4	1	20	0.85	0.42	-1.34	0.46
	2	16	0.56	0.23	-0.26	0.45
	3	14	0.86	0.33	-0.08	0.64
	4	5	0.91	0.4	0.2	0.55
5	1	13	0.91	0.25	-1.13	0.29
	2	24	0.63	0.32	0.11	0.7
	3	14	0.67	0.34	0.24	0.99
	4	5	0.76	0.38	-0.06	0.46

continued

Grade	Tier	Number of Items	a		b	
			Mean	SD	Mean	SD
6	1	12	1.02	0.41	-0.97	0.32
	2	21	0.77	0.25	-0.07	0.99
	3	14	0.73	0.21	-0.01	0.69
	4	7	0.69	0.44	0.81	0.6
7	1	17	1.15	0.67	-1.07	0.37
	2	18	0.79	0.2	-0.38	0.44
	3	18	0.59	0.22	0.31	0.78
	4	7	0.52	0.2	0.55	0.33
8	1	16	0.92	0.45	-1.07	0.24
	2	15	0.94	0.26	-0.36	0.54
	3	18	0.66	0.2	0.05	0.58
	4	5	0.64	0.3	0.24	0.37
11	1	21	1.03	0.44	-1.11	0.24
	2	19	0.96	0.34	-0.37	0.45
	3	12	0.75	0.28	0.06	0.85
	4	6	0.76	0.13	0.21	0.61

Table 9-4. 2017-18 MSAA: IRT Summary Statistics by Trait and Tier—ELA Writing Prompt Items

Trait	*Tier*	*Number of Items*	*a*		*B*		*d0*		*d1*		*b-d0*		*b-d1*	
			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
	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
	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
	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 9-5. 2017–18 MSAA: IRT Summary Statistics by Grade and Tier—Mathematics

Grade	*Tier*	*Number of Items*	*a*		*b*	
			Mean	*SD*	*Mean*	*SD*
3	1	10	0.87	0.3	-0.87	0.28
	2	24	0.87	0.3	-0.15	0.4
	3	22	0.74	0.29	0.34	0.54
	4	4	0.88	0.17	0.06	0.63
4	1	9	0.64	0.25	-0.95	0.28
	2	23	0.7	0.36	0.46	0.65
	3	24	0.81	0.42	0.3	0.45
	4	3	0.88	0.39	0.47	0.26
5	1	11	0.78	0.23	-0.78	0.23
	2	25	0.57	0.26	0.55	0.5
	3	23	0.63	0.3	0.81	0.64
	4	3	0.76	0.4	0.7	0.56
6	1	10	0.84	0.3	-1.09	0.19
	2	20	0.76	0.26	-0.09	0.39
	3	19	0.82	0.31	0.07	0.45
	4	6	0.71	0.27	0.42	0.47
7	1	10	0.87	0.32	-0.94	0.25
	2	21	0.74	0.27	-0.06	0.38
	3	20	0.73	0.25	0.18	0.3
	4	6	0.84	0.28	0.58	0.48
8	1	10	0.91	0.4	-0.84	0.46
	2	20	0.74	0.22	-0.08	0.26
	3	21	0.74	0.17	0.08	0.31
	4	6	0.75	0.17	0.36	0.38
11	1	9	1.06	0.61	-0.78	0.26
	2	22	1	0.4	0.17	0.33
	3	20	0.78	0.3	0.24	0.29
	4	5	0.75	0.36	0.66	0.33

First, we discuss the results for the dichotomously scored items. We examined the relationship between grade, tier and item difficulty. Item difficulty did not differ significantly by grade level for either ELA or mathematics. On the other hand, item difficulty tends to have a positive relationship with tier; as the tier increases, the items tend to be more difficult (as intended). In all cases, the average difficulty increased

from Tier 1 to Tier 2 and from Tier 2 to Tier 3, but the largest differences were clearly the Tier 1 to Tier 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 tier level as the factor. Separate ANOVAs were run for ELA and mathematics.

The ANOVAs indicated that tier level was statistically significant for both ELA and mathematics with R-squared values of 41.9% for ELA and 40.6% for mathematics. Further Tukey paired-comparison tests were also conducted. These results showed that for both ELA and mathematics, the Tukey tests indicated statistically significant differences between Tier 1 and each of the other tiers. For ELA, the Tukey comparisons for Tier 2 versus Tiers 3 and 4 were also significant. Only the Tier 3 versus Tier 4 difference was not significant. For mathematics, only the Tier 1 differences were statistically significant in the Tukey tests.

Next, we discuss the results for the polytomously scored writing prompt traits. For all three traits, the Tier 3 traits tend to be more difficult than the Tier 2 traits, but the difference is much greater for the Idea Development trait and the Organization trait than for the Conventions trait.

The IRT statistics were also summarized by different paths (Tables 9-6 and 9-7).

Table 9-6. 2017-18 MSAA: IRT Summary Statistics by Grade by Path—ELA Dichotomous Items

Grade	*Path*	*Number of Items*	*a*		*b*	
			Mean	*SD*	*Mean*	*SD*
3	A	38	0.69	0.27	-0.74	0.74
	B	38	0.63	0.26	-0.47	0.73
	C	38	0.68	0.27	-0.46	0.65
4	A	37	0.71	0.27	-0.80	0.57
	B	37	0.73	0.31	-0.66	0.62
	C	37	0.70	0.30	-0.44	0.64
5	A	37	0.65	0.30	-0.39	0.85
	B	37	0.64	0.33	-0.12	1.24
	C	37	0.62	0.31	-0.08	1.24
6	A	38	0.88	0.39	-0.18	0.69
	B	38	0.80	0.33	0.01	0.77
	C	38	0.79	0.32	0.07	0.78
7	A	38	0.91	0.40	-0.58	0.58
	B	38	0.69	0.29	-0.31	0.56
	C	38	0.70	0.29	-0.27	0.60
8	A	38	0.88	0.38	-0.67	0.50

continued

Grade	Path	Number of Items	a		b	
			Mean	SD	Mean	SD
8	B	38	0.84	0.36	-0.53	0.55
	C	38	0.84	0.37	-0.51	0.57
11	A	38	0.95	0.34	-0.73	0.45
	B	38	0.89	0.33	-0.66	0.45
	C	38	0.84	0.32	-0.44	0.60

Table 9-7. 2017-18 MSAA: IRT Summary Statistics by Grade by Path—Mathematics

Grade	Path	Number of Items	a		b	
			Mean	SD	Mean	SD
3	A	35	0.78	0.37	-0.26	0.85
	B	35	0.74	0.37	0.01	0.77
	C	35	0.79	0.36	0.02	0.70
4	A	34	0.77	0.39	0.12	0.91
	B	35	0.67	0.34	0.21	0.95
	C	35	0.72	0.36	0.28	0.77
5	A	35	0.63	0.31	0.06	1.12
	B	35	0.57	0.24	0.54	0.94
	C	35	0.65	0.23	0.44	0.79
6	A	35	0.73	0.26	-0.37	0.60
	B	35	0.76	0.25	-0.13	0.59
	C	35	0.84	0.25	0.02	0.63
7	A	35	0.69	0.25	-0.22	0.75
	B	35	0.68	0.23	-0.01	0.75
	C	35	0.79	0.27	0.03	0.56
8	A	35	0.76	0.31	-0.26	0.60
	B	35	0.74	0.23	-0.04	0.44
	C	35	0.77	0.24	-0.03	0.48
11	A	35	0.87	0.45	0.06	0.72
	B	35	0.83	0.35	0.24	0.63
	C	35	0.96	0.38	0.21	0.43

The average item difficulty substantially increased from Path A to Path B for all tests, as intended. While difficulty also usually increased from Path B to Path C, in most cases the difference was negligible.

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 -4.0 and 4.0. 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 -4 to 4), 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 do 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:

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

Compared to the tails, TIFs are often higher near the middle of the θ distribution where most students generally are located and where most items are sensitive by design. Appendix L shows graphs of the TCCs and TIFs for each grade and content area.

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.

While the 2016–17 MSAA was a pre-equated assessment, the 2017–18 MSAA was post-equated because of a standards validation that was conducted after the operational administration. The 2017–18 test

was still administered as a multistage test with the pre-equated model used to carry out the routing after the completion of Stage 1. However, the operational scale scores and performance-level classifications for the students were determined after all the items were post-equated and the standards validation was completed. In this section we describe the procedures that were used to conduct the equating of the operational items. In addition to these operational items, new non-operational items (field-test items) were also included on the forms. The operational items were used as a set of common items for transforming the item parameters of the nonoperational 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.

For any equating design, it is critical that rigorous procedures are implemented to monitor the quality of the equating and check that the assumptions underlying the equating are not violated. The equating data is analyzed in detail for scale drift through traditional delta analyses and *b*-*b* analyses. The delta analysis converts *p*-values to a type of *z*-score called delta scores using the inverse of the normal cumulative function, followed by a linear transformation to a metric with a mean of 13 and a standard deviation of 4 (Dorans & Holland, 1993; Michaelides, 2003). The delta analysis then compares the old delta to the new delta using linear regression analysis. A standardized perpendicular difference from the regression line is calculated for each item; any item with a difference of a magnitude of 3 or greater is flagged for drift. The *b*-*b* analyses are similar in nature, with the main difference being that the IRT *b*-parameters are used rather than transformed *p*-values.

Item parameter estimates for the 2017–18 MSAA were placed on the base-year scale (i.e., the item bank scale) by using the method of Stocking and Lord (1983), which is based on the IRT principle of item parameter invariance. According to this principle, the equating items for both the base year and current year tests should have the same item parameters. After the item parameters for each current year's test are estimated using PARSCALE (Muraki & Bock, 2003), the Stocking and Lord method is employed to find the linear transformation (slope and intercept) that adjusts the equating items' parameter estimates such that the current year's test characteristic curve (TCC) for the equating items is as close as possible to that of the base year's tests.

In addition, the calibrated and equated parameters are evaluated to further investigate drift at both the item and test levels. At the item level, the individual item parameters are compared and investigated, and at the test level the TCC, test information function (TIF), and raw score cuts are compared. Finally, the operational item parameters resulting from this process are updated in the item bank, and these updated parameters are used in field-test calibrations and in future test form development.

Item parameters for the 2017–18 operational administration were calibrated after the 2017–18 MSAA operational administration.

9.5 MSAA Performance Levels, Cut Scores, and Standards Validation

Cut scores for MSAA in ELA and mathematics were originally set in a standard setting 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, Measured Progress 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 necessary because of the addition of the open-response writing prompt scores to the existing ELA score scale.

The ELA standards validation process involved five steps: (a) vertically articulating the performance level 3 cut scores for both ELA and mathematics, to update the performance standards and provide a coherent basis for interpreting 2018 scores and performance and validating the ELA cut scores; (b) expanding the ELA PLDs to include references to direct open-response writing prompt knowledge and skills; (c) expanding the existing ELA score scale by linking the open-response writing prompt scores to the scale; (d) reviewing and possibly adjusting the existing cut scores to align the response demands of all items, including the open-response writing prompt scores, and knowledge and skill requirements in the expanded PLDs; and (e) reviewing the expanded ELA PLDs with stakeholders from various schools and districts to confirm the writing knowledge and skills were clear and included language around the open-response writing prompt expectations in each of the performance levels. Members of the MSAA Psychometric Subcommittee, including MSAA ELA Content Specialists, comprised the ELA standards validation panel. In addition, two members of the TAC attended to monitor the process and provide advice and support as needed. Panelists recommended no adjustments to cut scores for grades 5, 7, 8, and 11. They recommended small adjustments as follows: in grade 3 performance level 4, from scale score 1251 to 1254; in grade 4 performance level 4, 1258 to 1259; and in grade 6 performance level 4, 1253 to 1251. Resulting differences in impact data are modest. The upward adjustments to the performance level 4 cut score in grades 3 and 4 decreased the percentages of students in that level by 5.5% (grade 3) and 3.0% (in grade 4); the downward adjustment to the performance level 4 cut score in grade 6 increased the percentage of students in that level by 2.6% (in each case, the percentages of students changed in performance level 3 by the amount of change in performance level 4). These cut scores for mathematics and ELA were brought by each state before their respective Board

of Education/Superintendents. Each state individually obtained approval of the cut scores, which constituted collective approval by MSAA on August 16, 2018. A complete description of the standards articulation and validation processes appears in the 2018 MSAA Standards Validation Report (see Appendix M).

Final cut scores, after mathematics and ELA vertical articulation and ELA standards validation for the 2018 MSAA, appear in Table 9-8.

Table 9-8. 2017-2018 MSAA: Cut Scores on the Theta Metric and Reporting Scale

Content Area	*Grade*	*Theta*				*Scale Score*			
		Cut1	*Cut2*	*Cut3*	*Minimum*	*Cut1*	*Cut2*	*Cut3*	*Maximum*
ELA	3	-0.70	-0.22	0.98	1200	1234	1240	1254	1290
	4	-0.53	-0.01	1.53	1200	1234	1240	1259	1290
	5	-0.84	-0.13	1.16	1200	1232	1240	1256	1290
	6	-0.63	-0.11	1.03	1200	1231	1237	1251	1290
	7	-0.59	-0.20	0.95	1200	1236	1240	1255	1290
	8	-0.75	-0.16	0.78	1200	1230	1238	1250	1290
	11	-0.77	-0.41	0.90	1200	1236	1240	1255	1290
Mathematics	3	-0.70	-0.17	0.77	1200	1235	1242	1254	1290
	4	-0.64	-0.10	0.82	1200	1232	1239	1251	1290
	5	-0.76	-0.11	0.85	1200	1232	1240	1253	1290
	6	-0.68	-0.21	0.72	1200	1233	1239	1251	1290
	7	-0.75	-0.25	0.77	1200	1234	1240	1254	1290
	8	-0.66	-0.22	0.63	1200	1234	1240	1251	1290
	11	-0.61	-0.23	0.54	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 within each grade and content area are provided in the table.

Table 9-9. 2017-2018 MSAA: Percentage of Students by Performance-Level Categories

Content Area	*Grade*	*Number of Students*	*Levels*					*Average Scale Score*	*SD of Scale Score*
			Level 1	Level 2	Level 3	Level 4	Levels 3 & 4		
ELA	3	2882	36.81	12.49	31.51	19.19	50.70	1241.27	14.15
	4	3095	34.38	18.45	36.54	10.63	47.17	1240.65	13.88
	5	3280	26.52	24.27	34.63	14.57	49.20	1241.25	13.66
	6	3354	16.61	24.39	40.52	18.49	59.01	1240.74	11.76
	7	3259	28.17	16.02	34.61	21.20	55.81	1243.57	12.40
	8	3364	23.25	24.82	30.56	21.37	51.93	1239.29	12.12
	11	3117	23.84	13.70	42.19	20.28	62.47	1244.46	11.84

continued

Content Area	Grade	Number of Students	Levels					Average Scale Score	SD of Scale Score
			Level 1	Level 2	Level 3	Level 4	Levels 3 & 4		
Mathematics	3	2858	26.24	27.33	33.21	13.23	46.44	1242.16	12.84
	4	3222	18.62	28.06	39.42	13.90	53.32	1240.06	11.94
	5	3245	21.17	26.44	34.36	18.03	52.39	1242.00	13.86
	6	3434	25.19	19.74	31.83	23.24	55.07	1242.10	13.33
	7	3224	17.62	30.06	35.30	17.03	52.33	1242.96	12.40
	8	3395	26.80	20.41	30.66	22.12	52.78	1241.73	13.69
	11	3129	12.11	25.47	46.09	16.33	62.42	1242.69	10.21

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 increased as we move from Path A to Path C. This trend was expected due to the stage adaptive nature of the 2017–18 MSAA.

Table 9-10. 2017–18 MSAA: Performance-Level Distributions by Path—ELA

Grade	Path	Number of Students	Levels					Average Scale Score	SD of Scale Score
			Level 1	Level 2	Level 3	Level 4	Levels 3 & 4		
3	A	1,118	83.99	11.63	4.38		4.38	1228.76	6.81
	B	702	16.52	27.64	53.42	2.42	55.84	1240.10	6.03
	C	1,062	0.56	3.39	45.57	50.47	96.04	1255.20	10.55
4	A	1,195	80.33	17.66	2.01		2.01	1228.37	6.88
	B	985	10.46	35.13	53.81	0.61	54.42	1241.11	6.30
	C	915	0.11	1.53	63.06	35.30	98.36	1256.20	10.40
5	A	1,319	63.99	30.71	5.31		5.31	1229.54	6.39
	B	833	2.88	41.54	54.98	0.60	55.58	1240.85	5.71
	C	1,128	0.18	3.99	53.90	41.93	95.83	1255.23	10.73
6	A	1,149	46.74	42.04	11.23		11.23	1230.37	5.91
	B	1,202	1.58	26.96	68.47	3.00	71.47	1239.88	5.34
	C	1,003	0.10	1.10	40.58	58.23	98.81	1253.65	9.89
7	A	938	82.84	15.03	2.13		2.13	1230.65	5.91
	B	1,004	13.45	36.35	49.60	0.60	50.20	1240.76	5.47
	C	1,317	0.46	1.21	46.32	52.01	98.33	1254.91	8.90
8	A	1,171	63.02	34.84	2.13		2.13	1227.83	5.77
	B	829	5.19	45.24	47.89	1.69	49.58	1237.03	5.12

continued

Grade	Path	Number of Students	Levels					Average Scale Score	SD of Scale Score
			Level 1	Level 2	Level 3	Level 4	Levels 3 & 4		
8	C	1,364	0.07	3.81	44.43	51.69	96.12	1250.50	8.72
11	A	950	67.05	23.89	9.05		9.05	1232.87	6.50
	B	756	13.23	24.21	62.30	0.26	62.56	1241.03	4.65
	C	1,411	0.43	1.20	53.72	44.65	98.37	1254.09	8.94

Table 9-11. 2017–18 MSAA: Performance-Level Distributions by Path—Mathematics

Grade	Path	Number of Students	Levels					Average Scale Score	SD of Scale Score
			Level 1	Level 2	Level 3	Level 4	Levels 3 & 4		
3	A	921	71.44	27.36	1.19		1.19	1230.47	7.59
	B	982	9.37	51.12	38.59	0.92	39.51	1240.81	4.70
	C	955		2.83	58.53	38.64	97.17	1254.83	11.01
4	A	1,061	50.33	39.77	9.90		9.90	1230.05	8.71
	B	1,253	5.27	38.07	55.31	1.36	56.67	1239.23	4.79
	C	908		0.55	51.98	47.47	99.45	1252.90	10.12
5	A	899	67.74	31.03	1.22		1.22	1228.25	7.53
	B	1,238	6.30	43.70	48.30	1.70	50.00	1239.65	5.52
	C	1,108		3.43	45.67	50.90	96.57	1255.77	11.55
6	A	1,444	58.66	30.82	10.39	0.14	10.53	1231.46	6.63
	B	742	2.43	29.51	63.88	4.18	68.06	1241.20	4.82
	C	1,248		1.12	37.58	61.30	98.88	1254.93	11.31
7	A	605	75.54	22.81	1.65		1.65	1229.98	7.32
	B	1,541	7.20	52.89	39.26	0.65	39.91	1239.29	4.91
	C	1,078		1.48	48.52	50.00	98.52	1255.50	11.13
8	A	1,111	77.05	19.71	3.24		3.24	1228.83	7.96
	B	711	7.03	56.26	36.43	0.28	36.71	1239.08	3.98
	C	1,573	0.25	4.70	47.43	47.62	95.05	1252.04	11.20
11	A	797	42.66	51.82	5.52		5.52	1233.29	7.54
	B	1,198	3.26	31.72	63.94	1.09	65.03	1240.54	3.56
	C	1,134		0.35	55.73	43.92	99.65	1251.58	9.40

9.6 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 are developed such that they range from 1200 through 1290 for all grade/content-area combinations. The second cut 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 above.

By providing information that is more specific 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 2017–18 MSAA tests were translated to scale scores using a data analysis process called *scaling*. Scaling simply converts from one scale to another scale. In the same way that a given temperature can be expressed on either Fahrenheit or Celsius scales, or the same distance can be expressed in either miles or kilometers, student scores on the 2017–18 MSAA tests can be expressed in raw or scale scores.

It is important to note that converting from raw scores to scale scores does not change students' performance-level classifications. Given the relative simplicity of raw scores, it is fair to question why scale scores for MSAA are reported instead of raw scores. Scale scores make for more consistent reporting of results. The psychometric advantage of scale scores over raw scores is that scale scores are linear transformations of θ. 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 θ 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 linear equation:

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

where
m is the slope, and
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 9-12 shows the slope and intercept terms used to calculate the scale scores for each content area and grade. Note that the values in Table 9-12 will not change unless the standards are reset.

Table 9-12. 2017–18 MSAA: Scale score Slope and Intercept by Content Area and Grade

Content Area	*Grade*	*Slope*	*Intercept*
ELA	3	11.720	1242.054
	4	12.059	1240.091
	5	12.424	1241.615
	6	12.352	1237.813
	7	12.296	1242.433
	8	12.608	1239.457
	11	11.492	1244.224
Mathematics	3	13.055	1243.665
	4	13.100	1239.867
	5	13.077	1241.410
	6	12.820	1241.253
	7	12.909	1243.244
	8	13.021	1242.358
	11	12.990	1242.480

Appendix N contains raw score to scale score lookup tables for the 2017–18 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 2017–18 MSAA tests and the last two years are presented in Appendix N. The cumulative graphs show the proportion of students at or below each scale score.

CHAPTER 10 RELIABILITY

Although an individual item's performance is 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; that is, no measurement is perfect. 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. When tests have a high amount of measurement error, student scores are very unstable. Students with high ability may get low scores or vice versa. Consequently, one cannot reliably measure a student's true level of ability with such a test. 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."

There are a number of ways to estimate an assessment's reliability. The most common method is Cronbach's alpha, which assumes that all the students for a given assessment were administered the same fixed form. For the 2017–18 MSAA, there were three different paths (A, B, and C), each of which essentially corresponded to a different test form. Even though Cronbach's alpha could be applied to each form separately, this is not ideal for two reasons. First, the ability distributions for the three forms are very different from each other by design—essentially the standard deviation for any one form is much smaller than the standard deviation for the whole population, and the mean increases from paths A to B to C. The resulting restriction of ability range causes severe underestimation of reliability for each path. Second, we want a single measure of reliability for each grade-level assessment, not three values. Thus, we turned to an IRT-based formulation of reliability that results in a single value for each grade-level assessment.

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:

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

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 directly examining 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, the values all reached levels associated with adequate reliability (0.85 or more).

Table 10-1. 2017–18 MSAA: IRT Marginal Reliability by Grade—ELA

Grade	*Min*	*Max*	*Mean Scale Score*	*IRT Marginal Reliability*	*Scale Score SD*	*Mean Scale CSEM*
3	1200	1290	1241.27	0.91	14.15	4.13
4	1200	1290	1240.65	0.91	13.88	4.21
5	1200	1290	1241.25	0.90	13.66	4.29
6	1200	1290	1240.74	0.90	11.76	3.65
7	1200	1290	1243.57	0.90	12.40	3.89
8	1200	1290	1239.29	0.90	12.12	3.77
11	1200	1290	1244.46	0.90	11.84	3.65

Table 10-2. 2017–18 MSAA: IRT Marginal Reliability by Grade—Mathematics

Grade	*Min*	*Max*	*Mean Scale Score*	*IRT Marginal Reliability*	*Scale Score SD*	*Mean Scale CSEM*
3	1200	1290	1242.16	0.89	12.84	4.27
4	1200	1290	1240.06	0.87	11.94	4.26
5	1200	1290	1242.00	0.88	13.86	4.89
6	1200	1290	1242.10	0.90	13.33	4.24
7	1200	1290	1242.96	0.88	12.40	4.25
8	1200	1290	1241.73	0.90	13.69	4.38
11	1200	1290	1242.69	0.87	10.21	3.64

10.2 SUBGROUP RELIABILITY

The reliability coefficients discussed in the previous section were based on all students who took a particular 2017–18 MSAA test. Appendix P 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 where more than 25% of the students scored above the LOSS (lowest obtainable scale score, which was 1200).

For several reasons, the results relating to subgroup reliability should be interpreted with caution. First, inherent differences between tests preclude making valid inferences about the quality of a test based on statistical comparisons with other tests. Second, reliabilities are 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 P 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 for subgroups with little variability (Draper & Smith, 1998). Third, there is no industry standard to interpret the strength of a reliability coefficient. This is particularly true when the population of interest is a single subgroup. Again, the reliability statistics provided in the tables in Appendix P should be cautiously interpreted because of the restriction of range mentioned earlier (Section 8.1).

10.3 Reliability of Performance-Level Categorization

While related to reliability, the accuracy and consistency of classifying students 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, empirical 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. Accuracy must be estimated because errorless test scores do not exist. 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 2017–18 MSAA because it can be easily applied to data that is 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.

Readers are referred to Rudner (2001, 2005) for details of the Rudner method; here we briefly review the basic idea behind the method. Using an examinee's estimated scale score and standard error, assuming a normal probability distribution, the method first calculates for all examinees at a fixed value of true scale score, the expected proportion whose observed scale score is in an interval [a,b]. Then, by summing over all examinees whose true scale scores are in an interval [c,d], the method yields the expected proportion of all examinees whose true scale score is in [c,d] and whose observed scale score is in [a,b]. By setting [a,b] and [c,d] to correspond to the true score intervals defined by the cut scores yields the elements of a classification

table that shows the expected proportion of all examinees with observed and true scale scores in each cell. These proportions can then be used to calculate both classification accuracy and classification consistency estimates.

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 exactly 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.

10.3.1 Accuracy and Consistency

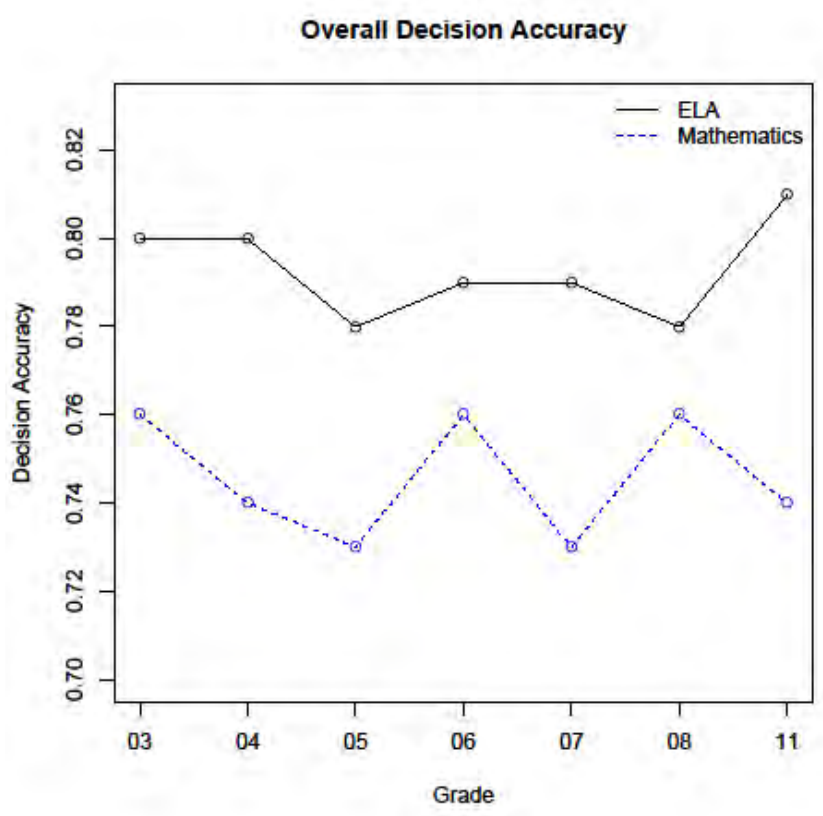
Figure 10-1 shows the overall decision accuracy for ELA and mathematics by grade level. More details on decision accuracy and consistency (DAC) are provided in Appendix Q. Table Q-1 in Appendix Q includes overall accuracy and consistency indices, along with kappa. Accuracy and consistency values conditional on performance level are also provided in Table Q-1. For these calculations, the denominator is

the proportion of students associated with a given performance level. We'll take an example from Table P-1, looking at Level 1 for grade 3 ELA.

- The conditional *accuracy* value was 0.87. This indicates that among the students whose *true scale scores* placed them in Level 1, 87% would be expected to be in this same level again when categorized according to their observed scale scores.
- The *consistency* value was 0.85. This indicates that among the students whose *observed scale scores* placed them in Level 1, 85% 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 around level thresholds. For example, in testing done for No Child Left Behind accountability purposes, the primary concern is distinguishing between students who are proficient and those who are not yet proficient. For the 2017–18 MSAA, Table Q-2 in Appendix Q provides accuracy and consistency estimates at each cutpoint, as well as false positive and false negative decision rates. A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.

Figure 10-1. 2017–18 MSAA: Overall Decision Accuracy by Content Area by Grade



CHAPTER 11 VALIDITY

Chapter 1 presents the primary intended score interpretations and three primary intended score uses. This chapter elaborates upon the assumptions that underlie these four score interpretations and uses (SIUs), elements of each assumption, and the evidence that supports each element.

11.1 PRIMARY INTENDED SCORE INTERPRETATION

The primary intended score interpretation is that the MSAA scores provide valid information about the grade-level literacy and numeracy that students with the most significant cognitive disabilities are attaining.

Assumption 1.1. The Content of the Test Represents the Content of the Standards (i.e., the Core Content Connectors)

- **Element 1.1.1.** The CCCs are aligned to the states' general education standards.
- **Element 1.1.2.** The 2018 MSAA items are aligned to the CCCs.

  Evidence: The evidence in support of these elements was generated in a series of alignment studies that were conducted between 2012 and 2015. Table 3-2 in section 3.2, Alignment and Linkages, summarizes each study, telling when the study was conducted and the alignment question that was addressed in each study. Details regarding the alignment reports and evidence of findings is available in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* (see http://ncscpartners.org/Media/Default/PDFs/Resources/NCSC15_NCSC_TechnicalManualNarrative.pdf).

- **Element 1.1.3**. The CCCs are the states' adopted alternate academic achievement standards for the AA-AAAS.

  Evidence: Membership in MSAA requires each member state to adopt the alternate academic achievement standards for the AA-AAAS.

Assumption 1.2. MSAA test items are construct relevant.

The elements corresponding to this assumption are concerned with the skills and cognitive processes required to understand and respond to an item, in particular, whether they correspond to the skills and processes required in the PLDs.

- **Element 1.2.1.** Items require application of the KSAs of the targeted construct.

- **Element 1.2.2.** Items are accessible for all students.
- **Element 1.2.3.** Appropriate accommodations are provided to meet student needs.
- **Element 1.2.4.** Item scoring rubrics and criteria focus on construct relevant aspects of student response.
- **Element 1.2.5.** Scaffolding is not a source construct-irrelevant variance.
- **Element 1.2.6.** Item rendering does not interfere with students' access to test content.
- **Element 1.2.7**. Platform does not interfere with students' ability to interact with test content.
- **Element 1.2.8**. Items are free of bias and sensitive issues.

Evidence: Depending on the element, two to five of the following strands of evidence are relevant here:

Alignment Studies (Element 1.2.1). Two of the alignment studies referred to above (under Assumption 1.1) are also applicable here: The Task/Item Alignment Study and the Item Mapping Study, both conducted in 2015.

Item Reviews (Elements 1.2.1, 1.2.2, 1.2.3, 1.2.4, 1.2.5, 1.2.8). During the item development process, the items followed a rigorous development cycle, including reviews by MSAA State Leads and by Item Content and Bias and Sensitivity panelists. See Chapter 3 for a detailed description of item review process.

Cognitive labs (Elements 1.2.1, 1.2.2, 1.2.3, 1.2.4, 1.2.6, 1.2.7). Cognitive labs (also referred to as Student Interaction Studies) were conducted by NCSC in the early stages of development of the assessment program to acquire detailed information about the cognitive processes used by students in responding to assessment tasks. The studies resulted in adjustments in the assessment program to ensure the construct validity of student response data. This information is outlined in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual*.

Usability studies (Elements 1.2.6 and 1.2.7). Three usability studies were conducted by NCSC in the early stages of development of the assessment program to evaluate how students and teachers interacted with items and gathered evidence related to item complexity and usability. The studies resulted in adjustments to ensure the assessment met all usability standards required to support the validity of the assessment program. This information is outlined in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual*.

Universal Design Principles (Elements 1.2.1, 1.2.2, 1.2.3). As mentioned in Chapter 4, Test Development, the item development and review process includes evaluation for adherence to the principles of Universal Design. The evaluation necessarily involves close collaboration between content specialists and special education specialists. The use of UD principles helps ensure the assessment is accessible for the widest range of student participation, providing access to the general education curriculum while also fostering higher expectations for

students with the most significant cognitive disabilities. Additional details about Universal Design can be found in the *National Center and State Collaborative 2015 Operational Assessment Technical Manual*.

Accommodations Policy (Elements 1.2.1, 1.2.2, 1.2.3). The allowable accommodations for the MSAA are defined as changes in the standard administration of the assessment that do not alter the construct being measured. Any accommodation required by a student must be included in the student's IEP prior to testing and should be used regularly during instruction. The allowable accommodations are listed and explained in Chapter 5. The Test Administration Manual (TAM) supplies still further detail about the implementation of the MSAA accommodations policies.

Observation checklists (Element 1.2.3). Observers were sent into the field to observe test administration and fill-in an observation checklist. The checklists include evidence as to whether an accommodation was properly implemented for appropriate students. The reports by the observers were a new feature of the 2017–18 test and have not yet been compiled into a document.

Differential item functioning (DIF) analyses (Element 1.2.8). Subgroup differences in performance are examined when sample sizes permit and actions are taken to ensure that differences in performance are due to construct-relevant, rather than irrelevant, factors. A detailed description of the DIF analysis procedures is given in Chapter 8 along with a summary of the results. Detailed results are presented in Appendix J.

Assumption 1.3. Test administration procedures in 2018 were sound.

- **Element 1.3.1**. Test administrators and School and District Coordinators understood and performed their roles appropriately.

Chapter 5, Training and Administration, provides detailed evidence in regard to ensuring the test administrators and test coordinators properly understood and performed their roles.

Evidence: Required training for test administrators.

Six online training modules address the specific responsibilities of the test administrators and provide information from the three documents they were required to use: Test Administrator Manual (TAM), the Directions for Test Administration (DTA), and the MSAA Online Assessment System User Guide for Test Administrators. After completing the training modules, test administrators were required to successfully complete a final quiz with a score of 80% or better.

Required training for test coordinators. Six online training modules address the responsibilities of the test coordinators. Test coordinators are also provided the following supporting documents: TAM, DTAs, the MSAA Online Assessment System User Guide for

Test Administrators, and the MSAA Online Assessment System User Guide for Test Coordinators.

Best-practice videos. Four best-practice videos are also provided to the test administrators.

MSAA Technical Support. A technical support chart provides examples of when and who to contact to obtain answers in regard to MSAA assessment or administration.

End-of-Test Survey. Each test administrator completes a survey. Results are evidence that address this element.

All the above evidence is described in detail in Chapter 5.

Observation checklists. Observers were sent into the field to observe test administration and fill-in an observation checklist. The checklists and any accompanying notes provide evidence as to whether the training was effectively followed by the test administrators and test coordinators. The reports by the observers were a new feature of the 2017–18 test and have not yet been compiled into a document.

- **Element 1.3.2**. Test security concerns were limited.

Chapter 5, section 5.6.11, Test Security and Test Irregularities, provides detailed evidence in regard to ensuring that test security policies and practices resulted in limited test security concerns.

Evidence:

Irregularity reports. Testing irregularities are defined in the TAM (p. 25) and what constitutes a testing irregularity is defined in the training. Summary reports were organized at the school, district, and state levels and included information on administration irregularities.

Service Center records. Before, during, and after testing, the service center operated to receive, respond to, and track reported issues, including routing issues to appropriate people for resolution. All activity was tracked and included in weekly status reports.

Monitoring activities. Numerous measures taken to ensure proper testing procedures and appropriate test practices were maintained during test administration. These measures are detailed in Section 5.7.20, "Monitoring and Quality Control."

Observation checklists. Observers were sent into the field to observe test administration and fill-in an observation checklist. The checklists provide evidence as to whether secure administration protocol was followed and whether secure storage of the testing materials occurred. The reports by the observers were a new feature of the 2017–18 test and have not yet been compiled into a document.

Assumption 1.4. Test scores on the 2018 MSAA are sound.

- **Element 1.4.1.** MSAA scores and categorizations into proficiency levels are adequately reliable for their intended purpose.

  Evidence:

  Internal consistency. Chapter 10 provides a description of reliability theory and interpretation, a review of the relevant equations, and a summary of the results. In particular, the reliability estimates can be interpreted as the correlation that would be obtained between scaled scores on two parallel forms.

  Scaled score standard errors. Chapter 9 provides a description of calculation and interpretation of the scaled scores, as well as a description of the calculation of the standard error for a scaled score. The average standard error for a reported scaled score is reported in Chapter 10. The scaled score standard error can be compared to the scaled score range and the scaled score standard deviation to provide some context for interpretation.

  Performance Level Classification consistency and accuracy estimates. Accuracy is an estimate of the probability that the observed classification is the true classification. Consistency is an estimate of the probability that students would receive the same classification if they tested twice on parallel forms. Chapter 11 describes the theory and equations underlying the estimation of classification accuracy and consistency, while also reporting summary statistics. Detailed results are provided in Appendix N.

- **Element 1.4.2.** Item characteristics are adequate for intended interpretations about all students who take MSAA.

  Evidence:

  Item parameter estimates. The psychometric characteristics most pertinent to evaluating the adequacy of individual items are the estimated item parameters. The item parameter estimates are summarized in tables in Chapter 9. For dichotomously scored items, the item parameters include the discrimination parameter and the difficulty level parameter. For polytomously scored items, namely the writing traits, the item parameter estimates include the discrimination parameter, the overall difficulty level parameter, and the step difficulty parameters for each of the possible non-zero scores. All the statistics for all the items are provided in Appendix I.

- **Element 1.4.3.** Appropriate test characteristics (Path A, Path B, Path C).

  Evidence:

  Dimensionality. Dimensionality analysis was conducted on each Path for each grade-level test. Section 8.3 gives a detailed description of the hypothesis testing and effect size estimation methods. Results are summarized in a table accompanied by a description of the results. Small to moderate violations of local independence were noted, and interpretations of these results were presented.

Test Information Function. Chapter 9 provides a detailed description of the psychometric model that was fitted to the data. In particular, it describes the test information function (TIF), the most pertinent product of the psychometric model in regard to evaluating the adequacy of the test. Appendix L shows the TIF graphs for all three paths for all the MSAA tests. By examining the value of TIF at the performance level cuts (given in the graphs), the psychometric appropriateness of each test can be evaluated.

- **Element 1.4.4.** Scaling is appropriate.

  Evidence:

  Differential item functioning (DIF) analyses. The scale used for reporting scores is assumed to be measuring only those constructs that are intended to be measured by each test. DIF analyses were conducted to detect items that may be measuring construct-irrelevant variance. Subgroup differences in item-level performance are examined when sample sizes permit. If an item is flagged, appropriate actions are taken to investigate whether the differences in performance are due to construct-irrelevant factors. A detailed description of the DIF analysis procedures is given in Chapter 8 along with a summary of the results. Detailed results are presented in Appendix J.

  Dimensionality. The scale used for reporting scores is a unidimensional scale. Dimensionality analysis was conducted on each Path for each grade-level test to examine the degree to which unidimensionality is evident. When the null hypothesis of unidimensionality is reject, the dimensionality analysis quantifies the violation of unidimensionality and attempts to describe what may be causing the violation. Section 8.3 gives a detailed description of the hypothesis testing and effect size estimation methods. Results are summarized in a table accompanied by a description of the results. Small to moderate violations of local independence were noted, and interpretations of these results were presented.

  Calibration. The unidimensional scale used for reporting scores is based on an underlying unidimensional IRT model. The initial form of the IRT model is established by an initial calibration of the item response data. The calibration must be conducted accurately in order for the scaling to be appropriately implemented. Section 9.2 provides evidence that can be used to evaluate the effectiveness of the calibration. The evidence provided for the calibration procedure includes discussion of the removal of stringers and a description of how the convergence of the statistical calibration was evaluated.

  Model fit. After the initial calibration converged to a solution, the fit of the model was evaluated. Section 9.2 described how model fit was evaluated and the criteria that were used.

- **Element 1.4.5.** Equating is appropriate.

  Evidence:

  Evaluation of equating items. The appropriateness of the equating is contingent upon the substantive and statistical quality of the equating items. Section 9.2 gives a detailed

description of the procedures used to ensure the quality of the equating items, in terms of both content representativeness, as well as statistical stability.

Third-party analysis. As a check on our equating procedures, a third-party vendor is contracted with to also conduct the equating analysis. The third-party results are not included in the technical report. A separate report is written up by the third party and is available upon request. The third-party results for the 2017–18 were essentially identical to the results reported in the technical report.

- **Element 1.4.6.** Stage 1 covers a broad enough range to route students into appropriate stage 2 levels.
- **Element 1.4.7.** Stage 2 routing is appropriate for students.
- **Element 1.4.8.** Stage 2 levels are sufficiently separable and targeted toward different ranges of achievement.

Evidence:

Test construction process. The evidence most pertinent to the stages is the report *MSAA Test Construction Process for 2019* (which also reflects the process and criteria for the 2018 MSAA). This detailed report describes how item and test information is targeted for the various stages, including the determination of the routing rules. A brief description is given in Chapter 4. Refer to the full report for a more detailed description.

Performance Level Distributions by Path. The test-level statistical results in the technical report are focused on the test as a whole. Thus, the statistical results are focused on Paths A, B, and C, rather than on the stages. Still, some of the path results are especially pertinent to evaluating the psychometric characteristics of the stages. In particular, Section 9.5 reports the Performance Level Distributions across the different paths. If the stages are properly constructed and the routing is properly implemented, the performance level distributions should differ across paths in reasonable ways. The results presented in section 9.5 can be evaluated in this regard.

Assumption 1.5. Item and test scoring in 2018 were implemented soundly.

- **Element 1.5.1.** Assurances that machine scored items scored accurately.

Evidence:

Key verification process. As mentioned in Chapter 6, all the operational multiple-choice items are examined prior to score reporting to ensure that the option that was designated as the key was indeed the correct response.

- **Element 1.5.2.** Constructed response item scoring training and monitoring procedures met industry standards.

Evidence:

Administration training and monitoring procedures. Section 6.2, Scoring Processes and Rules for Operational Writing Open-Response Items, is predominantly devoted to describing all the procedures that are used to ensure the accuracy of the scoring for the constructed response items, including: administrator training and monitoring (6.2.1.2), benchmarking and identification of scoring materials (6.2.1.3), scorer recruitment and qualifications (6.2.2), scoring leadership (6.2.3), qualification (6.2.4), specific scoring rules to ensure accuracy (6.2.5), monitoring of quality control (6.2.6), quality reports (6.2.7), and inter-rater reliability (6.2.8).

Assumption 1.6. MSAA scores correlate appropriately with external indicators of student proficiency (i.e., concurrent evidence).

- **Element 1.6.1.** MSAA scores correlate as expected with other measures of student proficiency.

  Evidence: This evidence does not exist; this element and assumption are not supported. Evidence to support this assumption could include the consistency of MSAA scores with classroom achievement indicators (e.g., classroom grades, teacher evaluations of student proficiency) and other external measures.

11.1.1 Primary Intended Score Use 1

The MSAA and its results will be used to help schools and LEAs (a) monitor trends in student performance, and (b) design professional development for teachers

Assumption 2.1. MSAA scores enable teachers and school, district, and state leaders to monitor trends in student proficiency.

- **Element 2.1.1.** MSAA scale scores for groups of students are adequately reliable and valid to help school, district, and state leaders monitor changes in means, standard deviations, and proficiency level percentages for classroom, school, district, and state groups.

  Evidence: Evidence for the soundness (reliability and validity) of the scores and the corresponding scoring processes is presented above under Assumptions 1.4 and 1.5. Assumption 1.6, having to do with external validity, currently lacks supporting evidence.

- **Element 2.1.2.** MSAA scores and proficiency level categorizations of groups of students are adequately reliable and valid to enable monitoring of grade level performance and student cohort performance.

Evidence: Evidence for the soundness (reliability and validity) of the proficiency level categorizations is presented above under Assumption 1.4, with the most pertinent evidence being as follows:

Performance Level Classification consistency and accuracy estimates. Accuracy is an estimate of the probability that the observed classification is the true classification. Consistency is an estimate of the probability that students would receive the same classification if they tested twice on parallel forms. Chapter 11 describes the theory and equations underlying the estimation of classification accuracy and consistency, while also reporting summary statistics. Detailed results are provided in Appendix N.

Performance Level Distributions by Path. Section 9.5 reports the Performance Level Distributions across the different paths. If the stages are properly constructed and the routing is properly implemented, the performance level distributions should differ across paths in reasonable ways. The results presented in section 9.5 can be evaluated in this regard.

Assumption 1.6. having to do with external validity, currently lacks supporting evidence.

Assumption 2.2. Results can be used to design professional development for teachers.

- **Element 2.2.1.** MSAA and its results can be used to help schools and LEAs design professional development for teachers.

  Evidence: States provide guidance to local districts to promote and guide development of teacher professional development. For example, the Arizona Department of education provides a document called *How to Teach the State Standards to Students Who Take Alternate Assessments* https://cms.azed.gov/home/GetDocumentFile?id=5866dbe1aadebe085c4de5b4

11.1.2 Primary Intended Score Use 2

The MSAA and its results can be used to help teachers to integrate MSAA scores and other information with their instructional planning.

Assumption 2.1. Teachers find MSAA scores and other information useful for instructional planning.

- **Element 2.1.1.** Teachers find the performance level descriptors and their students' performance levels useful for planning instruction, especially students in levels 1 and 2.

  Evidence: Annual compliance monitoring of IEPs in all states indicates that special education teachers refer to PLDs to establish present levels of performance and to inform

goals. For example, the Arizona Department of Education guidance on IEP required components requires that "The IEP includes measurable annual goals, including academic and functional goals that reflect the needs identified in the PLAAFP and current assessment data" (p. D40; see https://cms.azed.gov/home/GetDocumentFile?id=5b2a897d1dcb250f1c55e5b3).

- **Element 2.1.2.** Teachers find their students' scale score information useful for planning instruction, especially students in levels 1 and 2.

  Evidence: No systematic evidence exists.

Assumption 2.2. Teachers use MSAA scores and other information for instructional planning.

- **Element 2.2.1.** Teachers use MSAA scores and other information for planning instruction.

  Evidence: Annual compliance monitoring of IEPs in all states indicates that special education teachers refer to PLDs to establish present levels of performance and to inform goals. For example, the Arizona Department of Education guidance on IEP required components requires that "The IEP includes measurable annual goals, including academic and functional goals that reflect the needs identified in the PLAAFP and current assessment data" (p. D40; see https://cms.azed.gov/home/GetDocumentFile?id=5b2a897d1dcb250f1c55e5b3).

11.1.3 Primary Intended Score Use 3

The MSAA and its results can be used to give parents information about (a) what their child knows and can do, and (b) their child's progress from year to year

Assumption 3.1. Parents find MSAA scores and other information useful for understanding what their child knows and can do.

- **Element 3.1.1.** Parents understand and interpret correctly MSAA scores and other information to understand what their child knows and can do.

  Evidence:

  MSAA provides information to guide parents in interpreting and using MSAA scores and other information about their child's achievement and learning needs. For example, the Arizona Education Department sends to districts a *Parent Overview* to accompany each child's Individual Score Report. The overviews are available online in both English and Spanish; see http://www.azed.gov/assessments/parents/. Similarly the Maryland State Department of Education provides *Curriculum and Instruction Resources for Families*; see

http://marylandpublicschools.org/programs/Documents/Special-Ed/IEP/CurriculumInstructionalResourcesFamilies.pdf.

- **Element 3.1.2.** Parents use MSAA scores and other information appropriately to understand what their child knows and can do.

  Evidence:

  MSAA provides information to guide parents in interpreting and using MSAA scores and other information about their child's achievement and learning needs. For example, the Arizona Education Department sends to districts a *Parent Overview* to accompany each child's Individual Score Report in both English and Spanish; see http://www.azed.gov/assessments/parents/.

Assumption 3.2. Parents find MSAA scores and other information useful for understanding their child's progress from year to year.

- **Element 3.2.1.** Parents understand and interpret correctly MSAA scores and other information to understand their child's progress from year to year.

  Evidence:

  MSAA provides information to guide parents in interpreting and using MSAA scores and other information about their child's achievement and learning needs. For example, the Arizona Education Department sends to districts a *Parent Overview* to accompany each child's Individual Score Report in both English and Spanish; see http://www.azed.gov/assessments/parents/.

- **Element 3.2.2.** Parents use MSAA scores and other information appropriately to understand their child's progress from year to year.

  Evidence:

  MSAA provides information to guide parents in interpreting and using MSAA scores and other information about their child's achievement and learning needs. For example, the Arizona Education Department sends to districts a *Parent Overview* to accompany each child's Individual Score Report in both English and Spanish; see http://www.azed.gov/assessments/parents/.

REFERENCES

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

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

Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Baker, F. B., & Kim, S-H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York: Marcel Dekker, Inc.

Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth: Holt, Rinehart, and Winston.

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

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.

Gulliksen, H. (1950). *Theory of Mental Tests*. New York, NY: John Wiley and Sons.

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.

Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). New York, NY: Routledge.

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.[4]

Michaelides, P. M. (2003). *Sensitivity of IRT equating to the behavior of test equating item.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Muraki, E., & Bock, R. D. (2003). PARSCALE 4.1 [Computer software]. Lincolnwood, IL: Scientific Software International.

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.

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.

Samejima, F. (1994). Estimation of reliability coefficients using the test information function and its modifications. *Applied Psychological Measurement, 18,* 229–244.

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

Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, *52*, 589–617.

Stout, W. F., Froelich, A. G., & Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duign, & T. A. B. Snijders (Eds.), *Essays on item response theory* (pp. 357–375). New York, NY: Springer-Verlag.

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.

[4] See chapter 7: http://ncscpartners.org/Media/Default/PDFs/Resources/NCSC15_NCSC_TechnicalManualNarrative.pdf

Yan, D., von Davier, A. A., & Lewis, C. (Eds.). (2014). Computerized multistage testing: Theory and applications. Boca Raton: CRC Press.

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—LIST OF ACRONYMS

Table A-1. 2017–18 MSAA: Terms and Acronyms

TERMS AND ACRONYMS USED IN THE MSAA 2017-2018 TECHNICAL REPORT	
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
AD	assistant director
AERA	American Educational Research Association
ANOVA	analysis of variance
APA	American Psychological Association
APIP	Accessible Portable Item Protocol
CCC	Core Content Connector
CCSS	Common Core State Standards
CNMI	Commonwealth of Northern Mariana Islands
CR	constructed response
CSEM	conditional standard error of measurement
DAC	decision accuracy and consistency
DETECT	Dimensionality Evaluation to Enumerate Contributing Traits
DIF	differential Item functioning
DIMTEST	computer program used by Measured Progress
DNU	do not use
DOK	depth of knowledge
DTA	Directions for Test Administration
ELA	English language arts
EOTS	end-of-test survey
ESR	early stopping rule
ESSA	Every Student Succeeds Act
EU	essential understanding
FCIP	fixed common item parameter
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
IDEA	Individuals with Disabilities Education Act
IEP	individualized education program
IRC	Item Review Committee (includes Content Review Committee and Bias-Sensitivity Review Committee)
IRT	Item Response Theory
ISR	Individual Student Report
IT	information technology

continued

TERMS AND ACRONYMS USED IN THE MSAA 2017-2018 TECHNICAL REPORT	
IUA	Interpretation/use argument
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
NCSC	National Center and State Collaborative
OR	open response
PAC	Pacific Assessment Consortium
PAC-6	Guam, the Commonwealth of Northern Mariana Islands, South Dakota, Tennessee, the U.S. Virgin Islands, and Washington D.C.
PARSCALE	Item response theory (IRT) software program that can perform item analysis and test scoring for dichotomous and polytomous IRT models
PLD	performance level descriptor
R9-stringer	student who responds to nine (or more) consecutive multiple-choice items with the exact same option
RF	reading foundation
SD	standard deviation
SEM	standard error of measurement
SES	socioeconomic status
SPED	Special Education
SR	selected response
SRC	student response check
SRR	School Roster Report
SSR	School Summary Report
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
USVI	U.S. Virgin Islands
UWC	use with caution

APPENDIX B—ACCOMMODATION FREQUENCIES

Table B-1. 2017–18 MSAA: Accommodation Frequencies—Mathematics

Accommodations	*Grades*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
LCI_Vision[1]	184	183	174	179	167	191	164
SAR_Assistive_Presentation_After[2]	-	-	-	-	-	-	-
SAR_Assistive_Response_After[2]	276	299	285	397	327	310	324
SAR_No_Accomm_Needed_After[3]	506	630	723	744	851	921	971
SAR_Paper_Version_After[4]	214	218	189	171	123	136	140
SAR_Scribe_After[5]	959	1012	1043	1001	885	856	514
SAR_Sign_Interpretation_After[6]	27	58	41	37	30	21	30

[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 B-2. 2017–18 MSAA: Accommodation Frequencies—ELA

Accommodations	*Grades*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
LCI_Vision[1]	185	184	172	180	167	194	165
SAR_Assistive_Presentation_After[2]	-	-	-	-	-	-	-
SAR_Assistive_Response_After[2]	277	302	281	397	325	312	325
SAR_No_Accomm_Needed_After[3]	502	628	721	745	847	923	971
SAR_Paper_Version_After[4]	216	219	189	173	124	139	142
SAR_Scribe_After[5]	960	1013	1042	1002	884	855	513
SAR_Sign_Interpretation_After[6]	27	59	41	37	30	21	31

[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 B-3. 2017–18 MSAA: Accommodation Summary

Content Area	*Grade*	*Number of Students Tested*	
		With	*Without*
ELA	3	1,673	1,129
	4	1,866	1,254
	5	1,973	1,203
	6	2,012	1,328
	7	1,955	1,412
	8	2,014	1,403
	11	1,795	1,329
Mathematics	3	1,675	1,129
	4	1,866	1,262
	5	1,980	1,210
	6	2,011	1,336
	7	1,962	1,421
	8	2,010	1,403
	11	1,793	1,335

APPENDIX C—PARTICIPATION RATES

Table C-1. 2017–18 MSAA: Summary of Participation by Demographic Category—Mathematics

Description	*Tested*		*Total Tested*	*Total Percent*
	# Complete	*# No Observable Mode of Communication*[1]		
All Students	21,472	921	22,393	100.00
Female	7,477	388	7,865	35.12
Male	13,946	530	14,476	64.65
Gender Undefined	49	3	52	0.23
Hispanic or Latino	4,522	210	4,732	21.13
American Indian or Alaska Native	705	45	750	3.35
Asian	600	16	616	2.75
Black or African American	5,177	176	5,353	23.90
Native Hawaiian or Pacific Islander	104	6	110	0.49
White (non-Hispanic)	9,562	429	9,991	44.62
Two or More Races (non-Hispanic)	614	17	631	2.82
No Primary race/Ethnicity Undefined	188	22	210	0.94
Currently receiving LEP services	805	12	817	3.65
Not receiving LEP services	8,780	455	9,235	41.24
LEP: All Other Students	11,887	454	12,341	55.11
Economically Disadvantaged Students	7,126	238	7,364	32.89
Non-economically Disadvantaged Students	6,994	289	7,283	32.52
SES: All Other Students	7,352	394	7,746	34.59
Migrant	5	0	5	0.02
Non- migrant	8,656	443	9,099	40.63
Undefined Migrant Status	12,811	478	13,289	59.34
Augmentative Communication	3,653	297	3,950	17.64
No Augmentative Communication	17,708	621	18,329	81.85
Undefined Augmentative Communications	111	3	114	0.51
Hearing Loss	569	169	738	3.30
Within Normal Limits	20,814	751	21,565	96.30
Undefined Hearing Loss	89	1	90	0.40
Visual Impairment	882	360	1,242	5.55
Within Normal Limits	20,451	556	21,007	93.81
Undefined Visual Impairment	139	5	144	0.64

continued

Description	Tested		Total Tested	Total Percent
	# Complete	# No Observable Mode of Communication[1]		
Sensory Stimuli Response	1,492	751	2,243	10.02
Follow Directions	19,971	170	20,141	89.94
Undefined Receptive Language	9	0	9	0.04
Special School	2,278	217	2,495	11.14
Regular School Self-contained	13,633	659	14,292	63.82
Regular School Resource Room	3,442	32	3,474	15.51
Regular School Primarily Self-contained	1,512	7	1,519	6.78
Regular School General Education	598	6	604	2.70
Undefined Classroom Setting	9	0	9	0.04
Student Communicates Primarily Through Cries	1,209	710	1,919	8.57
Uses Intentional Communication	4,475	166	4,641	20.73
Uses Symbolic Language	15,779	45	15,824	70.66
Undefined Expressive Communication	9	0	9	0.04

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

Table C-2. 2017–18 MSAA: Summary of Participation by Demographic Category—ELA

Description	*Tested*		*Total Tested*	*Total Percent*
	# Complete	*# No Observable Mode of Communication[1]*		
All Students	21,423	923	22,346	100.00
Female	7,468	390	7,858	35.17
Male	13,906	530	14,436	64.60
Gender Undefined	49	3	52	0.23
Hispanic or Latino	4,497	211	4,708	21.07
American Indian or Alaska Native	707	45	752	3.37
Asian	601	16	617	2.76
Black or African American	5,156	176	5,332	23.86
Native Hawaiian or Pacific Islander	105	6	111	0.50
White (non-Hispanic)	9,562	429	9,991	44.71
Two or More Races (non-Hispanic)	609	18	627	2.81
No Primary race/Ethnicity Undefined	186	22	208	0.93
Currently receiving LEP services	802	12	814	3.64
Not receiving LEP services	8,770	455	9,225	41.28
LEP: All Other Students	11,851	456	12,307	55.07
Economically Disadvantaged Students	7,116	238	7,354	32.91
Non-economically Disadvantaged Students	6,988	289	7,277	32.57
SES: All Other Students	7,319	396	7,715	34.53
Migrant	5	0	5	0.02
Non- migrant	8,649	443	9,092	40.69
Undefined Migrant Status	12,769	480	13,249	59.29
Augmentative Communication	3,655	298	3,953	17.69
No Augmentative Communication	17,658	622	18,280	81.80
Undefined Augmentative Communications	110	3	113	0.51
Hearing Loss	570	170	740	3.31
Within Normal Limits	20,763	752	21,515	96.28
Undefined Hearing Loss	90	1	91	0.41
Visual Impairment	887	360	1,247	5.58
Within Normal Limits	20,396	558	20,954	93.77
Undefined Visual Impairment	140	5	145	0.65

continued

Description	Tested		Total Tested	Total Percent
	# Complete	# No Observable Mode of Communication[1]		
Sensory Stimuli Response	1,485	753	2,238	10.02
Follow Directions	19,929	170	20,099	89.94
Undefined Receptive Language	9	0	9	0.04
Special School	2,273	217	2,490	11.14
Regular School Self-contained	13,597	661	14,258	63.81
Regular School Resource Room	3,436	32	3,468	15.52
Regular School Primarily Self-contained	1,510	7	1,517	6.79
Regular School General Education	598	6	604	2.70
Undefined Classroom Setting	9	0	9	0.04
Student Communicates Primarily Through Cries	1,200	712	1,912	8.56
Uses Intentional Communication	4,470	166	4,636	20.75
Uses Symbolic Language	15,744	45	15,789	70.66
Undefined Expressive Communication	9	0	9	0.04

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

Table C-3. 2017–18 MSAA: Participation Rates by Subgroup

Description	Total Tested	Invalidated	Did Not Test
ELA	22,346	167	1,348
Mathematics	22,393	140	1,327

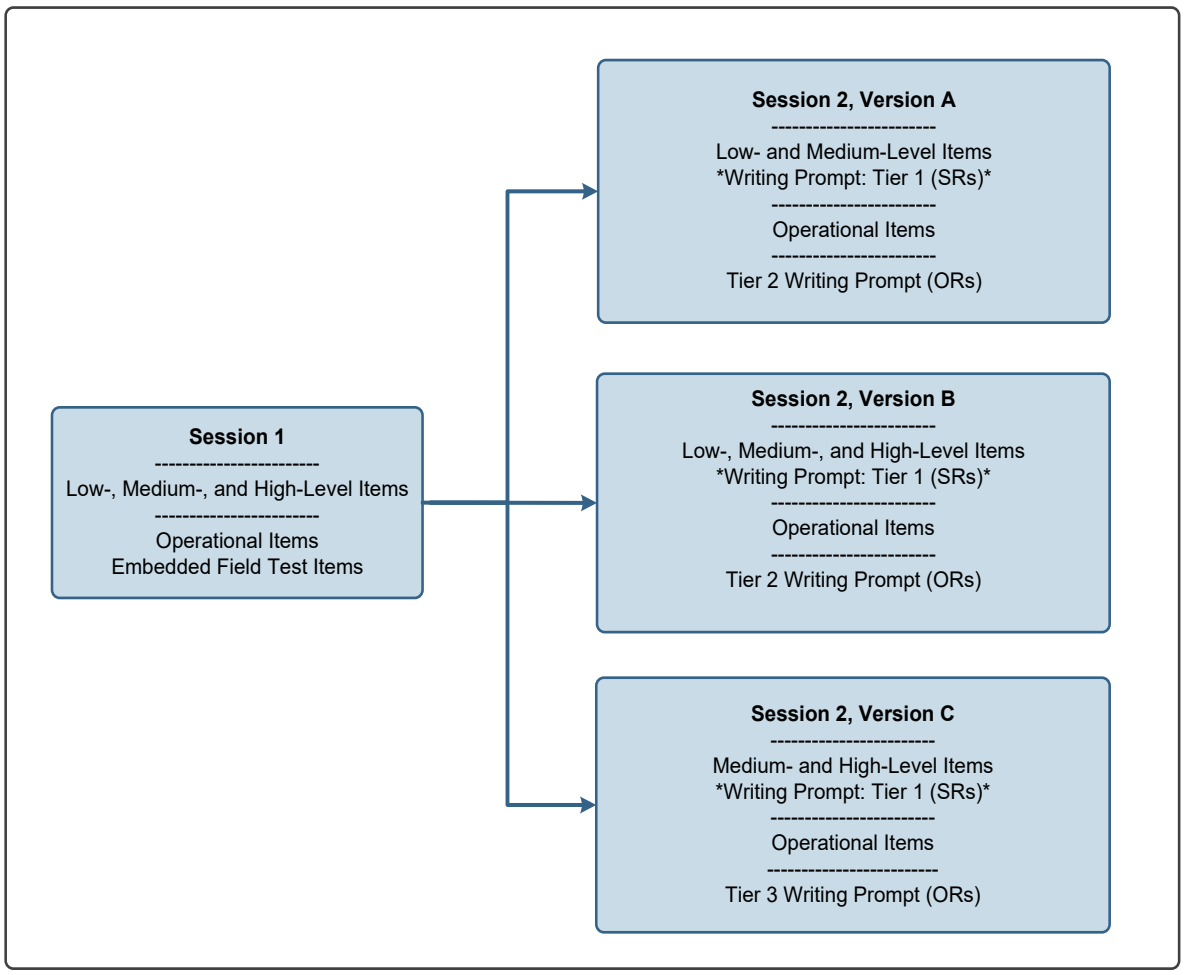
APPENDIX D—TEST BLUEPRINTS

2018-2019 MSAA Test Design and Blueprint

I. Overview of Test Design

- Three Forms in Session 1
- Operational Test selection will use Stat data instead of Tiers for Test Construction.
- Items used to generate student scores are from the 2015 -2018 administrations
- Item sets may overlap within Session 2Refresh rate 33% for both Math & ELA
- All Field Test items will appear in Session 1

  - 30 FT items in Math (10 FT in each form)
  - ~27-30 FT items in ELA (~9-10 FT in each form) (Each grade must be consistent in # of FT items per form but can vary across grades)



Guidelines for MSAA Content Category Distribution

(Number of items)

Content Category	GR 3	GR 4	GR 5	GR 6	GR 7	GR 8	GR 11
Operations and Algebraic Thinking	10-11	10-11	3-4				
Number and Operations Base Ten	7	3-4	14				
Number and Operations Fractions	7	10-11	7				
Measurement and Data	7	7	7				
Geometry	3-4	3-4	3-4	3-4	7	10-11	3-4
Ratio and Proportions				10-11	14		
Expressions and Equations				7	3-4	7	
The Number System				10-11	7	3-4	
Statistics and Probability				3-4	3-4	7	7
Functions						7	
Algebra and Functions							17-18
Number and Quantity							7
TOTAL	**35**	**35**	**35**	**35**	**35**	**35**	**35**

Guidelines for MSAA CR items

(Number of items)

Grade	3	4	5	6	7	8	11
Number of CR items	1 – 2	2 – 4	2 – 3	1 – 2	1 – 2	1 – 2	1 – 2

Note: MSAA would like to see more CRs developed for Math.

Field Test Positions: 5, 6, 9, 10, 14, 15, 19, 20, 24, 25

II. Blueprint Guidelines ELA

When the item pool allows, these are the blueprint guidelines that will inform test construction.
*Chart reflects inclusion of 3 Operational Foundational Items in grades 3 and 4
The ELA chart below specifies where the Operational passage sets, operational writing items, and field test slots will be in the sessions.

ELA Content Category	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	Gr 11
Reading Literary	29-38%	32%	34%	45%	28-38%	36%	33-36%
Reading Informational	22-31%	32%	39%	25%	21-31%	36%	26-28%
Reading Vocabulary and Foundational (G3 and G4)	17%	17%	5%	8%	11-13%	10%	8-13%
Writing	26%	20%	21%	23%	28%	26%	28%

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.
    - Goal is to move toward:
        - 2A: difficulty range-low
        - 2B: difficulty range-medium
        - 2C: difficulty range-high
- Writing standalones in Session 1.
- Writing Prompt-SRs (Tier 1) in Session 2 versions.
- Writing Prompt-OR Tier 2 in Session 2A and B versions. Writing Prompt -OR Tier 3 in Session 2C versions.
- FT Foundational items are added to Session 1, Form 1 for grades 3 and 4. Grades 3-8 & 11 will have FT writing items, L1 Writing Prompts or shortened passage set.
- Item Types:

SR- Independent item that is not connected to any other items
Two-Part SR- answering one item is not dependent on answering the previous item. Students could reference the previous item without impacting their score.
MSR- indicates dependency (EBSR). Students should not be able to reference previous item because the answer to it is in the second item.

	Session One Gr 3& 4 OP Points* = 21 Gr 5-8 &11 OP Points* = 18 FT Points** = ~8-10 Difficulty Range-low to high (with focus on medium)	Session Two A OP Points –SRs = 18-20 OP Points–OR = 0-9 Difficulty Range-low	Session Two B OP Points –SRs = 18-20 OP Points–OR = 0-9 Difficulty Range-medium	Session Two C OP Points –SRs = 18-20 OP Points–OR = 0-9 Difficulty Range-high
Form 1	**Passage set**	**Passage set**	**Passage set**	**Passage set**
	Passage set	**Passage set**	**Passage set**	**Passage set**
	Passage set	**Passage set**	**Passage set**	**Passage set**
	Gr 3&4 Foundational Items only			
	3-4 Writing Standalones	**Writing Prompt - SRs (Tier 1)**	**Writing Prompt - SRs (Tier 1)**	**Writing Prompt - SRs (Tier 1)**
	FT passage set 1	**Writing Prompt – OR (Tier 2)**	**Writing Prompt – OR (Tier 3)**	**Writing Prompt – OR (Tier 3)**
	FT Foundational Items for grades 3 & 4 FT writing stand-alone items OR shortened passage set grade 5-8 & 11 OR Level 1 WP			
Form 2	**Passage set**	same as above	same as above	same as above
	Passage set			
	Passage set			
	Gr 3&4 Foundational Items only			
	3-4 Writing Standalones			
	FT passage set 2			
	FT writing stand-alone items OR shortened passage set OR Level 1 WP			
Form 3	**Passage set**	same as above	same as above	same as above
	Passage set			
	Passage set			
	Gr 3&4 Foundational Items only			
	3-4 Writing Standalones			
	FT passage set 3			
	FT writing stand-alone items OR shortened passage set OR Level 1 WP			

*Session 1 will contain three operational passage sets. They will all be the same regardless of form.
**Field test slots are color-coded peach, orange, and green.

Options

1. Add 3 foundational items to operational tests and increase length of test

Remove items from Info or Literary passage sets and keep overall t

2018-2019 MSAA Math Blueprint

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

Grade 3 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 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	2019 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 x 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 or 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	2019 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	2019 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 fractions or decimals	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	2019 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	2019 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

Grade 9 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 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

2018-2019 MSAA ELA Blueprint

Notes:

- Measured Progress psychometricians have analyzed each proposed test as a whole to show how well they differentiate between stages 2A, 2B, and 2C using IRT stats.
    - Goal is to move toward:
        - 2A: difficulty range-low
        - 2B: difficulty range-medium
        - 2C: difficulty range-high
- Overlapping passage sets will occur in Session 2A, B & C, but they will vary based on how well they differentiate based on IRT stats
- Writing standalones are positioned in Session 1 representing varied levels of difficulty.
- Tier 1 writing prompts are positioned in Session 2 and are common across all forms.
- Writing Prompt-OR Tier 2 in Session 2A and B versions. Writing Prompt -OR Tier 3 in Session 2C versions.
- Operational Foundational items are positioned in Session 1 for grades 3 and 4.
- Session 1 tests and Form 2A will have a less challenging passage set in the first position.
- Item Types:
  SR- Independent item that is not connected to any other items
  Two-Part SR- answering one item is not dependent on answering the previous item. Students could reference the previous item without impacting their score.
  MSR- indicates dependency (EBSR). Students should not be able to reference previous item because the answer to it is in the second item.

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	Gr 11
Reading Literary	29-38%	32%	34%	45%	28-38%	36%	33-36%
Reading Informational	22-31%	32%	39%	25%	21-31%	36%	26-28%
Reading Vocabulary and Foundational (G3 and G4)	17%	17%	5%	8%	11-13%	10%	8-13%
Writing	26%	20%	21%	23%	28%	26%	28%

*Chart reflects inclusion of 3 Operational Foundational Items in grades 3 and 4

Grade 3 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 Item Range
Reading: Literary	29-38%	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
		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	22-31%	3.RI.h1** Identify the purpose of a variety of text features NOT 2-PART	SR	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	17%	3.RWL.i2 Use sentence context as a clue to the meaning of a new word, phrase, or multiple meaning word	SR	7
		3.RWL.i1 Use context to confirm or self-correct word recognition.	SR	
Writing	26%	3.WI.l4 Sort evidence (e.g., graphic organizer) collected from print and/or digital sources into provided categories	SR	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	6 MSR 1 CR
Total	100%			42

* Percentages are approximate with the total equaling 100%
** CCCs require a multipart item to assess

Grade 4 Targets by Standards

Content Category	Weight	Core Content Connector	Item Type	2019 Item Range
Reading: Literary	32%	4.RL.i1 Refer to details and examples in a text when explaining what the text says explicitly	SR	13
		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	32%	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	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	17%	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	7
		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	20%	4.WI.q1 Provide a concluding statement or section to support the information presented	SR	3
		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	4-6 MSR 1 CR
Total	100%			41

* Percentages are approximate with the total equaling 100%
** CCCs require a multipart item to assess.

Grade 5 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 Item Range
Reading: Literary	34%	5.RL.b1 Refer to details and examples in a text when explaining what the text says explicitly	SR	12-13
		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	39%	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	15
		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	5%	5.RWL.a2 Use context to determine the meaning of unknown or multiple meaning words or phrases	SR	2
Writing	21%	5.WI.b3 Organize ideas, concepts, and information (using definition, classification, comparison/contrast, and cause/effect)	SR	3
		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	4-6 MSR 1 CR
Total	100%			38

* Percentages are approximate with the total equaling 100%
** CCCs require a multipart item to assess.

Grade 6 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 Item Range
Reading: Literary	45%	6.RL.b2 Refer to details and examples in a text when explaining what the text says explicitly	SR	18
		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	25%	6.RI.b4 Summarize information gained from a variety of sources including media or texts	SR	9-10
		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	8%	6.RWL.a1 Use context to determine the meaning of unknown or multiple meaning words or phrases	SR	2-3
		6.RWL.c1 Use general academic and domain specific words and phrases accurately	SR	
Writing	23%	6.WL.c1 Organize ideas and event so that they unfold naturally	SR	3
		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	4-6 MSR 1 CR
Total	100%			39

* Percentages are approximate with the total equaling 100%
** CCCs require a multipart item to assess.

Grade 7 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 Item Ranges
Reading: Literary	28-38%	7.RL.i2** Use two or more pieces of textual evidence to support inferences, conclusions, or summaries of text	SR, SR two-part	11-15
		7.RL.j1 Analyze the development of the theme or central idea over the course of the text	SR	
Reading: Informational	21-31%	7.RI.j1** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR, SR two-part	8-12
		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	11-13%	7.RWL.g1 Use context as a clue to determine the meaning of a grade appropriate word or phrase	SR	4-5
Writing	28%	7.WL.o1 Select or provide a concluding statement or paragraph that follows from the narrated experiences or events.	SR	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	6 MSR 1 CR
Total	100%			39

* Percentages are approximate with the total equaling 100%
** CCCs require a multipart item to assess.

Grade 8 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 Item Ranges
Reading: Literary	36%	8.RL.i2** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR, SR two-part	14
		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	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	14
		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	10%	8.RWL.g1 Use context as a clue to the meaning of a grade-appropriate word or phrase	SR	4
		8.RWL.i1 Use general academic and domain specific words and phrases accurately	SR	
Writing	26%	8.WP.k2 Create an organizational structure in which ideas are logically grouped to support the writer's claim	SR	3
		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	6 MSR 1 CR
Total	100%			39

* Percentages are approximate with the total equaling 100%
** CCCs require a multipart item to assess.

Grade 11 Targets by Standard

Content Category	Weight	Core Content Connector	Item Type	2019 Item Ranges
Reading: Literary	33-36%	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	13-14
		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	26-28%	1112.RI.b1** Use two or more pieces of evidence to support inferences, conclusions, or summaries or text	SR, SR two-part	9-11
		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	8-13%	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	3-5
		1112.RWL.c3 Develop and explain ideas for why authors made specific word choices within text	SR	
Writing	28%	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	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	6 MSR 1 CR
Total	100%			39

* Percentages are approximate with the total equaling 100%

** CCCs require a multipart item to assess.

APPENDIX E—ITEM REVIEW, BIAS/SENSITIVITY REVIEW, AND TECHNICAL ADVISORY COMMITTEE MEMBERS

Table E-1. 2017-18 MSAA: Technical Advisory Committee Members

Name	*Organization*	*Expertise*
Derek Briggs	University of Colorado	Assessment / Growth / Psychometrics
Joseph Martineau	The National Center for the Improvement of Educational Assessment	Psychometrics / Computer Adaptive Testing
Rachel Quenemoen	National Center on Educational Outcomes	Students with Significant Cognitive Disabilities / NCSC Awareness
Michael Russell	Boston College	Technology / Accessibility
Martha Thurlow	University of Minnesota / NCEO	Special Education / Accessibility

MSAA 2018 Item Content and Bias Review Meeting—Final Panelist List

ELA Content Grades 3-6	
Name	**State**
Alison Wilhelm	TN
Clarissa Wright	DC
Elsbeth Falk	SD
Shelly Bohy	SD
Rhonda Gross	PAC6
Amy Ashline	AZ
Melissa Adams	ME
Sabre Aldrette	MT

Math Content Grades 3-6	
Name	**State**
Elmie G. Manley	PAC6
Laura Prullage	DC
Heather Hinners	SD
Christy Callahan	ME
Lizabeth B. Hofschneider	PAC6
Lisa Oliver	AZ
Mark Dennett	ME
Darla Stone	MT
Krista Bolen	TN
Sara Kempler	MD

ELA Content Grades 7,8,11	
Name	**State**
Melody Maitland	DC
Leslie Hoffman	TN
Tracy Lynn Del Rosario	PAC6
Kayla Bucciarelli	MD
Jami Kesling	SD
Raquel Payton	AZ
Brittany Garst	MD
Mechelle Ganglfinger	ME
Heidi Foreman	MT
Abigail Trask	ME

Math Content Grades 7,8,11	
Name	**State**
Eric Hoffman	TN
Erin Stabnow	SD
Alexis Dion	AZ
Fasefulu Tigilau	PAC6
Jenny Zephier	SD
Nicole Hash	MT
Jules O'Herron	AZ
Sarah Stare	MD
Kaitlyn Dove	MD

ELA Bias All Grades	
Name	**State**
Denise Johnson	TN
Valerie T. Guerrero	PAC6
Tammi Waltjer-Haverly	SD
Shelby Thibodeau	ME
Harvey Hart	AZ
Mary Brewer	TN
Bess Cropper	MD
Lora Travers Moncure	ME
Nicole Greenplate	MD
Gaye McNeil	MT

Math Bias All Grades	
Name	**State**
Helene S. A. Cruz	PAC6
Becky Whitlock	SD
Sarah Mester	MD
Jennifer Lowe	MD
Johanna Connell	ME
Melanie Home Gun	MT
Rebecca Coons	AZ
Maureen Fox	ME
Tabatha King	ME
Sherry Kelley	MT

MSAA 2018 Accessibility, Bias and Sensitivity Passage Review—Final Panelist List

ELA Grades 3-5 Reading-Writing	
Name	**State**
Elsbeth Falk	SD
Donald Drake	ME
Machelle Enright	AZ
JJ Walker	AR
Cathy Dunnigan	MD
Jill Bahti	AZ
Ashley Anders	AR
Lesa Warrick	DC

ELA Grades 6-7 Reading-Writing	
Michelle Bohy	SD
Alyssa Rasnick	AR
Sara Kempler	MD
Jami Kesling	SD
Christy Callahan	ME
Nicole Ugel	DC
Katie DiTullio	AZ
Joe Benamati	MD

ELA Grades 8,11 Reading-Writing	
Name	**State**
Erin Stabnow	SD
Lakeya Keynerd	DC
Gerald Neal	AZ
Meredith Verrill	ME
Johanna Connell	ME
Gina Wood	MD
Diane Shifflet	AZ
Sherema Copes	DC

APPENDIX F—IT-REPORTING PROCESSING AND REPORTING BUSINESS REQUIREMENTS DOCUMENTS

Analysis and Reporting Decision Rules

Multi-State Alternate Assessment (MSAA)			
Spring 2018 Administration			
Prepared Date: March 5, 2018		**Updated Date:** October 17, 2018	
Version Number	**Date**	**Updated Content Description**	**Updated By Name**
1.0	03/01/2018	Initial Document	Keira Nevers
1.1	03/07/2018	Review with Reports Subcommittee	Keira Nevers
2.0	04/12/2018	Review and Update rules after report design approval	Keira Nevers
2.1	04/19/2018	Update to Student Data and Student Report after review with RC	Keira Nevers
2.2	05/01/2018	Review with DA updates	Keira Nevers
2.3	05/08/2018	Adding clarification of AR Reporting and Data Deliverables	Keira Nevers
3.0	05/10/2018	Client Review/Signoff Edits	Keira Nevers
4.0	05/14/2018	Final clean published document	Keira Nevers

Approval
I acknowledge that I have read this Decision Rules document and been informed of the contents of this document. By entering my name, title and date approved, I certify my approval. I have received a copy of this document for my records and understand that any further changes will require additional approvals as necessary.

Printed Name	Title	Date Approved
Bethany Zimmerman	MSAA Report Subcommittee AZ	05/09/2018
Karla Marty	MSAA Report Subcommittee MD	05/09/2018
Melissa Flor	MSAA Report Subcommittee SD	05/09/2018

Table of Contents

Analysis and Reporting Decision Rules 1
Multi-State Alternate Assessment (MSAA) Spring 2018 3
I. Contract Overview 3
A. Test Administration(s) 3
B. Deliverables 3
II. Data Sources 4
A. Student Demographic Cleanup 4
B. Student Test Cleanup 4
C. Student Build Outs 5
D. Organization Cleanup 5
E. Scoring 5
III. Student Participation and Reporting Status 6
A. Overview 6
B. Student Test Attemptedness 6
C. Participation Status Assignment Hierarchy (by subject: Math, ELA) 6
D. Participation Status Summary 7
IV. Calculations 8
A. Raw Score 8
B. Writing Trait Raw Scores 8
C. Scaling and Equating 8
D. Performance Levels and Cut Scores 9
E. Aggregate Calculations 9
V. Data and Reporting Deliverable Requirements 9
A. General (all deliverables) 9
B. Student Report Specifics 10
C. School Roster Report Specifics 13
D. Student Results Data File Specifics 15
E. State Duplicate / Void Data File 15
F. Early Release Data File-Applicable to Arkansas (AR) Only 16
VI. Glossary of Terms 16
VII. Document Reference 17
A. Student Report Mock-up 17
B. AR Approved Change Order 17

Multi-State Alternate Assessment (MSAA) Spring 2018

This document details rules for analysis and reporting of the Multi-State Alternate Assessment (MSAA). This document is considered a draft until sign off has been granted. If there are rules that need to be added or modified after said sign-off, subsequent sign off will be obtained for each rule.

MSAA General Contract Information

Grades and Content Areas Assessed	Math and ELA, Grades 3-8 and 11
Students per Grade (approximately)	4,200
States/Entities	AR, AZ, GU, CNMI, DC, MD, ME, MT, SD, TN, VI
Total Districts	2,157
Total Schools	9,547

I. Contract Overview

A. Test Administration(s)

1. Eligible students are expected to test in Math and ELA in grades 03-08 and 11
2. ELA includes reading and writing; scores for the writing prompt shall contribute to the overall ELA score and shall be reported with a separate Reading and Writing scaled percentage of possible points
3. Tests are a Staged Adaptive design
    a. The raw score from core items in Session I determines the version(s) of Session II that is presented

B. Deliverables

1. Student Report: online with print options below
    a. State Option - Print Copies (Parent): AR, TN and VI*
    *Not administering 2018 due to catastrophe)
    b. State Option – Print Copies (Parent & School): MD
    c. State Option – AR reports will use the 16-17 templates, all others will use the new 17-18 report templates
2. Roster Report (School): online
    a. State of AR reports will use the 16-17 templates, all others will use the new 17-18 report templates
3. Summary Report (School, District, State): online
    a. State of AR reports will use the 16-17 templates, all others will use the new 17-18 report templates
4. Student Results Data File (School, District, State): online
    a. State of AR datafile will use the 16-17 templates, all others will use the new 17-18 datafile templates
5. Duplicate/Void Test Data (State): online
6. State Option - Early Release Data File (State): AR Only
    a. Online delivered results file via FTP no later than June 29, 2018

i. 16-17 data files will be used for AR Data deliverables

- MSAA1617StudentResultsLayout.xls

II. Data Sources

A. Student Demographic Cleanup

1. For the purpose of performing demographic cleanup, including identification of the final set of students to be reported via the MSAA, states are provided the complete list of all students registered in the MSAA system.
2. The demographic cleanup process enables states to:
   a. Identify and resolve instances of duplicate or erroneous registration records. States may indicate records to "merge" in order to resolve duplicates, "remove", or add, as necessary
   b. Update and add up-to-date demographic data
   c. Update the school and district a student should be at for reporting and aggregations
   d. Provide state-supplied test status information, such as exemptions and invalidations
   e. Confirm or update the grade level expected for testing for each student. The grade level returned by states is the grade level the student is expected to be reported in
3. See the *Demographic Clean up Instructions* document for additional details.

B. Student Test Cleanup

1. All tests associated with a final student (demographic row of record), including tests from student records merged during the demographic cleanup process, will be compiled for the test cleanup process.
2. The test cleanup process will independently determine the <u>final</u> Math test and the <u>final</u> ELA test to be used for analysis and reporting for each student. These tests are considered the *Analysis and Reporting Dataset*. All other tests are considered *Duplicate/Void* tests and are provided to states separately for informational purposes only.
3. **Off-Grade Tests**:
   a. If a student's expected grade level for testing from the demographic file does not match the test grade the test is "off-grade"
   b. Off-grade tests are classified as *Duplicate/Void* and are excluded from the *Analysis and Reporting Dataset* prior to completion of additional test cleanup steps
   c. Measured Progress will create a discrepancy alert for states of any case(s) where the tests associated with a student are off-grade. For these cases, the state may:
      i. Leave the data as-is
         - The off-grade test will be considered *Duplicate/Void*
         - The student will be included in the *Analysis and Reporting dataset* without a test, see *C. Student Build Outs*
      ii. Update the student's expected grade level for testing to match the test grade, if appropriate
         - The test will be considered on-grade and processed per normal rules at the test grade level
      iii. In either case above the state may also provide Measured Progress with an updated state-supplied status code for the student if they determine one is applicable for reporting while reviewing the scenario

4. Duplicate/Multiple Test Reconciliation
   a. After off-grade tests have been resolved, if a student still has more than one associated test for the same subject, the final test for *Analysis and Reporting* is determined using the following hierarchy:
      i. Submitted/Completed Test
      ii. Closed – Early Stopping Rule Applied
      iii. In Progress
   b. If two or more tests have the same status above the test associated with the last (latest) date-time stamp will be used
5. States should provide Measured Progress with all unique test-clean up scenarios that need to be handled outside of the process defined above. This "Bull Pen" file will be handled manually to ensure the correct test, as identified by the state, is used for analysis and reporting.

C. Student Build Outs

Student demographic rows of record from the state that do not have an on-grade test for one or both subjects are included in the *Analysis and Reporting dataset* with no test data.

D. Organization Cleanup

1. The schools and districts returned by states for each demographic row of record in the demographic cleanup file are considered the final school and district codes to be used for analysis and reporting, regardless of where a student's test was taken.
2. Measured Progress will work with states to identify the complete set of these school and district organizations, along with organization names for reporting, during the demographic file acceptance and organization cleanup process with each state.
3. The complete set of organizations in the *Analysis and Reporting dataset* will be loaded in MSAA System to enable access to the reports. States may restrict access through control of the user-accounts associated with each organization.
   a. New or revised Organization data shall be updated in both iCore and Breakthrough reporting platforms

E. Scoring

1. The Level 2 or Level 3 writing prompt is scored by Measured Progress resulting in a final score or score-condition code for each of the three traits:
   a. Traits:
      i. Organization
      ii. Idea Development
      iii. Conventions
   b. Writing prompt scores are operational and shall be included in the overall ELA score
   c. Valid Scores for each trait:

Original Score / Code	Description	Reported Value	Translated Score Value
0, 1-3	Final Score	0, 1-3	0 = No Evidence of Trait 1 = Limited Evidence 2 = Partial Evidence 3 = Full Evidence

Original Score / Code	Description	Reported Value	Translated Score Value
B	Blank Prompt	0	B = No Evidence Submitted
U	Unreadable	0	U = Unreadable
F	Foreign Language	0	F = Foreign Language
P	Copy of Prompt	0	P = Copy of Prompt
N	No Score	0	N = No Score
5	Off-Topic	0	O = Off Topic
6	Section is Blank	0	B = Section is Blank

2. All other item scores are taken from the MSAA testing system. Non-responses (blank responses) to any item are scored as 0 points.

III. Student Participation and Reporting Status

A. Overview

1. Participation statuses are assigned independently for Math and ELA for each student in the final *Analysis and Reporting dataset* using state-supplied test status information in conjunction with test submission and closure data, per the hierarchy.

B. Student Test Attemptedness

1. For non-writing prompt items, if the item response variable has a valid response value, the student responded to the item
   a. A valid response value is any value except NULL or X, where X is an internal code to exclude an item from a student's score due to an issue
2. For writing prompts, if a dimension score is '0','1','2','3','F', or 'P', The student responded to the writing prompt. All other dimension scores are considered a non-response
3. A student has a recorded response for the ELA test if any item on the ELA test (including the writing prompt) has a response
4. A student has a recorded response for the Math test if any item on the Math test has a response

C. Participation Status Assignment Hierarchy (by subject: Math, ELA)

1. If the state has supplied a test status code for the subject, then the Participation Status is the state provided code:
   a. Administration Irregularity
   b. Invalidated
   c. Parental Refusal
   d. ELL Exempt (ELA tests only)
   e. Exempt
   f. Withdrew
   g. No Longer Eligible
2. Otherwise, if the test is *Submitted* then the Participation Status is **Tested**, regardless of the number of item responses
3. Otherwise, if the test is *Closed – No Observable Communication Mode*:

a. And no item responses are recorded then the Participation Status is **Early Stopping Rule Applied**

b. And has one or more item responses recorded then the Participation Status is **Early Stopping Rule Misadministration**

4. Otherwise, if the test is *In Progress*:

a. And has no item responses recorded then the Participation Status is **Did Not Test**

b. And has one or more item responses recorded then the Participation Status is **Tested – Incomplete**

5. Otherwise the Participation Status is **Did Not Test**

6. Duplicate/Void tests, including off-grade tests, are not assigned participation statuses and are excluded from the *Analysis and Reporting dataset*.

7. Adjust participation status as follows:

a. If the participation status for one subject is ESR, but the student responded to at least one item in either ELA or Math, set the ESR participation status to ESM

b. If the participation status for a subject is ESR and the other subject has a participation status of DNT, then set the DNT participation status to ESR

D. Participation Status Summary

Participation Status	Description	Abbrev.	MP Code	State Data File (All Scores[1])	School, District Data Files:			In Agg. Calcs
					Scaled Score	R/W Percent	Perf. Level	
Tested	Tested	TES	A	Yes	Yes	0-100	Yes	Yes
Early Stopping Rule	Closed – No Observable Communication Mode, no responses.	ESR	B	Yes	Yes	N/A	Yes	Yes
Early Stopping Rule Misadministration	Closed – No Observable Communication Mode with at least 1 response.	ESM	C	Yes	Yes	0-100	No	No
Incomplete	In Progress with at least 1 response.	INC	D	Yes	Yes	0-100	No	No
* Administration Irregularity	Administration Irregularity was reported but does not necessitate an invalidation. Scores should be interpreted with caution.	IRR	E	Yes	Yes	0-100	Yes	Yes
* Invalidated	Student-based *or* administration-based irregularity resulting in invalidation.	INV	F	Yes	No	N/A	No	No

Participation Status	Description	Abbrev.	MP Code	State Data File (All Scores[1])	School, District Data Files:			In Agg. Calcs
					Scaled Score	R/W Percent	Perf. Level	
* Parental Refusal	Parental Refusal	PRF	G	No	No	No	No	No
* ELL Exempt (ELA Only)	Student meets the requirements for ELL 1st Year in the U.S. exemption from ELA.	ELL	H	No	No	N/A	No	No
* Exempt (Emergency, Medical, Other)	Student meets the requirements for exemption from the test.	EXE	I	No	No	N/A	No	No
Did Not Test	No test or an In Progress test with no responses.	DNT	J	No	No	N/A	No	No
* Withdrew	Student withdrew	WDR	K	No	No	N/A	No	No
* No Longer Eligible	Student is no longer eligible for testing.	NLE	L	No	No	N/A	No	No
Void/Duplicate	Test is a Duplicate or Void; excluded from Analysis and Reporting Dataset.	N/A	M	Separate File from Student Results; raw (unscored) data only.				
REMOVE	Student demographic record marked by state as REMOVE	These students and all associated tests are excluded from the analysis and reporting dataset entirely and are not provided to the state.						

IV. Calculations

A. Raw Score

1. Overall raw scores are calculated based on scores to items that are classified as "core" items for the test form. All other item response scores are excluded.
2. The "core" item list was determined in collaboration with the states
   a. The writing prompt is eligible and shall be included as a "core" item

B. Writing Trait Raw Scores

1. Student level writing trait scores are not included in reporting, and an overall writing score is not calculated or reported.
2. Measured Progress will work closely with states during and after scoring to provide feedback on the writing prompt results to inform item selection and for instructional purposes. This feedback will be defined outside the scope of this document.

C. Scaling and Equating

1. Psychometrics provides the raw score to scaled score lookup for each grade and subject and adaptive version of the test.

D. Performance Levels and Cut Scores

1. The following performance levels are used for MSAA Reporting:

Level	Title
1 (lowest)	Level 1
2	Level 2
3	Level 3
4 (highest)	Level 4

   a. MSAA cut scores for each performance level shall be validated 2018 using "Empirical Standards Validation Process" as recommended by TAC.

E. Aggregate Calculations

1. Eligible Students:
   a. For school, district, and state level aggregate calculations all students are eligible to be included based on their participation status.
   b. For MSAA level aggregate calculations (technical report, item statistics) all students are eligible to be included based on their participation status.
2. Participation Counts:
   a. All eligible students are included in participation summaries based on participation status for the subject if their participation status is reported.
   b. Classification of participation statuses into reported groupings (i.e.: "the number of *Tested* students") is documented for each individual report deliverable as necessary.
3. Results Aggregations:
   a. Results-based aggregations include, but are not limited to:
      i. Min, Max, Average Raw Score and SEM
      ii. Min, Max, Average Scaled Score and SEM
      iii. Number and percent of students by performance level
   b. Eligible students with the following participation statuses are included in results-based aggregate calculations for reporting:
      i. Tested
      ii. Early Stopping Rule
      iii. Administration Irregularity
   c. Only eligible students with a participation status of Tested (A) are included in item statistic calculations for the technical report.
   d. Aggregations with less than 10 students included in the denominator will be suppressed from state level reports only.

V. Data and Reporting Deliverable Requirements

A. General (all deliverables)

1. Only tests included in the *Analysis and Reporting Dataset* are eligible for final reporting.
   a. *Duplicate/Void* tests, although not reported, are provided to States in the State Duplicate/Void data file hand off, which will include off-grade tests.

2. Students classified as "Withdrew" or "No Longer Enrolled" for *both* ELA and Math are excluded from the Roster Report, Student Report, and Summary Report entirely. They are included in the Student Results data files still.
3. Final reports and data files are generated by Measured Progress for all organizations with reported students in the *Analysis and Reporting dataset*, as applicable for their organization level.
4. Access to reports for specific schools or districts can be restricted via management of the log-in credentials through the Breakthrough system.
5. All school and district level reports are marked "Confidential" on all pages.
6. N-size suppression is done on state level reports only. Any aggregations with less than 10 students included are suppressed from the state level reports.

B. Student Report Specifics

1. Each student report consists of a two page report (front and back). Page One shall be noted as "Confidential".
2. Student reports are generated for all students in the *Analysis and Reporting dataset* **earning a performance level** in at least one content are:
   a. Tested (A)
   b. Early Stopping Rule (B)
   c. Administration Irregularity (E).
3. Since both content areas are always displayed, alternate text is provided for each status that does not receive a student report in the event that a student receives a report for the other content area, see the Participation Status Summary.
4. For all statuses that have scaled scores but are not receiving a reported Performance Level [Early Stopping Rule Misadministration (C), Tested-Incomplete (D), ~~and Invalid (F)~~], the school and district will have access to the earned scaled score in the student results data file. Since these statuses do not earn a performance level they do not receive student reports.
5. For statuses receiving a report:
   a. "What is in this Report?" Section.
      i. Page one and two shall be described via agreed upon text, separated into two lines
   b. Performance Summary
      i. The Performance Summary descriptive shall display as agreed upon text. Student Name will be variable.
      ii. The scaled score and performance level earned are printed in the appropriate Performance Summary section(s), English Language Arts and/or Mathematics
      iii. If only one section is applicable, the other section will not have any scores displayed
   c. The sentence explaining variation in test scores associated with the student's scaled score is displayed.
   d. English Language Arts (ELA) shall have a percentage of possible point earned in each area displayed if applicable
   e. The performance level description associated with the earned performance level is printed below the Performance Summary

f. The Student's PLD Text shall appear below the Performance Level Descriptors for each subject as applicable (see PLD Lookup Text document)

i. For students classified as Early Stopping Rule (ESR):

- An asterisk (*) is added to the earned performance level at the top: Level 1*
- The asterisk corresponds to the alternate text to be displayed below the bar graph – see Participation Status Summary Table: Alternate Text.

ii. The PLD text for Level 1 is not shown.

6. For statuses that do not receive a report but must appear because the other content area is reported:

a. The sentences for "Your child's scaled score" and "Your child's performance level" are not shown.

b. The graph is replaced with alternate text directing parents to contact their school or teacher, see the Participation Status Summary Table: Alternate Text.

c. Because the reports are displaying both Math and ELA on one page, Alternate Text where No Student Report may appear if, in fact, the student has a score in one of the Math or ELA categories.

7. Page Two of the report shall begin with the static text Parent letter, followed by the "What skills can be worked on next".

8. The "What skills can be worked on next" section shall have lookup text printed based on the earned performance level for each section of ELA and/or Math

a. In cases where there is no performance level to report, the area shall remain blank

9. The report shall include the "What Now" static text section at the bottom of page 2

10. States electing to receive printed student reports will receive report packages packed by school and shipped to the district. If a state is receiving parent and school copies, two identical packages per school are created and shipped.

11. Participation Status Summary – Full List Available to States - **Student Report**:

Participation Status	Description	Abbrev.	MP Code	Student Report Specifics			
				Scaled Score	R/W Percent	Perf Level	Alternate Text
Tested	Tested	TES	A	Yes	0-100	Yes	
Early Stopping Rule	Closed – No Observable Communication Mode, no responses.	ESR	B	Yes (lowest)	N/A	Yes Level 1	PLD Text: 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.

Participation Status	Description	Abbrev.	MP Code	Student Report Specifics			
				Scaled Score	R/W Percent	Perf Level	Alternate Text
Early Stopping Rule Misadministration	Closed – No Observable Communication Mode with at least 1 response.	ESM	C	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
Incomplete	In Progress with at least 1 response.	INC	D	No Student Report			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
* Administration Irregularity	Administration Irregularity was reported but does not necessitate an invalidation. Scores should be interpreted with caution.	IRR	E	Yes	0-100	Yes	
* Invalidated	Student-based *or* administration-based irregularity resulting in invalidation.	INV	F	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
* Parental Refusal	Parental Refusal	PRF	G	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
* ELL Exempt (ELA Only)	Student meets the requirements for ELL 1st Year in the U.S. exemption from ELA.	ELL	H	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
* Exempt (Emergency, Medical, Other)	Student meets the requirements for exemption from the test.	EXE	I	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
Did Not Test	No test or an In Progress test with no responses.	DNT	J	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
* Withdrew	Student withdrew	WDR	K	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.
* No Longer Eligible	Student is no longer eligible for testing.	NLE	L	No Student Report.			Your child did not receive a score in this content area. Please contact your child's teacher/school for more information.

C. School Roster Report Specifics

1. Rosters are generated for each school in the *Analysis and Reporting dataset* and will list all students, regardless of participation status, except:
   a. Student's classified as "Withdrew" or "No Longer Enrolled" for *both* ELA and Math.
2. Comparison to State
   a. The state average scaled score is calculated using the earned scaled score for all students included in aggregations calculations: Tested (A), Early Stopping Rule (B), and Administration Irregularity (E).
   b. The standard error of measurement (SEM) associated with the student's obtained score is used to identify the range around the state average scaled score to classify the student as above, similar to, or below the state average:

Classification	Performance	Display
Student Score < (State Average – Student SEM)	Lower than the state average	-
(State Average – Student SEM) <= Student Score <= (State Average + Student SEM)	Similar to the state average	=
Student Score > (State Average + Student SEM)	Above the state average	+

3. For Test Status print the "Test Status" column from the Participation Status Summary – Roster Report table.
4. For participation statuses that do not receive a state comparison, scaled score, or performance level (listed as "No" in the Participation Status Summary Table) these fields appear blank on the roster.
5. School Summary Table on the Roster:
   a. School, District, and State Summary data are displayed at the top of the report. Since reports are marked "confidential" there is no suppression rules applied.
   b. The number **Enrolled** is equal to the total number of students listed on the roster. This includes all students in the Reporting and Analysis dataset except those that are "Withdrew" or "No Longer Eligible" in *both* Math and ELA and are therefore not listed on the roster.
   c. The number **Tested** is equal to the set of students receiving a reported *performance level*: Tested (A), Early Stopping Rule (B), and Administration Irregularity (E).
6. The Average Scaled Score and Percent of Students by Performance level calculations are based on the number of **Tested** students.
7. Participation Status Summary – Full List Available to States - **Roster Report**:

Participation Status	Abbrev.	MP Code	Roster Report Specifics:			
			Display Test Status	State Compare	Scaled Score	PerfLevel
Tested	TES	A		Yes	Yes	Yes

Participation Status	Abbrev.	MP Code	Roster Report Specifics:			
			Display Test Status	State Compare	Scaled Score	PerfLevel
Early Stopping Rule	ESR	B	ESR	Yes	Yes	Yes (Level 1)
Early Stopping Rule Misadministration	ESM	C	MIS	Yes	Yes	No
Tested – Incomplete	INC	D	INC	Yes	Yes	No
Administration Irregularity	IRR	E	IRR	Yes	Yes	Yes
Invalidated	INV	F	INV	No	No	No
Parental Refusal	PRF	G	PRF	No	No	No
ELL Exempt (ELA Only)	ELL	H	ELL	No	No	No
Exempt (Emergency, Medical, Other)	EXE	I	EXE	No	No	No
Did Not Test	DNT	J	DNT	No	No	No
Withdrew	WDR	K	WDR	Not Included on Roster Reports. If appearing for 1 content area, then State Compare, Scaled Score, and PerfLevel are blank,		
No Longer Eligible	NLE	L	NLE			

C. Summary Report Specifics

1. Summary Reports are generated for each school, district, and state in the *Analysis and Reporting dataset* with at least one student who is not classified as "Withdrew" or "No Longer Enrolled" in *both* ELA and Math.

2. The number **Enrolled** is equal to the total number of students listed on the roster. This includes all students in the Reporting and Analysis dataset except those that are "Withdrew" or "No Longer Eligible" in *both* Math and ELA (same as Roster).

3. The number **Tested** is equal to the set of students receiving a reported *performance level*: Tested (A), Early Stopping Rule (B), and Administration Irregularity (E). (Same as Roster).

4. The number that **Did Not Test** is equal to the number of students classified as: Did Not Test (J), Parental Refusal (G), ELL Exempt (H), Exempt (I), Withdrew (K), No Longer Eligible (L), Invalidated (F), Tested-Incomplete (D) or Early Stopping Rule Misadministration (C).

a. Note: Withdrew and No Longer Eligible students are only included if they are included in the number Enrolled, as a result of being reported in the other content area.

5. The number and percent at each performance level calculations are based on the number of **Tested** students.

D. Student Results Data File Specifics

1. All students in the *Analysis and Reporting Dataset* are included in the Student Results data files for their school, district, and state, per the Student Results Data File Layout. One file is created containing all grades for each entity with reporting results.
2. Refer to the file layout for specific data elements and valid values, as well as identification of which fields are included in the school and district files. All fields are included in the state file.
3. Student Results Data Files are comma delimited (CSV).
   a. Measured Progress will remove embedded commas from character fields in the data prior to exporting.
4. There will be one (1) record per student containing the final Math and ELA test results used for reporting.
   a. The ELA score will also contain two additional fields consisting of the Percentage of possible points correct for Reading and Writing items
5. For students with reporting statuses that do not receive item scores, raw score, scaled scores, and performance levels, these fields will be set to blank in the school, district, and state student results data files. See the Participation Status Summary Table (Pg. 5).
6. School, District Files – Additional Notes:
   a. All fields marked as "No" in the Student Results Data File layout for the "School or District data" column shall be excluded from school and district data files.
   b. Raw scores, scaled scores, and performance levels are set to blank for students with a participation status showing "No" for these scores in the Participation Status summary table.
7. Item responses to core items (items that contribute to a student's raw score for reporting) are included in the state file for Math and ELA, following MP's "+-data" format. See the layout for specific value details.
8. The operational Level 2 and Level 3 writing prompt scores available at the time of reporting shall be populated.
   a. AR Early Reporting state file will exclude the writing prompts (Change request #6027-09)
9. The Test_Proctor_ID associated with each test is included in the state file. This ID corresponds to an additional lookup file that will be delivered to states (via MP FTP) with Test Proctor information for the District of Columbia only (DC).
   a. This has been included in the past and has been deleted from the final file delivered to all states, except for DC. Future deliverables are expected to only provide this for DC

E. State Duplicate / Void Data File

1. One file is created per state containing all non-reported tests classified as Duplicate/Void, including off-grade tests.
2. The Duplicate/Void data file will follow the same layout as the State Student Results data file layout, however, there may be several records per student depending on the number of Duplicate/Void tests. Each record may contain results for only one or both subjects.

3. The grade will reflect the grade level of the test. For off-grade tests this will differ from the grade level the student is reported under and may differ for a single student within this file if they took tests at multiple grades.
4. Scores and performance levels are not calculated for these tests and may be blank. All available data will be provided as-is and is provided to states for informational purposes only and should be interpreted with caution as it has not been through the full cleanup process that is applied to reported data.

F. Early Release Data File-Applicable to Arkansas (AR) Only

1. An early results file will be generated for Arkansas this year. The file will be produced after removing students moved to the Do Not Process School "9999".
2. The following issues may be present in the early results, and will be resolved through the standard MP cleanup and processing rules defined by this document for final reporting:
   a. Duplicate student records.
   b. Duplicate tests.
   c. Incorrect and/or incomplete demographics, missing demographics will be left blank.
   d. Incorrect school/district assignments.
   e. No state-supplied invalidations or exemptions applied. All tests will be assigned one of the MP-calculated participation statuses.
   f. No writing scores.
   g. Blank or invalid values for fields expected to be resolved during cleanup.
   h. The grade level will reflect the grade level of the test.
   i. Scaled scores, performance levels are assigned based on available information and calculated statuses.
   j. The same blanking rules of scores and results that are defined for the state student results file based on test status are applied per the calculated test statuses available.
3. The State is required to follow the standard demographic cleanup process (separate from this early results file), and early results are subject to change as a result of cleanup.
4. The early results file will follow the same layout as the State Student Results data file layout, however, as a result of the data being incomplete and the capacity for a student to have multiple tests per content area; fields may contain blank or invalid values.
5. Hand-scored writing prompts shall be excluded from AR early reporting only.
6. AR Early Reporting will utilize 2016-2017 cut scores and report shells in generating the static PDF and CSV reports. Shells must be updated with 2018 date information only.

VI. Glossary of Terms

Glossary	
MSAA	Multi-State Alternate Assessment
ELA	English Language Arts
TAO	Testing Assisté par Ordinateur (in French) / Computer Based Testing
TAC	Technical Advisory Committee
SEM	Standard Error of Measurement
OAT	Open Assessment Technologies

VII. Document Reference

A. Student Report Mock-up

MSAAReportDesign Final.pdf

B. AR Approved Change Order

CO 6027-09 (AR Early Reporting)_Exec

APPENDIX G—MSAA 2018 GUIDE FOR SCORE REPORT INTERPRETATION

Multi-State Alternate Assessment

(MSAA)

2018 Guide
for
Score Report Interpretation

State Specific Information

Listed below is the contact information for each state's MSAA State Lead:

Arizona Bethany Zimmerman 602-542-4061 Bethany.Zimmerman@azed.gov Cindy Sandner 602-542-3059 Cindy.Sandner@azed.gov Audra Ahumada 602-542-5450 Audra.Ahumada@azed.gov	**Arkansas** Ann Finch 501-682-5303 Ann.Finch@arkansas.gov Debbie Young 501-682-4946 Debbie.Young@arkansas.gov	**District of Columbia** Michael Craig 202-257-3371 Michael.craig@dc.gov
Maine Sue Nay 207-624-6774 Sue.Nay@maine.gov	**Maryland** Marsie Torchon 410-767-2498 martha.torchon@maryland.gov Nancy Schmitt 410-767-0743 Nancy.Schmitt@maryland.gov Ann Herrmann 410-767-0086 Ann.Herrmann@maryland.gov	**Montana** Yvonne Field 406-444-0748 yfield@mt.gov
South Dakota Jan Martin 605-773-3246 Jan.Martin@state.sd.us Chris Booth 605-773-6156 Christina.Booth@state.sd.us	**Tennessee** Megan Sellers 615-906-1548 Megan.sellers@tn.gov	**United States Virgin Islands** Alexandria Baltimore-Hookfin 340-773-1095 ext. 7084 alexandria.baltimore@vide.vi
PAC-6 June De Leon (Guam / CNMI) 671-735-2494 June.DeLeon@guamcedders.org Terese Crisostomo (Guam) 671-300-1323 tdcrisostomo@gdoe.net Fasefulu Tigilau (CNMI) 670-237-3199 Fasefulu.Tigilau@cnmipss.org		

Table of Contents

Introduction to the MSAA

Purpose 1
Student Participation 2
Overview of the MSAA Format 3
Scoring 3

MSAA Score Reports

Overview 4
Interpreting and Using the MSAA Scores 5
Talking to Parents and Guardians 5
Special Reporting Codes and Messages 6
Types of Score Reports 7
Student Results File CSV 7
Testing Participation Requirements by Content Area 7
District Summary Report 8

Reports for the School

School Summary Report 9
School Roster Report 10
Individual Student Report 11

Appendix A: Performance Level Descriptors 12

Appendix B: Individual Student Report 28

Appendix C: Writing Scoring Rubrics 32

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 significant cognitive disabilities in preparation for a broader array of post-secondary outcomes. The MSAA is designed to measure academic content that is aligned to and derived from your state's content standards. This test 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 areas of English Language Arts (ELA) and mathematics in grades 3-8 and 11.

This assessment was developed through the research and development completed by the National Center and State Collaborative (NCSC) and has been carried forward by the MSAA State Partners. MSAA is currently being administered by ten participating states: Arizona, Arkansas, Maine, Maryland, Montana, the Pacific Assessment Consortium (PAC-6)[1], South Dakota, Tennessee, U.S. Virgin Islands, and Washington, D.C.

This guide provides information regarding the administration and results of the spring 2018 MSAA to district and school personnel.

[1] The Pacific Assessment Consortium (including the entities of American Samoa, Commonwealth of the Northern Mariana Islands, Federated States of Micronesia, Guam, Republic of Palau, and Republic of the Marshall Islands) are collectively considered one state, led by the University of Guam Center for Excellence in Developmental Disabilities Education, Research, and Service (CEDDERS).

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 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 (derived from) the State's Content Standards.	Goals and instruction listed in the IEP for this student are linked to the enrolled grade-level State's 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 the 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 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 Toolkit developed by NCSC to help these students develop a mode of communication. The Toolkit can be found here: https://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 11 and is aligned to the State's Content Standards and the MSAA Core Content Connectors. The MSAA is a computer–based, on demand, stage adaptive assessment consisting mostly of selected response and some constructed response items written at four levels of complexity. These complexity levels represent different levels of skill acquisition by students.

Students with significant cognitive disabilities often need materials and instructional strategies that are substantially adapted, 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 (CBT) designed to be 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, object replacement, translating the test into ASL, among others.

Each content area consists of 30-40 items that are mostly selected response. The writing portion of the ELA test contains a scaffolded writing prompt at each grade level. Each content test is divided into test sessions. Test administrators have substantial leeway in developing a testing schedule with the ability to start and stop a test depending on the engagement of the student.

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. Other constructed response items are scored by the Test Administrator 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 C.

MSAA Score Reports

Overview

This guide describes the types of score reports provided for the 2017-18 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. The performance level descriptors are broad and general statements regarding skills and abilities of students who have attained each level. Performance levels for the MSAA were established by committees of educators after the first NCSC administration of the assessment in 2015 and were updated in 2018. Performance level descriptors for each content area and grade level can be found in Appendix A of this document. The scale score ranges that make up each performance level can be found in Appendix B.
- *Scale scores* report the performance level the student achieved. Scale scores are more precise than performance levels and may be used to make comparisons between groups of students, schools, and districts. In Appendix B, Table 1 shows the scale score ranges for each performance level, content area, and grade level.
- *Descriptive and informative reports*. In addition to including student demographic information, performance level, and scale scores, the Individual Student Report contains supportive information about student performance and MSAA measures.
    - *Reading and Writing Scores* – the percent of items answered correctly for reading and writing separately. The writing items consisted of selected response and constructed response (or multiple choice and the writing prompt).
    - *What skills can be worked on next* – skills that may be tested in the following grade of standards at a specific grade level.
    - *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 in ELA and mathematics based on State's Content Standards. The student's performance level is based on alternate 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 Individualized Education Program (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 Performance Level Descriptors 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. Contact your MSAA State Lead to locate these materials.

When talking to parents and guardians about their child's score, it may be helpful to keep the following in mind:

- MSAA assessment 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 students learning.
- MSAA scores alone should not be used to make placement or eligibility decisions.

Special Reporting Codes and Messages

In some cases, students were assigned a special reporting code. A complete list of special reporting codes and their associated descriptions is provided below. For additional information or interpretation of special reporting codes, contact your MSAA State Lead.

Test Status		
Code	**Test Status**	**Description**
ESR	Early Stopping Rule	If the TA did not observe a student response after the presentation of 4 items, the test was closed by the TC.
ESM	Early Stopping Rule Misadministration	Testing may have ended early on the basis that a consistent mode of communication was not observed. 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.
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 assessment.
WDR	Withdrew	The student withdrew.
NLE	No Longer Eligible	The student is not eligible to test via the MSAA assessment.

Types of Score Reports

Below are the types of MSAA score reports that will be available on the MSAA Reporting Portal. Only District testing coordinators using their current MSAA username and password may access the MSAA reports here: https://www.msaaassessment.org under the Reporting Tab. All MSAA score reports are confidential documents.

- Reports for the District
  - District Summary Report
  - Student Results File CSV
- Reports for the School
  - School Summary report
  - School Roster Report
  - Student Results File CSV
  - Individual Student Report

If you have any questions about accessing these MSAA reports, contact your MSAA State Lead. Contact information can be found at the beginning of this document.

Student Results File CSV

A CSV file of all student results will be available to District Test Coordinators through the MSAA Reporting Portal. For information regarding this file, contact your MSAA State Lead.

Testing Participation Requirements by Content Area

All students in grades 3 – 8 and 11 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 item responses.

For additional information regarding the reported test status, contact your MSAA State Lead. Contact information can be found at the beginning of this document.

Reports for District

District Summary Report

The *District Summary Report* (DSR) provides district staff with a summary of student participation and performance by district and school. 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 A

		3				4							
						Performance Level							
		Enrolled	Tested	Did Not Test	Average Scale Score	Level 1		Level 2		Level 3		Level 4	
						N	%	N	%	N	%	N	%
Grade 03	State	25	17	8	1233	10	59	1	6	4	24	2	12
	District	19	15	4	1232	9	60	1	7	3	20	2	13
Grade 04	State	25	25	0	1231	15	60	4	16	5	20	1	4
	District	19	19	0	1230	12	63	3	16	4	21	0	0
5 Grade 05	State	25	21	4	1228	15	71	4	19	1	5	1	5
	District	19	16	3	1228	11	69	4	25	1	6	0	0
Grade 06	State	25	25	0	1228	14	56	5	20	6	24	0	0
	District	18	18	0	1230	11	61	2	11	5	28	0	0
Grade 07	State	25	18	7	1233	9	50	3	17	6	33	0	0
	District	19	15	4	1231	9	60	2	13	4	27	0	0
Grade 08	State	25	23	2	1227	12	52	5	22	6	26	0	0
	District	19	18	1	1228	9	50	4	22	5	28	0	0
Grade 11	State	25	24	1	1241	9	38	1	4	11	46	3	13
	District	19	18	1	1244	5	28	0	0	10	56	3	17

The District Summary Report contains the following features, highlighted above:

1. Content Area of the report.
2. State and District included in the report.
3. Number of students Enrolled, Tested, Did Not Test, and Average Scale Score by State, District and School. Refer to the Special Reporting Codes and Messages for information regarding test status.
4. The number and percentage of students at each performance level by grade in the state, district.
5. Summary of results by Grade Level. The state and district data shown here are other third graders in the state and district

Reports for the School

School Summary Report

Figure 2 – Sample School Summary Report

CONFIDENTIAL

msaa
Multi-State Alternate Assessment

1 Mathematics

2 **SUMMARY REPORT**
Demonstration State
Demonstration District A
Demonstration School 1

		3 Enrolled	Tested	Did Not Test	Average Scale Score	4 Performance Level: Level 1 N	Level 1 %	Level 2 N	Level 2 %	Level 3 N	Level 3 %	Level 4 N	Level 4 %
5 Grade 03	State	25	17	8	1239	6	35	0	0	8	47	3	18
	District	19	15	4	1237	6	40	0	0	7	47	2	13
	School	9	6	3	1243	2	33	0	0	2	33	2	33
Grade 04	State	25	25	0	1229	10	40	9	36	5	20	1	4
	District	19	19	0	1229	8	42	7	37	3	16	1	5
	School	8	8	0	1226	5	63	2	25	0	0	1	13
Grade 05	State	25	21	4	1236	6	29	9	43	5	24	1	5
	District	19	16	3	1236	5	31	6	38	5	31	0	0
	School	8	6	2	1233	3	50	2	33	1	17	0	0
Grade 06	State	25	25	0	1234	11	44	5	20	5	20	4	16
	District	18	18	0	1236	7	39	5	28	2	11	4	22
	School	8	8	0	1237	2	25	4	50	1	13	1	13
Grade 07	State	25	18	7	1236	6	33	3	17	7	39	2	11
	District	19	15	4	1234	6	40	2	13	6	40	1	7
	School	9	7	2	1228	4	57	1	14	1	14	1	14
Grade 08	State	25	23	2	1235	10	43	5	22	3	13	5	22
	District	19	18	1	1236	8	44	3	17	3	17	4	22
	School	8	7	1	1237	4	57	0	0	1	14	2	29
Grade 11	State	25	24	1	1242	6	25	5	21	6	25	7	29
	District	19	18	1	1244	5	28	2	11	4	22	7	39
	School	9	9	0	1242	3	33	2	22	1	11	3	33

The School Summary Report contains the following features, highlighted above:

1. Content Area of the report.
2. State, District and School included in the report.
3. Number of students Enrolled, Tested, Did Not Test, and Average Scale Score by State, District and School. Refer to the Special Reporting Codes and Messages for information regarding test status.
4. The number and percentage of students at each performance level by grade in the state, district and school.

School Roster Report

The School Roster Report provides student performance information at the school level for each grade, including each student's test status, scale score and performance level. See Figure 3 below.

Figure 3 – Sample School Roster Report

CONFIDENTIAL

msaa
Multi-State Alternate Assessment

SCHOOL ROSTER REPORT
1
Demonstration State
Demonstration District A
Demonstration School 1
Grade 03

2

3		English Language Arts						Mathematics					
	Enrolled	Tested	Average Scale Score	Level 1 (%)	Level 2 (%)	Level 3 (%)	Level 4 (%)	Tested	Average Scale Score	Level 1 (%)	Level 2 (%)	Level 3 (%)	Level 4 (%)
State	25	17	1233	59	6	24	12	17	1239	35	0	47	18
District	19	15	1232	60	7	20	13	15	1237	40	0	47	13
School	9	6	1241	33	0	33	33	6	1243	33	0	33	33

Spring 2018

4

Student Name Student ID	English Language Arts				Mathematics			
	Test Status*	State Compare	Scale Score	Performance Level	Test Status*	State Compare	Scale Score	Performance Level
LastName1, First1 10021		+	1251	Level 3		+	1258	Level 4
LastName11, First11 10104	ESR	-	1200	Level 1	ESR	-	1200	Level 1
LastName13, First13 10107		=	1231	Level 1		-	1234	Level 1
LastName17, First17 10127		+	1259	Level 4		+	1253	Level 3
LastName19, First19 10138	ESM	-	1223		ESM	-	1200	
LastName23, First23 10165		+	1245	Level 3		+	1250	Level 3
LastName25, First25 10182		+	1259	Level 4		+	1260	Level 4
LastName5, First5 10048	ESM	+	1251		ESM	+	1249	
LastName7, First7 10087	DNT				DNT			

5

6

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.

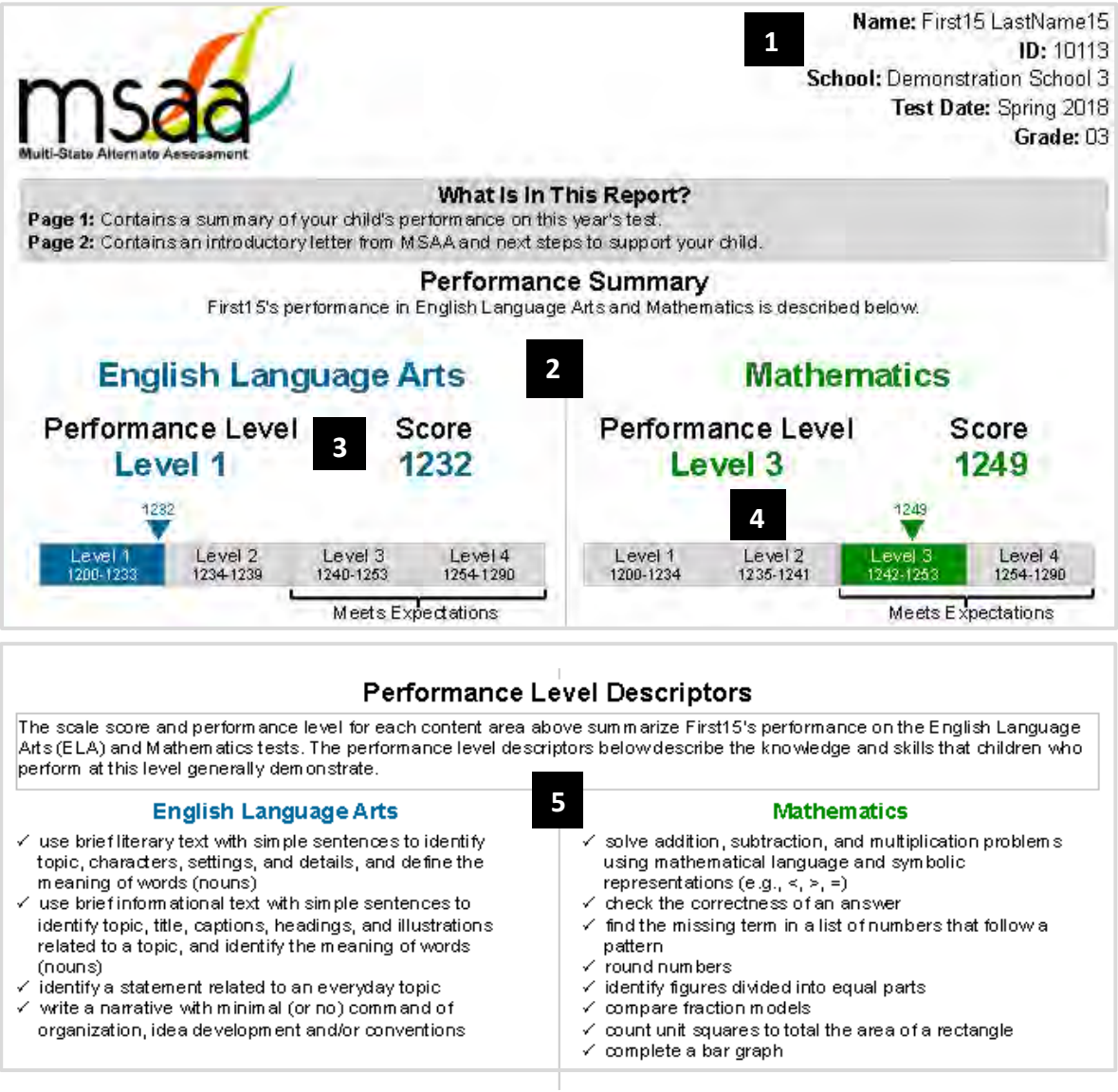
The School Roster Report contains the following features, highlighted above:

1. The state, district and school included in the report. The results are displayed by Content Area.
2. A summary of enrolled and tested students and the average scale score for the state, district and reported school.
3. For each content area the student's test status, comparison to other students in the same grade level in the state, scale score and performance level is displayed.
4. This section of the report includes all students tested at the school for the specified grade.
5. This key shows symbols used in the "State Compare" column.

Individual Student Report

The Individual Student Report (ISR) provides scale score and performance level information for a specific student. Figure 4 shows page 1 of the ISR. A full sample ISR is included in Appendix B.

Figure 4 – Sample Individual Student Report



The Individual Student Report contains the following features, highlighted above:

1. The report header includes the student's full name, student ID, Grade and School.
2. The results for each content area are displayed separately on the report.
3. The student's scale score and performance level for each content area is shown.
4. This display shows the student's score compared to the performance level scale.
5. This text shows the performance level descriptor for the student's performance level.

Appendix A

Performance Level Descriptors

Performance Level Descriptors for ELA and Mathematics

MSAA developed Performance Level Descriptors for ELA and mathematics at grades 3-8 and 11 through an iterative process involving multiple stakeholder groups. 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 knowledge, skills and abilities of the preceding performance levels.

The performance descriptors included in Appendix A 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 of 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 - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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 - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 supportive 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 supportive 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 supportive 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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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 - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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 - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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

Grade 7 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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 - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 which 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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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.

Grade 11 ELA Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low text complexity - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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 which 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.

Grade 3 Mathematics Performance Level Descriptors

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
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 in 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 which represents factors in a problem • solve multiplication equations in which both numbers are equal to or less than five • 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 five • 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 five • 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 - *Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• identify geometric figures which are divided into equal parts	• round numbers to nearest 10 • identify geometric figures which 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 - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
The student is able to: • identify an array with the same number of objects in each row • identify values rounded to 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 an multiplication expression using two single digit numbers • identify a model of a multiplicative comparison • show division of objects into equal groups • round numbers to nearest 10, 100 or 1000 • 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 nearest 10, 100, or 1000 • compare two fractions with different denominators • sort a set of 2-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 nearest 10, 100 or 1000 • compare two fractions with different denominators • sort a set of 2-dimensional shapes • compute the perimeter of a rectangle • transfer data to a graph
	AND with Moderate task complexity - *Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• identify equivalent fractions • select a 2-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 - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
The student is able to: • solve one-step subtraction word problems • divide sets (no greater than 6) 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 plan • 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 - *Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• compare the values of two products based upon multipliers • round decimals to nearest whole number	• compare the values of two products based upon multipliers • round decimals to 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 - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
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 zero 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 zero • 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 *- Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• 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 - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
The student is able to: • identify a representation which 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 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 - *Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• 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 - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
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 which 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 which 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 which 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 - *Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• 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.

Grade 11 Mathematics Performance Level

Level 1	Level 2*	Level 3*	Level 4*
Low task complexity - *Simple problems using common mathematical terms and symbols*	**Low task complexity -** *Simple problems using common mathematical terms and symbols*	**Moderate task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	**High task complexity -** *Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements*
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 which 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 which 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 which 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 - *Common problems presented in mathematical context using various mathematical terms and symbols*	**AND with High task complexity -** *Common problems presented in mathematical context using various mathematical terms and symbols*	
	• 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 which represents a provided data set	

*Levels 2, 3, and 4 include demonstration of skills described in previous performance levels.

Appendix B

Individual Student Report

CONFIDENTIAL

Multi-State Alternate Assessment

Name: First15 LastName15
ID: 10113
School: Demonstration School 3
Test Date: Spring 2018
Grade: 03

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

First15's performance in English Language Arts and Mathematics is described below.

English Language Arts

Performance Level	Score
Level 1	1232

Level 1	Level 2	Level 3	Level 4
1200-1233	1234-1239	1240-1253	1254-1290

1232

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 1229 and 1235.

English Language Arts consists of Reading and Writing. See below for percent of possible points earned in each area.

Reading 61% **Writing 25%**

Mathematics

Performance Level	Score
Level 3	1249

Level 1	Level 2	Level 3	Level 4
1200-1234	1235-1241	1242-1253	1254-1290

1249

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 1245 and 1253.

Performance Level Descriptors

The scale score and performance level for each content area above summarize First15'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 text with simple sentences to identify topic, characters, settings, and details, and define the meaning of words (nouns)
- ✓ use brief informational text with simple sentences to identify topic, title, captions, headings, and illustrations related to a topic, and identify the meaning of words (nouns)
- ✓ identify a statement related to an everyday topic
- ✓ write a narrative with minimal (or no) command of organization, idea development and/or conventions

Mathematics

- ✓ solve addition, subtraction, and multiplication problems using mathematical language and symbolic representations (e.g., <, >, =)
- ✓ check the correctness of an answer
- ✓ find the missing term in a list of numbers that follow a pattern
- ✓ round numbers
- ✓ identify figures divided into equal parts
- ✓ compare fraction models
- ✓ count unit squares to total the area of a rectangle
- ✓ complete a bar graph

© 2018 MSAA. All rights reserved.

Dear Parents and Guardians,

This report summarizes your child's performance on the online 2018 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 11 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

+ Determine the main idea and supporting details of text
+ Use information from charts, graphs, diagrams to answer questions
+ Use general academic words in reading and writing
+ Produce writing that expresses real or imaginary experiences and ideas

Mathematics

+ Use mathematical terms and symbols (<, >, =)
+ Round numbers to the nearest ten, hundred, or thousand
+ Solve single digit multiplication problems including arrays, models, and word problems
+ Compare fractions with different denominators; identify equivalent fractions
+ Identify perimeter of a rectangle
+ Analyze data in bar graphs
+ Sort two-dimensional shapes

What now?

Bring this report to your next conference with First15's teachers.
You can ask First15's teachers:

- What is First15 learning in ELA and Mathematics this year?
- How is First15 doing?
- How can I use this information to work with First15 this year?
- What resources should I use to support First15?

© 2018 MSAA. All rights reserved.

Table 1

2018 Performance-Level Scale Score Ranges by Content Area and Grade

Performance Level	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
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

Appendix C

Writing Scoring Rubrics

Grade 3 Writing Scoring Rubrics
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 Ex: "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	

February 28, 2018

Grade 3 Writing Scoring Rubrics
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 **and** 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 Ex: "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	

February 28, 2018

Grade 4 Writing Scoring Rubrics
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 or setting) and includes a character. The response provides a conclusion.	The narrative includes at a minimum: ☐ 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) ☐ 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 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** of the events include a detail related to 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., 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** complex thought unit that expresses a complete idea **with** subject-verb agreement Ex: "The dog runs" or "the dog runs"	The narrative includes at a minimum: ☐ 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 (end punctuation for **one** thought unit **or one** thought unit **with or without** subject-verb agreement)	0 ☐ **no** evidence of standard English conventions	

February 28, 2018

Grade 4 Writing Scoring Rubrics
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 Ex: "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	

February 28, 2018

Grade 5 Writing Scoring Rubrics
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 ☐ establish a 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** Ex.: I said "No, I want to play."	The narrative includes at a minimum: ☐ **two** events related to a characters' 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 which **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 narrative 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 Ex: "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	

February 28, 2018

Grade 5 Writing Scoring Rubrics
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 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 Ex.: 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 **majority** of thought units ☐ **one** complete sentence that expresses an idea **with** subject-verb agreement Ex: "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	

February 28, 2018

Grade 6 Writing Scoring Rubrics
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, 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 Ex: "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	

February 28, 2018

Grade 6 Writing Scoring Rubrics
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, 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 **majority** of thought units ☐ **one** complete sentence that expresses an idea **with** subject-verb agreement Ex: "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	

February 28, 2018

Grade 7 Writing Scoring Rubrics
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, 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 Ex: "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	

February 28, 2018

Grade 7 Writing Scoring Rubrics
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 (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, 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 **majority** of thought units ☐ **one** complete sentence that expresses an idea **with** subject-verb agreement Ex: "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	

February 28, 2018

Grade 8 Writing Scoring Rubrics
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, 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 Ex: "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	

February 28, 2018

Grade 8 Writing Scoring Rubrics
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 (problem/solution).	The essay includes at a minimum: ☐ an introduction that states **both** parts of the problem ☐ 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, 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(s) that connects the problem to the solution	The essay includes at a minimum: ☐ **one** problem **or** solution **with** a relevant detail ☐ **one** transitional word(s) 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 **majority** of thought units ☐ **one** complete sentence that expresses an idea **with** subject-verb agreement Ex: "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	

February 28, 2018

Grade 11 Writing Scoring Rubrics
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 Ex: "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	

February 28, 2018

Grade 11 Writing Scoring Rubrics
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 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 ☐ 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 **majority** of thought units ☐ **one** complete sentence that expresses an idea **with** subject-verb agreement Ex: "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	

February 28, 2018

APPENDIX H—MSAA QUALIFICATION RATES

Table H-1 summarizes the qualifications rates for this MSAA. Rates of success during qualification varied. Multiple factors determine the success of a scorer during qualification. These include familiarity with the assessment, the grade levels, and the variation of item types.

Table H-1. 2017–18 MSAA: Qualification Summary

Grade	*Passed or Failed*	*WRCC002 Qual 1*	*WRCC002 Qual 2*	*Scorers Qualified WRCC002*	*WRCC003 Qual 1*	*WRCC003 Qual 2*	*Scorers Qualified WRCC003*
Grade 3	Passed	3	5	8	5	14	19
	Failed	15	10		24	10	
Grade 4	Passed	9	3	12	7	15	22
	Failed	8	5		20	5	
Grade 5	Passed	11	0	11	14	5	19
	Failed	4	4		10	2	
Grade 6	Passed	7	8	15	23	1	24
	Failed	10	2		1	0	
Grade 7	Passed	10	2	12	3	3	6
	Failed	6	4		20	17	
Grade 8	Passed	7	7	14	5	1	6
	Failed	9	2		12	10	
Grade 11	Passed	14	0	14	8	10	18
	Failed	0	0		18	8	

Note: For identification purposes in *iScore,* Tier 2 prompts were designated as WRCC002 across all grades and Tier 3 prompts were designated WRCC003. Qual 1 = Qualification set 1; Qual 2 = Qualification set 2.

APPENDIX I—ITEM-LEVEL CLASSICAL STATISTICS

Table I-1. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 3

Item ID	*Item Type*	*p-values*	*Item-Total Correlation*	*Omit Rates*
113681A	MC	0.75	0.34	0
113682A	MC	0.69	0.32	0
114957A	MC	0.81	0.41	1
114958A	MC	0.83	0.38	0
114960A	MC	0.60	0.22	1
115985A	MC	0.73	0.41	0
115986A	MC	0.71	0.41	0
115987A	MC	0.66	0.28	0
116009A	MC	0.52	0.18	1
116010A	MC	0.60	0.28	1
116011A	MC	0.55	0.18	1
116012A	MC	0.67	0.23	1
116202A	MC	0.54	0.26	1
116203A	MC	0.49	0.43	1
116204A	MC	0.62	0.52	1
116205A	MC	0.73	0.38	1
117686A	MC	0.59	0.26	1
117687A	MC	0.59	0.41	1
117688A	MC	0.31	0.05	1
120785A	MC	0.69	0.27	1
120786A	MC	0.42	0.18	1
120787A	MC	0.64	0.26	1
120879A	MC	0.45	0.27	1
120880A	MC	0.52	0.08	0
120902A	MC	0.58	0.21	1
120912A	MC	0.70	0.40	0
120914A	MC	0.58	0.26	0
120922A	MC	0.49	0.28	1
120967A	MC	0.86	0.36	1
121194A	MC	0.45	0.18	1
121423A	MC	0.65	0.24	3
122070A	MC	0.39	0.29	1
125942A	MC	0.68	0.38	1
125943A	MC	0.63	0.19	0
125945A	MC	0.71	0.36	0
125947B	MC	0.68	0.43	2
125948A	MC	0.58	0.39	0
125949B	MC	0.67	0.32	2
448821	MC	0.74	0.41	0
448950	MC	0.68	0.45	3
449494	MC	0.43	0.32	0
449541	MC	0.32	0.28	1
451136	MC	0.80	0.31	0
451148	MC	0.83	0.30	0
451160	MC	0.82	0.36	0
451172	MC	0.77	0.40	0
451186	MC	0.81	0.30	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
451474	MC	0.69	0.30	1
451486	MC	0.61	0.25	1
451498	MC	0.55	0.32	3
451521	MC	0.34	0.25	1
451534	MC	0.55	0.45	1

MSAA MC items have either 2 or 3 options.

Table I-2. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 4

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
113087A	MC	0.86	0.28	0
113088A	MC	0.88	0.25	0
113089A	MC	0.85	0.28	0
113090A	MC	0.69	0.25	0
113091A	MC	0.69	0.32	0
113092A	MC	0.72	0.30	0
113093A	MC	0.65	0.25	0
113094A	MC	0.65	0.27	0
113097A	MC	0.56	0.20	1
113098A	MC	0.65	0.17	1
113099A	MC	0.64	0.23	1
113100A	MC	0.46	0.25	1
113280A	MC	0.65	0.38	0
113281A	MC	0.58	0.38	0
113283A	MC	0.65	0.37	1
114053A	MC	0.89	0.28	0
114054A	MC	0.75	0.27	0
114055A	MC	0.64	0.20	0
114056A	MC	0.93	0.23	0
116574A	MC	0.81	0.38	1
116576A	MC	0.83	0.36	1
116577A	MC	0.80	0.33	1
116618A	MC	0.57	0.32	2
116620A	MC	0.57	0.28	1
116621A	MC	0.59	0.38	1
121279A	MC	0.80	0.43	0
121426A	MC	0.76	0.46	0
121550A	MC	0.60	0.35	1
121551A	MC	0.57	0.21	1
121570A	MC	0.64	0.21	1
121580A	MC	0.42	0.30	1
121985A	MC	0.50	0.11	0
121987A	MC	0.38	0.10	0
122582A	MC	0.64	0.05	0
126141A	MC	0.70	0.37	0
126142A	MC	0.81	0.32	0
126143A	MC	0.71	0.27	1
126144B	MC	0.62	0.44	2
449648	MC	0.84	0.30	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
449662	MC	0.71	0.30	0
449675	MC	0.57	0.38	0
451867	MC	0.62	0.24	0
451881	MC	0.56	0.42	0
451895	MC	0.50	0.31	1
451913	MC	0.47	0.23	1
451925	MC	0.55	0.42	0
455543	MC	0.67	0.39	1
455556	MC	0.70	0.38	1
455569	MC	0.64	0.25	1
455581	MC	0.77	0.28	1
455593	MC	0.74	0.18	1
512069	MC	0.47	0.23	0

MSAA MC items have either 2 or 3 options.

Table I-3. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 5

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
114072A	MC	0.87	0.27	0
114320A	MC	0.51	0.25	0
114322A	MC	0.56	0.14	0
114323A	MC	0.71	0.37	0
114329A	MC	0.39	0.12	1
114331A	MC	0.39	0.11	1
114332A	MC	0.30	0.22	1
115053A	MC	0.75	0.47	0
115054A	MC	0.85	0.36	0
115055A	MC	0.81	0.28	1
115056A	MC	0.73	0.46	0
117523A	MC	0.57	0.31	0
117524A	MC	0.45	0.23	0
117525A	MC	0.48	0.36	0
120209A	MC	0.66	0.32	0
120210A	MC	0.55	0.47	0
120211B	MC	0.35	0.13	1
120212A	MC	0.36	0.14	0
120909A	MC	0.59	0.38	0
120910A	MC	0.61	0.36	0
121222A	MC	0.77	0.40	1
121325A	MC	0.52	0.34	0
121326B	MC	0.36	0.25	3
121457A	MC	0.70	0.20	0
121458A	MC	0.54	0.08	0
121459A	MC	0.74	0.30	0
121720B	MC	0.32	-0.03	0
121730A	MC	0.63	0.26	0
121733A	MC	0.41	0.07	0
121735A	MC	0.41	0.12	1
126984B	MC	0.74	0.39	1

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
126985B	MC	0.64	0.42	1
126986A	MC	0.78	0.26	0
126987B	MC	0.71	0.46	1
449342	MC	0.47	0.20	0
449348	MC	0.41	0.12	0
449385	MC	0.40	0.19	0
449387	MC	0.51	0.37	0
449391	MC	0.48	0.37	0
449781	MC	0.71	0.37	0
449796	MC	0.53	0.28	0
449808	MC	0.58	0.43	0
451036	MC	0.51	0.10	0
452013	MC	0.41	0.23	0
452025	MC	0.86	0.26	0
452038	MC	0.68	0.26	0
455685	MC	0.60	0.21	0
455697	MC	0.46	0.29	1
455709	MC	0.61	0.25	0
455721	MC	0.70	0.18	0

MSAA MC items have either 2 or 3 options.

Table I-4. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 6

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
113536A	MC	0.42	0.32	1
113537A	MC	0.49	0.24	2
113612A	MC	0.56	0.34	0
113614A	MC	0.56	0.39	0
114380A	MC	0.76	0.41	0
114382A	MC	0.73	0.40	1
115502A	MC	0.76	0.24	0
115503A	MC	0.84	0.24	1
119997A	MC	0.39	0.10	1
119998A	MC	0.47	0.21	1
119999A	MC	0.38	0.23	1
120000A	MC	0.44	0.16	1
120389A	MC	0.76	0.23	1
120390A	MC	0.75	0.22	1
120391A	MC	0.60	0.35	1
120392A	MC	0.68	0.24	1
121225A	MC	0.60	0.32	1
121226A	MC	0.57	0.33	1
121349A	MC	0.30	0.09	1
121353A	MC	0.57	0.45	1
121482A	MC	0.65	0.30	1
121483A	MC	0.52	0.31	1
121764A	MC	0.53	0.37	1
121768A	MC	0.59	0.35	1
121802A	MC	0.66	0.33	1

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
121803A	MC	0.76	0.38	1
121804A	MC	0.61	0.39	1
122258A	MC	0.60	0.26	0
122263A	MC	0.61	0.24	0
124257A	MC	0.53	0.27	1
127272B	MC	0.65	0.39	4
127273A	MC	0.70	0.40	0
127274A	MC	0.76	0.25	1
127276A	MC	0.60	0.18	2
127277B	MC	0.66	0.46	2
448846	MC	0.70	0.36	0
449305	MC	0.66	0.38	1
452219	MC	0.62	0.41	1
452231	MC	0.48	0.38	1
452243	MC	0.53	0.28	1
452257	MC	0.51	0.39	1
452269	MC	0.69	0.54	0
452282	MC	0.83	0.33	0
452299	MC	0.76	0.30	0
452311	MC	0.60	0.33	0
452335	MC	0.26	0.12	0
452348	MC	0.72	0.39	0
452360	MC	0.76	0.36	0

MSAA MC items have either 2 or 3 options.

Table I-5. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 7

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
113801A	MC	0.55	0.32	1
113802B	MC	0.59	0.28	3
114482A	MC	0.74	0.30	0
114483A	MC	0.82	0.35	0
114484A	MC	0.58	0.18	0
114593A	MC	0.74	0.46	1
114594A	MC	0.68	0.35	1
114596A	MC	0.66	0.41	0
114643A	MC	0.70	0.48	1
114644A	MC	0.58	0.42	0
114645A	MC	0.60	0.45	0
114646A	MC	0.75	0.41	0
115372A	MC	0.49	0.11	1
115373A	MC	0.53	0.34	1
115431A	MC	0.66	0.50	0
115432A	MC	0.59	0.42	0
115433A	MC	0.63	0.48	0
120060A	MC	0.63	0.17	0
120061A	MC	0.65	0.24	0
120072A	MC	0.48	0.18	0
120073A	MC	0.42	0.12	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
120098A	MC	0.69	0.25	3
120099A	MC	0.69	0.30	1
121313A	MC	0.63	0.45	1
121343A	MC	0.49	0.21	0
121347A	MC	0.53	0.31	1
121490A	MC	0.61	0.21	0
121491A	MC	0.41	0.13	1
121493A	MC	0.54	0.33	1
121494A	MC	0.47	0.23	0
121495A	MC	0.32	0.13	0
121497A	MC	0.62	0.16	1
121505A	MC	0.72	0.26	2
121507A	MC	0.48	0.16	1
121509A	MC	0.58	0.30	1
121513A	MC	0.59	0.21	0
121871A	MC	0.66	0.40	0
121874A	MC	0.62	0.33	0
121997A	MC	0.57	0.32	1
122037A	MC	0.54	0.26	1
122038A	MC	0.65	0.26	1
122380A	MC	0.63	0.11	0
123641A	MC	0.58	0.16	0
123649A	MC	0.47	0.11	0
127690A	MC	0.48	0.24	3
127691A	MC	0.73	0.34	1
127692B	MC	0.71	0.32	2
127693A	MC	0.65	0.34	0
127694A	MC	0.72	0.32	1
127695B	MC	0.75	0.36	1
449566	MC	0.83	0.40	1
449584	MC	0.75	0.39	1
449607	MC	0.38	0.14	1
449624	MC	0.44	0.05	0

MSAA MC items have either 2 or 3 options.

Table I-6. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 8

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
114228A	MC	0.47	0.23	1
114229A	MC	0.50	0.19	1
114230A	MC	0.57	0.35	0
114231A	MC	0.46	0.31	0
114796A	MC	0.73	0.35	0
114797A	MC	0.87	0.34	0
114798A	MC	0.87	0.41	1
114799A	MC	0.86	0.42	0
114876A	MC	0.65	0.43	1
114877A	MC	0.41	0.17	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
114879A	MC	0.63	0.46	0
115285A	MC	0.58	0.45	0
115286A	MC	0.64	0.44	0
115288A	MC	0.70	0.33	0
118798A	MC	0.75	0.38	0
118800A	MC	0.50	0.33	0
121030A	MC	0.66	0.32	2
121031A	MC	0.40	0.19	1
121032A	MC	0.63	0.14	1
121033A	MC	0.52	-0.03	1
121036A	MC	0.51	0.20	0
121037A	MC	0.69	0.29	0
121038A	MC	0.67	0.29	0
121040A	MC	0.37	0.13	1
121041A	MC	0.42	0.21	0
121042A	MC	0.55	0.26	0
121075A	MC	0.78	0.38	1
121078A	MC	0.79	0.48	0
121107A	MC	0.52	0.27	1
121148A	MC	0.38	0.04	0
121149A	MC	0.60	0.13	0
121164A	MC	0.64	0.12	0
121165A	MC	0.56	0.16	0
121202A	MC	0.51	0.47	1
121203A	MC	0.69	0.44	1
121205A	MC	0.65	0.47	0
121805A	MC	0.50	0.25	1
122082A	MC	0.54	0.31	1
122562A	MC	0.70	0.48	0
127781A	MC	0.69	0.26	3
127782A	MC	0.78	0.37	1
127783A	MC	0.82	0.43	1
127784A	MC	0.56	0.16	0
127785A	MC	0.69	0.34	0
127786B	MC	0.83	0.43	2
449868	MC	0.85	0.38	0
449882	MC	0.60	0.44	0
449900	MC	0.62	0.45	1

MSAA MC items have either 2 or 3 options.

Table I-7. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 11

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
113726A	MC	0.51	0.22	4
113727A	MC	0.47	0.15	0
113728A	MC	0.71	0.47	0
114166A	MC	0.34	0.23	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
114167A	MC	0.69	0.18	0
114193A	MC	0.56	0.15	0
114194A	MC	0.66	0.24	1
114205A	MC	0.66	0.33	1
114207A	MC	0.63	0.27	1
114208A	MC	0.51	0.23	1
116323A	MC	0.91	0.31	0
116324A	MC	0.86	0.33	0
116326A	MC	0.68	0.25	3
116348A	MC	0.74	0.33	1
116349A	MC	0.66	0.40	1
116350A	MC	0.72	0.38	0
116351A	MC	0.77	0.32	1
119078A	MC	0.64	0.55	6
119079A	MC	0.67	0.52	1
119080A	MC	0.72	0.46	0
119081A	MC	0.73	0.49	0
120148A	MC	0.84	0.36	0
120149A	MC	0.73	0.36	0
120150A	MC	0.72	0.35	1
120151A	MC	0.82	0.34	1
121229A	MC	0.74	0.37	0
121695A	MC	0.53	0.30	0
121702A	MC	0.46	0.06	0
121703A	MC	0.65	0.29	1
121711A	MC	0.65	0.26	0
121714A	MC	0.70	0.32	0
121718A	MC	0.59	0.11	1
121719A	MC	0.50	0.19	1
121745A	MC	0.53	0.29	1
121746A	MC	0.69	0.43	0
121875A	MC	0.58	0.16	1
122000A	MC	0.63	0.44	0
122538A	MC	0.56	0.10	0
126773A	MC	0.80	0.43	1
126774B	MC	0.81	0.46	2
126775A	MC	0.67	0.30	1
126776B	MC	0.59	0.30	5
126777B	MC	0.78	0.39	4
126778B	MC	0.76	0.44	2
449987	MC	0.82	0.34	0
450006	MC	0.64	0.38	0
450027	MC	0.41	0.15	1
450048	MC	0.62	0.39	0
453006	MC	0.60	0.39	5
453019	MC	0.65	0.38	1
453033	MC	0.54	0.31	0
453047	MC	0.65	0.41	1

MSAA MC items have either 2 or 3 options.

Table I-8. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 3

Item ID	*Item Type*	*p-values*	*Item-Total Correlation*	*Omit Rates*
110842A	MC	0.59	0.43	1
110855A	MC	0.25	0.20	1
110871A	MC	0.23	0.18	1
110873A	MC	0.32	0.28	1
110876A	MC	0.54	0.31	0
110919A	MC	0.26	0.18	1
110920A	MC	0.43	0.28	0
110923A	MC	0.56	0.38	1
110928A	MC	0.37	0.23	1
110959A	MC	0.72	0.20	1
110964A	MC	0.49	0.30	1
110966A	MC	0.63	0.17	0
110974A	MC	0.34	0.24	1
110975A	MC	0.57	0.29	0
111377A	MC	0.51	0.24	2
111382A	MC	0.55	0.20	0
111386A	MC	0.52	0.41	0
111387A	MC	0.47	0.34	0
111390A	MC	0.41	0.18	1
111397A	MC	0.56	0.20	1
111400A	MC	0.64	0.26	0
111416A	MC	0.60	0.23	0
111420A	MC	0.61	0.33	0
111426A	MC	0.48	0.36	0
111432A	MC	0.22	0.13	2
111434A	MC	0.74	0.36	2
111650A	MC	0.48	0.30	0
111883A	MC	0.23	0.11	2
112551A	MC	0.65	0.33	1
112552A	MC	0.76	0.29	0
112555A	MC	0.71	0.40	1
112559A	MC	0.36	0.17	1
112564A	MC	0.64	0.35	1
112565A	MC	0.33	0.08	0
112566A	MC	0.43	0.28	0
112569A	MC	0.41	0.23	1
112570A	MC	0.40	0.16	0
112575A	MC	0.60	0.29	1
112576A	MC	0.28	0.27	1
112585A	MC	0.33	0.16	0
112586A	MC	0.32	0.27	1
112595A	MC	0.54	0.46	1
112600A	MC	0.35	0.25	1
112601A	MC	0.76	0.33	0
112615A	MC	0.56	0.28	0
112616A	MC	0.35	0.25	0
120682A	MC	0.85	0.21	0
122090A	MC	0.62	0.45	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
122091A	MC	0.77	0.30	0
442037	MC	0.87	0.31	0
442130	MC	0.59	0.30	1
442166	MC	0.23	0.10	2
442226	MC	0.68	-0.07	1
442402	MC	0.83	0.29	0
442416	MC	0.30	0.30	2
451071	MC	0.21	0.13	2
451090	MC	0.40	-0.05	1
451107	MC	0.29	0.21	0
451600	MC	0.70	-0.04	0
463237	MC	0.41	0.38	1

MSAA MC items have either 2 or 3 options.

Table I-9. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 4

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
111123A	MC	0.49	0.31	1
111124A	MC	0.34	0.34	1
111135A	MC	0.58	0.16	1
111136A	MC	0.51	0.39	0
111148A	MC	0.32	0.26	1
111162A	MC	0.28	0.25	2
111166A	MC	0.50	0.36	1
111185A	MC	0.32	0.06	2
111658A	MC	0.75	0.20	0
111663A	MC	0.76	0.33	1
111672A	MC	0.74	0.18	0
111676A	MC	0.40	0.31	1
111678A	MC	0.66	0.42	0
111682A	MC	0.46	0.34	1
111685A	MC	0.59	0.24	2
111686A	MC	0.43	0.29	1
111688A	MC	0.62	0.20	0
111696A	MC	0.48	0.29	0
111707B	MC	0.47	0.28	0
111711A	MC	0.43	0.07	1
111712A	MC	0.47	0.22	0
111715A	MC	0.74	0.29	1
111716A	MC	0.45	0.31	1
111717A	MC	0.74	0.27	0
111721A	MC	0.38	0.28	0
111731A	MC	0.40	0.30	1
112783A	MC	0.67	0.36	0
112788A	MC	0.57	0.22	0
112794A	MC	0.44	0.25	1
112803A	MC	0.67	0.26	0
112812A	MC	0.40	0.29	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
112817A	MC	0.41	0.27	0
112824A	MC	0.40	0.30	1
112832B	MC	0.36	0.05	1
112833A	MC	0.55	0.20	0
112837B	MC	0.49	0.20	0
112838B	MC	0.54	0.35	0
121661A	MC	0.42	0.17	0
121663A	MC	0.35	0.28	1
121665A	MC	0.30	0.35	2
121691A	MC	0.38	0.20	1
121737A	MC	0.54	0.16	0
122153A	MC	0.75	0.23	0
122267A	MC	0.50	0.40	0
122368A	MC	0.36	0.35	0
122432A	MC	0.27	0.34	1
445588	MC	0.45	0.28	1
446106	MC	0.41	0.19	0
446131	MC	0.76	0.18	1
446178	MC	0.32	0.17	0
446443	MC	0.65	0.23	1
446780	MC	0.27	0.05	2
446795	MC	0.36	0.01	1
454850	MC	0.66	0.37	0
454860	MC	0.33	0.06	1
454863	MC	0.38	-0.03	1
455016	MC	0.46	0.08	2
463034	MC	0.10	0.10	1

MSAA MC items have either 2 or 3 options.

Table I-10. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 5

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
111234A	MC	0.51	-0.11	0
111242A	MC	0.47	0.22	1
111243A	MC	0.47	0.41	1
111244A	MC	0.55	0.43	0
111258A	MC	0.23	0.11	1
111263A	MC	0.44	-0.16	1
111276A	MC	0.48	0.33	0
111277A	MC	0.53	0.35	0
111294A	MC	0.72	0.40	0
111295A	MC	0.44	0.35	0
111298A	MC	0.32	0.25	1
111299A	MC	0.39	0.21	0
111303A	MC	0.53	0.28	1
111308A	MC	0.33	0.14	0
112335A	MC	0.45	0.25	1
112342A	MC	0.46	0.05	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
112346A	MC	0.67	0.31	1
112348A	MC	0.58	0.41	0
112352A	MC	0.40	0.12	2
112354A	MC	0.41	0.30	0
112358A	MC	0.43	0.34	0
112359A	MC	0.32	0.27	0
112363A	MC	0.38	0.29	0
112364A	MC	0.38	0.31	0
112368A	MC	0.35	0.30	1
112369A	MC	0.26	0.26	1
112372A	MC	0.72	0.33	0
112373A	MC	0.69	0.23	0
112377A	MC	0.73	0.26	1
112384A	MC	0.54	0.42	0
112385A	MC	0.65	0.33	0
112392A	MC	0.36	0.27	0
112408A	MC	0.33	0.21	0
112410A	MC	0.69	0.35	0
113843B	MC	0.30	0.15	3
113853A	MC	0.28	0.23	1
113856A	MC	0.45	0.23	0
113862A	MC	0.24	0.11	1
113863A	MC	0.41	-0.21	1
113867A	MC	0.29	0.03	0
113872A	MC	0.49	0.39	0
113877A	MC	0.36	0.16	0
113883A	MC	0.48	0.09	0
113884B	MC	0.68	0.25	2
113889A	MC	0.46	0.24	1
113892A	MC	0.52	0.38	0
113899A	MC	0.73	0.21	1
113902A	MC	0.43	0.35	0
120724A	MC	0.08	0.17	1
120737A	MC	0.39	0.46	0
120739A	MC	0.25	0.45	1
121514A	MC	0.55	0.30	0
450055	MC	0.37	0.14	0
450093	MC	0.33	0.33	0
450111	MC	0.72	0.01	0
450129	MC	0.25	0.12	0
450143	MC	0.35	-0.09	0
450180	MC	0.23	0.17	0
450200	MC	0.31	0.15	1
450210	MC	0.46	0.19	0
450274	MC	0.27	0.16	0
450339	MC	0.26	0.11	0

MSAA MC items have either 2 or 3 options.

Table I-11. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 6

Item ID	*Item Type*	*p-values*	*Item-Total Correlation*	*Omit Rates*
110891A	MC	0.74	0.35	0
110944A	MC	0.67	0.43	1
110977A	MC	0.41	0.26	0
110980A	MC	0.63	0.27	1
110981A	MC	0.63	0.38	0
110984A	MC	0.35	0.19	0
110986A	MC	0.61	0.33	1
110991A	MC	0.44	0.32	1
110993A	MC	0.64	0.28	0
111022A	MC	0.52	0.46	0
111025A	MC	0.53	0.32	1
111036A	MC	0.43	0.16	1
111445A	MC	0.65	0.45	0
111455A	MC	0.69	0.39	0
111456A	MC	0.50	0.39	0
111465A	MC	0.59	0.35	0
111479A	MC	0.73	0.41	0
111482A	MC	0.50	0.20	1
111487A	MC	0.48	0.16	2
111508A	MC	0.42	0.19	1
111514A	MC	0.56	0.38	0
111518A	MC	0.57	0.37	0
111630A	MC	0.78	0.34	0
112632A	MC	0.83	0.29	0
112633A	MC	0.34	0.21	1
112645A	MC	0.63	0.35	0
112655A	MC	0.45	0.24	0
112656A	MC	0.59	0.28	3
112658A	MC	0.64	0.30	0
112663A	MC	0.36	0.11	1
112676A	MC	0.60	0.01	1
112678A	MC	0.42	0.25	1
112679A	MC	0.82	0.35	0
112691A	MC	0.71	0.08	1
112697A	MC	0.49	0.33	1
112699A	MC	0.62	0.34	0
120494A	MC	0.59	0.35	1
120855A	MC	0.38	0.13	1
121487A	MC	0.67	0.44	0
442356	MC	0.76	0.37	0
442369	MC	0.45	0.25	1
442538	MC	0.58	0.37	0
442566	MC	0.73	0.38	1
442628	MC	0.37	0.32	1
442631	MC	0.36	0.21	1
442641	MC	0.68	0.36	0
442711	MC	0.70	0.29	1
442785	MC	0.39	0.06	1

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
442813	MC	0.76	0.32	1
450368	MC	0.51	0.11	1
450436	MC	0.57	0.26	2
450459	MC	0.32	0.22	1
453664	MC	0.68	0.30	0
453707	MC	0.44	0.32	1
453730	MC	0.27	0.07	1

MSAA MC items have either 2 or 3 options.

Table I-12. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 7

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
111046A	MC	0.74	0.17	1
111048A	MC	0.58	0.36	0
111055A	MC	0.46	0.35	1
111067A	MC	0.41	0.21	0
111069A	MC	0.60	0.20	1
111071A	MC	0.55	0.37	0
111075A	MC	0.83	0.41	0
111076A	MC	0.38	0.32	1
111080A	MC	0.44	0.33	1
111085A	MC	0.55	0.20	0
111092A	MC	0.59	0.06	2
111093A	MC	0.51	0.29	1
111100A	MC	0.33	0.12	3
111104A	MC	0.51	0.26	1
111105A	MC	0.51	0.37	0
111106A	MC	0.36	0.33	0
111113A	MC	0.43	0.33	1
111119A	MC	0.34	0.13	1
111127A	MC	0.47	0.38	1
111130A	MC	0.56	0.10	1
111734A	MC	0.60	0.38	1
111744A	MC	0.49	0.07	2
111748A	MC	0.74	0.16	1
111754B	MC	0.51	0.29	1
111759B	MC	0.47	0.28	1
111764B	MC	0.73	0.36	0
111765A	MC	0.49	0.37	0
111766A	MC	0.66	0.43	0
111769A	MC	0.63	0.33	0
111779A	MC	0.42	0.24	1
111780A	MC	0.27	0.03	2
111795A	MC	0.44	0.41	0
111796A	MC	0.63	0.35	0
111799A	MC	0.38	0.06	1
111804A	MC	0.45	0.19	1
112523B	MC	0.59	0.26	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
112605A	MC	0.45	0.27	0
112849A	MC	0.49	0.15	1
112852A	MC	0.48	0.08	1
112853A	MC	0.34	0.06	4
112887A	MC	0.51	0.36	0
112899A	MC	0.63	0.26	0
112909B	MC	0.73	0.30	1
112910A	MC	0.70	0.32	0
112911A	MC	0.50	0.33	0
113101A	MC	0.66	0.42	0
446543	MC	0.46	0.43	0
446673	MC	0.78	0.32	0
446831	MC	0.26	-0.09	6
446838	MC	0.40	0.11	2
446891	MC	0.33	0.00	0
451801	MC	0.42	0.21	1
451854	MC	0.36	-0.10	1
452439	MC	0.67	0.19	1
453976	MC	0.34	0.10	1
454049	MC	0.62	0.37	0
454054	MC	0.67	0.17	0

MSAA MC items have either 2 or 3 options.

Table I-13. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 8

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
111247A	MC	0.70	0.44	0
111281A	MC	0.58	0.19	0
111286A	MC	0.29	0.23	1
111335A	MC	0.37	0.25	1
111339A	MC	0.48	0.29	0
111352A	MC	0.74	0.39	0
111560A	MC	0.44	0.39	1
111565A	MC	0.65	0.41	1
111581A	MC	0.41	0.27	1
111583A	MC	0.43	0.28	0
111594A	MC	0.60	0.31	0
111615A	MC	0.59	0.35	1
111622A	MC	0.57	0.41	0
112460A	MC	0.43	0.31	1
112466A	MC	0.48	0.34	1
112470A	MC	0.63	0.28	1
112475A	MC	0.60	0.13	1
112486A	MC	0.43	0.33	0
112490A	MC	0.66	0.31	0
112491A	MC	0.33	0.16	1
112499A	MC	0.53	0.39	4
112500A	MC	0.36	0.07	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
112506A	MC	0.61	0.26	0
112509A	MC	0.55	0.26	1
112516A	MC	0.47	0.45	0
113908A	MC	0.33	0.27	1
113909A	MC	0.48	0.39	1
113917A	MC	0.77	0.39	0
113918A	MC	0.71	0.46	1
113931A	MC	0.28	0.05	1
113932A	MC	0.51	0.38	0
113937A	MC	0.45	0.38	0
113942A	MC	0.30	0.25	4
113943A	MC	0.39	0.21	0
113952A	MC	0.28	0.23	0
113959A	MC	0.40	0.28	0
113963A	MC	0.42	0.46	1
113964A	MC	0.65	0.42	0
113968A	MC	0.27	0.14	1
113973A	MC	0.56	0.38	1
117071A	MC	0.61	-0.15	1
117072A	MC	0.44	0.40	1
120568A	MC	0.64	0.30	0
120571A	MC	0.29	0.19	1
122051A	MC	0.58	0.41	0
122099A	MC	0.29	0.12	1
446956	MC	0.60	0.13	0
446979	MC	0.65	0.32	0
447047	MC	0.55	0.36	0
447054	MC	0.48	0.40	1
447109	MC	0.36	0.17	1
447166	MC	0.61	0.37	2
454069	MC	0.25	0.26	1
454255	MC	0.67	0.31	2
462215	MC	0.63	0.41	0
462289	MC	0.30	0.26	1
471660	MC	0.45	-0.12	1

MSAA MC items have either 2 or 3 options.

Table I-14. 2017–18 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 11

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
110843A	MC	0.48	0.38	0
110858A	MC	0.56	0.25	0
110881A	MC	0.48	0.32	1
110913A	MC	0.48	0.38	0
110914A	MC	0.67	0.40	0
110921A	MC	0.51	0.32	0
110936A	MC	0.62	0.28	1
111000A	MC	0.63	0.35	0

continued

Item ID	Item Type	p-values	Item-Total Correlation	Omit Rates
111002A	MC	0.59	0.30	0
111016A	MC	0.36	0.12	1
111024A	MC	0.72	0.17	5
111533A	MC	0.58	0.18	1
111537A	MC	0.50	0.31	0
111538A	MC	0.43	0.41	0
111539A	MC	0.47	0.37	1
111548A	MC	0.39	0.22	0
111553A	MC	0.41	0.32	1
111557A	MC	0.34	0.09	0
111810A	MC	0.72	0.40	0
111813A	MC	0.69	0.22	1
111815A	MC	0.47	0.42	1
111818A	MC	0.49	0.36	2
111819A	MC	0.27	0.02	1
111824A	MC	0.38	0.12	2
111828A	MC	0.38	0.29	1
111829A	MC	0.44	0.34	0
111830A	MC	0.61	0.15	2
111833A	MC	0.43	0.24	1
111840A	MC	0.39	0.12	1
112701A	MC	0.56	0.28	1
112702A	MC	0.58	0.43	0
112708A	MC	0.38	0.16	1
112709A	MC	0.37	0.40	0
112717A	MC	0.37	0.16	2
112722A	MC	0.46	0.25	1
112727A	MC	0.60	0.36	1
112733A	MC	0.46	0.27	0
112743A	MC	0.43	0.03	2
112744A	MC	0.65	0.35	0
112929A	MC	0.53	-0.03	2
112930B	MC	0.31	-0.01	3
112940A	MC	0.54	0.33	0
112946A	MC	0.49	0.26	2
122055A	MC	0.36	-0.03	0
442881	MC	0.48	0.12	0
442911	MC	0.32	-0.08	1
442914	MC	0.65	0.02	5
443287	MC	0.43	0.15	0
443312	MC	0.37	0.07	1
443494	MC	0.30	0.09	2
443575	MC	0.70	0.33	1
454925	MC	0.48	0.15	1
454980	MC	0.28	0.02	1
454987	MC	0.42	0.25	1
462343	MC	0.42	0.26	1
462630	MC	0.56	0.18	0

MSAA MC items have either 2 or 3 options.

APPENDIX J—DIFFERENTIAL ITEM FUNCTIONING RESULTS

Table J-1. 2017–18 MSAA: Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored—Mathematics

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
03	Male	Female	MC	60	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	60	7	3	4	0	0	0
	Non-LEP	LEP	MC	30	7	4	3	1	1	0
	White	American Indian or Alaska Native	MC	15	6	3	3	0	0	0
		Black or African American	MC	60	6	1	5	0	0	0
		Hispanic or Latino	MC	60	4	2	2	0	0	0
04	Male	Female	MC	58	3	3	0	0	0	0
	Non-EconDis	EconDis	MC	58	2	0	2	1	1	0
	Non-LEP	LEP	MC	31	7	4	3	1	0	1
	White	American Indian or Alaska Native	MC	15	4	2	2	1	1	0
		Asian	MC	15	5	4	1	2	2	0
		Black or African American	MC	58	11	7	4	0	0	0
		Hispanic or Latino	MC	58	11	6	5	0	0	0
		Two or More Races	MC	15	2	1	1	1	1	0
05	Male	Female	MC	62	3	3	0	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	62	5	1	4	2	1	1
	Non-LEP	LEP	MC	28	6	3	3	0	0	0
	White	American Indian or Alaska Native	MC	15	3	1	2	1	1	0
		Black or African American	MC	62	5	1	4	0	0	0
		Hispanic or Latino	MC	62	6	3	3	2	1	1
06	Male	Female	MC	55	3	2	1	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	55	4	2	2	0	0	0
	Non-LEP	LEP	MC	15	0	0	0	0	0	0
	White	Asian	MC	15	4	3	1	2	2	0
		Black or African American	MC	55	5	3	2	0	0	0
		Hispanic or Latino	MC	55	2	1	1	0	0	0
07	Male	Female	MC	57	3	2	1	2	1	1
	Non-EconDis	Economically Disadvantaged	MC	57	7	1	6	1	1	0
	Non-LEP	LEP	MC	24	7	3	4	1	0	1

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
07	White	American Indian or Alaska Native	MC	15	3	2	1	1	0	1
		Black or African American	MC	57	9	3	6	0	0	0
		Hispanic or Latino	MC	57	3	3	0	0	0	0
		Two or More Races	MC	15	7	4	3	3	3	0
08	Male	Female	MC	57	3	2	1	0	0	0
	Non-EconDis	EconDis	MC	57	4	2	2	0	0	0
	Non-LEP	LEP	MC	16	4	3	1	0	0	0
	White	Black or African American	MC	57	6	2	4	0	0	0
		Hispanic or Latino	MC	57	2	1	1	1	0	1
11	Male	Female	MC	56	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	56	3	1	2	0	0	0
	Non-LEP	LEP	MC	15	5	4	1	0	0	0
	White	Black or African American	MC	56	7	5	2	0	0	0
		Hispanic or Latino	MC	56	8	7	1	0	0	0

Table J-2. 2017–18 MSAA: Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored—ELA

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
03	Male	Female	MC	52	1	0	1	0	0	0
			OR	6	0	0	0	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	52	2	1	1	0	0	0
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	2	2	0	0	0	0
	White	American Indian or Alaska Native	MC	24	7	3	4	2	1	1
		Black or African American	MC	52	4	2	2	0	0	0
			OR	6	0	0	0	0	0	0
		Hispanic or Latino	MC	52	2	2	0	1	1	0
			OR	6	0	0	0	0	0	0
04	Male	Female	MC	52	6	6	0	0	0	0
			OR	6	0	0	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
	Non-EconDis	Economically Disadvantaged	MC	52	2	0	2	0	0	0
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	6	6	0	0	0	0
04	White	American Indian or Alaska Native	MC	27	9	5	4	2	1	1
		Asian	MC	27	8	6	2	6	6	0
		Black or African American	MC	52	4	2	2	0	0	0
			OR	6	0	0	0	0	0	0
		Hispanic or Latino	MC	52	6	4	2	0	0	0
			OR	6	0	0	0	0	0	0
		Two or More Races	MC	27	8	3	5	1	1	0
05	Male	Female	MC	50	1	1	0	1	1	0
			OR	6	0	0	0	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	50	2	0	2	0	0	0
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	7	5	2	0	0	0
	White	American Indian or Alaska Native	MC	28	5	4	1	3	3	0
		Black or African American	MC	50	3	2	1	0	0	0
			OR	6	0	0	0	0	0	0
		Hispanic or Latino	MC	50	6	5	1	0	0	0
			OR	6	0	0	0	0	0	0
06	Male	Female	MC	48	2	1	1	0	0	0
			OR	6	1	0	1	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	48	0	0	0	0	0	0
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	38	11	7	4	1	1	0
			OR	3	1	0	1	0	0	0
	White	Asian	MC	28	8	7	1	3	3	0
		Black or African American	MC	48	8	3	5	0	0	0
			OR	6	0	0	0	0	0	0
		Hispanic or Latino	MC	48	4	0	4	0	0	0
			OR	6	1	1	0	0	0	0
07	Male	Female	MC	54	0	0	0	1	1	0
			OR	6	0	0	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
07	Non-EconDis	Economically Disadvantaged	MC	54	3	2	1	0	0	0
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	3	2	1	4	3	1
	White	American Indian or Alaska Native	MC	28	7	3	4	0	0	0
		Black or African American	MC	54	4	3	1	1	1	0
			OR	6	0	0	0	0	0	0
		Hispanic or Latino	MC	54	3	3	0	0	0	0
			OR	6	0	0	0	0	0	0
		Two or More Races	MC	28	10	7	3	0	0	0
08	Male	Female	MC	48	0	0	0	0	0	0
			OR	6	0	0	0	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	48	1	1	0	1	1	0
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	9	5	4	0	0	0
	White	American Indian or Alaska Native	MC	33	6	3	3	2	1	1
		Black or African American	MC	48	3	1	2	0	0	0
			OR	6	1	1	0	0	0	0
		Hispanic or Latino	MC	48	2	2	0	0	0	0
			OR	6	0	0	0	0	0	0
11	Male	Female	MC	52	1	0	1	0	0	0
			OR	6	0	0	0	0	0	0
	Non-EconDis	Economically Disadvantaged	MC	52	4	2	2	1	0	1
			OR	6	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	7	5	2	0	0	0
	White	Black or African American	MC	52	4	3	1	0	0	0
			OR	6	0	0	0	0	0	0
		Hispanic or Latino	MC	52	7	5	2	0	0	0
			OR	6	1	1	0	0	0	0

Table J-3. 2017–18 MSAA: DIF by Path— Mathematics Grade 3

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	3	1	2	0	0	0
	Non-LEP	LEP	MC	21	5	2	3	1	1	0
	White	American Indian or Alaska Native	MC	15	6	3	3	0	0	0
		Black or African American	MC	35	2	0	2	0	0	0
		Hispanic or Latino	MC	35	2	2	0	0	0	0
B	Male	Female	MC	35	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	35	3	2	1	0	0	0
	Non-LEP	LEP	MC	30	7	4	3	1	1	0
	White	American Indian or Alaska Native	MC	15	6	3	3	0	0	0
		Black or African American	MC	35	3	1	2	0	0	0
		Hispanic or Latino	MC	35	2	0	2	0	0	0
C	Male	Female	MC	35	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	35	3	1	2	0	0	0
	Non-LEP	LEP	MC	24	5	3	2	1	1	0
	White	American Indian or Alaska Native	MC	15	6	3	3	0	0	0
		Black or African American	MC	35	5	1	4	0	0	0
		Hispanic or Latino	MC	35	2	0	2	0	0	0

Table J-4. 2017–18 MSAA: DIF by Path— Mathematics Grade 4

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	34	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	34	0	0	0	0	0	0
	Non-LEP	LEP	MC	25	7	4	3	1	0	1
	White	American Indian or Alaska Native	MC	15	4	2	2	1	1	0
		Asian	MC	15	5	4	1	2	2	0
		Black or African American	MC	34	4	3	1	0	0	0
		Hispanic or Latino	MC	34	4	2	2	0	0	0
		Two or More Races	MC	15	2	1	1	1	1	0
B	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	31	7	4	3	1	0	1
	White	American Indian or Alaska Native	MC	15	4	2	2	1	1	0
		Asian	MC	15	5	4	1	2	2	0
		Black or African American	MC	35	7	5	2	0	0	0
		Hispanic or Latino	MC	35	7	4	3	0	0	0
		Two or More Races	MC	15	2	1	1	1	1	0
C	Male	Female	MC	35	2	2	0	0	0	0
	Non-EconDis	EconDis	MC	35	2	0	2	1	1	0
	Non-LEP	LEP	MC	21	3	3	0	0	0	0
	White	American Indian or Alaska Native	MC	15	4	2	2	1	1	0
		Asian	MC	15	5	4	1	2	2	0
		Black or African American	MC	35	6	3	3	0	0	0
		Hispanic or Latino	MC	35	5	3	2	0	0	0
		Two or More Races	MC	15	2	1	1	1	1	0

Table J-5. 2017–18 MSAA: DIF by Path— Mathematics Grade 5

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	35	2	2	0	0	0	0
	Non-EconDis	EconDis	MC	35	2	1	1	1	0	1
	Non-LEP	LEP	MC	20	3	2	1	0	0	0
	White	American Indian or Alaska Native	MC	15	3	1	2	1	1	0
		Black or African American	MC	35	1	0	1	0	0	0
		Hispanic or Latino	MC	35	1	0	1	1	0	1
B	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	2	0	2	1	1	0
	Non-LEP	LEP	MC	28	6	3	3	0	0	0
	White	American Indian or Alaska Native	MC	15	3	1	2	1	1	0
		Black or African American	MC	35	2	1	1	0	0	0
		Hispanic or Latino	MC	35	2	1	1	0	0	0
C	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	2	0	2	0	0	0
	Non-LEP	LEP	MC	23	4	2	2	0	0	0
	White	American Indian or Alaska Native	MC	15	3	1	2	1	1	0
		Black or African American	MC	35	3	1	2	0	0	0
		Hispanic or Latino	MC	35	3	2	1	1	1	0

Table J-6. 2017–18 MSAA: DIF by Path— Mathematics Grade 6

Path	*Group*		*Item Type*	*Number of Items*	*Number "Low"*			*Number "High"*		
	Reference	*Focal*			*Total*	*Favoring*		*Total*	*Favoring*	
						Reference	*Focal*		*Reference*	*Focal*
A	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	3	1	2	0	0	0
	Non-LEP	LEP	MC	15	0	0	0	0	0	0
	White	Asian	MC	15	4	3	1	2	2	0
		Black or African American	MC	35	3	2	1	0	0	0
		Hispanic or Latino	MC	35	1	0	1	0	0	0
B	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	2	0	2	0	0	0
	Non-LEP	LEP	MC	15	0	0	0	0	0	0
	White	Asian	MC	15	4	3	1	2	2	0
		Black or African American	MC	35	3	1	2	0	0	0
		Hispanic or Latino	MC	35	0	0	0	0	0	0
C	Male	Female	MC	35	2	1	1	0	0	0
	Non-EconDis	EconDis	MC	35	2	1	1	0	0	0
	Non-LEP	LEP	MC	15	0	0	0	0	0	0
	White	Asian	MC	15	4	3	1	2	2	0
		Black or African American	MC	35	3	1	2	0	0	0
		Hispanic or Latino	MC	35	1	1	0	0	0	0

Table J-7. 2017–18 MSAA: DIF by Path— Mathematics Grade 7

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	35	3	2	1	0	0	0
	Non-EconDis	EconDis	MC	35	4	1	3	1	1	0
	Non-LEP	LEP	MC	15	3	2	1	1	0	1
	White	American Indian or Alaska Native	MC	15	3	2	1	1	0	1
		Black or African American	MC	35	4	2	2	0	0	0
		Hispanic or Latino	MC	35	3	3	0	0	0	0
		Two or More Races	MC	15	7	4	3	3	3	0
B	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	3	0	3	0	0	0
	Non-LEP	LEP	MC	24	7	3	4	1	0	1
	White	American Indian or Alaska Native	MC	15	3	2	1	1	0	1
		Black or African American	MC	35	3	0	3	0	0	0
		Hispanic or Latino	MC	35	2	2	0	0	0	0
		Two or More Races	MC	15	7	4	3	3	3	0
C	Male	Female	MC	35	1	1	0	2	1	1
	Non-EconDis	EconDis	MC	35	2	0	2	0	0	0
	Non-LEP	LEP	MC	24	7	3	4	1	0	1
	White	American Indian or Alaska Native	MC	15	3	2	1	1	0	1
		Black or African American	MC	35	5	1	4	0	0	0
		Hispanic or Latino	MC	35	2	2	0	0	0	0
		Two or More Races	MC	15	7	4	3	3	3	0

Table J-8. 2017–18 MSAA: DIF by Path— Mathematics Grade 8

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	35	3	2	1	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	16	4	3	1	0	0	0
	White	Black or African American	MC	35	2	0	2	0	0	0
		Hispanic or Latino	MC	35	0	0	0	1	0	1
B	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	3	2	1	0	0	0
	Non-LEP	LEP	MC	15	4	3	1	0	0	0
	White	Black or African American	MC	35	1	1	0	0	0	0
		Hispanic or Latino	MC	35	1	1	0	0	0	0
C	Male	Female	MC	35	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	2	0	2	0	0	0
	Non-LEP	LEP	MC	16	4	3	1	0	0	0
	White	Black or African American	MC	35	3	1	2	0	0	0
		Hispanic or Latino	MC	35	1	0	1	0	0	0

Table J-9. 2017–18 MSAA: DIF by Path— Mathematics Grade 11

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	1	1	0	0	0	0
	Non-LEP	LEP	MC	15	5	4	1	0	0	0
	White	Black or African American	MC	35	4	4	0	0	0	0
		Hispanic or Latino	MC	35	4	4	0	0	0	0
B	Male	Female	MC	35	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	15	5	4	1	0	0	0
	White	Black or African American	MC	35	5	4	1	0	0	0
		Hispanic or Latino	MC	35	3	3	0	0	0	0
C	Male	Female	MC	35	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	35	2	0	2	0	0	0
	Non-LEP	LEP	MC	15	5	4	1	0	0	0
	White	Black or African American	MC	35	4	2	2	0	0	0
		Hispanic or Latino	MC	35	5	4	1	0	0	0

Table J-10. 2017–18 MSAA: DIF by Path— ELA Grade 3

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	2	2	0	0	0	0
	White	Am. Indian or Alaska Native	MC	24	7	3	4	2	1	1
		Black or African American	MC	38	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	2	0	0	0	0
			OR	3	0	0	0	0	0	0
B	Male	Female	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	2	2	0	0	0	0
	White	Am. Indian or Alaska Native	MC	24	7	3	4	2	1	1
		Black or African American	MC	38	4	2	2	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	2	0	0	0	0
			OR	3	0	0	0	0	0	0
C	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	24	0	0	0	0	0	0
	White	Am. Indian or Alaska Native	MC	24	7	3	4	2	1	1
		Black or African American	MC	38	3	2	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	1	1	0	1	1	0
			OR	3	0	0	0	0	0	0

Table J-11. 2017–18 MSAA: DIF by Path— ELA Grade 4

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring		Total	Favoring	
	Reference	Focal				Reference	Focal		Reference	Focal
A	Male	Female	MC	37	4	4	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	37	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	6	6	0	0	0	0
	White	Am. Indian or Alaska Native	MC	27	9	5	4	2	1	1
		Asian	MC	27	8	6	2	6	6	0
		Black or African American	MC	37	3	2	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	37	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
		Two or More Races	MC	27	8	3	5	1	1	0
B	Male	Female	MC	37	3	3	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	37	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	6	6	0	0	0	0
	White	Am. Indian or Alaska Native	MC	27	9	5	4	2	1	1
		Asian	MC	27	8	6	2	6	6	0
		Black or African American	MC	37	4	2	2	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	37	2	0	2	0	0	0
			OR	3	0	0	0	0	0	0
		Two or More Races	MC	27	8	3	5	1	1	0
C	Male	Female	MC	37	3	3	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	37	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	27	6	6	0	0	0	0
	White	Am. Indian or Alaska Native	MC	27	9	5	4	2	1	1
		Asian	MC	27	8	6	2	6	6	0
		Black or African American	MC	37	3	2	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	37	3	3	0	0	0	0
			OR	3	0	0	0	0	0	0
		Two or More Races	MC	27	8	3	5	1	1	0

Table J-12. 2017–18 MSAA: DIF by Path— ELA Grade 5

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	37	0	0	0	1	1	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	37	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	7	5	2	0	0	0
	White	Am. Indian or Alaska Native	MC	28	5	4	1	3	3	0
		Black or African American	MC	37	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	37	2	2	0	0	0	0
			OR	3	0	0	0	0	0	0
B	Male	Female	MC	37	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	37	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	7	5	2	0	0	0
	White	Am. Indian or Alaska Native	MC	28	5	4	1	3	3	0
		Black or African American	MC	37	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	37	5	4	1	0	0	0
			OR	3	0	0	0	0	0	0
C	Male	Female	MC	37	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	37	2	0	2	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	7	5	2	0	0	0
	White	Am. Indian or Alaska Native	MC	28	5	4	1	3	3	0
		Black or African American	MC	37	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	37	3	3	0	0	0	0
			OR	3	0	0	0	0	0	0

Table J-13. 2017–18 MSAA: DIF by Path— ELA Grade 6

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	38	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	8	6	2	1	1	0
	White	Asian	MC	28	8	7	1	3	3	0
		Black or African American	MC	38	5	3	2	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	0	2	0	0	0
			OR	3	1	1	0	0	0	0
B	Male	Female	MC	38	2	1	1	0	0	0
			OR	3	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	38	11	7	4	1	1	0
			OR	3	1	0	1	0	0	0
	White	Asian	MC	28	8	7	1	3	3	0
		Black or African American	MC	38	3	2	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
C	Male	Female	MC	38	2	1	1	0	0	0
			OR	3	1	0	1	0	0	0
	Non-EconDis	EconDis	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	34	8	4	4	1	1	0
			OR	3	1	0	1	0	0	0
	White	Asian	MC	28	8	7	1	3	3	0
		Black or African American	MC	38	5	2	3	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	0	2	0	0	0
			OR	3	0	0	0	0	0	0

Table J-14. 2017–18 MSAA: DIF by Path— ELA Grade 7

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	3	2	1	4	3	1
	White	Am. Indian or Alaska Native	MC	28	7	3	4	0	0	0
		Black or African American	MC	38	3	2	1	1	1	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	3	3	0	0	0	0
			OR	3	0	0	0	0	0	0
		Two or More Races	MC	28	10	7	3	0	0	0
B	Male	Female	MC	38	0	0	0	1	1	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	3	2	1	4	3	1
	White	Am. Indian or Alaska Native	MC	28	7	3	4	0	0	0
		Black or African American	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	2	0	0	0	0
			OR	3	0	0	0	0	0	0
		Two or More Races	MC	28	10	7	3	0	0	0
C	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	28	3	2	1	4	3	1
	White	Am. Indian or Alaska Native	MC	28	7	3	4	0	0	0
		Black or African American	MC	38	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	2	0	0	0	0
			OR	3	0	0	0	0	0	0
		Two or More Races	MC	28	10	7	3	0	0	0

Table J-15. 2017–18 MSAA: DIF by Path— ELA Grade 8

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
A	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	9	5	4	0	0	0
	White	Am. Indian or Alaska Native	MC	33	6	3	3	2	1	1
		Black or African American	MC	38	1	0	1	0	0	0
			OR	3	1	1	0	0	0	0
		Hispanic or Latino	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
B	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	0	0	0	1	1	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	9	5	4	0	0	0
	White	Am. Indian or Alaska Native	MC	33	6	3	3	2	1	1
		Black or African American	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
C	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	9	5	4	0	0	0
	White	Am. Indian or Alaska Native	MC	33	6	3	3	2	1	1
		Black or African American	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0

Table J-16. 2017–18 MSAA: DIF by Path— ELA Grade 11

Path	Group		Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring		Total	Favoring	
	Reference	Focal				Reference	Focal		Reference	Focal
A	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	2	1	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	7	5	2	0	0	0
	White	Black or African American	MC	38	3	2	1	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
B	Male	Female	MC	38	1	0	1	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	2	1	1	1	0	1
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	7	5	2	0	0	0
	White	Black or African American	MC	38	2	2	0	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	2	1	1	0	0	0
			OR	3	1	1	0	0	0	0
C	Male	Female	MC	38	0	0	0	0	0	0
			OR	3	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	38	3	1	2	0	0	0
			OR	3	0	0	0	0	0	0
	Non-LEP	LEP	MC	29	7	5	2	0	0	0
	White	Black or African American	MC	38	1	1	0	0	0	0
			OR	3	0	0	0	0	0	0
		Hispanic or Latino	MC	38	6	5	1	0	0	0
			OR	3	1	1	0	0	0	0

APPENDIX K—ITEM RESPONSE THEORY PARAMETERS

Table K-1. 2017–18 MSAA: IRT Parameters for ELA Grade 3

IREF	*a*	*SE (a)*	*b*	*SE (b)*
114957A	1.0820	0.0525	-1.3729	0.0390
448821	0.8543	0.0394	-1.1367	0.0402
448950	0.8589	0.0369	-0.8439	0.0346
449494	0.4683	0.0241	0.3503	0.0548
449541	0.4158	0.0237	1.0922	0.0810
451474	0.5007	0.0277	-1.1918	0.0659
451486	0.3960	0.0240	-0.8393	0.0689
451498	0.4855	0.0251	-0.3906	0.0500
451521	0.3631	0.0226	1.0831	0.0910
451534	0.7748	0.0318	-0.3372	0.0336
116009A	0.5700	0.0668	-1.0599	0.0639
114958A	1.0570	0.0528	-1.4538	0.0423
116010A	0.9464	0.0819	-1.2813	0.0438
116011A	0.8119	0.0762	-1.1546	0.0470
116012A	0.9242	0.0825	-1.5130	0.0540
117686A	0.6420	0.0471	-1.0421	0.0500
117687A	1.1530	0.0626	-0.9486	0.0289
117688A	0.2271	0.0369	1.4591	0.3656
120785A	1.1328	0.0922	-1.5128	0.0457
120786A	0.5526	0.0674	-0.6291	0.0806
120787A	1.0546	0.0873	-1.3953	0.0435
120902A	0.8114	0.0763	-1.2585	0.0494
114960A	0.3346	0.0227	-0.8813	0.0821
121194A	0.4671	0.0422	-0.4054	0.0662
121423A	1.0048	0.0858	-1.4811	0.0489
122070A	0.6940	0.0482	-0.2682	0.0499
125942A	0.6235	0.0304	-0.9757	0.0484
125943A	0.2908	0.0222	-1.2128	0.1090
125945A	0.6016	0.0305	-1.1279	0.0541
125947B	0.7793	0.0350	-0.9367	0.0391
125948A	0.5817	0.0274	-0.4831	0.0434
125949B	0.5086	0.0276	-1.1071	0.0621
116202A	0.3834	0.0230	-0.3624	0.0616
113681A	0.7059	0.0490	-0.6095	0.0711
113682A	0.5979	0.0427	-0.4173	0.0716
115985A	0.9206	0.0552	-0.3644	0.0474
115986A	0.9288	0.0542	-0.2681	0.0442
115987A	0.5016	0.0388	-0.3379	0.0798
120879A	0.4567	0.0347	0.7961	0.0679
120880A	0.1795	0.0273	0.2618	0.1622
120912A	0.8560	0.0514	-0.2831	0.0478
116203A	0.6700	0.0286	-0.0880	0.0378
120914A	0.4471	0.0356	0.0719	0.0719
451136	0.7740	0.0742	-0.3019	0.1049
451148	0.8642	0.0824	-0.3666	0.1028
451160	1.1231	0.0933	-0.1089	0.0666
451172	1.2039	0.0913	0.1184	0.0514
451186	0.8827	0.0801	-0.2336	0.0897
116204A	1.0521	0.0412	-0.5850	0.0273
116205A	0.7288	0.0350	-1.1560	0.0465
120922A	0.3924	0.0229	-0.0349	0.0595
120967A	1.1258	0.0599	-1.6048	0.0462

Table K-2. 2017–18 MSAA: IRT Parameters for ELA Grade 4

IREF	*a*	*SE (a)*	*b*	*SE (b)*
113280A	0.7088	0.0316	-0.6638	0.0388
121551A	0.3307	0.0225	-0.5145	0.0734
449648	0.6973	0.0394	-1.6676	0.0748
449662	0.5102	0.0283	-1.1402	0.0668
449675	0.6158	0.0279	-0.3205	0.0394
451867	0.3520	0.0233	-0.8356	0.0795
451881	0.7569	0.0311	-0.2510	0.0328
451895	0.4888	0.0247	0.0510	0.0469
451913	0.3829	0.0228	0.2734	0.0600
451925	0.7779	0.0315	-0.2096	0.0320
113097A	0.9256	0.0815	-1.0216	0.0408
113281A	0.6958	0.0298	-0.3230	0.0355
113098A	0.6203	0.0699	-1.4490	0.0812
113099A	0.8635	0.0800	-1.2857	0.0529
113100A	1.0758	0.0873	-0.7595	0.0354
116618A	0.4346	0.0243	-0.4226	0.0552
116620A	0.3727	0.0232	-0.4654	0.0645
116621A	0.5658	0.0271	-0.4077	0.0433
121570A	0.7709	0.0762	-1.3406	0.0612
121580A	0.4480	0.0239	0.5439	0.0563
126141A	0.6840	0.0323	-0.9134	0.0453
126142A	0.6383	0.0355	-1.5475	0.0728
113283A	0.7148	0.0315	-0.6251	0.0379
126143A	0.4414	0.0267	-1.2523	0.0809
126144B	0.8886	0.0356	-0.4878	0.0301
455543	1.3222	0.0643	-0.9047	0.0260
455556	1.1343	0.0593	-1.0327	0.0322
455569	0.5505	0.0406	-1.0812	0.0625
455581	0.9195	0.0556	-1.3539	0.0501
455593	0.4798	0.0414	-1.7606	0.1149
116574A	0.9684	0.0470	-1.2771	0.0436
512069	0.3340	0.0220	0.2651	0.0678
113091A	1.3788	0.1072	-0.3594	0.0384
113092A	1.2179	0.1032	-0.4998	0.0490
113093A	0.9138	0.0860	-0.3873	0.0563
113094A	1.0676	0.0917	-0.3301	0.0466
121987A	0.3581	0.0613	0.9162	0.1735
113087A	1.2030	0.1185	-0.0298	0.0835
116576A	0.8738	0.0447	-1.4147	0.0526
113088A	1.1291	0.1206	-0.1876	0.1040
113089A	1.2181	0.1162	0.0357	0.0769
113090A	0.6276	0.0674	0.3423	0.0980
114053A	1.2446	0.1325	-0.1814	0.0976
114054A	0.7895	0.0789	0.1982	0.0911
114055A	0.5288	0.0614	0.4581	0.1045
114056A	1.0773	0.1400	-0.5590	0.1629
121985A	0.3563	0.0512	1.1767	0.1118
122582A	0.2527	0.0472	-0.2231	0.2964
116577A	0.7437	0.0388	-1.4176	0.0593
121279A	1.2578	0.0562	-1.0735	0.0302
121426A	1.2592	0.0530	-0.9089	0.0270
121550A	0.6179	0.0284	-0.4497	0.0405

Table K-3. 2017–18 MSAA: IRT Parameters for ELA Grade 5

IREF	*a*	*SE (a)*	*b*	*SE (b)*
115053A	1.3407	0.0541	-0.9317	0.0243
121222A	0.9286	0.0424	-1.1801	0.0391
449342	0.3218	0.0220	0.1649	0.0678
449348	0.2162	0.0205	0.9393	0.1325
449385	0.2855	0.0216	0.8195	0.0967
449387	0.6284	0.0275	-0.1142	0.0364
449391	0.6229	0.0272	0.0344	0.0371
449781	0.7739	0.0348	-0.9702	0.0395
449796	0.4391	0.0242	-0.2546	0.0502
449808	0.7940	0.0319	-0.3970	0.0307
114329A	0.3801	0.0588	-0.1800	0.1364
115054A	1.0211	0.0515	-1.4707	0.0466
114331A	0.3676	0.0583	-0.0889	0.1516
114332A	0.9170	0.0797	-0.2564	0.0597
120209A	0.5575	0.0281	-0.8609	0.0498
120210A	0.8834	0.0336	-0.2666	0.0278
120211B	0.1956	0.0205	1.8665	0.2215
120212A	0.2042	0.0205	1.6198	0.1916
121325A	0.5429	0.0259	-0.1704	0.0412
121326B	0.3748	0.0230	0.9285	0.0794
121733A	0.3678	0.0579	-0.2768	0.1288
121735A	0.5109	0.0637	-0.4379	0.0836
115055A	0.6213	0.0362	-1.7343	0.0815
126984B	0.7402	0.0350	-1.1253	0.0454
126985B	0.8252	0.0339	-0.6446	0.0318
126986A	0.4937	0.0310	-1.7612	0.0975
126987B	1.1186	0.0447	-0.8476	0.0269
455685	0.6932	0.0700	-1.2466	0.0585
455697	1.1407	0.0835	-0.7781	0.0319
455709	1.1391	0.0856	-1.1581	0.0349
455721	0.6555	0.0718	-1.6781	0.0936
115056A	1.1986	0.0483	-0.8989	0.0261
114320A	0.4875	0.0363	0.4676	0.0574
114322A	0.2578	0.0314	-0.0763	0.1244
114323A	0.9288	0.0534	-0.2408	0.0424
121720B	0.0895	0.0197	5.4198	1.1170
121730A	0.4985	0.0384	-0.1680	0.0707
451036	0.4221	0.0710	-0.0875	0.0983
117523A	0.5179	0.0259	-0.4262	0.0447
452013	0.8624	0.0895	0.1998	0.0560
452025	1.2984	0.1406	-1.0500	0.0842
452038	1.1935	0.1090	-0.5286	0.0485
114072A	1.0476	0.0993	-0.4202	0.0960
121457A	0.4941	0.0565	-0.1754	0.1340
121458A	0.2478	0.0415	0.5627	0.1555
121459A	0.8972	0.0744	0.0699	0.0671
117524A	0.3405	0.0223	0.3433	0.0675
117525A	0.5922	0.0266	0.0506	0.0388
120909A	0.6388	0.0286	-0.4729	0.0375
120910A	0.6048	0.0283	-0.6044	0.0411

Table K-4. 2017–18 MSAA: IRT Parameters for ELA Grade 6

IREF	*a*	*SE (a)*	*b*	*SE (b)*
113612A	0.6834	0.0322	-0.0619	0.0334
121226A	0.6894	0.0323	-0.0667	0.0332
121349A	0.2201	0.0248	2.4915	0.2669
121353A	1.0490	0.0405	-0.0476	0.0232
121802A	0.7181	0.0350	-0.4762	0.0390
121803A	1.0319	0.0468	-0.6885	0.0335
121804A	0.8541	0.0366	-0.1950	0.0287
124257A	0.5336	0.0291	0.0431	0.0408
448846	0.8856	0.0400	-0.5264	0.0335
449305	0.8820	0.0385	-0.3693	0.0304
113536A	0.9552	0.0538	0.0815	0.0300
113614A	0.8586	0.0359	-0.0357	0.0273
113537A	0.6037	0.0461	-0.1466	0.0419
120389A	1.1286	0.1053	-1.1495	0.0594
120390A	1.0702	0.1022	-1.1460	0.0616
120391A	1.7941	0.1255	-0.6735	0.0236
120392A	1.0944	0.0999	-0.9541	0.0478
121482A	1.5439	0.1169	-0.7971	0.0299
121483A	1.6421	0.1176	-0.5268	0.0240
121764A	1.1590	0.0589	-0.2414	0.0236
121768A	1.1266	0.0589	-0.3836	0.0253
127272B	0.8406	0.0373	-0.3519	0.0314
114380A	1.1772	0.0514	-0.6517	0.0291
127273A	1.0015	0.0428	-0.4594	0.0288
127274A	0.5632	0.0348	-1.0972	0.0766
127276A	0.3379	0.0265	-0.4796	0.0786
127277B	1.1982	0.0469	-0.3096	0.0226
452219	0.8633	0.0369	-0.2101	0.0286
452231	0.7620	0.0329	0.2957	0.0301
452243	0.5106	0.0286	0.0723	0.0423
452257	0.7340	0.0326	0.1801	0.0306
452269	1.7311	0.0643	-0.3150	0.0170
114382A	1.0457	0.0455	-0.5644	0.0299
452282	1.1468	0.0725	-0.4889	0.0505
452299	0.8056	0.0544	-0.4186	0.0614
452311	0.7776	0.0464	0.2012	0.0385
452335	0.2950	0.0367	2.7710	0.2747
452348	1.1084	0.0602	-0.0805	0.0342
119997A	0.2420	0.0243	1.3010	0.1362
452360	1.1494	0.0643	-0.2012	0.0370
115502A	0.9375	0.0893	0.2347	0.0753
115503A	1.0561	0.1062	-0.0222	0.0910
122258A	0.7919	0.0736	0.7480	0.0557
122263A	0.8038	0.0744	0.7165	0.0561
119998A	0.4175	0.0268	0.3788	0.0519
119999A	0.4427	0.0274	0.9197	0.0627
120000A	0.3056	0.0251	0.7236	0.0794
121225A	0.6660	0.0324	-0.2146	0.0361

Table K-5. 2017–18 MSAA: IRT Parameters for ELA Grade 7

IREF	*a*	*SE (a)*	*b*	*SE (b)*
114593A	1.2156	0.0492	-0.7565	0.0262
121343A	0.3363	0.0237	0.1350	0.0635
121347A	0.5573	0.0274	-0.0951	0.0403
121871A	0.8170	0.0349	-0.5710	0.0333
121874A	0.6132	0.0295	-0.4861	0.0413
121997A	0.6112	0.0288	-0.2820	0.0385
449566	1.2301	0.0568	-1.0939	0.0338
449584	0.9839	0.0426	-0.8551	0.0335
449607	0.2929	0.0233	1.1047	0.1061
449624	0.1449	0.0199	1.0750	0.1977
113801A	1.5013	0.1257	-0.9388	0.0289
114594A	0.7417	0.0338	-0.7062	0.0390
113802B	1.4693	0.1257	-1.0446	0.0312
115431A	1.1313	0.0430	-0.4875	0.0245
115432A	0.7658	0.0322	-0.2919	0.0318
115433A	0.9827	0.0381	-0.3975	0.0266
120098A	1.3051	0.1235	-1.3066	0.0463
120099A	1.4701	0.1304	-1.2343	0.0383
121313A	0.8461	0.0347	-0.4261	0.0304
121493A	1.8282	0.1405	-0.9196	0.0244
121497A	0.7072	0.0925	-1.2605	0.0731
121505A	1.4609	0.1322	-1.3046	0.0422
114596A	0.9353	0.0378	-0.5422	0.0292
121509A	1.5407	0.1283	-0.9994	0.0291
122037A	1.3293	0.1182	-0.9383	0.0320
122038A	1.1001	0.1111	-1.2162	0.0477
127690A	0.3940	0.0245	0.2098	0.0554
127691A	0.7156	0.0347	-0.9392	0.0467
127692B	0.6279	0.0320	-0.9243	0.0518
127693A	0.6352	0.0306	-0.6312	0.0429
127694A	0.6338	0.0323	-0.9373	0.0518
127695B	0.7539	0.0368	-1.0352	0.0478
114643A	1.2018	0.0469	-0.6303	0.0247
114482A	0.7491	0.0472	-0.5575	0.0593
114483A	1.0214	0.0619	-0.7072	0.0532
114484A	0.3587	0.0331	-0.0855	0.0829
120072A	0.5198	0.0758	-0.0173	0.0749
120073A	0.5367	0.0768	0.2335	0.0867
121491A	0.4324	0.0720	0.3621	0.1183
121495A	0.5529	0.0804	0.7320	0.1371
121507A	0.4662	0.0734	-0.0149	0.0831
114644A	0.8638	0.0343	-0.2533	0.0285
122380A	0.2131	0.0300	-1.0659	0.2371
123649A	0.3890	0.0690	0.0574	0.1022
120060A	0.4867	0.0541	0.2221	0.0960
120061A	0.6697	0.0605	0.2999	0.0676
121490A	0.5771	0.0563	0.4119	0.0711
121494A	0.5110	0.0523	1.0536	0.0678
121513A	0.6067	0.0568	0.4882	0.0644
114645A	0.9916	0.0377	-0.2991	0.0258
123641A	0.4100	0.0503	0.4456	0.0955
114646A	0.9779	0.0423	-0.8514	0.0335
115372A	0.2174	0.0218	0.2491	0.0978
115373A	0.6286	0.0287	-0.0899	0.0363

Table K-6. 2017–18 MSAA: IRT Parameters for ELA Grade 8

IREF	*a*	*SE (a)*	*b*	*SE (b)*
114228A	0.3988	0.0250	0.1770	0.0546
115286A	0.9709	0.0385	-0.5399	0.0267
115288A	0.6272	0.0322	-0.9527	0.0500
121075A	1.0153	0.0463	-1.0679	0.0362
121078A	1.5050	0.0631	-0.9308	0.0234
121805A	0.4229	0.0255	-0.0031	0.0503
122562A	1.2197	0.0474	-0.6822	0.0236
449868	1.2100	0.0592	-1.2631	0.0382
449882	0.9172	0.0363	-0.4180	0.0269
449900	0.9087	0.0364	-0.4755	0.0276
114876A	0.8388	0.0355	-0.6004	0.0311
114229A	0.3485	0.0244	-0.0517	0.0601
114877A	0.3043	0.0238	0.6949	0.0870
114879A	0.9771	0.0383	-0.4851	0.0261
118798A	0.8383	0.0393	-1.0195	0.0408
118800A	0.6057	0.0285	-0.0426	0.0365
121030A	1.5847	0.1217	-1.1698	0.0305
121031A	0.8682	0.0935	-0.5360	0.0515
121032A	0.6931	0.0852	-1.3214	0.0734
121033A	0.2525	0.0559	-0.9929	0.1395
121107A	1.2178	0.1037	-0.8974	0.0306
121202A	0.9249	0.0350	-0.0782	0.0258
114230A	0.6296	0.0295	-0.3395	0.0364
121203A	1.0122	0.0411	-0.6904	0.0277
121205A	0.9945	0.0393	-0.5610	0.0264
122082A	0.5107	0.0272	-0.2543	0.0431
127781A	0.4890	0.0293	-1.0883	0.0683
127782A	0.8555	0.0410	-1.1143	0.0433
127783A	1.3609	0.0611	-1.0792	0.0290
127784A	0.2808	0.0237	-0.5357	0.0835
127785A	0.6651	0.0327	-0.8648	0.0447
127786B	1.3713	0.0633	-1.1369	0.0306
114231A	0.5792	0.0278	0.1613	0.0390
121040A	0.4818	0.0884	0.5076	0.1482
121041A	0.8522	0.1044	0.0643	0.0568
121042A	0.9439	0.1092	-0.3141	0.0473
121148A	0.3695	0.0786	0.6045	0.1968
121149A	0.6131	0.0969	-0.5631	0.0889
114796A	0.7540	0.0364	-1.0024	0.0441
121036A	0.5286	0.0539	0.7408	0.0625
121037A	0.8452	0.0692	0.1117	0.0582
121038A	0.9506	0.0713	0.2466	0.0464
121164A	0.3329	0.0499	-0.3098	0.1828
121165A	0.3926	0.0505	0.3754	0.0953
114797A	1.0898	0.0575	-1.4261	0.0488
114798A	1.6660	0.0819	-1.2229	0.0292
114799A	1.5605	0.0751	-1.1999	0.0298
115285A	0.9085	0.0357	-0.3480	0.0267

Table K-7. 2017–18 MSAA: IRT Parameters for ELA Grade 11

IREF	*a*	*SE (a)*	*b*	*SE (b)*
113726A	0.3987	0.0254	-0.0932	0.0554
121746A	0.9324	0.0400	-0.6907	0.0313
449987	0.8831	0.0461	-1.3105	0.0521
450006	0.7336	0.0335	-0.5770	0.0365
450027	0.2996	0.0238	0.7674	0.0923
450048	0.7676	0.0338	-0.4673	0.0335
453006	0.7456	0.0329	-0.3923	0.0335
453019	0.7789	0.0349	-0.6074	0.0351
453033	0.5760	0.0287	-0.1850	0.0401
453047	0.8230	0.0359	-0.5693	0.0328
114205A	1.6452	0.1386	-1.1381	0.0313
113727A	0.2645	0.0231	0.3283	0.0860
114207A	1.4143	0.1258	-1.1090	0.0343
114208A	1.0841	0.1087	-0.8734	0.0377
116348A	1.3659	0.0870	-1.1432	0.0336
116349A	1.4379	0.0845	-0.9285	0.0260
116350A	1.5627	0.0931	-1.0550	0.0273
116351A	1.3875	0.0910	-1.2400	0.0373
119078A	1.2307	0.0468	-0.4739	0.0228
119079A	1.2119	0.0472	-0.5608	0.0237
119080A	1.0733	0.0450	-0.7382	0.0285
119081A	1.1701	0.0486	-0.7603	0.0268
113728A	1.1082	0.0460	-0.7240	0.0275
121718A	0.7253	0.0911	-1.1484	0.0629
121719A	0.8715	0.0991	-0.8431	0.0460
122000A	0.8994	0.0372	-0.4680	0.0293
126773A	1.1379	0.0524	-1.0376	0.0337
126774B	1.4300	0.0657	-1.0383	0.0281
126775A	0.6151	0.0315	-0.7728	0.0474
126776B	0.5384	0.0286	-0.4651	0.0459
126777B	1.0083	0.0476	-1.0723	0.0382
126778B	1.0757	0.0481	-0.9383	0.0325
120148A	0.9892	0.0516	-1.3283	0.0489
114193A	0.5578	0.0888	-0.5013	0.0866
114194A	0.9850	0.1108	-0.6973	0.0621
121702A	0.4036	0.0786	-0.0008	0.1176
121703A	1.0315	0.1119	-0.6403	0.0561
121875A	0.6834	0.0951	-0.5399	0.0740
120149A	0.8396	0.0392	-0.9142	0.0394
114166A	0.5936	0.0518	1.4636	0.0787
114167A	0.5260	0.0539	-0.2220	0.1067
116323A	1.7697	0.1428	-0.4024	0.0524
116324A	1.3098	0.1027	-0.3647	0.0607
116326A	0.6554	0.0572	0.0073	0.0730
121695A	0.6893	0.0541	0.6136	0.0491
121711A	0.7382	0.0588	0.1596	0.0582
121714A	0.9381	0.0676	0.0910	0.0497
122538A	0.2751	0.0416	0.2120	0.1389
120150A	0.7819	0.0373	-0.9065	0.0417
120151A	0.8630	0.0457	-1.3460	0.0546
121229A	0.8818	0.0407	-0.9177	0.0379
121745A	0.5061	0.0273	-0.1383	0.0447

Table K-8. 2017–18 MSAA: IRT Parameters for Mathematics Grade 3

IREF	*a*	*SE (a)*	*b*	*SE (b)*
110842A	1.0789	0.0460	-0.4419	0.0251
112564A	0.8249	0.0398	-0.6405	0.0342
112575A	0.5222	0.0322	-0.6624	0.0519
112576A	0.4852	0.0324	1.1571	0.0901
112586A	0.5009	0.0319	0.8650	0.0730
112595A	1.1017	0.0458	-0.2651	0.0238
442130	0.5504	0.0326	-0.5857	0.0475
110855A	0.8243	0.1021	0.0101	0.1075
110871A	0.6824	0.0980	0.2827	0.1625
110919A	0.7074	0.0952	0.0916	0.1306
110923A	1.5128	0.0774	-0.6452	0.0217
110873A	0.5030	0.0318	0.8301	0.0711
110964A	1.2636	0.1049	-0.8321	0.0343
110974A	0.7073	0.0567	0.1166	0.0599
111377A	0.9462	0.0595	-0.5460	0.0311
111390A	0.6919	0.0542	-0.1458	0.0484
111432A	0.6351	0.0955	0.3650	0.1834
111883A	0.5453	0.0889	0.4930	0.2232
112559A	0.9014	0.0954	-0.4260	0.0605
112569A	1.3287	0.1111	-0.6600	0.0348
112600A	1.2905	0.1133	-0.5129	0.0403
442166	0.5695	0.0903	0.4330	0.2071
110928A	0.4302	0.0299	0.6783	0.0744
442226	0.2431	0.0528	-2.4512	0.3679
442416	1.5361	0.1296	-0.4423	0.0374
451071	0.6109	0.0951	0.4547	0.2034
451090	0.1082	0.0247	1.7066	0.5710
451600	0.3064	0.0599	-2.3014	0.2959
463237	1.4511	0.0768	-0.3070	0.0228
110966A	0.3736	0.0417	-0.6737	0.1178
110975A	0.6175	0.0466	-0.0875	0.0489
111382A	0.7159	0.0903	-0.3952	0.0584
111387A	0.7487	0.0484	0.3223	0.0394
110959A	0.4141	0.0321	-1.5290	0.1111
111416A	0.6603	0.0895	-0.5982	0.0765
111426A	0.8288	0.0506	0.2772	0.0357
111650A	1.4199	0.1150	-0.1630	0.0295
112565A	0.3552	0.0728	0.9847	0.2609
112566A	0.5751	0.0438	0.5266	0.0539
112570A	0.3019	0.0373	0.9994	0.1296
112585A	0.6787	0.0912	0.4365	0.0978
112616A	0.5321	0.0431	0.9346	0.0749
122090A	1.6605	0.0816	-0.0962	0.0209
451107	0.3994	0.0412	1.5966	0.1502
111386A	0.8485	0.0389	-0.2020	0.0293
110876A	0.7589	0.0734	0.4943	0.0545
110920A	0.5875	0.0652	0.9807	0.0761
111400A	0.6150	0.0714	0.0440	0.0883
111420A	0.8734	0.0807	0.2786	0.0532
112552A	1.0803	0.1020	-0.1340	0.0653
112601A	1.1957	0.1076	-0.0861	0.0570
112615A	0.6660	0.0702	0.4226	0.0630
120682A	0.9715	0.1141	-0.6051	0.1217
122091A	1.3209	0.1146	-0.0580	0.0509
442037	1.6343	0.1588	-0.3362	0.0627
111397A	0.3503	0.0283	-0.5537	0.0713
442402	1.3178	0.1259	-0.2711	0.0664
111434A	0.9859	0.0472	-0.9776	0.0369
112551A	0.7557	0.0382	-0.7103	0.0384
112555A	1.1747	0.0516	-0.7768	0.0277

Table K-9. 2017–18 MSAA: IRT Parameters for Mathematics Grade 4

IREF	*a*	*SE (a)*	*b*	*SE (b)*
111135A	0.3342	0.0284	-0.5663	0.0779
112824A	0.6185	0.0335	0.4350	0.0410
121691A	0.5026	0.0314	0.6249	0.0554
122267A	0.8275	0.0377	0.0197	0.0278
122432A	0.7460	0.0380	0.9320	0.0487
446131	0.4255	0.0338	-1.6938	0.1328
446443	0.5420	0.0332	-0.7542	0.0567
111123A	1.1882	0.0608	-0.2455	0.0232
111124A	1.4487	0.1072	-0.1812	0.0343
111136A	1.4302	0.0677	-0.2780	0.0200
111162A	1.0591	0.0938	0.0722	0.0574
111148A	0.5273	0.0326	0.9252	0.0650
111185A	0.1919	0.0353	2.0896	0.4528
111676A	1.2205	0.0633	-0.0187	0.0244
111685A	0.8267	0.0740	-0.8335	0.0510
111716A	1.1081	0.0591	-0.1341	0.0251
111721A	1.0747	0.0598	0.0558	0.0286
111731A	1.0393	0.0582	-0.0036	0.0283
112794A	0.9206	0.0544	-0.0926	0.0300
121663A	1.1865	0.0946	-0.1614	0.0412
121665A	1.5515	0.1145	-0.1184	0.0343
445588	1.0890	0.0586	-0.1385	0.0254
111166A	0.9450	0.0405	-0.0001	0.0249
446780	0.3077	0.0585	1.3868	0.3848
446795	0.1539	0.0308	2.0641	0.4968
454863	0.2119	0.0451	0.8681	0.3526
455016	0.3180	0.0541	-0.1910	0.1326
463034	0.5155	0.0981	2.1007	0.4851
111658A	0.7666	0.0861	-0.9555	0.1018
111712A	0.5350	0.0421	0.4193	0.0503
112803A	0.7222	0.0502	-0.3403	0.0515
112832B	0.2205	0.0496	1.4799	0.3732
121661A	0.4662	0.0406	0.7196	0.0660
111663A	0.9174	0.0443	-0.9057	0.0409
121737A	0.3779	0.0634	-0.2845	0.0969
122368A	0.7965	0.0485	0.7843	0.0430
446106	0.3922	0.0390	0.8487	0.0842
446178	0.3179	0.0383	1.7850	0.1901
454860	0.2204	0.0500	1.8499	0.4536
111678A	1.9017	0.1399	0.4341	0.0279
111688A	0.5729	0.0733	0.2183	0.0949
111696A	0.8289	0.0794	0.8515	0.0513
111707B	0.6351	0.0718	0.8846	0.0655
111717A	0.9222	0.0969	0.0132	0.0775
111672A	0.3961	0.0325	-1.6276	0.1345
112783A	1.5906	0.1235	0.3894	0.0332
112788A	0.5535	0.0705	0.4563	0.0813
112812A	0.6775	0.0731	1.1848	0.0728
112817A	0.6493	0.0720	1.1362	0.0731
112833A	0.5127	0.0680	0.5596	0.0825
112837B	0.4713	0.0652	0.8220	0.0853
112838B	0.9023	0.0838	0.6589	0.0473
122153A	0.7463	0.0903	-0.1852	0.1129
454850	1.6545	0.1262	0.4130	0.0316
111682A	0.8732	0.0388	0.1571	0.0272
111686A	0.6952	0.0349	0.2981	0.0346
111711A	0.2024	0.0252	0.9002	0.1502
111715A	0.7800	0.0402	-0.9090	0.0467

Table K-10. 2017–18 MSAA: IRT Parameters for Mathematics Grade 5

IREF	*a*	*SE (a)*	*b*	*SE (b)*
111243A	0.8098	0.0334	0.1138	0.0296
112377A	0.5380	0.0305	-1.1966	0.0695
113883A	0.1602	0.0200	0.3113	0.1346
113884B	0.5550	0.0299	-0.9320	0.0572
113899A	0.4100	0.0280	-1.5406	0.1074
120739A	0.9081	0.0377	0.9847	0.0387
450093	0.5288	0.0276	0.9476	0.0586
111234A	0.1052	0.0264	-0.8787	0.3736
111242A	1.0739	0.0898	-0.7269	0.0412
111258A	0.3363	0.0442	1.7439	0.2831
111263A	0.0549	0.0130	2.4642	0.8296
111276A	0.6220	0.0290	0.1139	0.0368
111308A	0.5290	0.0450	0.4278	0.0814
112335A	1.0801	0.0903	-0.7059	0.0411
112352A	0.6796	0.0751	-0.4842	0.0694
112368A	0.8437	0.0522	0.1008	0.0410
112369A	0.9864	0.0985	-0.0877	0.0719
113843B	0.8655	0.0885	-0.1956	0.0717
113853A	0.8386	0.0895	-0.0821	0.0826
113862A	0.5908	0.0821	0.3902	0.1688
113863A	0.0809	0.0204	2.0749	0.8894
113889A	1.2556	0.0982	-0.7229	0.0362
111295A	0.6265	0.0291	0.2940	0.0380
120724A	0.8696	0.1389	1.0043	0.2521
450111	0.2087	0.0469	-3.3976	0.5934
450180	0.4836	0.0489	1.1642	0.1561
450200	0.9112	0.0904	-0.2272	0.0665
450210	1.0910	0.0908	-0.7035	0.0408
450339	0.5168	0.0765	0.4238	0.1897
111244A	0.9895	0.0476	0.1647	0.0287
111277A	0.7146	0.0396	0.2375	0.0374
111299A	0.3568	0.0310	1.1866	0.0971
112342A	0.4044	0.0602	0.1253	0.0908
111298A	0.3997	0.0254	1.2685	0.0890
112358A	0.6208	0.0365	0.6536	0.0453
112364A	0.5334	0.0345	0.9513	0.0594
112384A	0.9329	0.0457	0.2064	0.0300
112408A	0.7861	0.0756	0.4665	0.0660
113867A	0.3255	0.0599	1.5542	0.3170
113892A	0.7706	0.0409	0.2627	0.0350
120737A	1.0107	0.0469	0.6903	0.0311
450055	0.4980	0.0653	0.5453	0.1065
450129	0.4398	0.0686	1.4058	0.2391
450143	0.1230	0.0308	2.8901	0.8012
111303A	0.4152	0.0251	-0.1631	0.0530
450274	0.5186	0.0708	1.0960	0.1682
111294A	1.1593	0.0856	0.2550	0.0456
112348A	0.9649	0.0700	0.6555	0.0428
112354A	0.5659	0.0531	1.3469	0.0759
112373A	0.5367	0.0583	-0.0774	0.1150
112385A	0.7365	0.0636	0.3235	0.0643
112392A	0.5191	0.0519	1.6097	0.0953
112410A	0.8167	0.0688	0.2119	0.0637
113856A	0.4276	0.0486	1.1911	0.0908
113872A	0.8190	0.0624	0.9432	0.0481
112346A	0.6942	0.0326	-0.7008	0.0411
113877A	0.3314	0.0456	1.9677	0.1755
113902A	0.7015	0.0577	1.1867	0.0586
121514A	0.7231	0.0598	0.7316	0.0539
112359A	0.4379	0.0261	1.1846	0.0785
112363A	0.4964	0.0266	0.6923	0.0543
112372A	0.9185	0.0399	-0.8250	0.0350

Table K-11. 2017–18 MSAA: IRT Parameters for Mathematics Grade 6

IREF	*a*	*SE (a)*	*b*	*SE (b)*
110891A	0.9555	0.0418	-0.8316	0.0336
120494A	0.6823	0.0312	-0.3571	0.0345
442566	0.9551	0.0417	-0.8212	0.0334
442628	0.6005	0.0285	0.6292	0.0438
442711	0.6317	0.0325	-0.8916	0.0499
442813	0.7936	0.0386	-1.0389	0.0460
453707	0.6151	0.0287	0.2908	0.0373
110977A	0.8027	0.0531	-0.1180	0.0378
111025A	0.9853	0.0568	-0.5038	0.0281
111482A	0.5084	0.0456	-0.4323	0.0510
111487A	0.7031	0.0686	-0.5577	0.0467
110944A	1.1380	0.0444	-0.5447	0.0242
111508A	0.8869	0.0756	-0.4013	0.0414
112633A	0.8444	0.0772	-0.1250	0.0573
112656A	1.3162	0.0886	-0.8557	0.0283
112663A	0.5985	0.0684	-0.0420	0.0846
112676A	0.2890	0.0514	-1.4415	0.1718
112678A	0.7266	0.0512	-0.1173	0.0414
112691A	0.4652	0.0619	-1.7589	0.1534
112697A	1.0364	0.0581	-0.3926	0.0268
120855A	0.7131	0.0713	-0.2156	0.0603
442369	0.7087	0.0505	-0.2159	0.0398
110980A	0.5431	0.0290	-0.6074	0.0477
442631	0.6068	0.0497	0.1775	0.0624
442785	0.2342	0.0380	0.7076	0.2130
450368	0.3389	0.0414	-0.4525	0.0753
450436	1.3192	0.0884	-0.7924	0.0272
450459	0.6830	0.0525	0.2772	0.0616
453730	0.4592	0.0677	0.6667	0.1998
110981A	0.8974	0.0531	0.0927	0.0370
110984A	0.3726	0.0373	1.5533	0.1194
110993A	0.6229	0.0458	-0.0732	0.0578
111022A	1.1308	0.0574	0.4508	0.0272
110986A	0.6455	0.0307	-0.4357	0.0374
111456A	0.8413	0.0483	0.5457	0.0348
111479A	1.2106	0.0688	-0.1529	0.0343
111518A	0.8735	0.0507	0.2868	0.0346
112645A	0.8393	0.0513	0.0834	0.0395
112658A	0.6767	0.0470	-0.0151	0.0513
121487A	1.1850	0.0639	0.0390	0.0300
111445A	1.3221	0.0870	0.4508	0.0331
111455A	1.0820	0.0790	0.3067	0.0436
111465A	0.8413	0.0651	0.5677	0.0460
112655A	0.5151	0.0525	1.1047	0.0711
110991A	0.6089	0.0286	0.3270	0.0381
112679A	1.2317	0.1004	-0.0842	0.0586
112699A	0.8667	0.0671	0.4730	0.0471
442356	1.1560	0.0884	0.0994	0.0501
442538	0.9091	0.0674	0.5862	0.0426
442641	0.9345	0.0727	0.2884	0.0504
453664	0.7256	0.0653	0.1614	0.0704
111036A	0.3201	0.0238	0.6047	0.0760
111514A	0.7931	0.0332	-0.2197	0.0291
111630A	1.0195	0.0460	-0.9837	0.0360
112632A	0.8710	0.0454	-1.3500	0.0558

Table K-12. 2017–18 MSAA: IRT Parameters for Mathematics Grade 7

IREF	*a*	*SE (a)*	*b*	*SE (b)*
111055A	0.7079	0.0335	0.1048	0.0331
111769A	0.8368	0.0389	-0.5259	0.0310
112899A	0.5851	0.0332	-0.6698	0.0467
112909B	0.8304	0.0424	-0.9539	0.0428
446673	1.0397	0.0509	-1.0200	0.0380
451801	0.3697	0.0271	0.5257	0.0703
452439	0.4560	0.0314	-1.0563	0.0767
111046A	1.2023	0.1338	-1.4051	0.0665
111069A	1.2284	0.1327	-1.0546	0.0467
111092A	0.7658	0.1061	-1.1051	0.0723
111100A	0.8327	0.1207	-0.2825	0.0948
111076A	0.6049	0.0313	0.4972	0.0447
111104A	1.0466	0.0622	-0.4457	0.0264
111119A	0.7787	0.1169	-0.2675	0.1018
111130A	0.8853	0.1133	-0.9755	0.0589
111744A	0.7697	0.1082	-0.7761	0.0657
111779A	0.8496	0.0581	-0.1650	0.0355
111780A	0.5680	0.1077	0.2606	0.2188
111799A	0.3746	0.0474	0.4139	0.1249
111804A	0.7920	0.0563	-0.2593	0.0356
112605A	0.9491	0.0601	-0.2760	0.0300
112849A	0.5443	0.0502	-0.3844	0.0482
111080A	0.6614	0.0324	0.1776	0.0358
112852A	0.5702	0.0959	-0.6782	0.0910
112853A	0.4886	0.0950	0.0129	0.1915
446831	0.3489	0.0822	0.9205	0.4367
446838	0.3848	0.0471	0.2098	0.1018
451854	0.1241	0.0281	2.4212	0.6796
453976	0.4092	0.0494	0.5791	0.1337
111048A	0.8283	0.0433	-0.1510	0.0318
111071A	0.8319	0.0427	-0.0453	0.0308
111105A	0.8168	0.0417	0.0754	0.0311
111754B	0.5720	0.0361	0.0892	0.0425
111093A	0.6120	0.0319	-0.1029	0.0363
111765A	0.8080	0.0412	0.1581	0.0316
111795A	0.8991	0.0429	0.3059	0.0301
112887A	0.7903	0.0410	0.0868	0.0320
112911A	0.6908	0.0385	0.1371	0.0360
446543	0.9592	0.0446	0.2361	0.0280
446891	0.1782	0.0418	2.1279	0.5864
454054	0.7051	0.0743	-0.8834	0.0726
111067A	0.4759	0.0550	1.2462	0.0947
111075A	1.7380	0.1381	-0.0921	0.0424
111085A	0.4620	0.0555	0.4815	0.0848
111113A	0.6650	0.0325	0.2464	0.0366
111106A	0.7662	0.0646	1.2406	0.0621
111764B	1.0802	0.0886	0.0351	0.0540
111766A	1.2842	0.0919	0.3144	0.0369
111796A	0.9118	0.0742	0.3312	0.0491
112523B	0.6127	0.0619	0.3371	0.0703
112910A	0.8736	0.0772	0.0456	0.0640
113101A	1.2046	0.0885	0.2932	0.0395
454049	0.9044	0.0735	0.3599	0.0486
111127A	0.8094	0.0358	0.0564	0.0292
111734A	0.9864	0.0420	-0.3861	0.0253
111748A	0.4577	0.0334	-1.4652	0.1023
111759B	0.5661	0.0307	0.0782	0.0398

Table K-13. 2017–18 MSAA: IRT Parameters for Mathematics Grade 8

IREF	*a*	*SE (a)*	*b*	*SE (b)*
111247A	1.1085	0.0444	-0.7409	0.0270
112516A	0.9144	0.0361	0.0190	0.0260
113963A	0.9255	0.0363	0.1863	0.0268
113973A	0.7089	0.0315	-0.3075	0.0324
117072A	0.7401	0.0318	0.1742	0.0319
447054	0.7105	0.0311	0.0233	0.0319
454255	0.5962	0.0304	-0.8885	0.0491
111286A	0.7312	0.0616	0.1832	0.0704
111335A	0.7951	0.0592	-0.1846	0.0456
111581A	0.9072	0.0803	-0.6164	0.0459
112491A	0.5327	0.0696	-0.0360	0.1249
111352A	0.9443	0.0406	-0.9249	0.0344
112499A	1.4544	0.0997	-0.9545	0.0289
113908A	0.8314	0.0622	-0.0473	0.0498
113909A	1.4906	0.1030	-0.8153	0.0278
113918A	0.9395	0.0417	-0.8290	0.0359
113931A	0.3938	0.0652	0.5955	0.2559
113952A	0.7920	0.0642	0.1806	0.0658
113968A	0.6385	0.0788	0.1129	0.1258
117071A	0.1456	0.0343	-2.3884	0.4194
120571A	0.6225	0.0582	0.2967	0.0896
122099A	0.4975	0.0702	0.2198	0.1651
111560A	0.7106	0.0311	0.1661	0.0330
447109	0.5039	0.0512	0.0847	0.0876
447166	1.4367	0.0980	-1.1020	0.0313
454069	0.9955	0.0966	-0.1184	0.0705
462289	0.8740	0.0652	0.0356	0.0520
471660	0.0913	0.0215	0.7497	0.4480
113942A	0.8689	0.0857	-0.2382	0.0684
111583A	0.5564	0.0381	0.6844	0.0521
111622A	1.0445	0.0530	0.0916	0.0274
112475A	0.6852	0.1127	-0.5730	0.0842
112486A	0.6520	0.0402	0.6298	0.0442
111565A	0.8572	0.0363	-0.6385	0.0313
112500A	0.4725	0.0995	0.5246	0.1812
112506A	0.5521	0.0409	-0.2057	0.0570
113917A	1.3635	0.0745	-0.4208	0.0309
113932A	0.8210	0.0452	0.2988	0.0331
113937A	0.8058	0.0441	0.4812	0.0349
113943A	0.9770	0.1267	0.0683	0.0585
122051A	1.0422	0.0533	0.0627	0.0277
446979	0.8018	0.0489	-0.2338	0.0413
111281A	0.4192	0.0449	0.1029	0.0869
111339A	0.6403	0.0492	0.6965	0.0501
111615A	0.6786	0.0311	-0.4554	0.0351
111594A	0.7063	0.0539	0.1871	0.0506
112490A	0.7746	0.0584	-0.0080	0.0534
113959A	0.5816	0.0473	1.0401	0.0633
113964A	1.2516	0.0746	0.1742	0.0306
120568A	0.7194	0.0555	0.0636	0.0540
446956	0.2908	0.0409	-0.2622	0.1553
447047	0.8352	0.0561	0.4326	0.0394
462215	1.1529	0.0700	0.2043	0.0323
112460A	0.5185	0.0271	0.3060	0.0452
112466A	0.5756	0.0282	0.0259	0.0382
112470A	0.4719	0.0272	-0.8252	0.0581
112509A	0.4422	0.0259	-0.3579	0.0499

Table K-14. 2017–18 MSAA: IRT Parameters for Mathematics Grade 11

IREF	*a*	*SE (a)*	*b*	*SE (b)*
110881A	1.0059	0.0472	0.0541	0.0234
111833A	0.5918	0.0381	0.3176	0.0417
112727A	1.3513	0.0577	-0.2201	0.0188
112946A	0.6735	0.0399	0.0474	0.0331
443312	0.2992	0.0330	1.1380	0.1405
443575	1.3301	0.0603	-0.4817	0.0229
462343	0.6414	0.0390	0.3410	0.0394
111016A	0.8532	0.0735	0.1413	0.0464
111533A	1.6310	0.1448	-0.6562	0.0297
111557A	0.3895	0.0596	0.7585	0.1728
111819A	0.4278	0.0915	0.8791	0.3193
110936A	0.9782	0.0486	-0.3419	0.0265
111824A	1.2481	0.1347	-0.2656	0.0439
111830A	1.1039	0.1188	-0.7722	0.0451
111840A	0.9533	0.0749	0.0209	0.0361
112701A	2.2221	0.1760	-0.6054	0.0227
112708A	0.6011	0.0661	0.2401	0.0716
112717A	0.6628	0.0680	0.2022	0.0628
112722A	1.3590	0.0835	-0.1938	0.0216
112743A	0.5219	0.0925	-0.1460	0.1089
112929A	0.4983	0.0880	-0.6269	0.0854
112930B	0.2539	0.0501	1.5076	0.3698
111024A	0.6047	0.0431	-0.9811	0.0715
442911	0.1706	0.0387	2.4706	0.6469
442914	0.5986	0.0944	-1.2268	0.1209
443494	1.2647	0.1442	-0.0740	0.0572
454925	0.6500	0.0647	-0.2075	0.0424
454980	0.4403	0.0925	0.7954	0.2966
454987	1.4228	0.0862	-0.1285	0.0216
110843A	1.1677	0.0619	0.2324	0.0235
110858A	0.6462	0.0502	-0.0164	0.0415
110913A	1.2011	0.0628	0.2442	0.0231
110921A	1.0006	0.0580	0.1560	0.0265
111539A	1.0369	0.0479	0.0999	0.0231
111002A	0.9691	0.0594	-0.0697	0.0293
111538A	1.1342	0.0603	0.3606	0.0255
111548A	0.5457	0.0462	0.7377	0.0630
111829A	0.8900	0.0540	0.3837	0.0315
112733A	0.7089	0.0499	0.3516	0.0378
122055A	0.2743	0.0653	1.1390	0.3251
110914A	1.5761	0.1160	0.1840	0.0297
111000A	1.1446	0.0941	0.2062	0.0386
111537A	0.9360	0.0798	0.5621	0.0402
111810A	2.0282	0.1454	0.1144	0.0258
111553A	0.8280	0.0429	0.3036	0.0308
112702A	1.5777	0.1088	0.3728	0.0262
112709A	1.2416	0.0890	0.8731	0.0373
112744A	1.1403	0.0948	0.1751	0.0399
112940A	0.9619	0.0821	0.4358	0.0395
442881	0.3929	0.0602	0.6871	0.0924
443287	0.4429	0.0624	0.9531	0.0966
462630	0.5345	0.0677	0.3042	0.0729
111813A	0.6803	0.0437	-0.7662	0.0532
111815A	1.4788	0.0596	0.0721	0.0172
111818A	1.0020	0.0472	0.0378	0.0233
111828A	0.7022	0.0404	0.4906	0.0405

APPENDIX L—TEST CHARACTERISTIC CURVES & TEST INFORMATION FUNCTIONS

Figure L-1. 2017–18 MSAA: Test Characteristic Curve for Grade 3 Mathematics – Paths A, B, and C

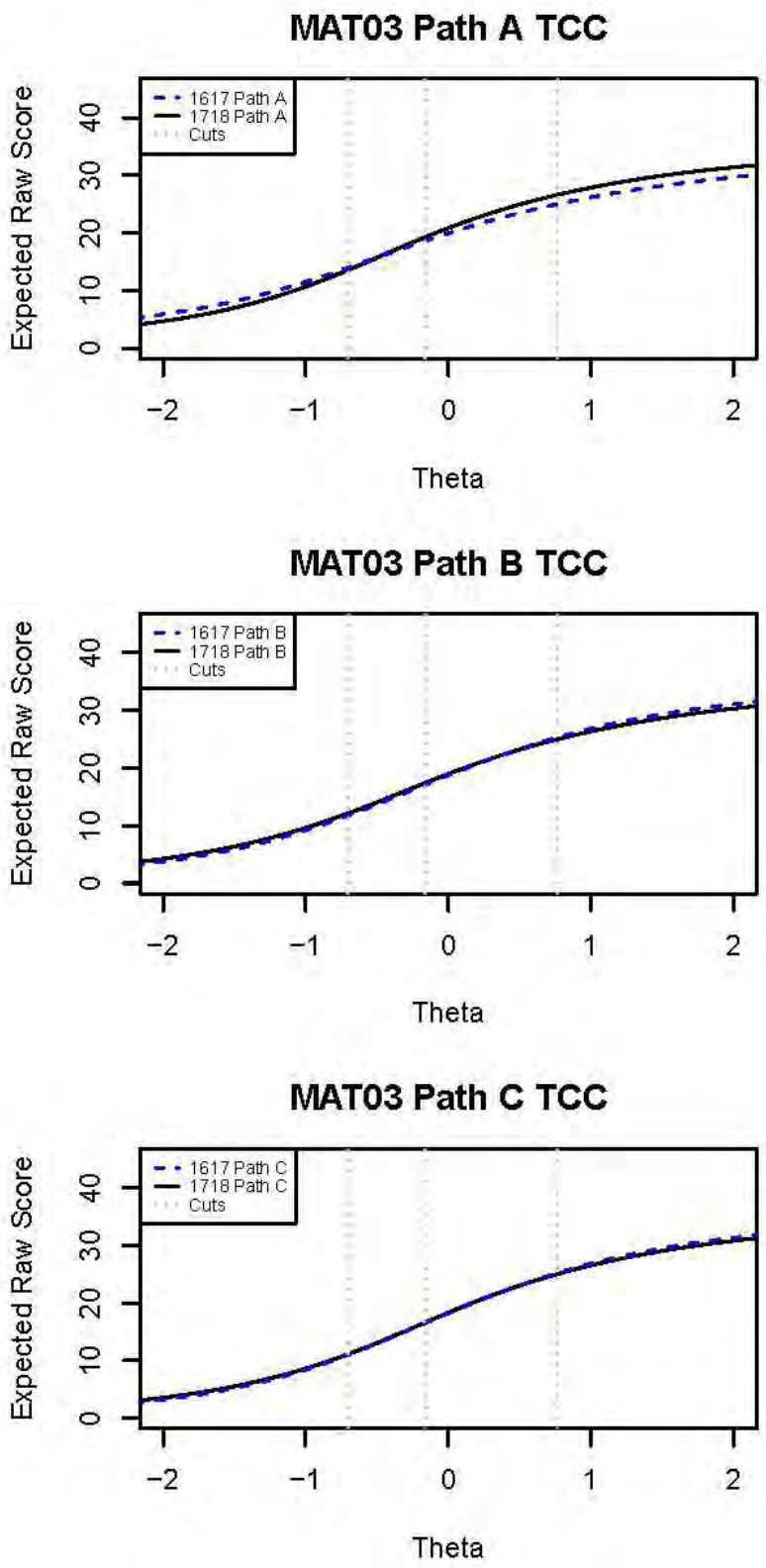


Figure L-2. 2017–18 MSAA: Test Information Function and Standard Error for Grade 3 Mathematics – Paths A, B, and C

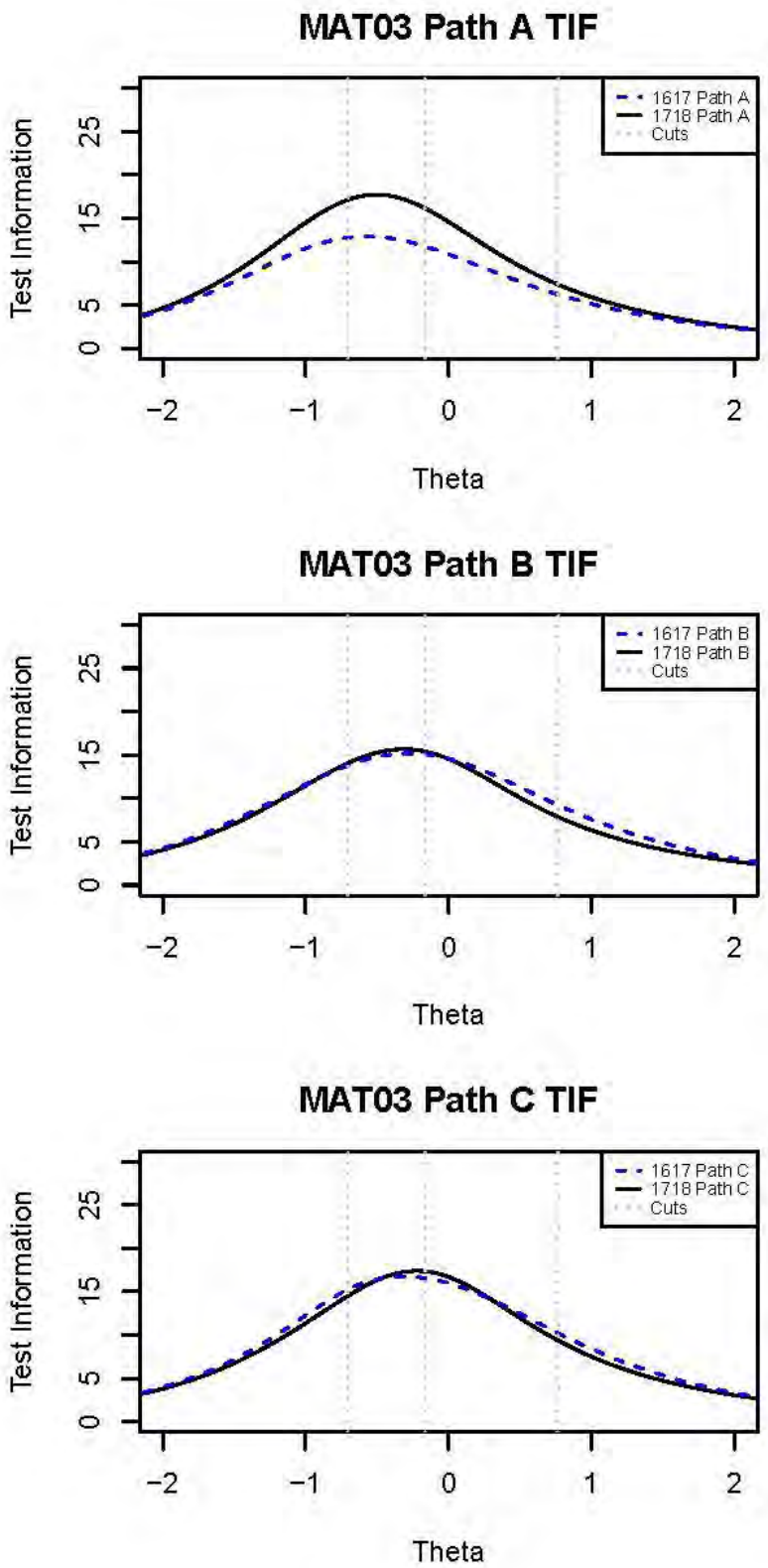


Figure L-3. 2017–18 MSAA: Test Characteristic Curve for Grade 4 Mathematics – Paths A, B, and C

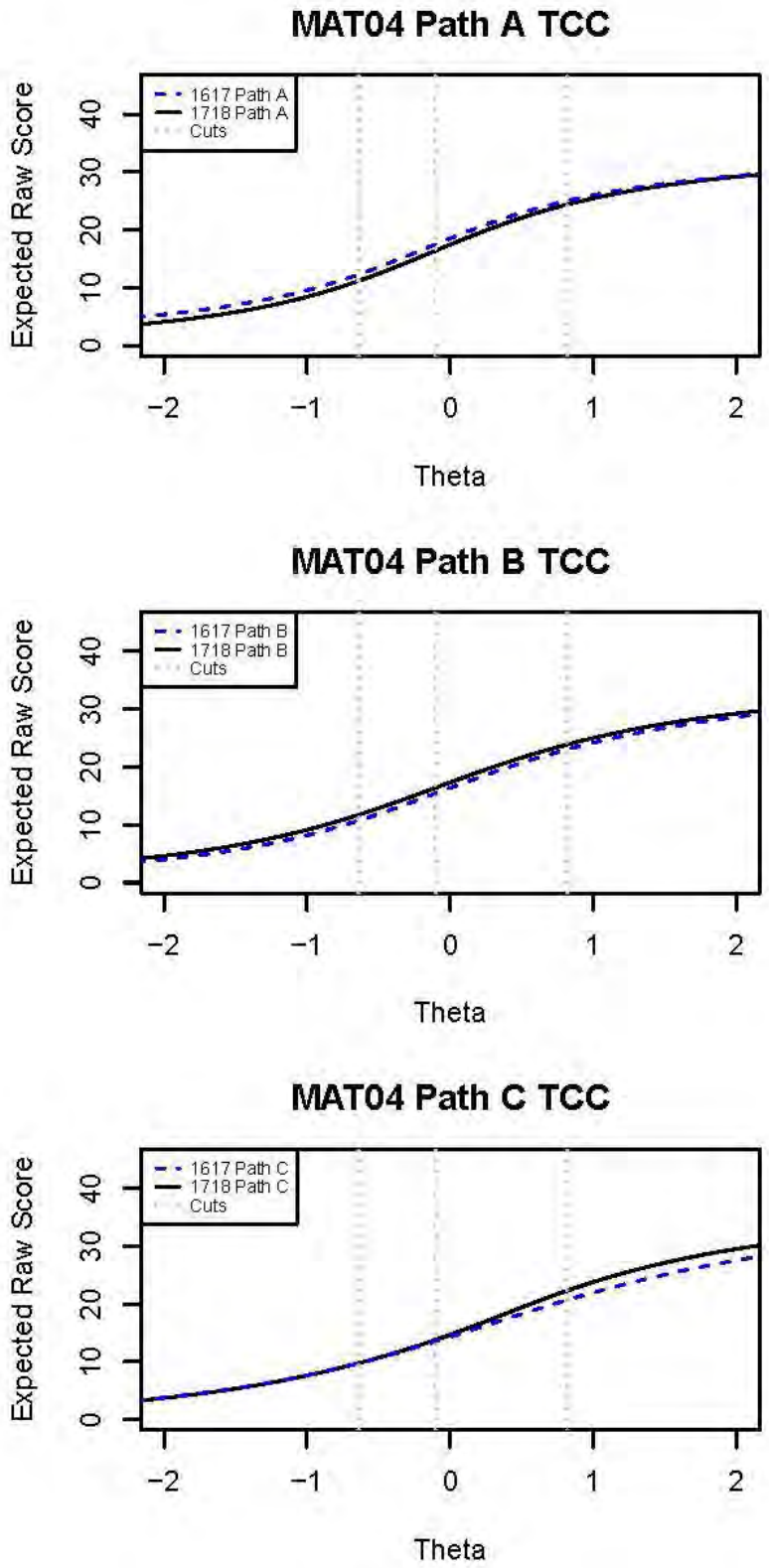


Figure L-4. 2017–18 MSAA: Test Information Function and Standard Error for Grade 4 Mathematics – Paths A, B, and C

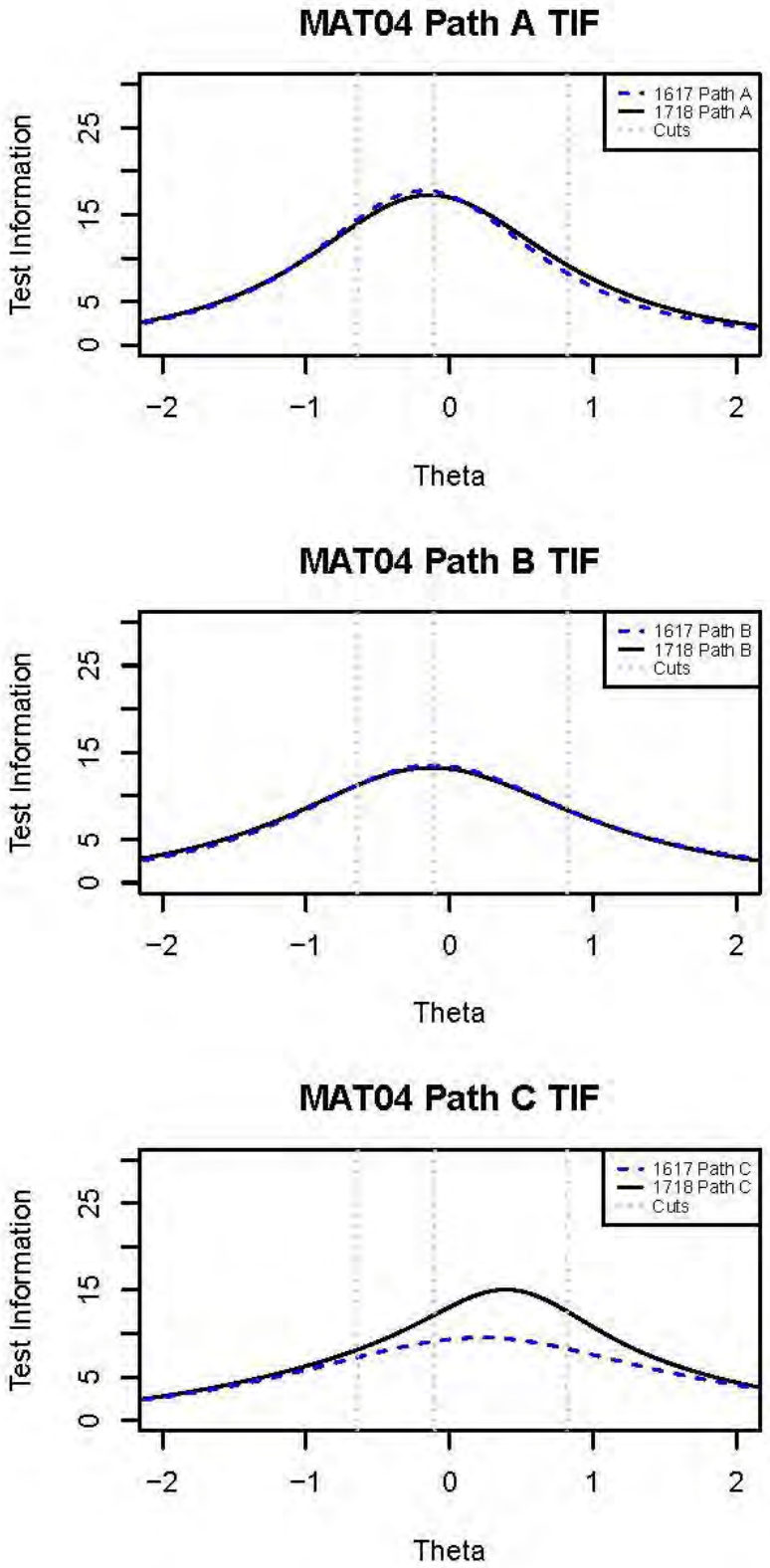
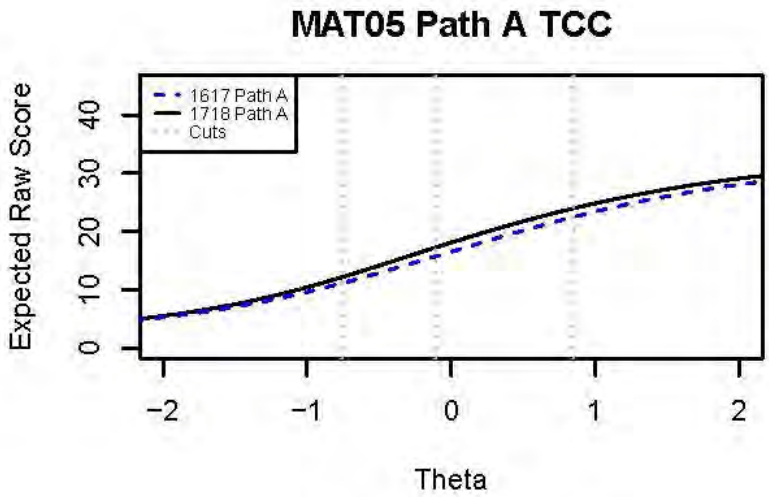
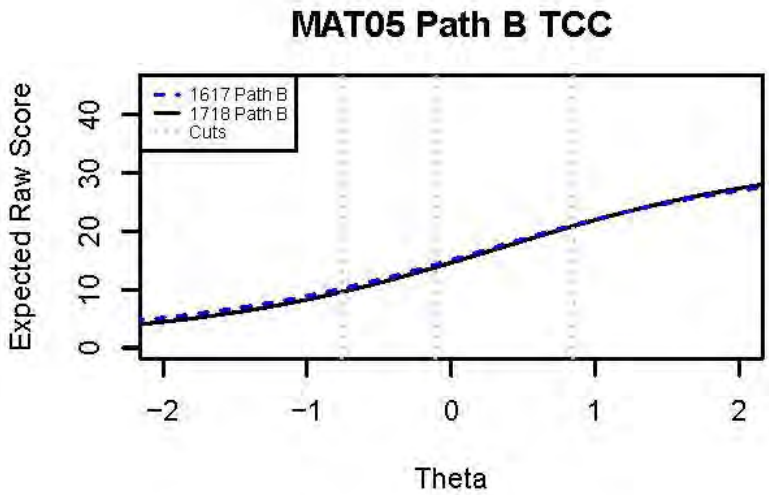


Figure L-5. 2017–18 MSAA: Test Characteristic Curve for Grade 5 Mathematics – Paths A, B, and C





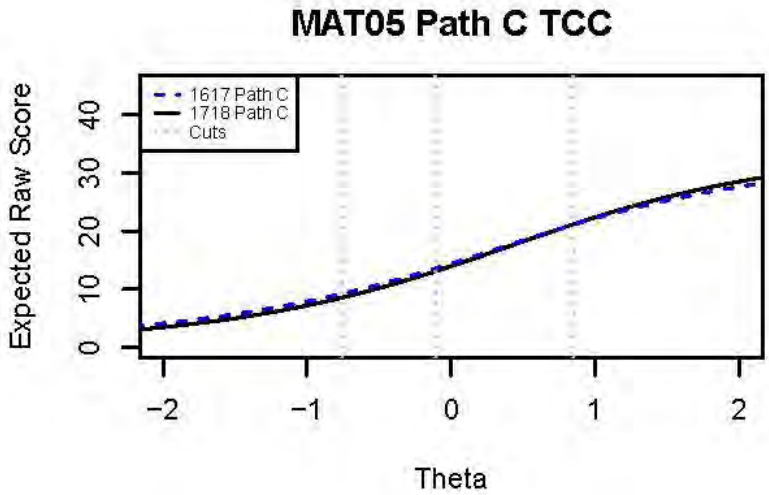


Figure L-6. 2017–18 MSAA: Test Information Function and Standard Error for Grade 5 Mathematics – Paths A, B, and C

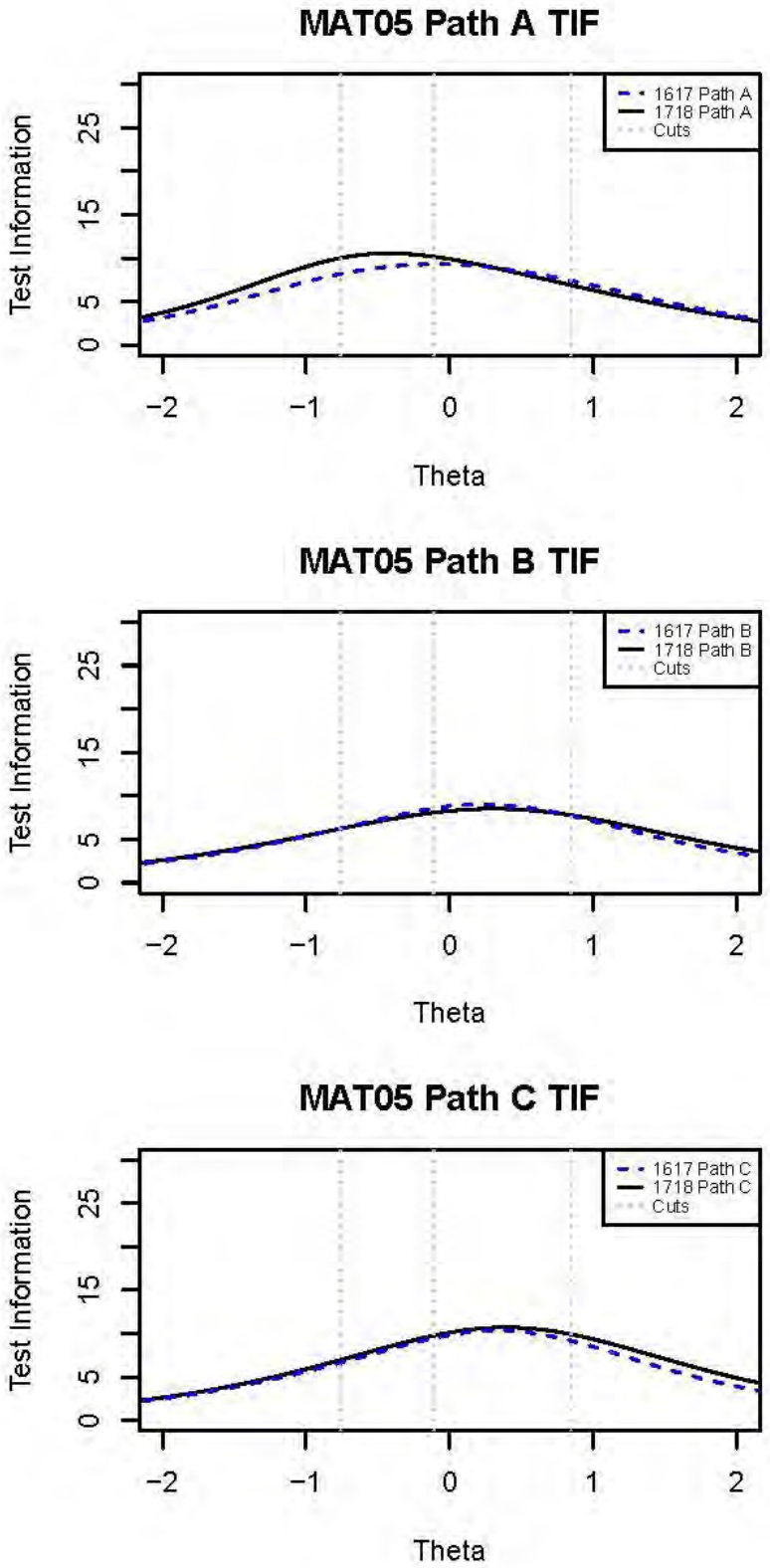


Figure L-7. 2017–18 MSAA: Test Characteristic Curve for Grade 6 Mathematics – Paths A, B, and C

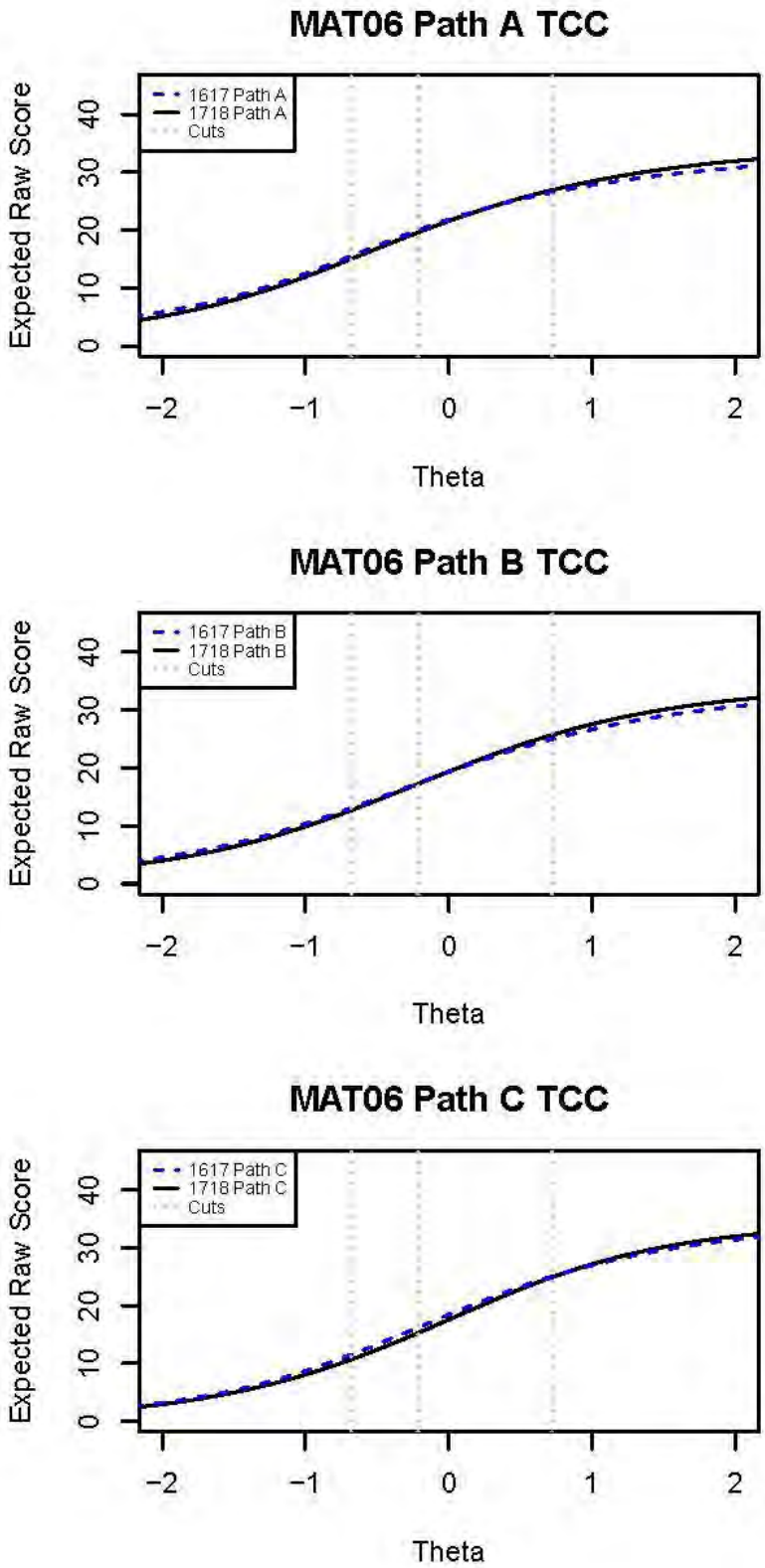


Figure L-8. 2017–18 MSAA: Test Information Function and Standard Error for Grade 6 Mathematics – Paths A, B, and C

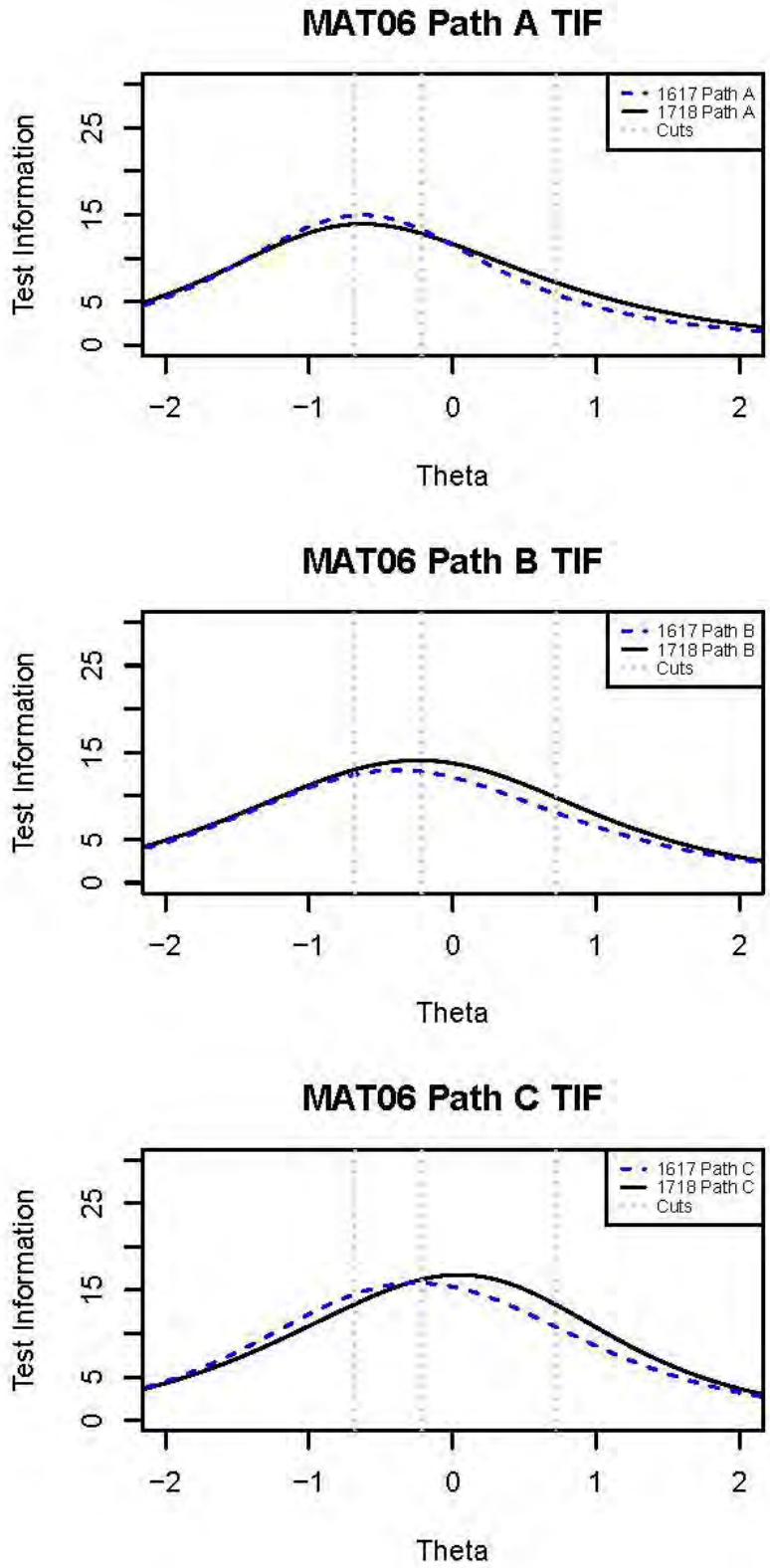
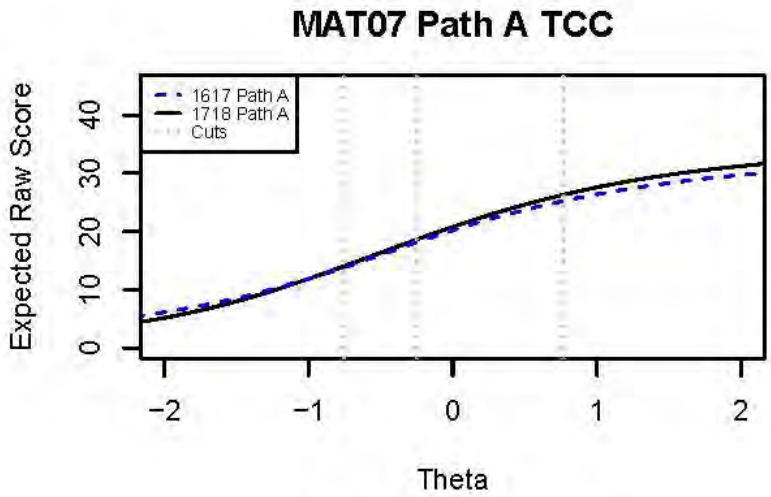
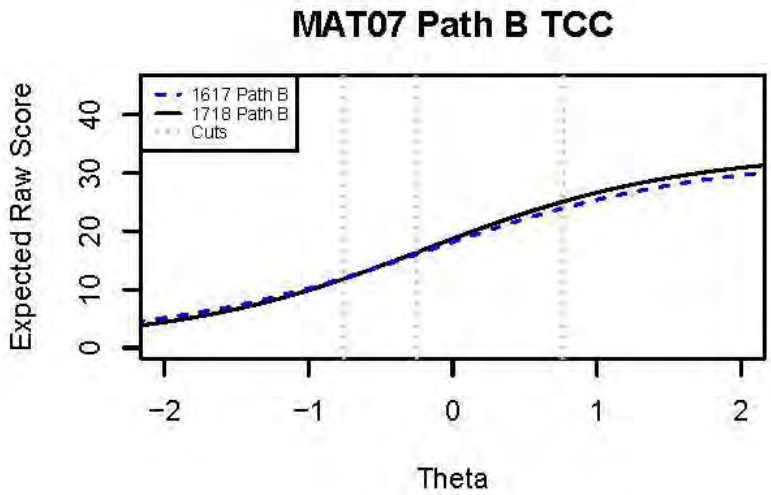


Figure L-9. 2017–18 MSAA: Test Characteristic Curve for Grade 7 Mathematics – Paths A, B, and C





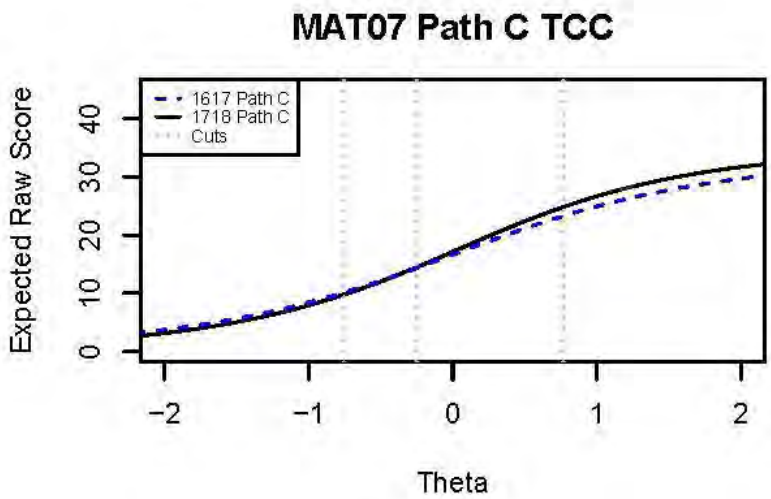
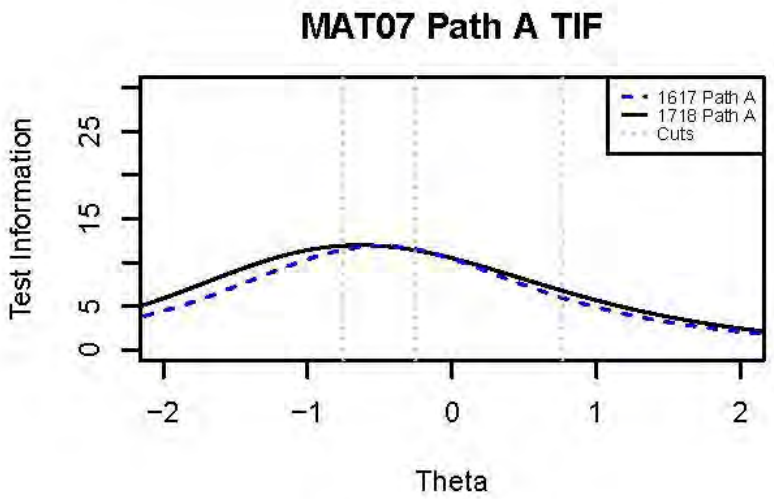
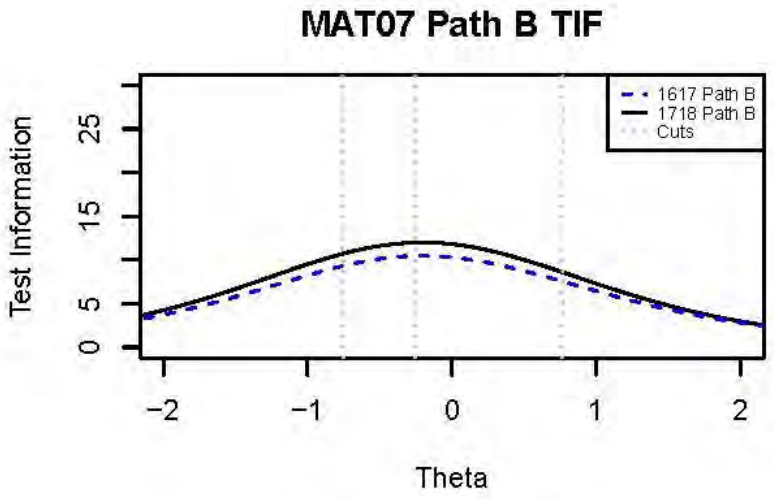


Figure L-10. 2017–18 MSAA: Test Information Function and Standard Error for Grade 7 Mathematics – Paths A, B, and C





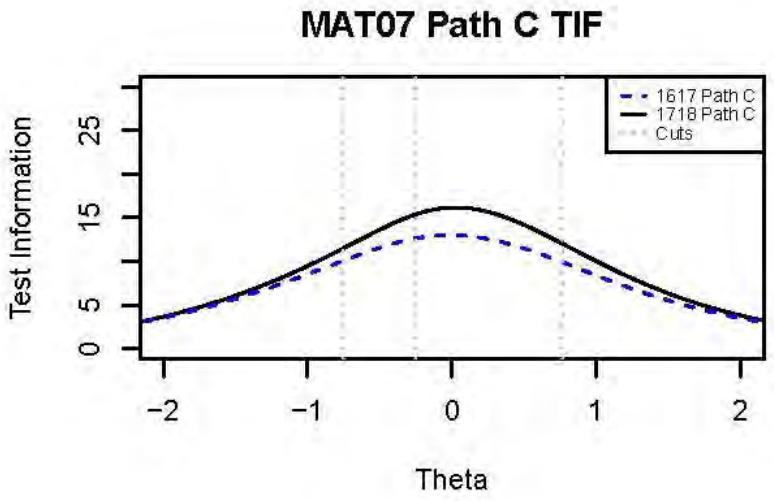
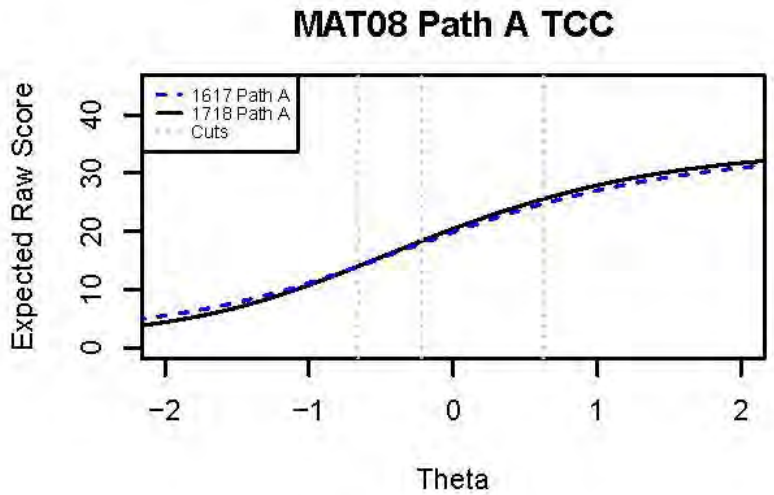
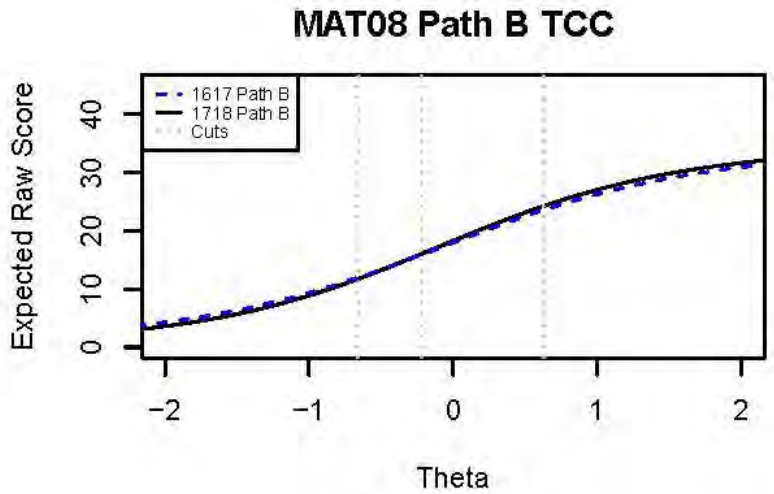


Figure L-11. 2017–18 MSAA: Test Characteristic Curve for Grade 8 Mathematics – Paths A, B, and C





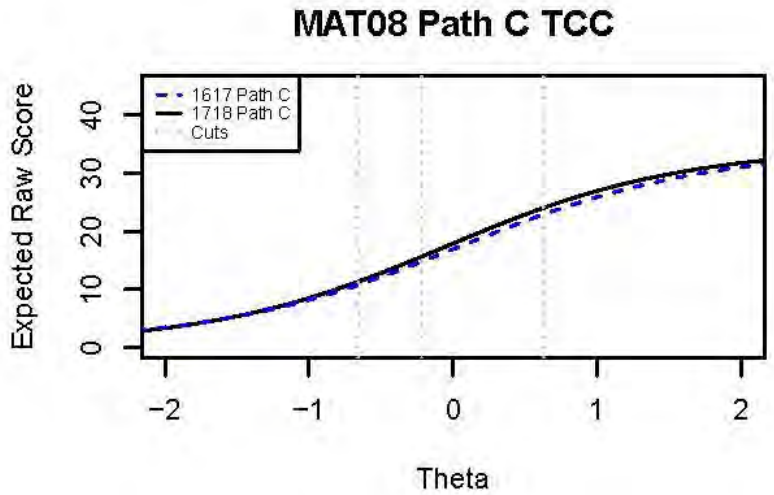


Figure L-12. 2017–18 MSAA: Test Information Function and Standard Error for Grade 8 Mathematics – Paths A, B, and C

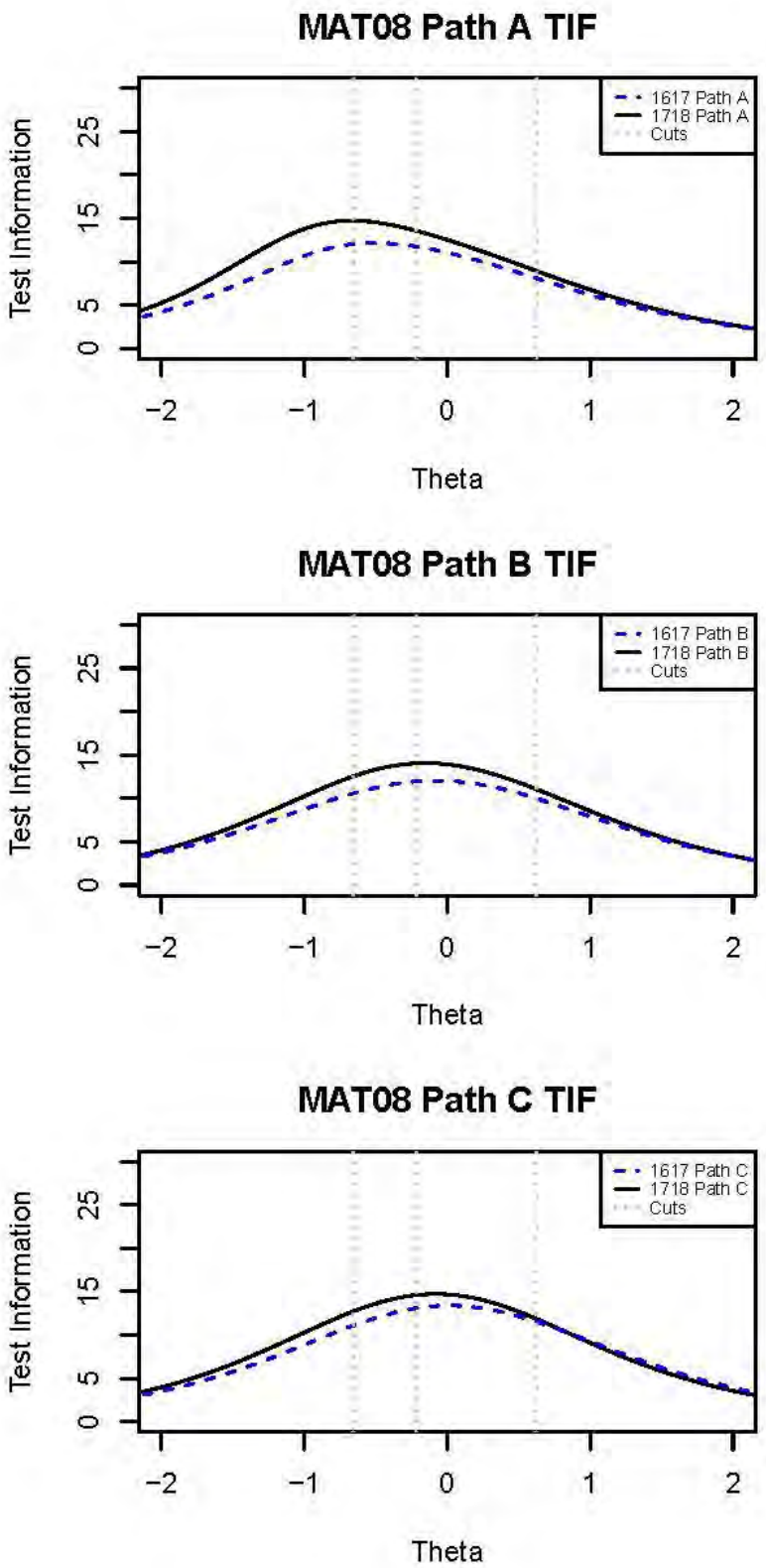
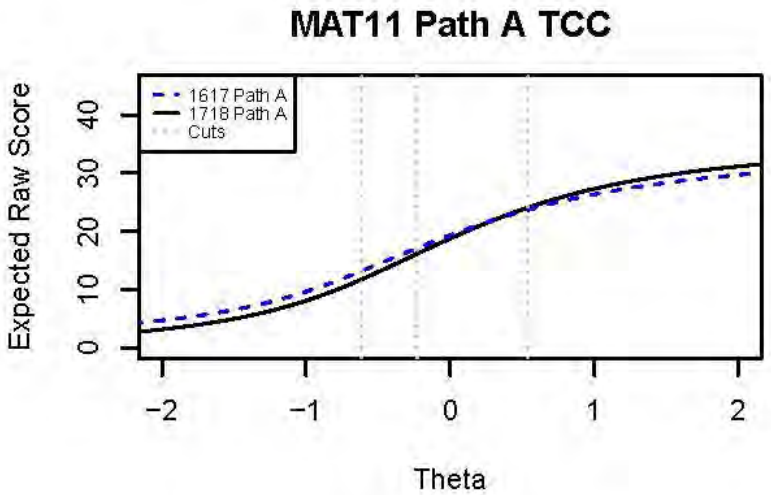
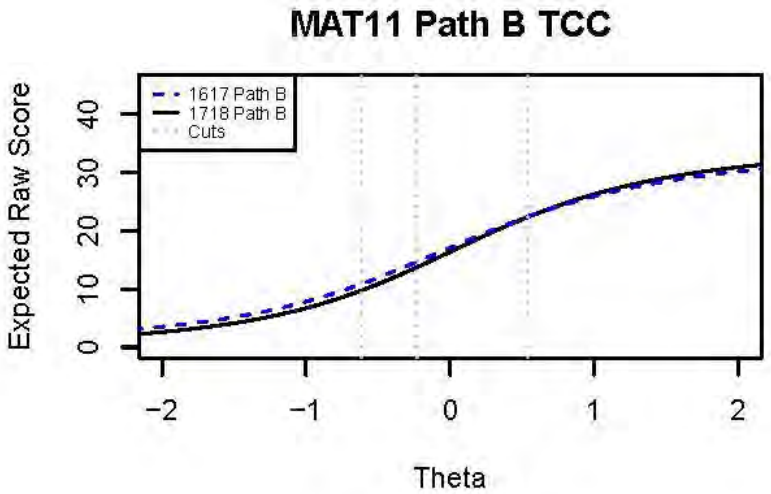


Figure L-13. 2017–18 MSAA: Test Characteristic Curve for Grade 11 Mathematics – Paths A, B, and C





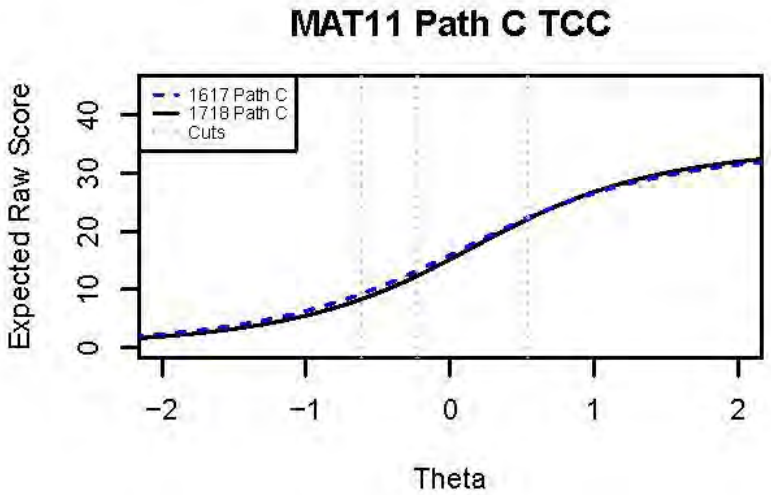


Figure L-14. 2017–18 MSAA: Test Information Function and Standard Error for Grade 11 Mathematics – Paths A, B, and C

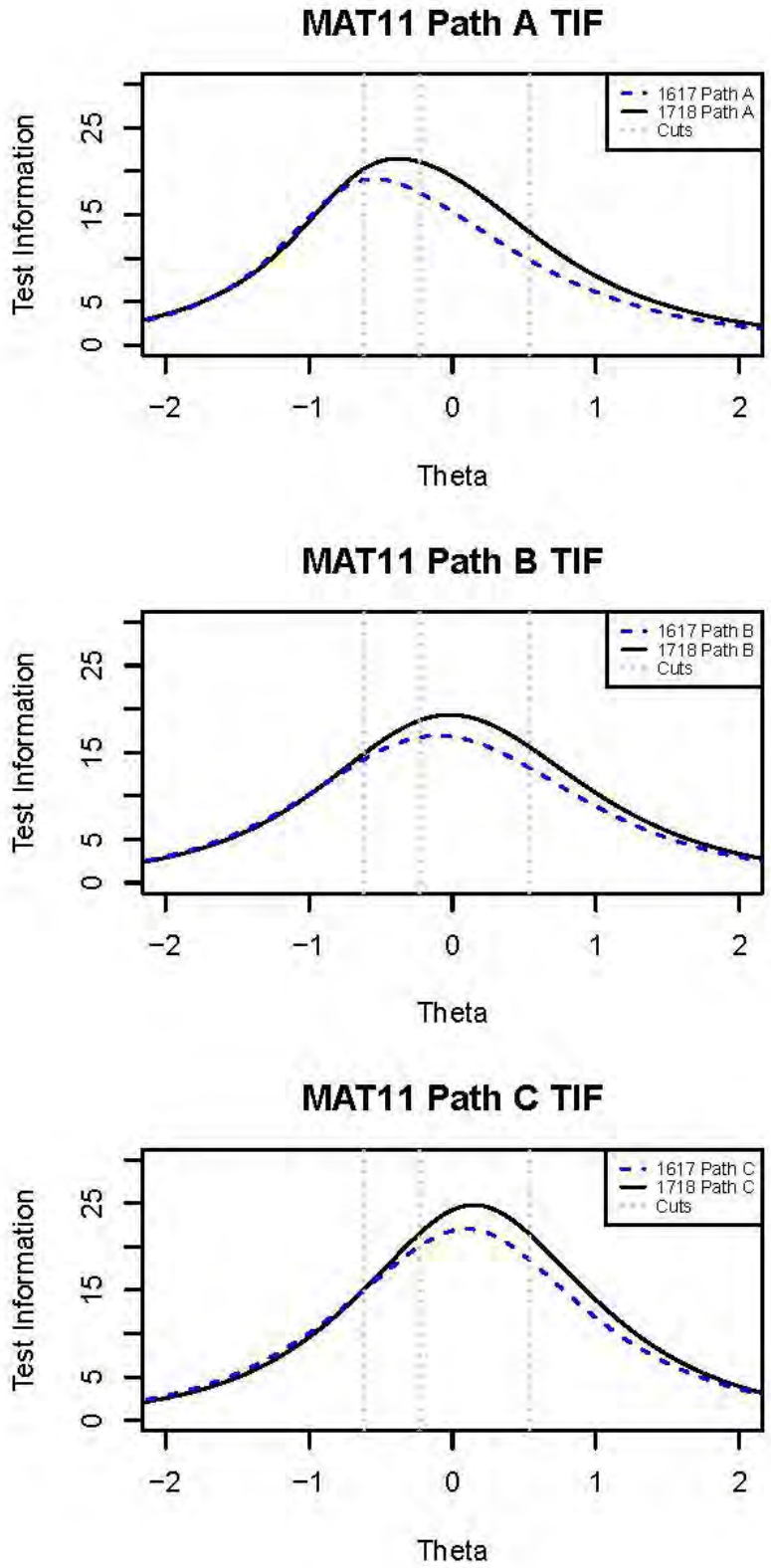


Figure L-15. 2017–18 MSAA: Test Characteristic Curve for Grade 3 ELA – Paths A, B, and C

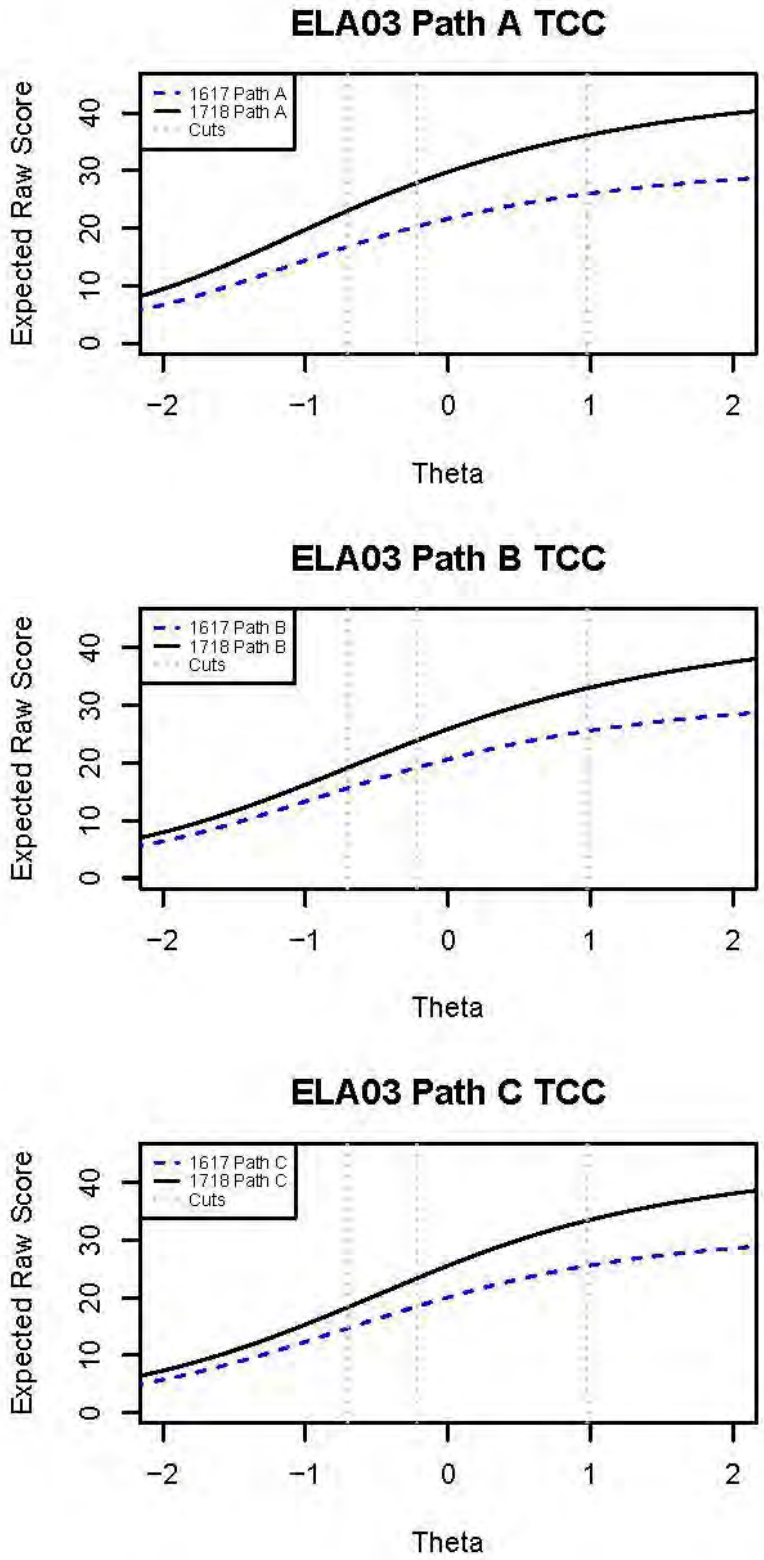


Figure L-16. 2017–18 MSAA: Test Information Function and Standard Error for Grade 3 ELA – Paths A, B, and C

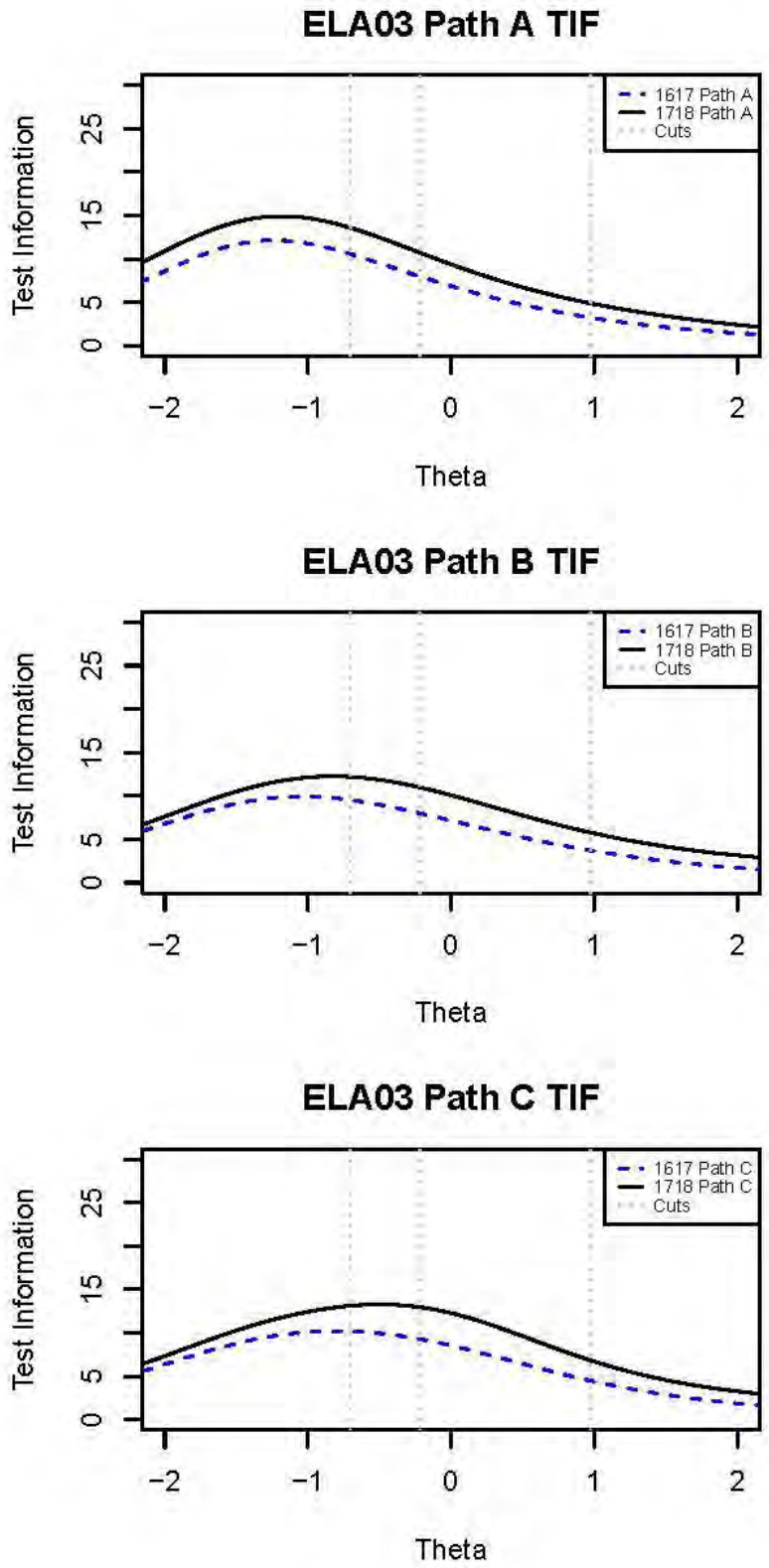
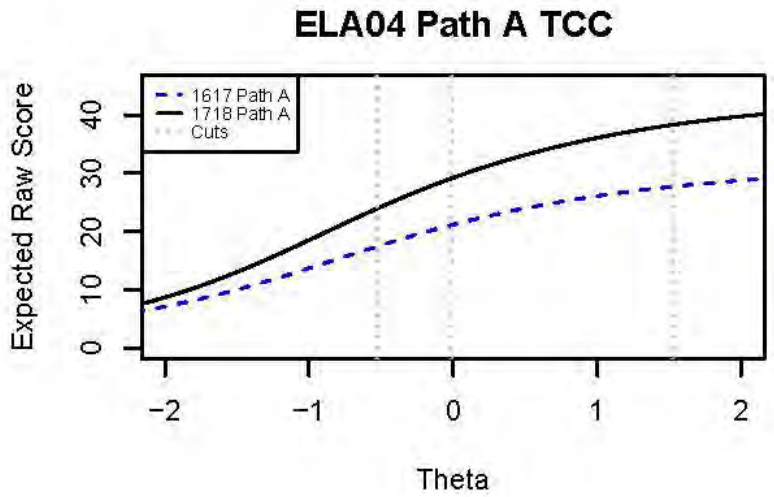
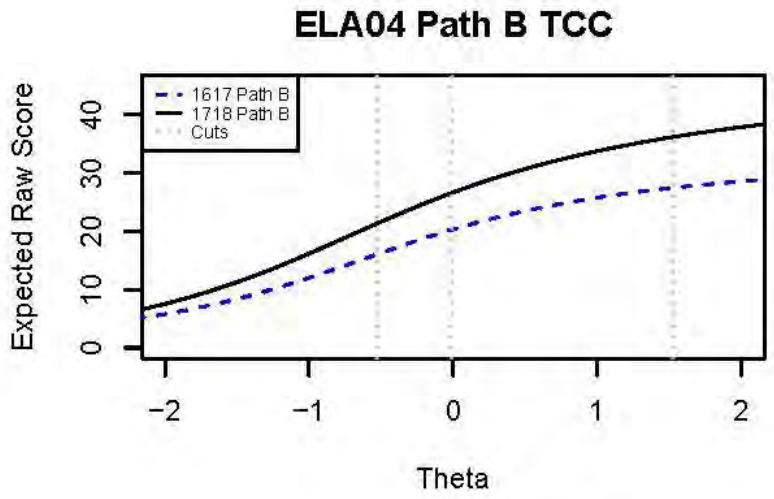


Figure L-17. 2017–18 MSAA: Test Characteristic Curve for Grade 4 ELA – Paths A, B, and C





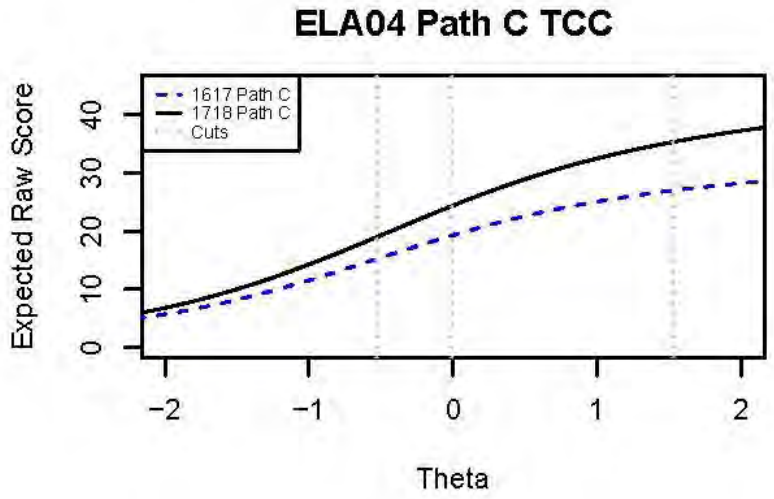


Figure L-18. 2017–18 MSAA: Test Information Function and Standard Error for Grade 4 ELA – Paths A, B, and C

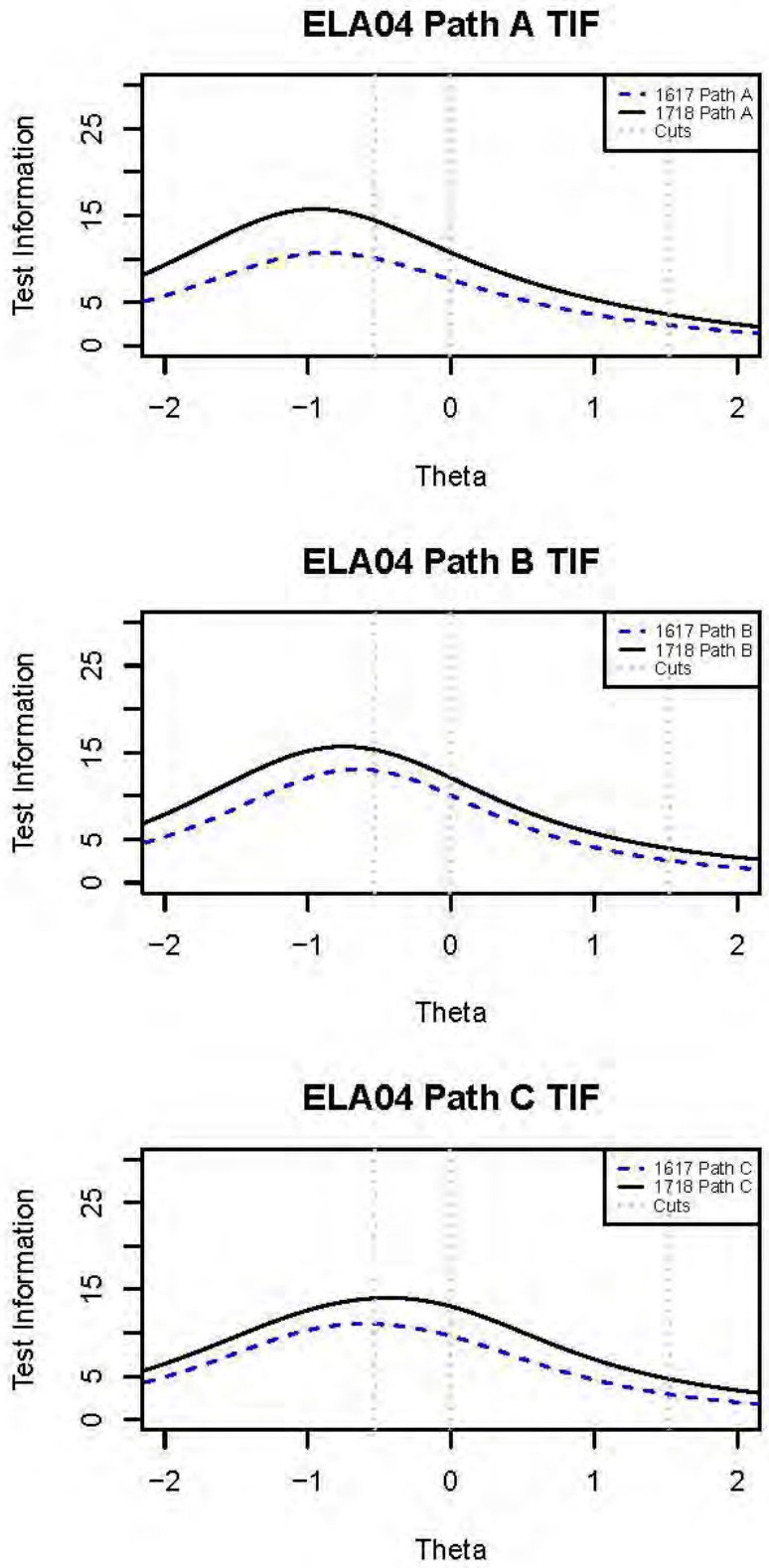
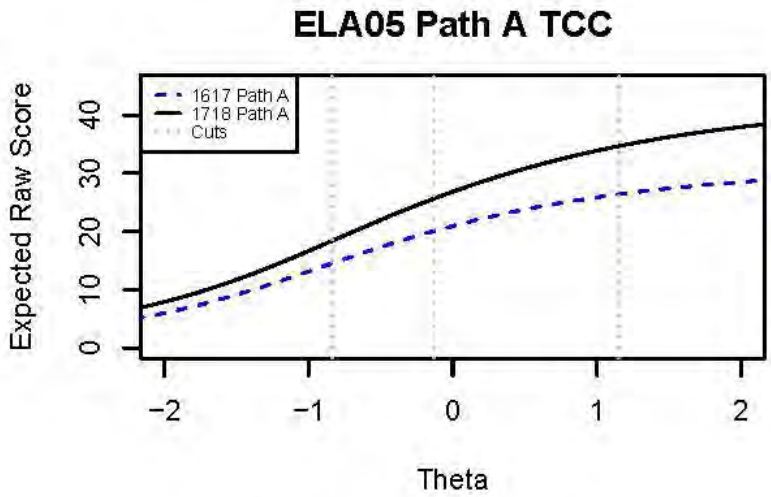
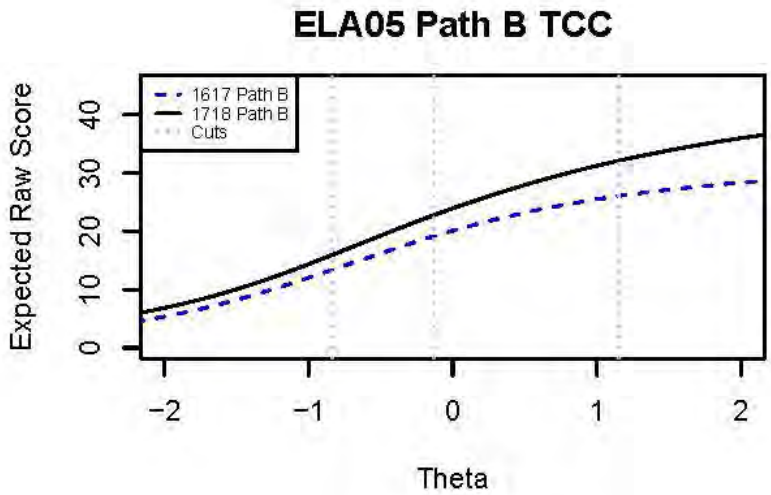


Figure L-19. 2017–18 MSAA: Test Characteristic Curve for Grade 5 ELA – Paths A, B, and C





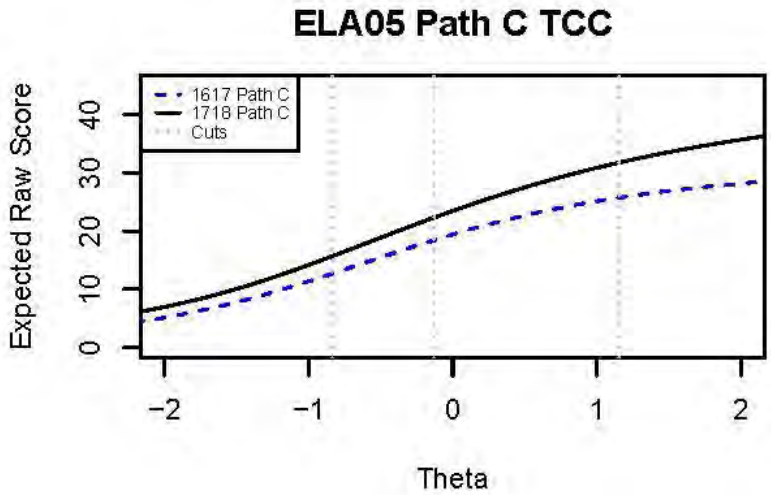


Figure L-20. 2017–18 MSAA: Test Information Function and Standard Error for Grade 5 ELA – Paths A, B, and C

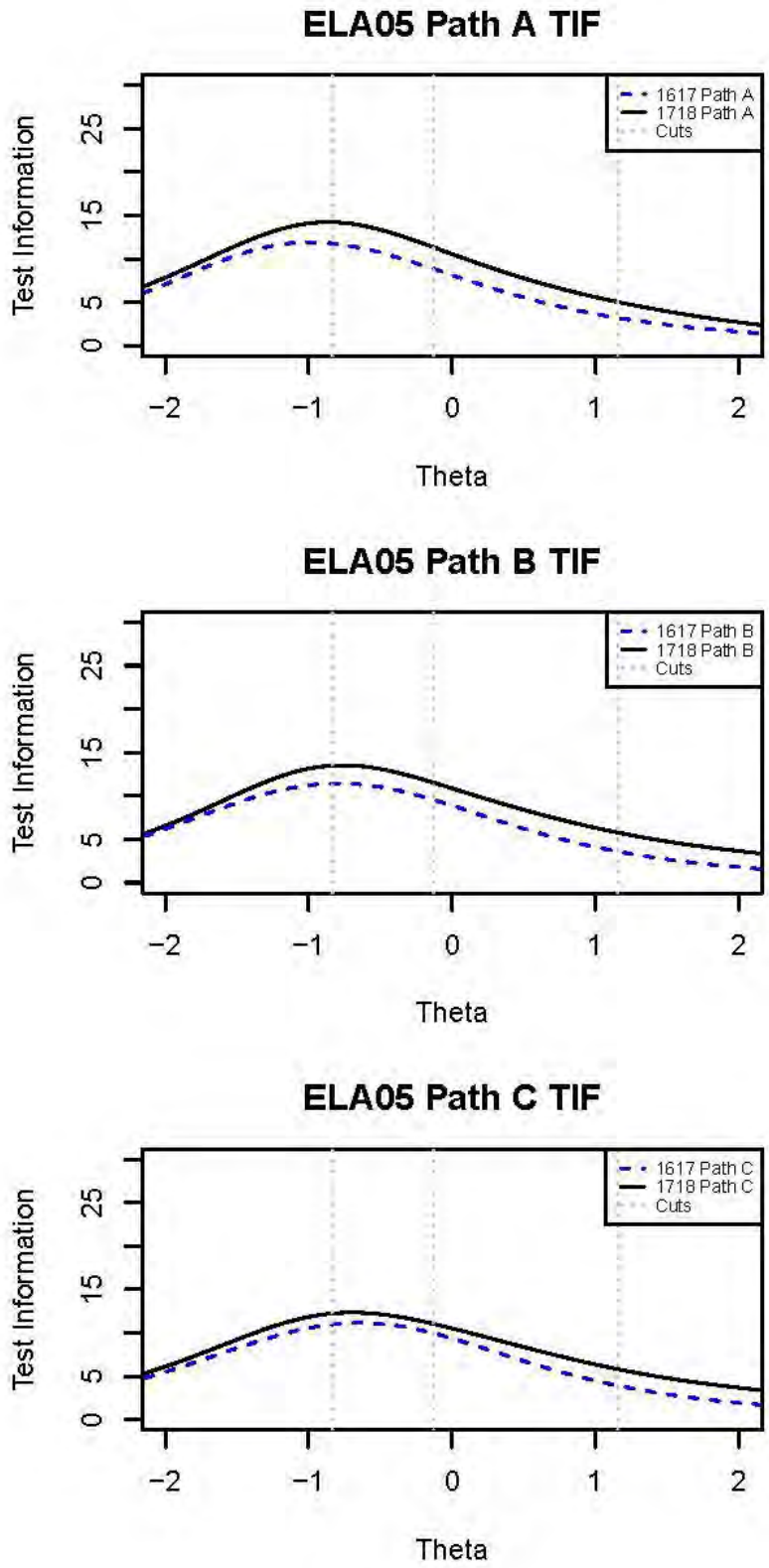


Figure L-21. 2017–18 MSAA: Test Characteristic Curve for Grade 6 ELA – Paths A, B, and C

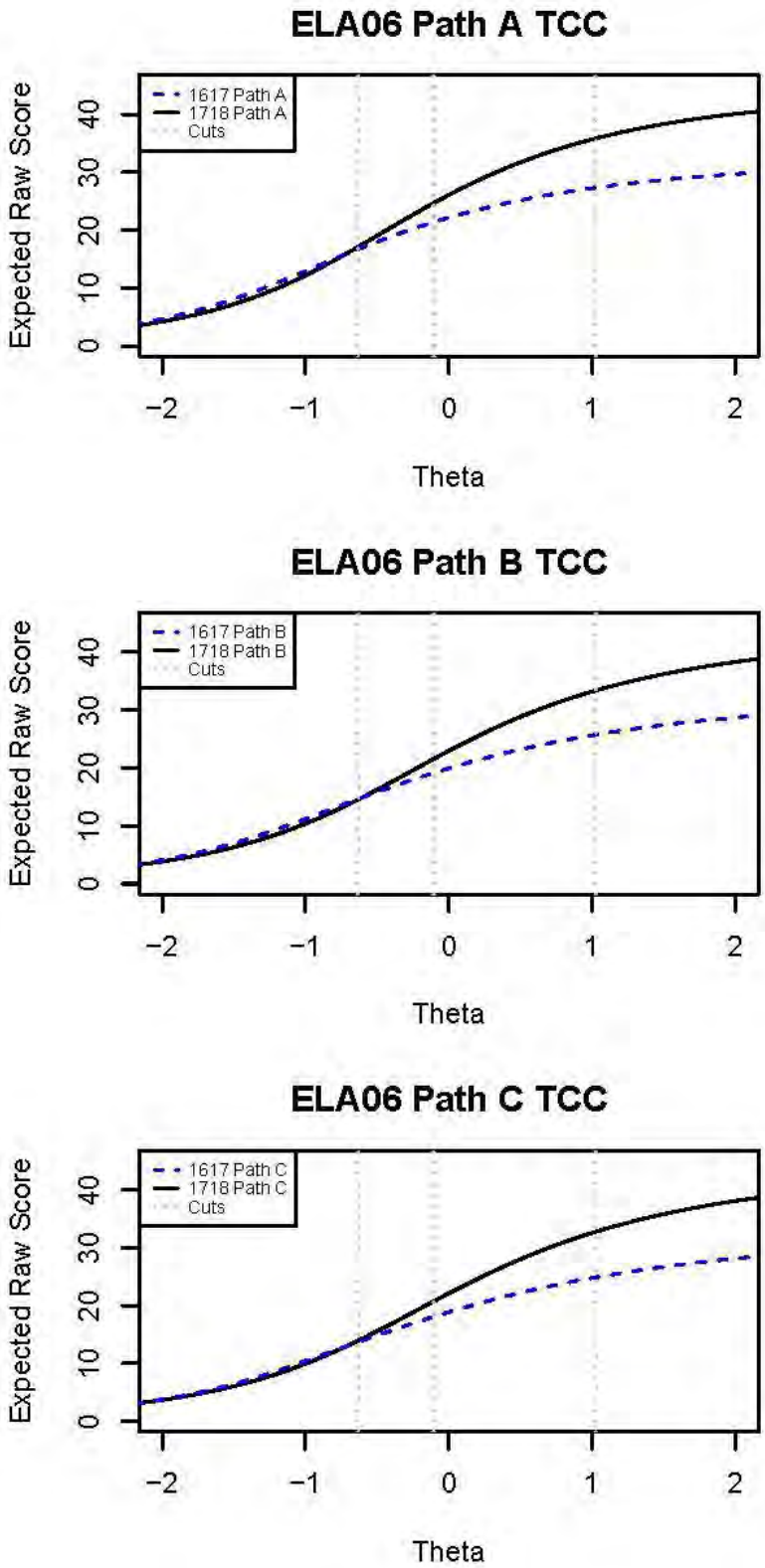


Figure L-22. 2017–18 MSAA: Test Information Function and Standard Error for Grade 6 ELA – Paths A, B, and C

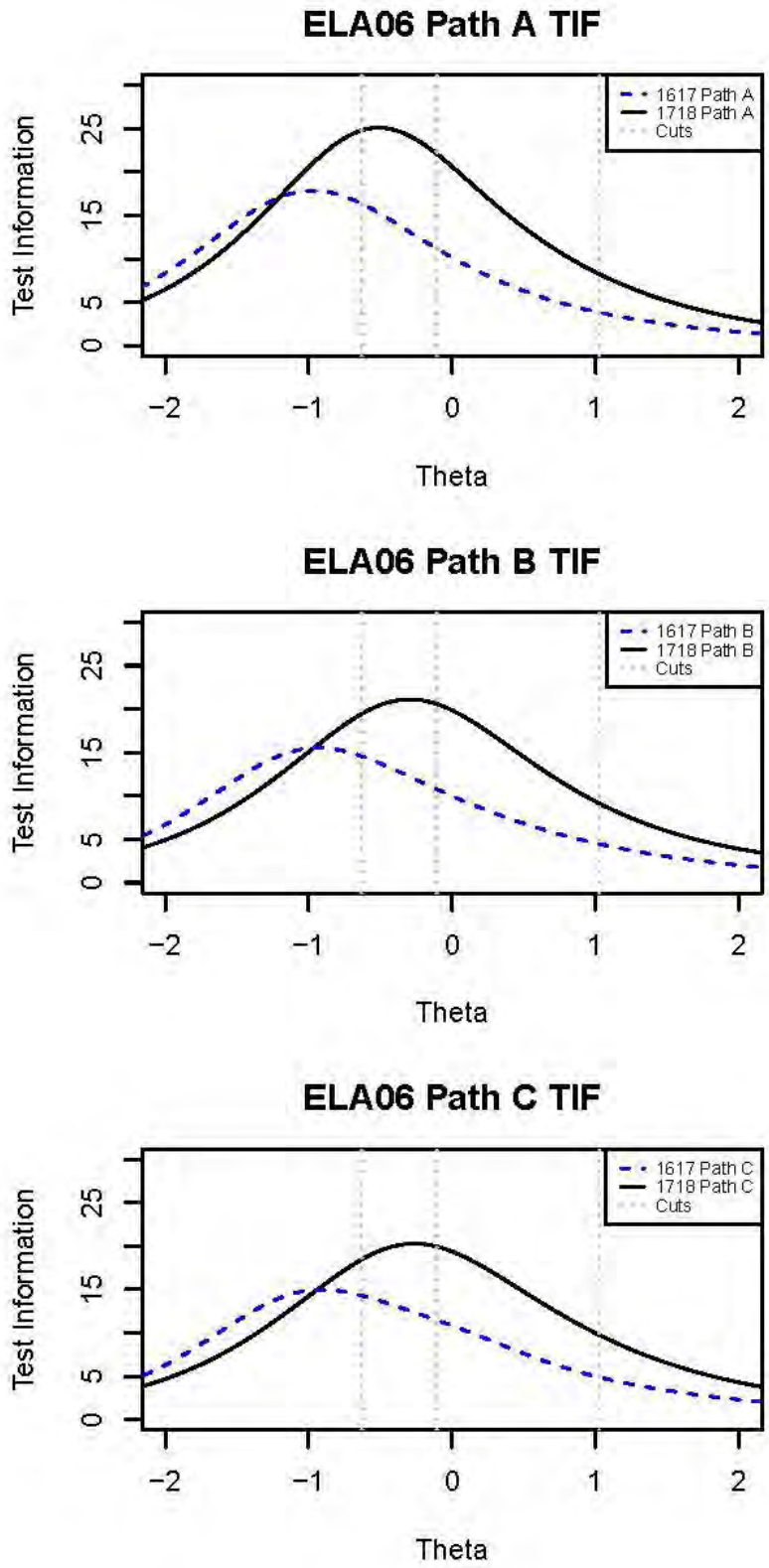


Figure L-23. 2017–18 MSAA: Test Characteristic Curve for Grade 7 ELA – Paths A, B, and C

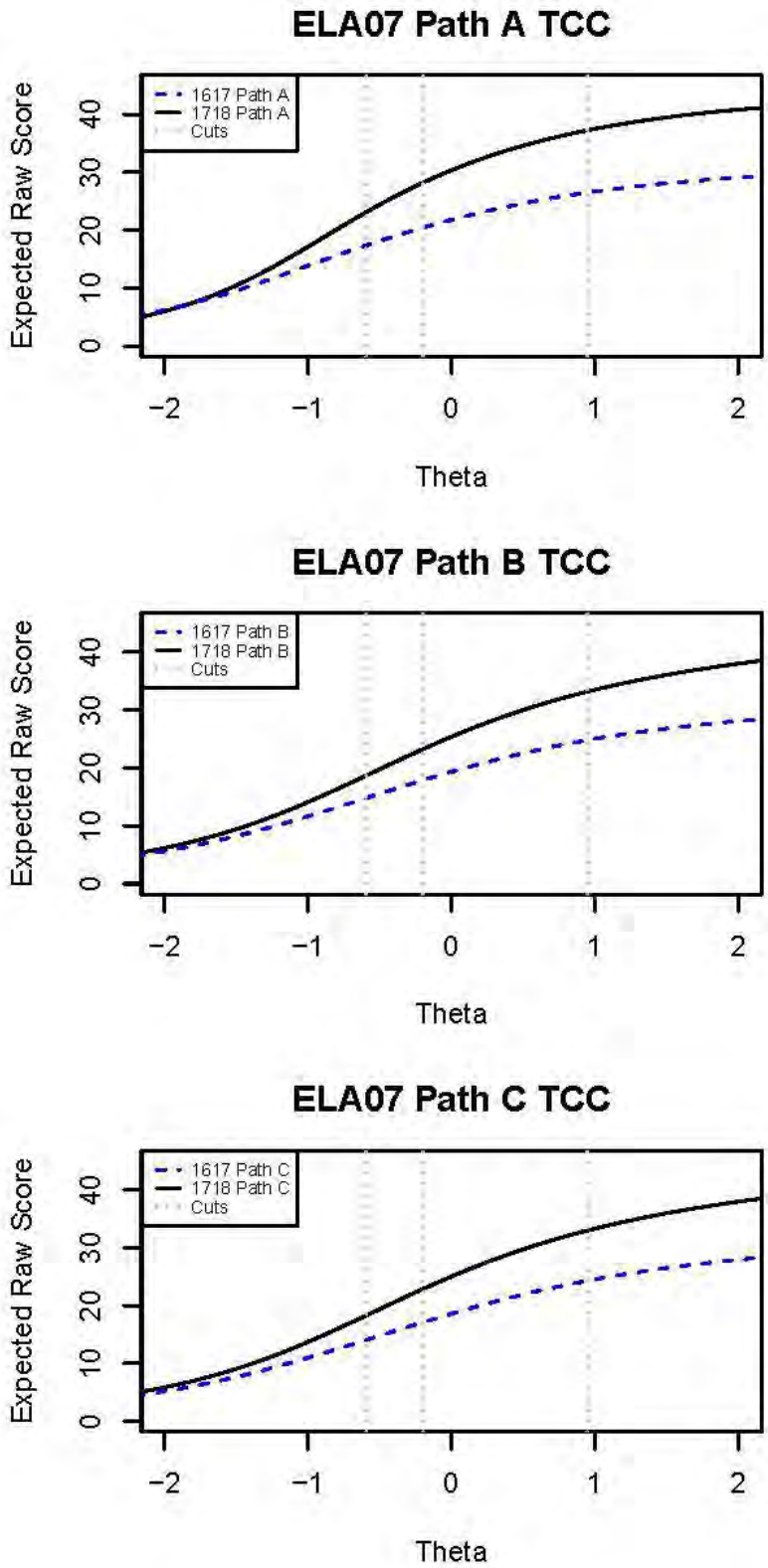


Figure L-24. 2017–18 MSAA: Test Information Function and Standard Error for Grade 7 ELA – Paths A, B, and C

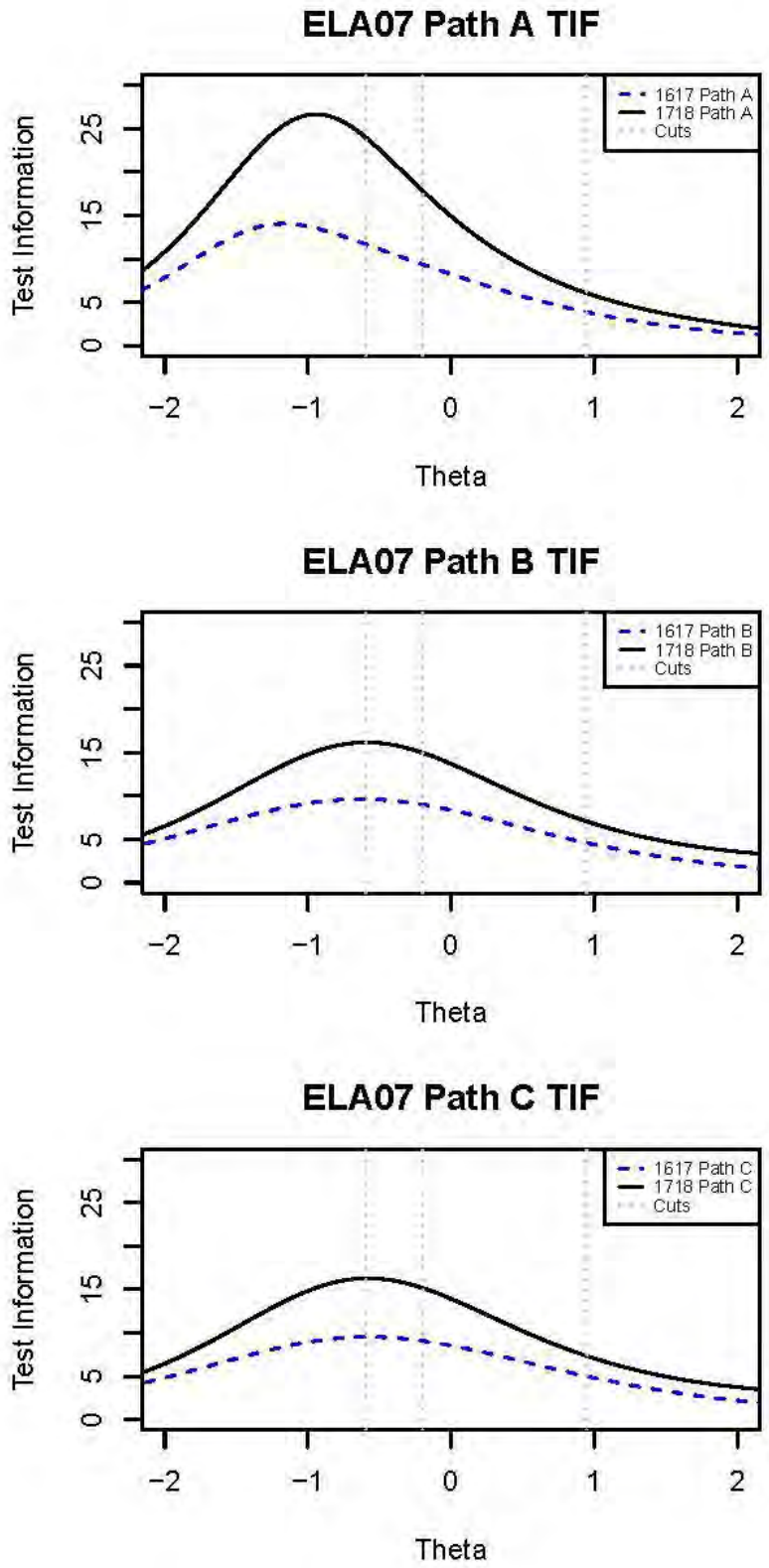
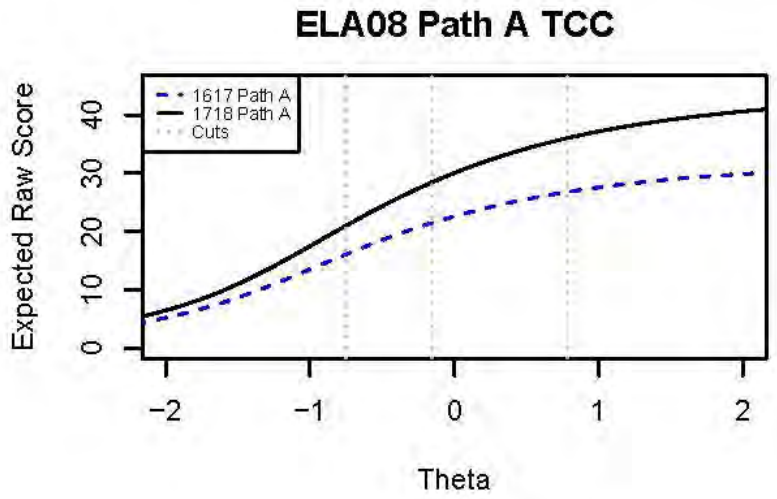
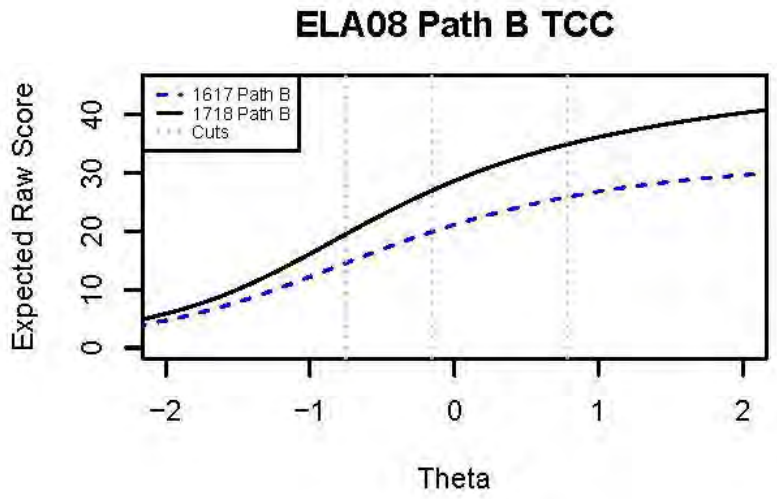


Figure L-25. 2017–18 MSAA: Test Characteristic Curve for Grade 8 ELA – Paths A, B, and C





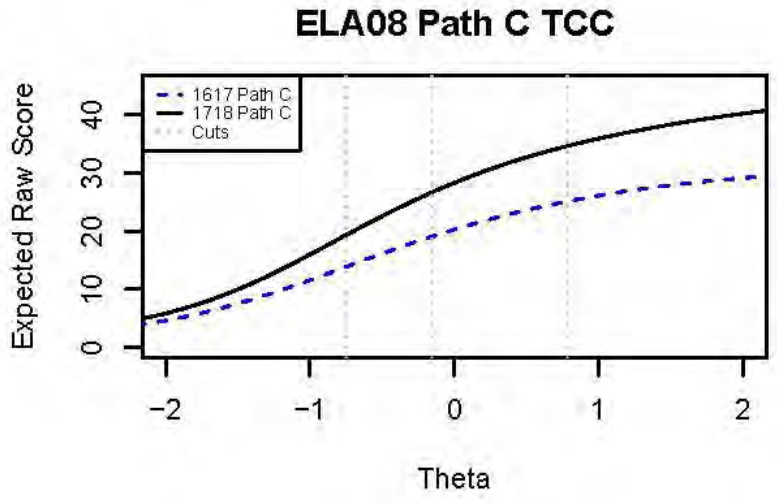


Figure L-26. 2017–18 MSAA: Test Information Function and Standard Error for Grade 8 ELA – Paths A, B, and C

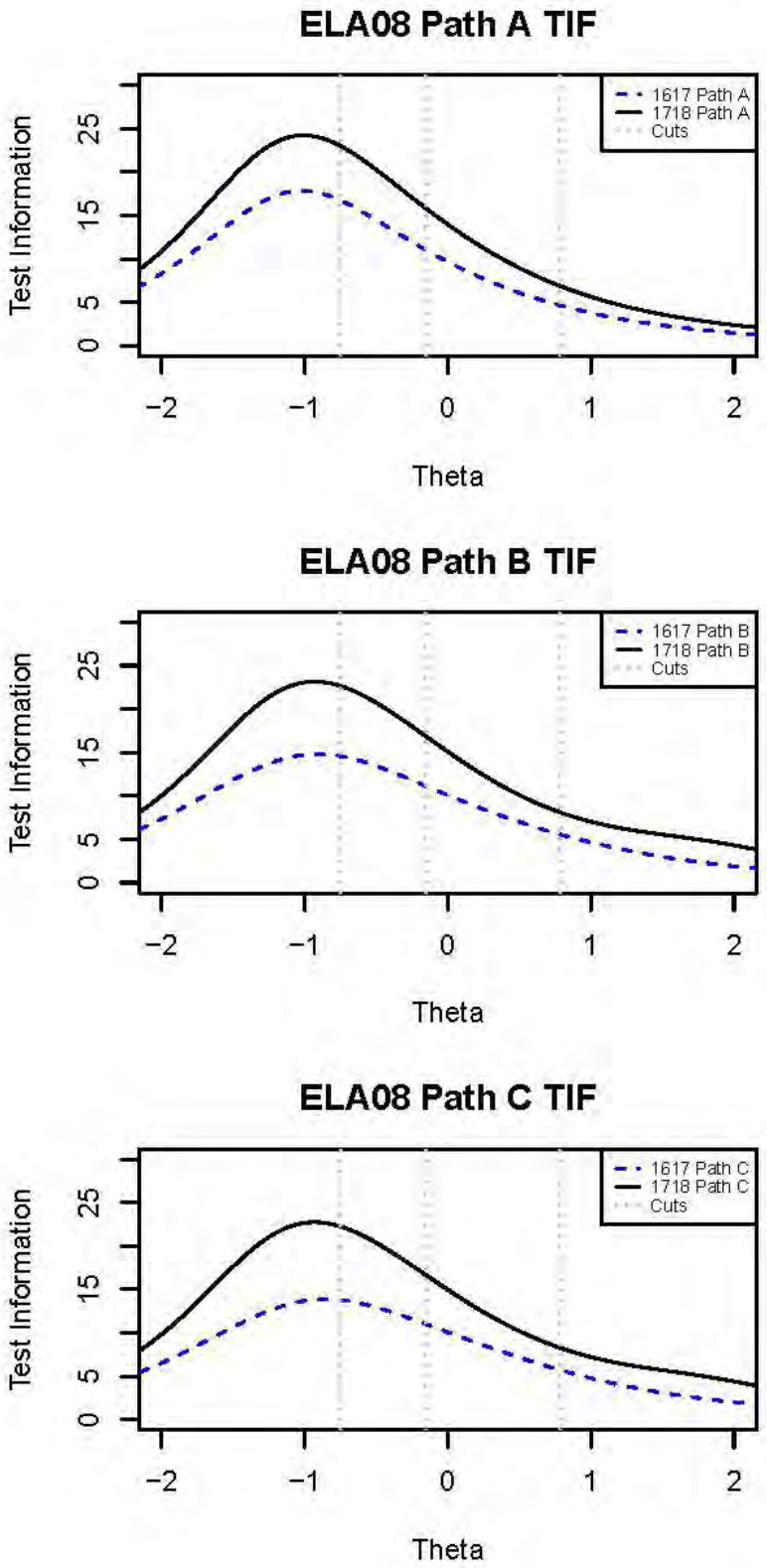
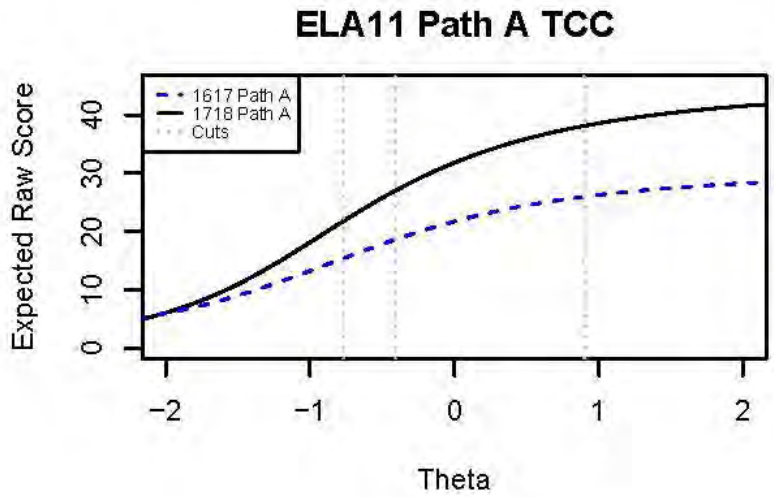
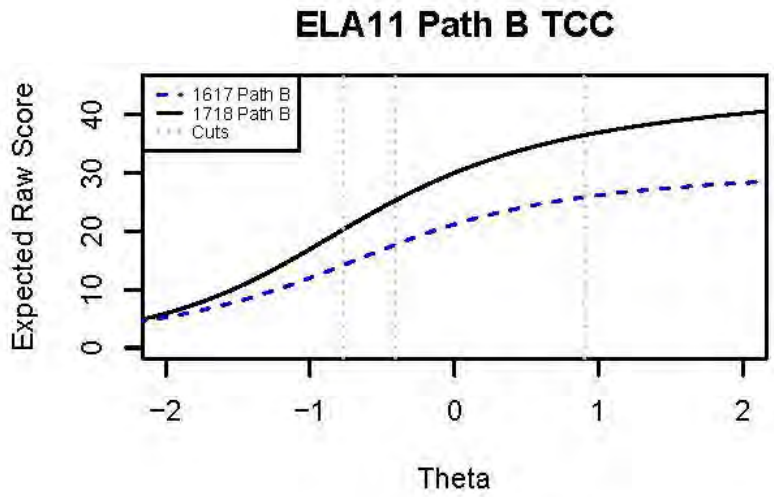


Figure L-27. 2017–18 MSAA: Test Characteristic Curve for Grade 11 ELA – Paths A, B, and C





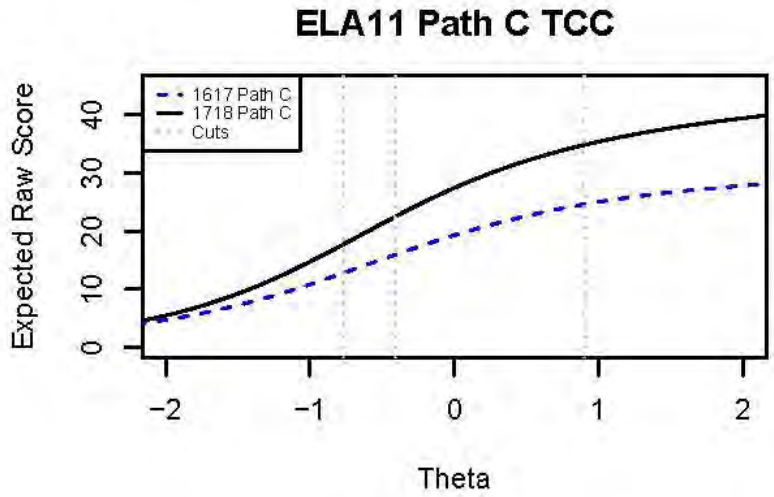
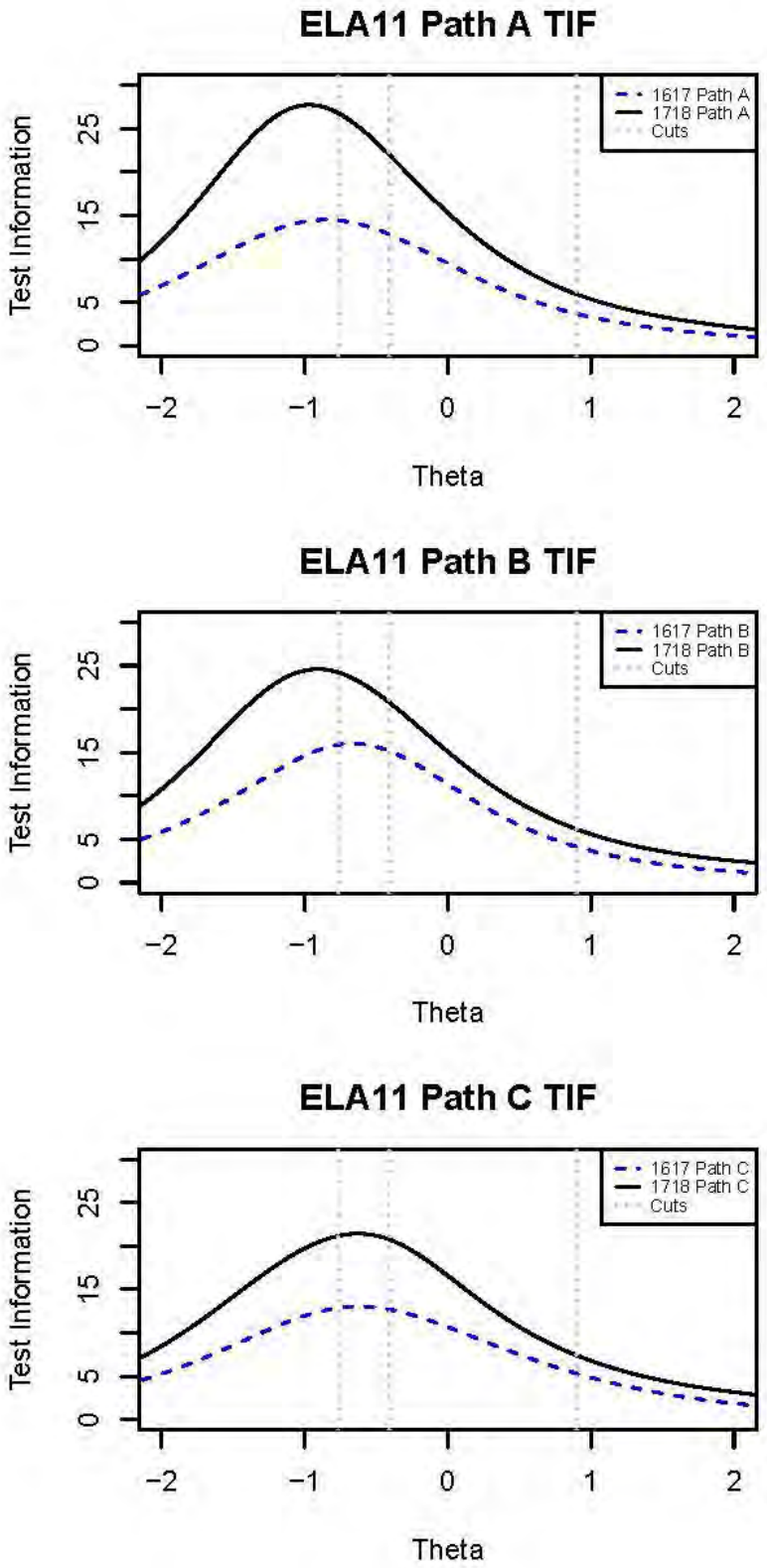


Figure L-28. 2017–18 MSAA: Test Information Function and Standard Error for Grade 11 ELA – Paths A, B, and C



APPENDIX M—STANDARDS VALIDATION REPORT

2018 Standards Validation Report

Prepared by: Measured Progress

October 2018

Copyright © 2018 Measured Progress. All rights reserved. Measured Progress and its logo are registered trademarks of Measured Progress, Inc.

Table of Contents

Chapter 1. Overview of the Standards Validation Process 1

Chapter 2. Tasks Completed Prior to the Cut Score Review Meeting 2

2.1 Vertical Articulation of the ELA and Mathematics Cut Scores 2

2.2 Expansion of the ELA PLDs to Reflect the Addition of the Writing Prompts 3

2.3 Expansion of the ELA Score Scale to Reflect the Addition of the Writing Prompts 4

2.4 Preparation for the ELA Cut Score Review Meeting 4

Chapter 3. Tasks Completed During the Cut Score Review Meeting 6

3.1 Cut Score Review Process 6

3.2 Final Recommendations on Cut Scores 8

Chapter 4. Tasks Completed After the Cut Score Review Meeting 10

4.1 Vertical Articulation of Mathematics Performance Level 2 and 4 Cut Scores 10

4.2 Vertical Articulation of ELA Performance Levels 2 and 4 11

4.3 Review of the Recommended Performance Level Cut Scores for ELA and Mathematics 12

4.4 ELA PLD Review Meeting Overview 13

4.4.1 Preparation for the ELA PLD Review Meeting 13

4.4.2 Panelist Selection 14

4.5 Overview of the ELA PLD Review Meeting Process 16

4.6 Final Approval of All Cut Scores 17

4.7 Preparation of the Standards Validation Process Report 18

References 19

Appendices 20

APPENDIX A **Vertical Articulation Report**

APPENDIX B **Participants and Roles in the Standards Validation Process**

APPENDIX C **ELA Cut Score Review Meeting Agenda**

APPENDIX D **MSAA Cut Score Review Meeting**

APPENDIX E **MSAA ELA CSR Summary**

APPENDIX F **MSAA ELA CSR Adjusted Cut Scores CBRs**

APPENDIX G **MSAA Mathematics Vertical Articulation Results by Grade**

APPENDIX H **MSAA 2018 ELA PLD Panelist Training Opening Presentation**

APPENDIX I **MSAA ELA PLD Review Meeting Process**

APPENDIX J **ELA PLD Panelist Feedback**

APPENDIX K **MSAA ELA PLD Review Evaluation Forms**

Chapter 1. Overview of the Standards Validation Process

This report summarizes the activities involved in reviewing and validating the cut scores for the 2018 MSAA English Language Arts (ELA) and mathematics assessments. This review 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 necessary because of the addition of the open-response writing prompt scores to the existing ELA score scale. All activities described in this report were recommended by the MSAA Technical Advisory Committee (TAC) in a February 2018 meeting with the Psychometric Subcommittee and Measured Progress.

The standards validation process involved five steps: (a) vertically articulating the performance level 3 cut scores for both ELA and mathematics, to update the performance standards and provide a coherent basis for interpreting 2018 scores and performance and validating the ELA cut scores; (b) expanding the ELA PLDs to include references to direct open-response writing prompt knowledge and skills; (c) expanding the existing ELA score scale by linking the open-response writing prompt scores to the scale; (d) reviewing and possibly adjusting the existing cut scores to align the response demands of all testing items, including the open-response writing prompt scoring rubrics and knowledge and skill requirements in the expanded PLDs; and (e) reviewing the expanded ELA PLDs with stakeholders from various schools and districts to confirm the writing knowledge and skills were clear and included language around the open-response writing prompt expectations in each of the performance levels.

This report is organized into three major sections: activities completed (a) prior to, (b) during, and (c) after the cut score review meeting.

Chapter 2. Tasks Completed Prior to the Cut Score Review Meeting

Chapter 2 details four activities that are part of the Standards Validation process:

(1) Vertical articulation of the ELA and mathematics cut scores
(2) Expansion of the ELA PLDs to reflect the addition of the writing prompts
(3) Expansion of the ELA score scale to reflect the addition of the writing prompts
(4) Preparation for the ELA cut score review meeting

2.1 Vertical Articulation of the ELA and Mathematics Cut Scores

The MSAA performance standards for ELA and mathematics were established in 2015, when the content standards and assessments were new to teachers and students in partner states. In discussion with the TAC, the MSAA states acknowledged the need to articulate the performance standards, to provide a coherent basis for interpreting 2018 scores and performance, and for validating the ELA cut scores. Vertically articulated standards for MSAA are reflected in similar percentages of students in performance levels across grade levels. The goal prior to the cut score review meeting was to articulate across grade levels the percentages of students at and above performance level 3 (i.e., levels 3 and 4 combined) in the 2017 data. (Performance levels 2 and 4 were articulated after the cut score review meeting for mathematics only. See section *4.1 Vertical Articulation of Mathematics Performance Level 2 and 4 Cut Scores*).

Measured Progress prepared for vertical articulation in two meetings with the Psychometric Subcommittee and TAC:

- May 4, 2018, to specify the vertical articulation approach as well as identifying grades and content areas that required focused attention; the TAC expressed concern regarding the ELA grades 6 and 8 standards and mathematics grade 6 standards, which appeared more difficult than standards in other grades
- A follow-up clarification meeting on June 15, 2018

Measured Progress also conducted internal working meetings on June 11 and 21, 2018, to complete the vertical articulation process and develop recommendations to MSAA.

Procedures and Recommendations

A Measured Progress team of Content Development, Psychometrics, and Program Management staff completed the vertical articulation process for the 2016–2017 performance data. The team reviewed the inverse cumulative percentages of students in each proficiency level in grades 3–8 and 11 in ELA and mathematics, and the locations in each distribution of the levels 2, 3, and 4 cut scores. The team focused on the areas of concern—performance level 3 cut scores in ELA grades 6 and 8 and mathematics grade 6—and viewed all grade level percentages together to consider the degree of cross-grade articulation. Based on previous discussions with the TAC members, we considered cut score adjustment for any differences in any pairs of percentages at/above performance level greater than 5%.

The team began by reviewing the impact data for ELA grades 6 and 8 in relation to the other grades and examined the effects on articulation of adjusting the cut scores in grades 6 and 8 by 1, 2, and 3 scale score points. The team then reviewed the impact data for mathematics grade 6 in relation to the other grades and examined the effects on articulation of adjusting the cut scores in grades 3, 4, and 6 by 1 scale score point. A senior Content Specialist at Measured Progress wrote content based rationales, based on the relationship between response demands of items at and above the adjusted cut scores and the knowledge and skill requirements in the corresponding PLD.

Measured Progress recommended the following to MSAA:

- Adjusting the ELA grade 8 performance level 3 cut score by 2 scale score points and the grade 6 ELA cut score by 3 points achieves vertical articulation. The final proposed cut scores display smooth articulated results across grades.
- Adjusting the mathematics performance level 3 cut scores in grades 3, 4, and 6 by 1 scale score point each achieves vertical articulation. The final proposed cut scores display smooth articulated results across grades.

Details of the process, recommendations, and content based rationales appear in the Vertical Articulation report in Appendix A.

2.2 Expansion of the ELA PLDs to Reflect the Addition of the Writing Prompts

The PLDs 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). The ELA PLDs were developed and finalized as part of the standard setting that took place in 2015. These ELA PLDs were used as the starting point for revision. The open-response writing prompt expectations were

added to each ELA PLD for all grades (3–8 and 11) without other expectation information being altered. In order to include the open-response writing prompt expectations in the ELA PLDs, the MSAA subcommittees and Measured Progress referenced the writing prompt definitions and emphasis from the National Center and State Collaborative (NCSC) document Building From the Ground Up: A Writing Assessment Story, as this document served as the basis for writing prompt development (both the selected response writing prompt and the open-response writing prompt), as well as the open-response writing prompt score rubrics. The ELA PLDs were revised to include the open-response writing prompt expectations through an iterative process among Measured Progress and members of the MSAA Psychometric, Item Development, and Scoring Subcommittees. A list of the subcommittee members is included in Appendix B. The revised ELA PLDs were used during the Cut Score Review meeting as a guiding document. In addition, the ELA PLDs were reviewed for clarity by stakeholders from across states, schools and districts during the ELA PLD Review meeting.

2.3 Expansion of the ELA Score Scale to Reflect the Addition of the Writing Prompts

Open-response writing prompts were incorporated as an operational component of the MSAA ELA assessment in spring 2018. These writing prompts were previously field tested. The purpose of this addition to the reading, language, and writing (i.e., selected response writing prompt and other writing skills) items was to assess content standards that cannot be addressed with selected response items and to expand the interpretation of ELA scores and performance.

After the reading, language, and writing items were equated to the operational scale, a second equating was completed to bring the open-response writing prompts onto scale. The second equating used a fixed common item parameter method of equating, where all reading, language, and writing item parameters were fixed to the operational scale, and the open-response writing prompts were then equated onto the operational scale. By using this method of equating, all reading, language, and writing item parameters were linked to the operational scale during the equating of the open-response writing prompt. The ELA scale was adjusted to reflect the linking of the open-response writing prompts, expanding the interpretation of ELA scores and performance.

2.4 Preparation for the ELA Cut Score Review Meeting

Two steps were taken to ensure that both MSAA and Measured Progress were well prepared for this meeting:

- On July 18, an internal run-through of the cut score review process was conducted to familiarize all participating Measured Progress staff with the process and time to test all materials needed for the meeting.
- On July 21, an e-mail was sent to MSAA members who would be participating in the July 24 cut score review meeting, detailing the meeting goals, the cut score review process, the schedule and materials that participants would review during the meeting.

Materials prepared for the ELA Cut Score Review Meeting:

- **Meeting PowerPoint slides**: Training was provided to participants in the cut score review process and decision criteria.
- **Expanded ELA Performance Level Descriptors**: The expanded ELA PLDs included new references to the content standards related to producing a written product. They served as the reference point for student expectations in each performance level as determined by the MSAA ELA assessment.
- **PDFs of all items that were included in the cut score review**: These were the 2018 items ordered as they were presented to students. Panelists were able to review items via item ID numbers as needed during the cut score reviews.
- **Open-response writing prompt level 2 and 3 rubrics**: These were the 2018 rubrics that were used to score student responses. Panelists were able to review the scoring rubrics as needed during the cut score reviews.
- **Online impact data tool**: This tool included an item map of all items included in the cut score review, with the scale locations of all selected response items and open-response writing prompt score level thresholds displayed, and all cut scores identified. This information indicated where the open-response writing prompt score level thresholds fell in relation to the cut scores and other items. The meeting facilitator managed the online tool for panelists. The tool enabled the facilitator to illustrate immediately changes in impact data (i.e., percentages of students in each performance level in 2018) as panelists considered adjusting cut scores.
- **A pre-formatted table**: Final, recommended cut scores and content based rationales for retaining or adjusting cut scores were recorded. This table was included in the report provided to MSAA State Partners to use with their Superintendents/Boards of Education for approval of the final recommended cut scores. (See appendices E and F)

Chapter 3. Tasks Completed During the Cut Score Review Meeting

3.1 Cut Score Review Process

The cut score review meeting was held on July 24, 2018, via webinar. Members of the MSAA Psychometric Subcommittee, including MSAA ELA Content Specialists, acted as review panelists. Two members of the TAC attended to monitor the process and provide advice and support as needed. A list of the cut score review meeting attendees is included in Appendix B. Measured Progress Psychometric, Content Design and Development, and Special Education-Client Services staff led and facilitated the meeting. In the meeting, panelists reviewed the alignment between the location of open-response writing prompt scores[1] on the expanded ELA scale and the PLDs. Recommendations were made to either (a) retain the existing 2015 cut scores, or (b) adjust the cut scores to improve alignment. Content Specialists provided rationales for adjusting cut scores. The agenda for the meeting appears in Appendix C.

The cut score review process:

- Measured Progress provided an overview for the meeting, including the need for a cut score review, purpose and goal of the meeting, and details of the meeting procedures and materials and introduced all meeting participants. Members included the MSAA Psychometric Subcommittee and MSAA Content Specialists who served as cut score review panelists, the two TAC members who monitored and supported the process, and Measured Progress staff who facilitated and supported the process. The PowerPoint slides used to train participants in the cut score review process and decision criteria and to manage the process appear in Appendix D.
- The facilitator reviewed each group's roles and responsibilities and laid out discussion ground rules. Specifically, panelists were instructed to (a) share insights about items and open-response writing prompts and scoring rubrics and avoid trying to persuade other panelists in round 1 about recommendations for cut scores; (b) work together to reach

[1] Specifically, the RP 67 location of all dichotomous items and open-response writing prompt score level thresholds for the three rubrics, consistent with the RP criterion used in the 2015 standard setting.

consensus on recommendations in round 2; (c) and collaborate in round 2 to articulate content based rationales when they recommended adjusting cut scores.

- Panelists were instructed on the process they should follow to make recommendations to adjust and validate cut scores, as described below.
- A Measured Progress Special Education Director led the panelists through a review of the expanded ELA PLDs, highlighting the information about content standards related to producing a written product that supplemented the information contained in the original PLDs.

Panelist Procedures and Judgmental Task

The meeting facilitator instructed the panelists to work as follows:

1) Review the locations of the prompt score level locations (i.e., the locations of threshold values for each rubric score: 0,1 and 1,2) (a) on the expanded ELA scale, (b) in relation to the current cut scores and corresponding PLDs.
2) Make content based judgments about the appropriateness of those relationships.
3) Write content based rationales for recommending adjustments to current cut scores.

The facilitator led the panelists through a modeling and practice session for round 1 of grade 11 by modeling the review process while panelists considered the alignment between the open-response writing prompt score level locations in relation to the cut scores and corresponding PLDs. Modeling of the cut score review process included considerations:

1) Do the rubric descriptions for each score level align with the corresponding PLD?
2) What reasonably can be adjusted—without causing undue disruption to impact data and interpretation using the PLDs?
3) Does the item-PLD alignment analysis support the adjusted cut score?

Panelists then took over the review process for round 2 of grade 11 and for grades 8, 7, and so forth through grade 3.

Based on a recommendation from a TAC member, panelists considered as reasonable, as defined in consideration (2) above, any cut score adjustments that were within one standard error (i.e., conditional standard error on the RP 67 theta scale). A Measured Progress psychometrician provided conditional standard errors for each cut score in each of the test form paths 1A, 1B, and 1C. Across grades 3–5, the lowest and highest standard errors are 3.2 and 6.2; 2.5 and 5.4 in grades 6–8; and 2.2 and 4.6 in grade 11. In all grades and paths, the lowest standard errors coincide with the cut score for performance level 1; in grades 3–8, the highest standard errors

coincide with the cut score for performance level 3 in path 1A; at grade 11, the highest standard error coincides with the cut score for performance level 3 in path 1B.

The round 1 judgmental task was stated to panelists as follows:

- Think about each score level threshold location and its corresponding PLD and any rubric-PLD misalignment.
- What recommendations do you want to make about each cut score, given those relationships?
- What is your content based rationale for each recommendation?
- Think independently.

In preparation for round 2, panelists shared insights and understandings about the round 1 judgmental task and their initial recommendations and rationales. In addition, they viewed a summary of the writing score level locations and the facilitator demonstrated effects on impact data of various cut score adjustments under consideration, using the impact data tool. In round 2, panelists reviewed locations independently one last time, completed discussions to achieve consensus on recommendations for all cut scores in a grade, and wrote content based rationales for any recommended cut score adjustments. The second meeting facilitator recorded recommendations and rationales for the panelists.

3.2 Final Recommendations on Cut Scores

Table 3-1 contains the cut scores recommended by panelists from the cut score review. The table includes the original cut scores from the 2015 standard setting and corresponding 2018 impact data and recommended adjusted cut scores and corresponding 2018 impact data. The panelists recommended no adjustments to cut scores for grades 5, 7, 8, and 11. They recommended small adjustments as follows: in grade 3 performance level 4, from scale score 1251 to 1254; in grade 4 performance level 4, 1258 to 1259; and in grade 6 performance level 4, 1253 to 1251. Resulting differences in impact data are modest. The upward adjustments to the performance level 4 cut score in grades 3 and 4 decreased the percentages of students in that level by 5.5% (grade 3) and 3.0% (in grade 4); the downward adjustment to the performance level 4 cut score in grade 6 increased the percentage of students in that level by 2.6%. (In each case, the percentages of students changed in performance level 3 by the amount of change in performance level 4.)

Table 3-1. Recommended ELA Cut Scores after Cut Score Review

Grade	Performance Level	Original Cut Score [1]	2018 Impact Data (%)	Adjusted Cut Score	Resulting Impact Data (%)
11	PL 4	1255	20.3	--	--
	PL 3	1240	40.9	--	--
	PL 2	1236	15.0	--	--
	PL 1	--	23.8	--	--
8	PL 4	1250	21.4	--	--
	PL 3	1238	27.6	--	--
	PL 2	1230	27.8	--	--
	PL 1	--	23.2	--	--
7	PL 4	1255	21.2	--	--
	PL 3	1240	34.6	--	--
	PL 2	1236	16.0	--	--
	PL 1	--	28.2	--	--
6	PL 4	1253	15.9	1251	18.5
	PL 3	1237	39.9	--	37.3
	PL 2	1231	27.6	--	--
	PL 1	--	16.6	--	--
5	PL 4	1256	14.6	--	--
	PL 3	1240	34.6	--	--
	PL 2	1232	24.3	--	--
	PL 1	--	26.5	--	--
4	PL 4	1258	13.6	1259	10.6
	PL 3	1240	33.6	--	36.5
	PL 2	1234	18.4	--	--
	PL 1	--	34.4	--	--
3	PL 4	1251	24.7	1254	19.2
	PL 3	1240	26.0	--	31.5
	PL 2	1234	12.5	--	--
	PL 1	--	36.8	--	--

[1] Before the cut score review, after vertical articulation of the performance level 3 cut scores on July 20, 2018

Chapter 4. Tasks Completed After the Cut Score Review Meeting

Upon conclusion of the cut score review meeting, MSAA State Partners and Measured Progress completed several follow-up tasks as part of the Standards Validation process:

- Vertical articulation of the performance level 2 and 4 cut scores for mathematics (only), following the procedures for vertical articulation of performance level 3 cut score prior to the cut score review meeting (see section *2.1 Vertical Articulation of the ELA and Mathematics Cut Scores* in this report)
- Review of the recommended cut scores by each partner state's Superintendent/Board of Education for approval
- An ELA PLD Review meeting with stakeholders from various states' schools and districts
- Final approval of all MSAA vertically articulated and validated cut scores
- Preparation of this report

4.1 Vertical Articulation of Mathematics Performance Level 2 and 4 Cut Scores

After the cut score review meeting, Measured Progress Psychometrics, Content Development, and Client Services staff reviewed the vertical articulation of the performance level 2 and 4 cut scores for mathematics (only). They replicated the procedures for articulating the performance level 3 scores; see section 2.1 Vertical Articulation of the ELA and Mathematics Cut Scores.

Following initial approval of the performance level 3 cut scores by the MSAA Psychometric Subcommittee, Measured Progress conducted a vertical articulation review of the cut scores for mathematics performance levels 2 and 4 to determine if any adjustment recommendations were needed. Following the procedures described in section 2.1, the team then reviewed the impact data for each grade in relation to each other. Measured Progress recommended the following to MSAA (as displayed in Table 4-1):

- For performance level 2, adjusting the cut scores in grades 3, 4, 5, 6, and 11 by 1 scale score point and in grade 7 by 2 scale score points achieves vertical articulation.
- For performance level 4, adjusting the cut scores in grade 11 by 1 scale score point and in grades 5, 6, and 8 by 2 scale score points achieves vertical articulation adjustments.

Table 4-1. Original and Vertically Articulated Mathematics Cut Scores and Impact Data

Grade	Performance Level	Original Cut Score [1]	2017 Impact Data by PL (%)	2017 Impact Data for PLs 3 & 4 (%)	Adjusted Cut Score	Resulting Impact Data (%)	2017 Impact Data for PLs 3 & 4 (%)	Adjustments made for Vertical Articulation
11	PL 4	1249	19.20	43.03	1250	17.92	--	Higher by 1
	PL 3	1240	23.83		--	25.11		
	PL 2	1234	26.81		1235	24.18		Higher by 1
	PL 1	--	30.16		--	32.79		
8	PL 4	1249	23.06	45.90	1251	18.78	--	Higher by 2
	PL 3	1240	22.84		--	27.12		
	PL 2	1234	23.41		--	--		
	PL 1	--	30.69		--	--		
7	PL 4	1254	17.67	47.00	--	--	--	
	PL 3	1240	29.33		--	--		
	PL 2	1232	32.36		1234	25.16		Higher by 2
	PL 1	--	20.64		--	27.84		
6	PL 4	1249	21.67	40.54	1251	17.53	46.50	Higher by 2
	PL 3	1240	18.87		1239	28.97		Lower by 1
	PL 2	1234	23.11		1233	21.49		Lower by 1
	PL 1	--	36.35		--	32.01		
5	PL 4	1255	13.60	46.54	1253	16.84	--	Lower by 2
	PL 3	1240	32.94		--	29.70		
	PL 2	1231	29.72		1232	24.08		Higher by 1
	PL 1	--	23.74		--	29.38		
4	PL 4	1251	17.40	41.01	--	17.40	44.13	
	PL 3	1240	23.61		1239	26.73		Lower by 1
	PL 2	1233	23.59		1232	24.83		Lower by 1
	PL 1	--	35.40		--	31.04		
3	PL 4	1254	19.31	49.46	--	19.31	44.66	
	PL 3	1240	30.15		1242	25.35		Higher by 2
	PL 2	1236	16.54		1235	22.55		Lower by 1
	PL 1	--	34.00		--	32.79		

The final proposed cut scores display smooth articulated results.

4.2 Vertical Articulation of ELA Performance Levels 2 and 4

The Psychometric Subcommittee agreed to a recommendation to postpone vertical articulation of the ELA performance level 2 and 4 cut scores. The TAC members who monitored the cut score review concurred. The rationale for this postponement is as follows:

- Based on past experience in other grade level and alternate assessment programs, open-response writing prompt score level thresholds are expected to remain in the upper end of the theta/scale score scale for the next year or two.
- As students are exposed to instruction and practice in direct writing in coming years, writing performance is expected to improve. Writing instruction is likely to be implemented in uneven ways across states. writing improvement is expected to be reflected as downward movement of the score level locations in unpredictable ways,

indicating open-response writing prompts that through increased instruction may become easier for students than they were in 2018.

- If ELA performance level 2 and 4 cut scores are adjusted now to articulate the current impact data, and as prompt locations move around in the next couple of years, the alignment among the performance levels, PLDs, and open-response writing prompt score level threshold locations are likely to disarticulate.

Consequently, it was determined by MSAA that it is reasonable and wise to hold off on articulating the ELA performance level 2 and 4 cut scores until MSAA writing performance stabilizes—or at least reveals how it will evolve—in the next couple of years.

4.3 Review of the Recommended Performance Level Cut Scores for ELA and Mathematics

Measured Progress developed four documents for each state to use to gain approval of the cut scores for ELA and mathematics. These were:

- **ELA Cut Score Review Summary**: This document contained the final adjusted cut scores from both the vertical articulation process and Cut Score Review meeting, as well as a brief overview of the cut score review process for each state to share with their Superintendents/Boards of Education (see Appendix E).
- **ELA Cut Score Review Summary with Content Based Rationales**: This document contained the final adjusted cut scores along with the specific content based rationales that were determined by the panelists during the Cut Score Review meeting (see Appendix F). This document was meant as a reference document for states to use should a question get raised by the Superintendent/Board of Education about an adjustment.
- **Mathematics Vertical Articulation Results Summary**: This document contained the final adjusted cut scores with figures showing the articulation lines and stacked bar charts showing impact data percentages before and after the vertical articulation process, as well as a brief overview of the vertical articulation process for each state to share with their Superintendents/Boards of Education (see Appendix G).
- **Mathematics Vertical Articulation Results Summary with Content Based Rationales**: This document contained the final adjusted cut scores from the vertical articulation process along with the specific content-based rationales that were determined (see Appendix DD of the Vertical Articulation Report). This document was

meant as a reference document for states to use should a question get raised by the Superintendent/Board of Education about an adjustment.

These documents were provided to the State Partners on July 26, 2018 (mathematics), and July 27, 2018 (ELA). Each state brought the ELA and mathematics document before their Superintendents/Boards of Education for official approval of the cut score adjustments. States individually sought approval of the cut score from their respective Superintendents/Boards of Education.

4.4 ELA PLD Review Meeting Overview

4.4.1 Preparation for the ELA PLD Review Meeting

Materials for the ELA PLD Review meeting were prepared by Measured Progress. An opening session PowerPoint presentation was developed to orient panelists to the MSAA design and administration, provide background information about the open-response writing prompt considerations and emphasis, and set the expectations of their review work on the ELA PLDs. A copy of the presentation is included in Appendix H The facilitators attended an initial training session, led by a Measured Progress Special Education Director, prior to the ELA PLD Review meeting. The purpose of the training was to prepare the facilitators for the panel activities and to ensure consistency in the implementation of PLD review process. A process document was created for group facilitators to refer to while working through each step of the PLD review process. A copy of the process document is included in Appendix I. In addition, the following materials were assembled for presentation to the panelists at the ELA PLD Review meeting:

- meeting agenda
- nondisclosure form
- ELA PLDs (the expanded ELA PLDs as used during the Cut Score Review meeting)
- writing definition and emphasis
- open-response writing prompt rubrics
- open-response writing prompts and student sample booklets
- ELA Blueprints
- MSAA 2016–17 Score Report Interpretation Guide
- MSAA 2017–18 Test Administration Manual
- evaluation survey

4.4.2 Panelist Selection

MSAA selected panelists for the ELA PLD Review meeting. Diverse panelists were chosen from the following criteria:

- experience with special education
- experience with general education
- experience as an administrator
- experience with special populations
- other demographic factors (e.g., race, geographic location, etc.)

Tables 4-2 to 4-4 provide the makeup of each panel.

Table 4-2. 2018 MSAA ELA PLD Review Meeting: State of Panelist by Grade Group

States	English Language Arts		
	3 5	6 8	11
Arkansas			
Arizona		1	1
Guam/CNMI	1	1	1
Maryland	1	2	3
Maine	1	2	1
Montana			
South Dakota	2		
Tennessee	1	1	1
US Virgin Islands			
Washington D.C.	2	1	1
	8	8	8

*Table 4-3. 2018 MSAA ELA PLD Review Meeting: Educator Type by Grade Group**

Educator Type	English Language Arts		
	3 5	6 8	11
Administrator	2	2	2
General Educator	5	4	5
Special Educator	6	5	5
Special Populations			3
Missing		1	

*Some individuals indicated more than one area of expertise/experience

Table 4-4. 2018 MSAA ELA PLD Review Meeting: Years of Experience by Grade Group

Years	English Language Arts		
	3 5	6 8	11
0–5	1		
5–10	1		1
10–15	1		1
15+	3	6	5
Missing	2	2	1
	8	8	8

Table 4-5. 2018 MSAA ELA PLD Review Meeting: Other Demographics by Grade Group

Other Demographic Information	English Language Arts		
	3 5	6 8	11
Setting			
Urban/Suburban	2	3	3
Rural	6	3	4
Missing		2	1
Race/Ethnicity			
Asian		1	
Black or African American	1	2	1
Native Hawaiian or Other Pacific Islander	1		1
White	5	3	6
Missing	1	2	
Gender			
Female	8	6	8
Missing		2	

Each panel consisted of 8 panelists. A list of the panelists by grade span is included in Appendix B.

4.5 Overview of the ELA PLD Review Meeting Process

The ELA PLD Review meeting took place August 9, 2018. The ELA PLD Review meeting began with an orientation training for all panelists. The purpose of the orientation was to ensure that all panelists received the same information about the MSAA, the goals of the ELA PLD Review meeting, and the expectations of panelists. Following introductions, Measured Progress provided an overview of the assessment, including administration, scoring, and participation criteria for the MSAA. Panelists were provided information related to the open-response writing prompts, and an overview of the PLD review process. Panelists were given an opportunity to ask questions.

Once the orientation was complete, each panel convened in a breakout room, where the panelists received more detailed information and orientation to the materials from their facilitator and completed the PLD review activities. The MSAA State Partner attendees floated among each of the panels to observe the process and answer questions related to administration and/or policy. Panelists were provided four guiding questions:

1. Does the open-response writing prompt information contained in the PLD for each level accurately account for what the open-response writing prompt is designed to measure?
2. Is the language clear and reflective of information that is understandable for administrators, teachers, and parents/guardians?
3. How might the open-response writing prompt measures and emphasis best be communicated to the field?
4. What avenues should be used to communicate information about the open-response writing prompts to administrators, teachers, and parents/guardians?

In each of the breakout groups, introductions occurred and the non-disclosure agreements were collected. The panelists were then provided with the open-response writing prompts and corresponding rubric for each level, as well as the ELA blueprint for reference. Students' responses were provided in the Open Response Writing Prompt and a Student Sample Booklet for each grade level. The student samples were acquired from a sampling of scoring anchor papers across the scoring rubric ranges. The purpose of this activity was to establish an understanding of the open-response writing prompts and what writing prompt evidence might look like from the students who take the test. Panelists were given time to reflect and ask questions. The panelists were then provided the PLD for a grade, and guided through the open-response writing prompt expectation

information in each performance level. Panelists were given time to reflect and ask questions. The four guiding questions became the framework for the discussion around the clarity, descriptiveness, and consistency with the MSAA writing definitions and emphasis of the ELA PLDs. Panelists were encouraged to provide consensus-based suggestions to MSAA involving possible edits to the ELA PLDs and suggestions on means of communicating the information to the field (including administrators, teachers, and parents/guardians).

In response to guiding questions 1 and 2, the panels agreed with the open-response writing prompt expectation information in performance levels 1 through 3. One edit was suggested in performance level 4. The consensus of all three panel groups indicated an approval to add a word to the performance description at the fourth level so that it read "…overall command…" instead of just "…command…" The panelists also responded to guiding questions 3 and 4 with suggestions that would be helpful to the field in understanding the open-response writing prompt measures and emphasis and the best avenues to distribute that information. Feedback provided from each of the panels to the guiding questions can be found in Appendix J.

Following the panel discussions, each panelist was asked to complete an evaluation survey. The evaluation provided panelists the opportunity to share their feedback on the training and overall process. Evaluation results for each panel can be found in Appendix K.

After completion of the panel work, Measured Progress facilitators and MSAA State Partners in attendance met to reconcile the panelist suggestion to the ELA PLDs and to review the suggestions provided by the panelists on distribution of information to the field about the open-response writing prompt measures and emphasis. Overall feedback on the meeting preparation and activities during the ELA PLD Review meeting was also discussed. MSAA State Partners in attendance were from a variety of states and were members of the MSAA Psychometric, Item Development, and/or Scoring Subcommittees. The MSAA State Partner attendees agreed to the recommended edit to performance level 4 for all grades. In addition, they requested minor updates to the PLDs for all grades for both ELA and mathematics for consistency and clarity (i.e., adding an asterisk and clarity statement to the header of performance levels 2, 3, and 4 and change "he/she" to "the student"). The MSAA State Partner attendees approved the ELA PLDs with the edits requested.

4.6 Final Approval of All Cut Scores

Measured Progress received final approval of all cut scores from MSAA on August 16, 2018. Each state individually obtained approval, which constituted collective approval by MSAA.

4.7 Preparation of the Standards Validation Process Report

This Standards Validation Report documents the procedures that were taken as part of the standards validation process. The five-step procedure used to develop the 2018 ELA cut scores are outlined within the document.

References

Ramist, L, Lewis, C, & McCamley-Jenkins, L (1994). *Student group differences in predicting college grades: Sex, language, and ethnic groups* (College Board Research Report 93-1). New York: College Board.

U.S. Department of Education. (2009). Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001. Washington, DC: U.S. Department of Education Office of Elementary and Secondary Education. Retrieved June 10, 2010, from www2.ed.gov/policy/elsec/guid/saaprguidance.pdf.

Appendices

Appendix A—Vertical Articulation Report

Vertical Articulation Report

Prepared by: Measured Progress

October 2018

Contents

MSAA English Language Arts (ELA) Vertical Articulation Process, Results, and Recommendations 4
Meetings to Agree on the Vertical Articulation Approach and Grades and Content Areas of Concern..... 4
Vertical Articulation Approach Meeting (May 4, 2018) 4
Standard Errors for the Performance-Level Classification Percentages 5
Cohort Analysis Approach to Vertical Articulation..... 5
Single-Year Approach to Vertical Articulation 6
Vertical Articulation Process Clarification Meeting (June 15, 2018)..... 7
Vertical Articulation Process, Recommendations, Results, and Approvals 7
ELA Performance Levels 3 and 4 8
Mathematics Performance Levels 3 and 4 9
Mathematics Performance Levels 2 and 4 10
Recommended Cut Scores 12
Content Based Rationales for the Adjusted Cut Scores 13
Vertical Articulation Report Appendices..... 15

MSAA English Language Arts (ELA) Vertical Articulation Process, Results, and Recommendations

In this report, Measured Progress describes the several steps we took to develop recommendations from conducting vertical articulation of the 2016–2017 MSAA results. This work was in preparation for the ELA Cut Score Review using 2018 MSAA data, which occurred on July 24, 2018. The steps in the vertical articulation process were:

- Meetings with the Psychometric Subcommittee and Technical Advisory Committee (TAC):
  - May 4, 2018, to specify the vertical articulation approach and grades and content areas that require focused attention
  - A follow-up clarification meeting on June 15, 2018
- Working meetings at Measured Progress in June and July 2018 to complete the vertical articulation process and develop recommendations
- Meetings with the Psychometric Subcommittee to review the vertical articulation results in ELA and mathematics:
  - July 20, 2018, to review the vertical articulation results for ELA and mathematics on performance level 3 and 4 combined
  - July 25, 2018, and July 26, 2018, to review the vertical articulation results for mathematics on performance levels 2 and 4

We provide this report to describe the vertical articulation process we followed, recommendations, results, and approvals provided by the Psychometric Subcommittee.

Meetings to Agree on the Vertical Articulation Approach and Grades and Content Areas of Concern

Vertical Articulation Approach Meeting (May 4, 2018)

In the May 4, 2018, meeting with the Psychometric Subcommittee and TAC, Measured Progress reviewed the background and purposes of vertically articulating the current ELA and mathematics cut scores. The purposes were to review (a) the ELA standards, to prepare for a cut score review after linking the writing prompt scores onto the existing ELA scale; (b) MSAA's cross-grade growth expectations, to define the desired vertical articulation pattern for ELA and mathematics; and (c) the amount of disarticulation that would require adjusting a cut score for ELA and mathematics (e.g., the difference in percentages of examinees at/above Level 3 in two adjacent years).

The Psychometric Subcommittee and TAC raised two issues in preparation for this meeting:

- The size of the standard errors associated with the performance-level classification percentages.
- The proposed "cohort approach" to pursuing vertical articulation. This is the degree of articulation that could be achieved if cut scores were adjusted in 2016–2017 cohorts by

making the percentages of positive and negative performance-level changers about equal.

Measured Progress proposed in the meeting to use a “single-year approach” and decision criteria for vertically articulating standards in the 2016–2017 data, prior to conducting an ELA cut score review in the 2018 data. The Psychometric Subcommittee and the TAC agreed to move forward with this plan.

Standard Errors for the Performance-Level Classification Percentages

One of the comments we received feedback on from the TAC was in regard to the performance-level classification percentages that were presented in our May 4, 2018, meeting. After noticing the differences in the percentages for grade 6 relative to other grades, the TAC asked what the standard errors were for those percentages so that they could more reliably interpret the observed differences.

We calculated the standard errors using two approaches: (1) the usual standard error for a binomial random variable: the square root of (pq/n), where p is the performance-level classification percentage, $q = 1 – p$, and n is the total number of students being classified; and (2) a bootstrap standard error. The two estimators gave essentially the same results.

- For the 2015–2016 mathematics tests, the standard errors (in terms of percentages) ranged from 0.5 percentage points to 1.1 percentage points, with a mean of 0.8. (For example, a classification rate of 37% would have a 95% confidence interval of 35% to 39%, if its standard error was on the high side, i.e., 1%.)
- For the 2016–2017 mathematics tests, the standard errors ranged from 0.6 percentage points to 0.8 percentage points, with a mean of 0.7.
- For the 2015–2016 ELA tests, the standard errors ranged from 0.5 percentage points to 0.9 percentage points, with a mean of 0.7.
- For the 2016–2017 ELA tests, the standard errors ranged from 0.5 percentage points to 0.8 percentage points, with a mean of 0.7.

These results indicate that the differences between grade 6 and the other grades are not due simply to the standard errors of the classification percentages. A detailed listing of standard errors of percentages of students in each performance level for all grades in 2015–2016 and 2016–2017 are provided in Appendix AA.

Cohort Analysis Approach to Vertical Articulation

This approach addresses whether there is statistical evidence that the cut scores for grade 6 ELA (ELA06) and mathematics (MAT06) and for grade 8 ELA (ELA08) are too stringent, beyond a reasonable doubt. There is no doubt that the standards seem harder to achieve for ELA06, MAT06, and ELA08. If there is a consistent pattern across the other grades that is missing only from ELA06, MAT06, and ELA08, then that would be the strongest statistical argument. To address this question, we calculated the conditional probabilities of moving to a year 2 Proficient level, conditional on a year 1 Proficient level. To make this still easier to interpret, we focused only on cut score 2. Thus, we looked at the probability that a student remained Proficient in year 2, given they were at a Proficient level in year 1, and the probability a student remained non-proficient in year 2, given they were in a non-proficient level in year 1. Here is what we found:

Grade 6 ELA: Most (but not all) the evidence suggests that the grade 6 ELA Proficient cut score may be more stringent than the Proficient cut score in other grades. We also investigated the size of this difference in stringency, i.e., whether the difference is large or small, as measured by how much of a change in the scaled score cut would result in ameliorating the statistical evidence. We determined that, at most, a change in the scaled score cut of 1 point (from 1240 to 1239) would be required, thus indicating that the size of difference in stringency is small.

Grade 8 ELA: None of the evidence from the probability calculation described above supported the hypothesis that the proficient cut for ELA08 is more stringent than for other tests. Indeed, the cohort analysis indicated that ELA07 students moving up to ELA08 exhibit performance-level probability changes that are very similar to ELA03 students moving up to ELA04.

Grade 6 Mathematics: Only about half the evidence suggests that the grade 6 mathematics Proficient cut score may be more stringent than the Proficient cut score in other grades. Indeed, the cohort analysis indicated that MAT06 students moving to MAT07 seem no different from students moving from MAT07 to MAT08, and only slightly different from students moving from MAT04 to MAT05.

Detailed descriptions of the analyses, interpretations, and logic that support these findings appear in Appendix BB.

Single-Year Approach to Vertical Articulation

In the May 4, 2018, meeting, we defined vertical articulation for this approach as the degree to which percentages of examinees in 2017 were roughly equivalent in each performance level across grades 3 through 8. The goal is to consider adjusting cut scores if evidence indicates disarticulation in a grade for one or more performance levels.

Measured Progress proposed the following in the May 4, 2018, meeting:

- Cross-grade articulation in ELA and mathematics in 2017 was flat; that is, percentages of students at and above Proficient are roughly equivalent.
- Some disarticulation is indicated in ELA, especially for performance levels 2 and 3.
- Disarticulation is indicated in mathematics for all performance levels.

Appendix CC contains the slides and visual displays from the May 4, 2018, meeting that provide the basis for these claims.

The question is how much grade-to-grade difference in the percentages at a performance level warrants consideration of adjusting cut scores to approximate vertical articulation. There is no statistical test or widely accepted criterion for making this judgment. It is a matter of judgment, with due consideration for policy concerns (e.g., How much of a difference in adjacent percentages in a performance level can teachers and principals tolerate?) and psychometric concerns (e.g., How much were cut scores adjusted as part of vertical articulation in the 2015 standard setting?). We proposed the following, applicable to both ELA and mathematics:

- Examine the impact of adjusting cut scores for Proficient and above only if a grade-to-grade difference is greater than 5%. (We had originally proposed 10%.)
- Consider the difference in the cut scores on the theta scale, where a theta difference of, say, 0.3 would be considered large.

- Make the adjustment to these cut scores if vertical articulation is improved to within 5% differences in adjacent grades and the cut score on the theta scale is less than 0.3.

These two criteria are arbitrary, of course. They also are consistent with other judgment-guiding rules of thumb used in evaluating year-to-year equating results (i.e., from the Maryland TAC in the 1990s) and consistent with the vertical articulation adjustments made in 2015.

As we consider this proposal, we might keep in mind that this effort at vertical articulation may become irrelevant in ELA when we link the writing prompt scores to the current ELA scale and apply cut scores to the new, enhanced ELA scale.

Vertical Articulation Process Clarification Meeting (June 15, 2018)

In this meeting, Measured Progress sought clarification and agreement with the Psychometric Subcommittee and the TAC on the process and criteria for completing vertical articulation and presenting recommendations to MSAA. The Psychometric Subcommittee and the TAC agreed to the following:

- First, articulate the percentages of students in levels 3 and 4 combined (referred to as "impact data" in the remainder of this report for simplicity).
- Then try to improve the articulation for levels 4 and 2 without upsetting the articulation of the impact data.
- Pay particular attention to ELA grades 6 and 8 and mathematics grade 6.
- Try to articulate cut scores so that impact data differences for any pair of grades is less than 5%.
- After proposing cut score adjustments to improve articulation of the impact data, review the alignment between the item response demands and performance level descriptions to determine if a reasonable content based rationale exists to support the adjustment.

Vertical Articulation Process, Recommendations, Results, and Approvals

A Measured Progress team of Content Development, Psychometrics, and Program Management staff completed the vertical articulation process to review the current MSAA cut scores and 2016–2017 performance data (i.e., percentages of all MSAA students in performance levels 1, 2, 3, and 4). The team reviewed the inverse cumulative percentages of students in each performance level in grades 3–8 and 11 in ELA and mathematics and the locations in each distribution of the levels 2, 3, and 4 cut scores. The percentages are based on all students who had valid and scorable response strings and received total test scale scores. Each grade and content area frequency distribution starts at scale score 1200 and ends at 1290. All level 3 cut scores are pegged to 1240; level 2 cut scores range between 1230 (ELA grade 8) to 1236 (ELA grades 7 and 11, mathematics grade 3); level 4 cut scores range between 1249 (mathematics grades 8 and 11) and 1259 (mathematics grade 8). The team focused on the areas of concern—ELA grades 6 and 8 and mathematics grade 6—and viewed all grade level percentages together to consider the degree of cross-grade articulation.

The results for ELA and mathematics, as outlined in the subsequent pages, were reviewed with the MSAA Psychometric subcommittee in two phases. The first phase was the work completed on performance levels 3 and 4 combined. During the meeting on July 20, 2018, the Psychometric subcommittee approved the adjustment recommendations to performance level 3 for both ELA and mathematics and approved postponing articulation of performance levels 2 and 4 for ELA due to the inclusion of the writing prompt on the ELA scale and its anticipated effect on performance over the next couple of years. The second phase was the articulation of performance levels 2 and 4 for mathematics. During the meeting on July 25 and 26, 2018, the Psychometric subcommittee approved the adjustment recommendations to performance levels 2 and 4 for mathematics. Subsequent to the vertical articulation process and ELA Cut Score Review, the ELA and mathematics cut scores were presented to each state partner's Board of Education/Superintendent for formal adoption across MSAA.

ELA Performance Levels 3 and 4

The team began by reviewing the impact data for ELA grades 6 and 8 in relation to the other grades, using the data in Table 1. We then reviewed impact data by adjusting the cut scores in grades 6 and 8 by 1, 2, and 3 scale score points.

Table 1: ELA Impact Data (in Percentages) and Proposed Cut Score Adjustments (in Scale Score Points)

Grades						
3	4	5	6	7	8	11
Performance Levels 3 and 4						
48.16	44.91	46.24	34.23	46.62	38.36	48.3
Proposed Level 3 Cut Score Adjustments (in Scale Scores)						
--	--	--	Lower by 1	--	Lower by 1	--
48.16	44.91	46.24	39.48	46.62	42.84	48.3
--	--	--	Lower by 2	--	Lower by 2	--
48.16	44.91	46.24	40.38	46.62	45.28	48.3
--	--	--	Lower by 3	--	Lower by 2	--
48.16	44.91	46.24	43.91	46.62	45.28	48.3

Adjusting the grade 8 cut score by 2 scale score points and the grade 6 cut score by 3 points achieves vertical articulation. As Table 1 indicates, the largest difference in impact data is 4.25% (grades 3 and 8). Figure 1 shows the articulation lines for the three proposed sets of cut score adjustments in Table 1. The final proposed cut scores display smooth articulated results, especially in comparison to the original cut scores and other proposed adjustments.

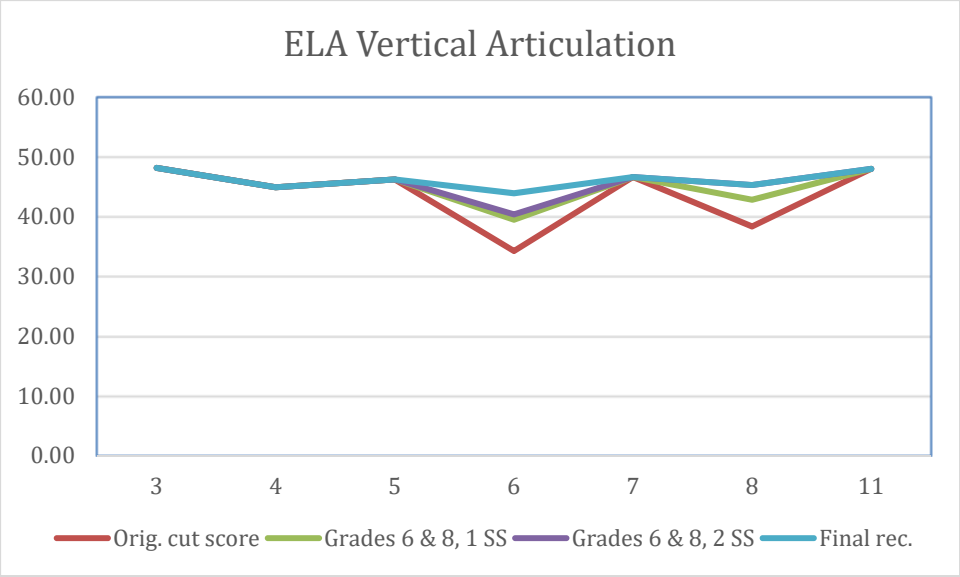


Figure 1. ELA vertical articulation lines for the original cut scores and all proposed adjustments, levels 3 and 4.

Mathematics Performance Levels 3 and 4

The team then reviewed the impact data for mathematics grade 6 in relation to the other grades, using the data in Table 2. We then reviewed impact data by adjusting the cut scores in grades 4 and 6 by 1 scale score point and grade 3 by 2 scale score points.

Table 2: Mathematics Impact Data (in Percentages) and Proposed Cut Score Adjustments in Scale Score Points

Grades						
3	4	5	6	7	8	11
Performance Levels 3 and 4						
49.46	41.01	46.54	40.54	47.00	45.90	43.03
Proposed Level 3 Cut Score Adjustments (in Scale Scores)						
--	--	--	Lower by 1	--	--	--
49.46	41.01	46.54	46.50	47.00	45.90	43.03
Higher by 1	Lower by 1	--	Lower by 1	--	--	--
49.46	44.13	46.54	46.50	47.00	45.90	43.03
Higher by 2	Lower by 1	--	Lower by 1	--	--	--
44.66	44.13	46.54	46.50	47.00	45.90	43.03

Adjusting the cut scores in grades 4 and 6 by 1 scale score point and in grade 3 by 2 scale score points achieves vertical articulation. As Table 2 indicates, the largest difference in impact data is 3.97% (grades 7 and 11). Figure 2 shows the articulation lines for the three proposed sets of cut score adjustments in Table 2. The final proposed cut scores display smooth articulated results, especially in comparison to the original cut scores and other proposed adjustments.

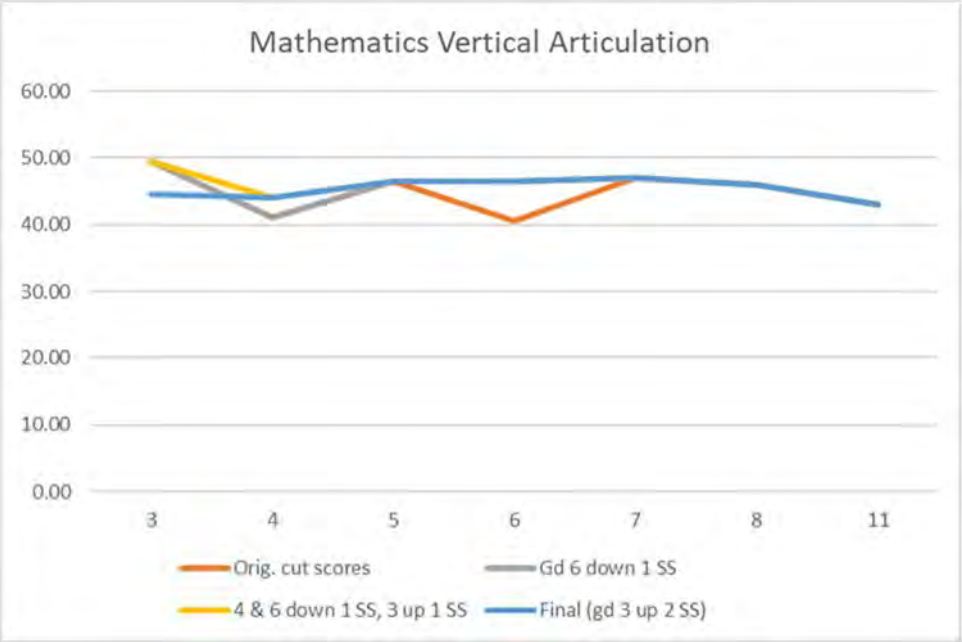


Figure 2. Mathematics vertical articulation lines for the original cut scores and all proposed adjustments, levels 3 and 4.

Mathematics Performance Levels 2 and 4

Following the work on the combination of performance levels 3 and 4 and initial approval by the MSAA Psychometric subcommittee, Measured Progress conducted the vertical articulation review to the cut scores for performance levels 2 and 4. The goal was to determine if any adjustment recommendations were needed. The team reviewed the impact data for each grade in relation to each other, resulting the data presented in Table 3.

Table 3: Mathematics Impact Data (in Percentages) and Proposed Cut Score Adjustments in Scale Score Points for Performance Levels 2 and 4

Grades						
3	4	5	6	7	8	11
Performance Level 2						
16.54	23.59	29.72	23.11	32.36	23.41	26.81
Proposed Level 2 Cut Score Adjustments (in Scale Scores)						
Lower by 1	Lower by 1	Higher by 1	Lower by 1	Higher by 2	--	Higher by 1
22.55	24.83	24.08	21.49	25.16	23.41	24.18
Grades						
3	4	5	6	7	8	11
Performance Level 4						
19.31	17.40	13.60	21.67	17.67	23.06	19.20
Proposed Level 4 Cut Score Adjustments (in Scale Scores)						
--	--	Lower by 2	Higher by 2	--	Higher by 2	Higher by 1
19.31	17.40	16.84	17.53	17.67	18.78	17.92

For performance level 2, adjusting the cut scores in grades 3, 4, 5, 6, and 11 by 1 scale score point and in grade 7 by 2 scale score points achieves vertical articulation at that performance level. For performance level 4, adjusting the cut scores in grade 11 by 1 scale score point and in grades 5, 6, and 8 by 2 scale score points achieves vertical articulation at that performance level. The largest difference in impact data between each of the performance levels are: performance level 1 (4.95%), performance level 2 (3.67%), performance level 3 (4.59%), and performance level 4 (2.47%). Figure 3 shows the articulation lines for the proposed sets of cut score adjustments in Table 3. The final proposed cut scores display smooth articulated results, especially in comparison to the original cut scores.

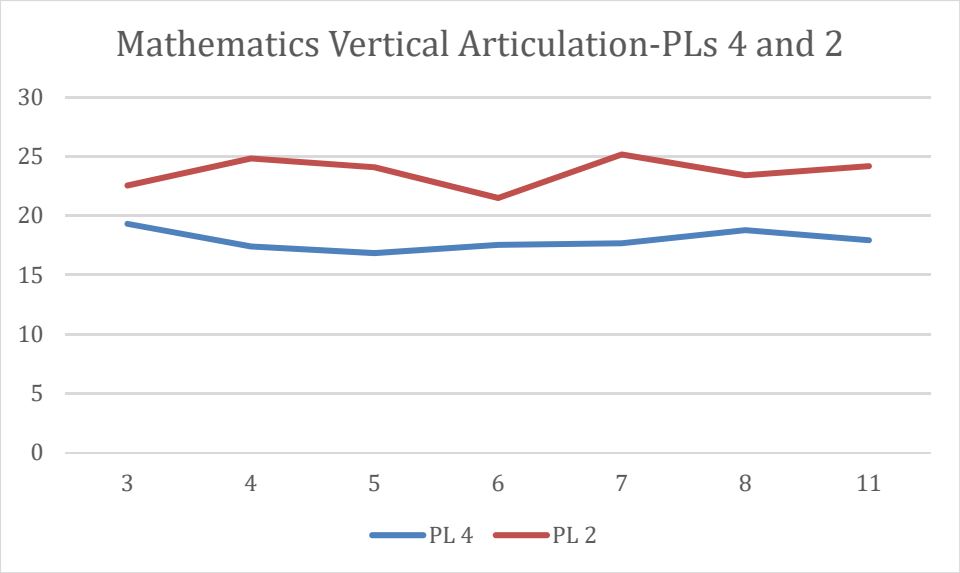


Figure 3. Mathematics vertical articulation lines for performance levels 2 and 4.

Recommended Cut Scores

We achieved the best vertical articulation results by making the adjustments labeled "Final" in Figures 1 and 2 and labeled "PL 4" and "PL 2" in Figure 3. The final, proposed articulated cut scores for performance level 3 for ELA and mathematics appear in Table 4. The final, proposed articulated cut scores for performance levels 2 and 4 for mathematics appear in Table 5.

Table 4. Final Articulated Cut Scores (Scale Scores) Performance Level 3

Grade	3	4	5	6	7	8	11
ELA	1240	1240	1240	1237	1240	1238	1240
Mathematics	1242	1239	1240	1239	1240	1240	1240

Table 5. Final Articulated Cut Scores (Scale Scores) Performance Levels 2 and 4

Grade	3	4	5	6	7	8	11
Mathematics-Level 2	1235	1232	1232	1233	1234	1234	1235
Mathematics-Level 4	1254	1251	1253	1251	1254	1251	1250

Content Based Rationales for the Adjusted Cut Scores

The Measured Progress Content Development Manager and Special Education Director conducted a review of item-Performance Level Descriptors (PLD) alignment[1] to determine if the response demands of items moved to the adjacent performance level after cut score adjustments are reasonably aligned with the knowledge and skill requirements of the corresponding PLD. Specific details of the item-PLD alignment analysis appear in Appendix DD.

Performance Level 3 Adjustments:

ELA grade 6: The item located just above the original cut score, 1240, aligns with the level 3 PLD. After adjusting the cut score to 1237, one item's response demands align with the level 3 PLD. A second item, located just above the articulated cut score, aligns most closely with the level 2 PLD, its associated performance level prior to the cut score adjustment. Items like this latter item are written to brief, straightforward text.

ELA grade 8: After adjusting the cut score to 1238, one item's response demands align with the level 3 PLD and one with the level 2 PLD, its associated performance level prior to the cut score adjustment. This latter item requires students to identify an idea relevant to a claim in a very short text, which is consistent with the level 2 PLD.

Mathematics grade 3: After adjusting the cut score to 1242, the items that are now located in level 2 are reasonably aligned to the level 2 PLD.

Mathematics grade 4: After adjusting the cut score to 1239, the item now located in level 3 aligns to the borderline of the level 3 PLD because it is of moderate complexity.

Mathematics grade 6: After adjusting the cut score to 1239, one item's response demands align with the level 3 PLD.

Performance Level 2 Adjustments:

Mathematics grade 5: After adjusting the cut score to 1232, the items that are now located in level 1 are reasonably aligned to the level 1 PLD.

Mathematics grade 6: After adjusting the cut score to 1233, the items that are now located in level 2 are reasonably aligned to the level 2 PLD.

Mathematics grade 7: After adjusting the cut score to 1234, one item that is now located in level 1 is reasonable aligned to the level 1 PLD. Three of the items now located in level 1 align to the borderline of levels 1 and 2 PLDs as these items all assess surface area.

Mathematics grade 11: After adjusting the cut score to 1235, the item that is now located in level 1 is reasonably aligned to the level 1 PLD.

---

[1] Ferrara, S. (2017 April 28). Aligning item response demands with knowledge and skill requirements in achievement level descriptors: An approach to achieving full alignment and engineering cut scores. In D. Lewis (Chair), *Engineered cut scores: Aligning standard setting methodology with contemporary assessment design principles*. Coordinated session conducted at the annual meeting of the National Council on Measurement in Education, San Antonio, TX.

Performance Level 4 Adjustments:

> **Mathematics grade 5:** After adjusting the cut score to 1253, two items that are now located in level 4 are reasonably aligned to the level 4 PLD. One of the items now located in level 4 aligns to the borderline of levels 3 and 4 PLDs as it is of moderate to high task complexity.
>
> **Mathematics grade 8:** After adjusting the cut score to 1251, four items that are now located in level 3 are reasonably aligned to the level 3 PLD. One item now located in level 3 aligns to the borderline of levels 3 and 4 PLDs as it is of moderate to high task complexity. One item now located in level 3 aligns with the level 4 PLD as it is of higher task complexity.

For performance level 3, the results indicate that after articulation in both ELA grade 6 and ELA grade 8 one item in level 3 aligns most closely with the level 2 PLD. For mathematics grade 3, four of the five items located in level 2 are well aligned with the level 2 PLD. However, the response demands of one item, 110959A, are aligned with the lower end of the level 3 PLD. For performance level 2, the results indicate that after articulation in mathematics grade 7, three of the four items located in level 1 are aligned with the lower end of the level 2 PLD and the upper end of the level 1 PLD. For performance level 4, the results indicate that after articulation in both mathematics grades 5 and 8 some items align to the adjusted performance level PLD; however, one item in grade 5 and one item in grade 8 align at the borderline with characteristics from both adjacent performance level PLDs, and one grade 8 item in level 3 aligns most closely with the level 4 PLD. This is not a surprising result, as item alignment misclassifications are highest around cut scores.[2] Two factors explain this common misalignment: the (a) standard errors of IRT item *b*-value estimates, and (b) widespread practice of not assembling test forms to maximize item-PLD alignment, which has been the case for MSAA. These mixed results do not undermine the reasonableness of the vertical articulation recommendations; nor do they undermine the validity of interpretation of student scores just above a performance level cut score, at least not any more than is the case in common practice. Overall, the content based rationales provide adequate content based support to adopt the recommended cut scores that achieve vertical articulation in the 2016–2017 data.

[2] See, for example, Ferrara, S., Svetina, D., Skucha, S., & Murphy, A. (2011). Test design with performance standards and achievement growth in mind. *Educational Measurement: Issues and Practice, 30*(4), 3–15.

Vertical Articulation Report
Appendices

Appendix AA—Standard Errors of Percentages of Students in Each Proficiency Level for All Grades in 2015–16 and 2016–17

Table AA-1. Standard Errors of Percentages of Students in Each Proficiency Level–ELA

SY	Grade	Performance Level	Freq. Students	Prop. Students	SE = sqrt(p*q/N)
1516	3	1	982	0.3049	0.0081
		2	502	0.1559	0.0064
		3	1,008	0.3129	0.0082
		4	729	0.2263	0.0074
		3 and 4 combined	1,737	0.5393	0.0088
	4	1	1,134	0.3315	0.0080
		2	625	0.1827	0.0066
		3	1,292	0.3777	0.0083
		4	370	0.1082	0.0053
		3 and 4 combined	1,662	0.4858	0.0085
	5	1	690	0.1945	0.0066
		2	1,051	0.2962	0.0077
		3	1,239	0.3492	0.0080
		4	568	0.1601	0.0062
		3 and 4 combined	1,807	0.5093	0.0084
	6	1	1,137	0.3179	0.0078
		2	1,018	0.2846	0.0075
		3	838	0.2343	0.0071
		4	584	0.1633	0.0062
		3 and 4 combined	1,422	0.3975	0.0082
	7	1	1,230	0.3313	0.0077
		2	635	0.1710	0.0062
		3	1,080	0.2909	0.0075
		4	768	0.2068	0.0066
		3 and 4 combined	1,848	0.4977	0.0082
	8	1	921	0.2558	0.0073
		2	1,247	0.3464	0.0079
		3	586	0.1628	0.0062
		4	846	0.2350	0.0071
		3 and 4 combined	1,432	0.3978	0.0082
	11	1	580	0.2294	0.0084
		2	485	0.1919	0.0078
		3	922	0.3647	0.0096
		4	541	0.2140	0.0082
		3 and 4 combined	1,463	0.5787	0.0098

Table AA-2. Standard Errors of Percentages of Students in Each Proficiency Level– ELA

SY	Grade	Performance Level	Freq. Students	Prop. Students	SE = sqrt(p*q/N)
1617	3	1	1,161	0.3369	0.0081
		2	703	0.2040	0.0069
		3	765	0.2220	0.0071
		4	817	0.2371	0.0072
		3 and 4 combined	1,582	0.4591	0.0085
	4	1	1,421	0.3865	0.0080
		2	648	0.1762	0.0063
		3	1,146	0.3117	0.0076
		4	462	0.1256	0.0055
		3 and 4 combined	1,608	0.4373	0.0082
	5	1	967	0.2581	0.0071
		2	1,028	0.2744	0.0073
		3	1,272	0.3395	0.0077
		4	480	0.1281	0.0055
		3 and 4 combined	1,752	0.4676	0.0082
	6	1	1,329	0.3431	0.0076
		2	1,101	0.2842	0.0072
		3	871	0.2248	0.0067
		4	573	0.1479	0.0057
		3 and 4 combined	1,444	0.3727	0.0078
	7	1	1,294	0.3365	0.0076
		2	602	0.1566	0.0059
		3	1,175	0.3056	0.0074
		4	774	0.2013	0.0065
		3 and 4 combined	1,949	0.5069	0.0081
	8	1	1,194	0.2911	0.0071
		2	1,264	0.3082	0.0072
		3	750	0.1829	0.0060
		4	893	0.2178	0.0064
		3 and 4 combined	1,643	0.4006	0.0077
	11	1	1,040	0.2988	0.0078
		2	684	0.1965	0.0067
		3	1,239	0.3559	0.0081
		4	518	0.1488	0.0060
		3 and 4 combined	1,757	0.5047	0.0085

Table AA-3. Standard Errors of Percentages of Students in Each Proficiency Level–Mathematics

SY	Grade	Performance Level	Freq. Students	Prop. Students	SE = sqrt(p*q/N)
1516	3	1	869	0.2695	0.0078
		2	544	0.1687	0.0066
		3	1,177	0.3650	0.0085
		4	635	0.1969	0.0070
		3 and 4 combined	1,812	0.5619	0.0087
	4	1	1,010	0.2941	0.0078
		2	802	0.2335	0.0072
		3	946	0.2755	0.0076
		4	676	0.1969	0.0068
		3 and 4 combined	1,622	0.4723	0.0085
	5	1	575	0.1619	0.0062
		2	1,057	0.2976	0.0077
		3	1,451	0.4085	0.0082
		4	469	0.1320	0.0057
		3 and 4 combined	1,920	0.5405	0.0084
	6	1	1,032	0.2881	0.0076
		2	944	0.2635	0.0074
		3	766	0.2138	0.0069
		4	840	0.2345	0.0071
		3 and 4 combined	1,606	0.4484	0.0083
	7	1	417	0.1127	0.0052
		2	1,393	0.3764	0.0080
		3	1,159	0.3132	0.0076
		4	732	0.1978	0.0065
		3 and 4 combined	1,891	0.5109	0.0082
	8	1	827	0.2304	0.0070
		2	851	0.2370	0.0071
		3	973	0.2710	0.0074
		4	939	0.2616	0.0073
		3 and 4 combined	1,912	0.5326	0.0083
	11	1	315	0.1827	0.0093
		2	547	0.3173	0.0112
		3	484	0.2807	0.0108
		4	378	0.2193	0.0100
		3 and 4 combined	862	0.5000	0.0120

Table AA-4. Standard Errors of Percentages of Students in Each Proficiency Level–Mathematics

SY	Grade	Performance Level	Freq. Students	Prop. Students	SE = sqrt(p*q/N)
1617	3	1	950	0.2739	0.0076
		2	650	0.1874	0.0066
		3	1,258	0.3627	0.0082
		4	610	0.1759	0.0065
		3 and 4 combined	1,868	0.5386	0.0085
	4	1	1,012	0.2743	0.0073
		2	849	0.2301	0.0069
		3	1,237	0.3352	0.0078
		4	592	0.1604	0.0060
		3 and 4 combined	1,829	0.4957	0.0082
	5	1	556	0.1476	0.0058
		2	1,321	0.3508	0.0078
		3	1,340	0.3558	0.0078
		4	549	0.1458	0.0058
		3 and 4 combined	1,889	0.5016	0.0081
	6	1	1,144	0.2939	0.0073
		2	1,081	0.2777	0.0072
		3	746	0.1917	0.0063
		4	921	0.2366	0.0068
		3 and 4 combined	1,667	0.4283	0.0079
	7	1	452	0.1171	0.0052
		2	1,458	0.3776	0.0078
		3	1,247	0.3230	0.0075
		4	704	0.1823	0.0062
		3 and 4 combined	1,951	0.5053	0.0080
	8	1	1,021	0.2487	0.0067
		2	986	0.2402	0.0067
		3	1,107	0.2697	0.0069
		4	991	0.2414	0.0067
		3 and 4 combined	2,098	0.5111	0.0078
	11	1	620	0.1883	0.0068
		2	1,033	0.3137	0.0081
		3	891	0.2706	0.0077
		4	749	0.2275	0.0073
		3 and 4 combined	1,640	0.4980	0.0087

Appendix BB—Cohort Analysis for Grades 6 and 8 English Language Arts (ELA) and Grade 6 Mathematics

We looked at the percentages of increases and decreases in proficiency levels for ELA and mathematics. First, definitions: the **percentage increase** is the percentage of a cohort that increased their proficiency-level categorization from 2016 to 2017. The **percentage decrease** is similarly defined.

This proposed method of vertical articulation is based on the theory that students should obtain the same proficiency from one year to the next. Assuming this expectation is appropriate for MSAA, the degree to which this does not occur is considered evidence that the standards are not properly articulated. The statistic used to evaluate this is the percentage of exact agreement in proficiency-level categorization from one year to the next. We found that, on average, the percentage of exact agreement was about 38% for mathematics and about 47% for ELA. Thus, this theory of vertical articulation does not seem appropriate for helping us evaluate vertical articulation for MSAA. The inappropriateness may be due to a variety of factors, such as borderline students shifting proficiency levels, changes in curriculum and instruction, or changes in test administration guidelines.

Based on the results we presented, the Technical Advisory Committee (TAC) asked us to do some further analysis focused on grade 6 for ELA (ELA06) and mathematics (MAT06) and focused on grade 8 for ELA (ELA08). Joseph Martineau suggested we look at the percentages of increase and decrease. The idea (for grade 6) is that the percentages of increase and decrease in going from grade 5 to grade 6 should be about the same in going from grade 6 to grade 7 for the 2016–2017 cohort. In particular, Joseph asked what change in the grade 6 cut scores would be needed to equalize the percentages of increase and decrease. We amended this slightly to suggest that the percentages may instead display a consistent pattern compared to what happened in other grades. Thus, we could alternatively find the change in the cut scores that result in obtaining a consistent pattern.

We then proceeded to conduct analyses in this regard.[1] We ran into a few issues that made this task problematic. First, equalizing the overall percentages of increase and decrease for a given grade-to-grade cohort can be done in many ways, since there are three cut scores. Secondly, if the task is redefined as equalizing the conditional percentages of increase and decrease, this is clearly not even possible for people who start in proficiency level 1 or 4. We could arbitrarily come up with a rule equalizing the percentage increase for proficiency level 1 with the percentage decrease for proficiency level 4, but the arbitrariness of the rule is then also problematic. Another problem that makes the task difficult is that the change in one cut score naturally affects the analysis for a neighboring cut score.

Returning to basics, the fundamental underlying question is whether there is statistical evidence that the cut scores for ELA06, MAT06, and ELA08 are “more stringent” beyond a reasonable doubt in comparison to the other tests. There is no doubt that the standards seem harder to achieve for ELA06, MAT06, and ELA08. If there is a consistent pattern across the other grades that is missing only from ELA06, MAT06, and ELA08, then that would be the strongest statistical argument. Note, however, that a strong statistical difference does not necessarily imply there is anything wrong with the cuts since there can certainly be non-statistical reasons for a particular test appearing to have more stringent standards. The analyses presented here address only the statistical evidence for a difference in stringency.

[1] Data for these analyses included only students who tested in both years and students whose test grade in 2017 was one year advanced from that for which they had a score in 2016.

The best way we could think of to address this question was to calculate the conditional probabilities of moving to a year 2 Proficient level, conditional on a year 1 Proficient level. To make this still easier to interpret, we focused only on cut score 2. Thus, we looked at two probabilities:

P(Proficient in Year 2 | Proficient in Year 1)
P(Not Proficient in Year 2 | Not Proficient in Year 1),

where Year 1 refers to the assessments administered for the 2015–16 school year, and Year 2 refers to the assessments administered for the 2016–17 school year.

ELA Grades 6 and 8

We proceeded to calculate these probabilities for ELA, using 2016 and 2017 data as Years 1 and 2, and obtained the following results:

For ELA03: P(not p in Y2 | not p in Y1) = 83%; P(p in Y2 | p in Y1) = 68%.
For ELA04: P(not p in Y2 | not p in Y1) = 76%; P(p in Y2 | p in Y1) = 72%.
For ELA05: P(not p in Y2 | not p in Y1) = 86%; P(p in Y2 | p in Y1) = 59%.
For ELA06: P(not p in Y2 | not p in Y1) = 69%; P(p in Y2 | p in Y1) = 82%.
For ELA07: P(not p in Y2 | not p in Y1) = 85%; P(p in Y2 | p in Y1) = 68%.

These probabilities are graphed in the figures below as the lines labeled “original,” meaning the probabilities on these lines correspond to the original scaled score cuts for the ELA tests. The other lines show how these probabilities change when the grade 6 scaled score cut for Cut 2 (the proficient cut) is lowered from its original value of 1240. These other lines will be discussed below.

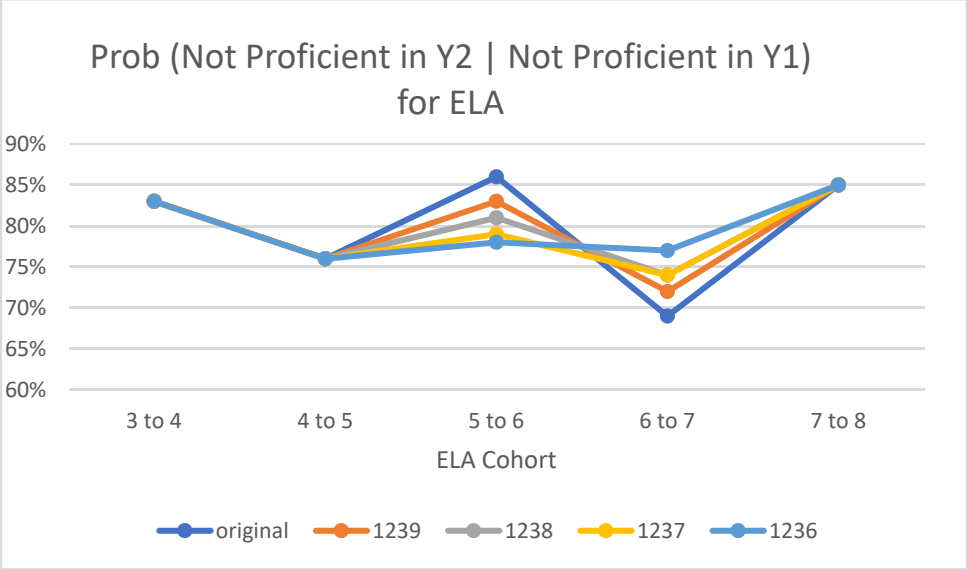


Figure 1. Probabilities for ELA—Not Proficient

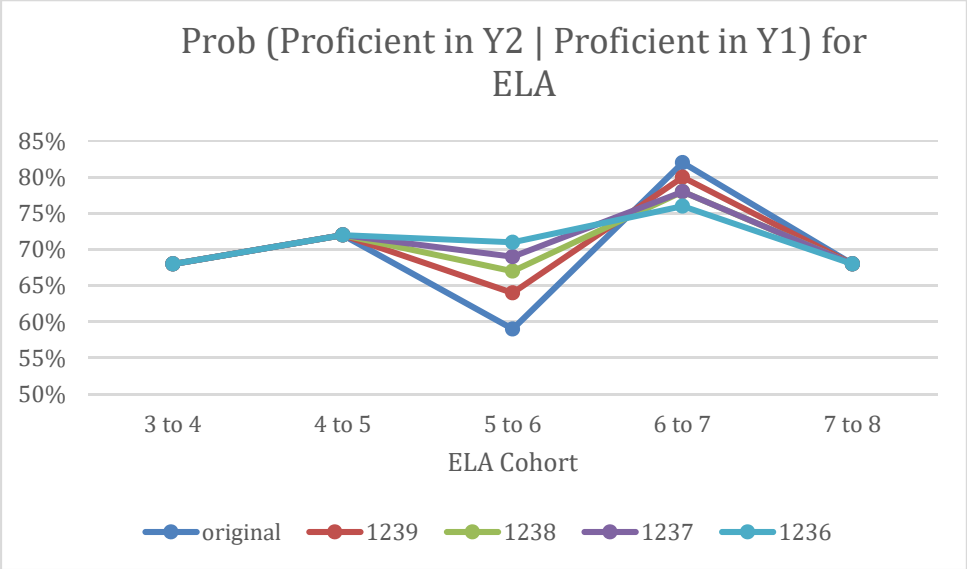


Figure 2. Probabilities for ELA—Proficient

If there is something unusual about ELA06 and ELA08, then we should see some evidence of that in the data for the grade 5, grade 6, and grade 7 2016 students.

If the ELA06 proficient cut is unreasonably high, while all the other grades are reasonable, we would expect:

(1) the probability of a non-proficient grade 5 student remaining non-proficient would be greater than for grades 3, 4, or 7; and
(2) the probability of a proficient grade 5 student remaining proficient would be lower than for grades 3, 4, or 7.

Also, we would expect:

(3) proficient grade 6 students to have a lower probability of remaining non-proficient in Year 2, as compared to grades 3, 4, or 7 students; and
(4) the probability of a proficient grade 6 student remaining proficient to be higher than for grades 3, 4, or 7.

Similarly, if the ELA08 proficient cut is unreasonably high, while all the other grades are reasonable, we would expect:

(1) the probability of a non-proficient grade 7 student remaining proficient would be greater than for grades 3, 4, 5, or 6; and
(2) the probability of a proficient grade 7 student remaining proficient to be lower than for grades 3, 4, 5, or 6.

Grade 6 ELA

First, let's check out the four hypotheses above for ELA06.

The data above indicate that, in going from 2016 to 2017, grade 5 proficient students were *not* significantly more likely to stay non-proficient, as compared to grade 3 or grade 7 students (86% compared to 83%, 76%, and 85%). So, the evidence did *not* support the first hypothesis.

And grade 5 proficient students were indeed less likely to stay proficient, as compared to grades 3, 4, or 7 (59% compared to 68%, 72%, and 68%). Thus, the evidence does support the second hypothesis.

Now let us look at the hypotheses regarding grade 6 2016 students.

The probability of a grade 6 non-proficient student remaining non-proficient is lower compared to grades 3, 4, and 7 (69% compared to 83%, 76%, and 85%). So, the evidence does support the third hypothesis.

The probability of a grade 6 proficient student remaining proficient is higher than for grades 3, 4, or 7 (82% compared to 68%, 72%, and 68%). So, the evidence also supports the fourth hypothesis.

In three out of four cases, the evidence supports the hypotheses that follow from the claim that the proficient cut for ELA06 is more stringently set than for the other grades.

We then delved into the data deeper. If the ELA06 proficiency cut is more stringently set relative to the observed performance of the students as compared to the other grades, the question remains, however, as to whether the size of the difference is large or small—in particular, whether a large or small adjustment to the grade 6 proficiency cut would ameliorate the observed differences that point to the possible over-stringency of the grade 6 proficiency cut.

We then investigated what would happen if we changed Cut 2 for ELA06 from 1240 to 1239.

For the first hypothesis, the 86% changed to 83%, resulting in, not surprisingly, no change in our conclusion that the evidence still does not support the first hypothesis.

For the second hypothesis, the 59% changed to 64%, which is now much closer to the other grades (68%, 72%, and 68%). Thus, the evidence now is weak in support of the second hypothesis.

For the third hypothesis, the 69% changed to 72%, which compares better with grade 4's 76% while still being notably lower than the 83% and 85% for grades 3 and 7. Thus, the third hypothesis is still supported but, again, only weakly.

Finally, for the fourth hypothesis, the 82% changed to 80%, which is still high compared to all the other grades (68%, 72%, and 68% for grades 3, 4, and 7, respectively).

All in all, changing the Cut 2 scaled score cut to 1239, makes it so that only one of the four probability hypotheses is strongly supported by the evidence.

The graphs above display the results for a Cut 2 scaled score cut of 1239 as well as for further reductions in Cut 2 to 1238, 1237, and 1236. As shown in the graphs, as Cut 2 is reduced, the probabilities for the 5-to-6 and 6-to-7 cohorts become closer to the probabilities for the other cohorts. After the adjustment to a 1239, three out of the four probabilities are now within four percentage points of at least one of the other cohorts.

Overall, the results point to, at most, a one-point scaled score cut adjustment for grade 6. When you take into account the fact that no grade 6 students in 2016 got a 1240 or 1241, and no grade 6 students in 2017 got a 1240, that reduces the differences between grade 6 and the other grades. So, if any adjustment is to be made, it is only a small one. Thus, we conclude the degree to which ELA06 is more stringent than the other grades is small.

Grade 8 ELA

Next, we checked out the two hypotheses for ELA08.

For ELA08, we have only the 2016 grade 7 cohort to judge from—that is why there are only two hypotheses concerning the cohort probabilities.

Recalling the hypotheses presented above:

If the ELA08 proficient cut is unreasonably high, while all the other grades are reasonable, we would expect:

> *(1) the probability of a PL1 or 2 grade 7 student remaining non-proficient would be greater than for grades 3, 4, 5, or 6; and*
>
> *(2) the probability of a PL3 or 4 grade 7 student remaining proficient to be lower than for grades 3, 4, 5, or 6.*

So, how does the evidence relate to the above hypotheses?

In regard to the first hypothesis, the probability of a non-proficient ELA07 student remaining non-proficient in grade 8 is 85%. For grades 3, 4, 5, and 6, the values are 83%, 76%, 86%, and 89%, respectively.

Even if you remove grade 5 (because we think moving from grade 5 to grade 6 might have the same problem as moving from grade 7 to grade 8), the hypothesis still does not hold up.

In regard to the second hypothesis, the probability of a proficient ELA07 student remaining proficient in grade 8 is 68%. For grades 3, 4, 5, and 6, the values are 68%, 72%, 59%, and 82%, respectively. Again, the hypothesis is not supported by the data.

Indeed, in general ELA07 students moving up to ELA08 display changes in proficiency probabilities that look exceedingly similar to ELA03 students moving up to ELA04. Thus, the cohort analysis for investigating ELA08 proficiency cut stringency indicates the cut is not unusually stringent compared to the other grades.

Mathematics Grade 6

We next proceeded to calculate the probabilities of interest for mathematics and obtained the following results:

For MAT03: P(not p in Y2 | not p in Y1) = 68%; P(p in Y2 | p in Y1) = 63%.
For MAT04: P(not p in Y2 | not p in Y1) = 65%; P(p in Y2 | p in Y1) = 67%.
For MAT05: P(not p in Y2 | not p in Y1) = 75%; P(p in Y2 | p in Y1) = 57%.
For MAT06: P(not p in Y2 | not p in Y1) = 66%; P(p in Y2 | p in Y1) = 72%.
For MAT07: P(not p in Y2 | not p in Y1) = 66%; P(p in Y2 | p in Y1) = 70%.

These probabilities are graphed in the figures below as the lines labeled “original,” meaning the probabilities on these lines correspond to the original scaled score cuts for the mathematics tests. The other lines show how these probabilities change when the grade 6 scaled score cut for Cut 2 (the proficient cut) is lowered from its original value of 1240 to a value of 1239. These other lines will be discussed below.

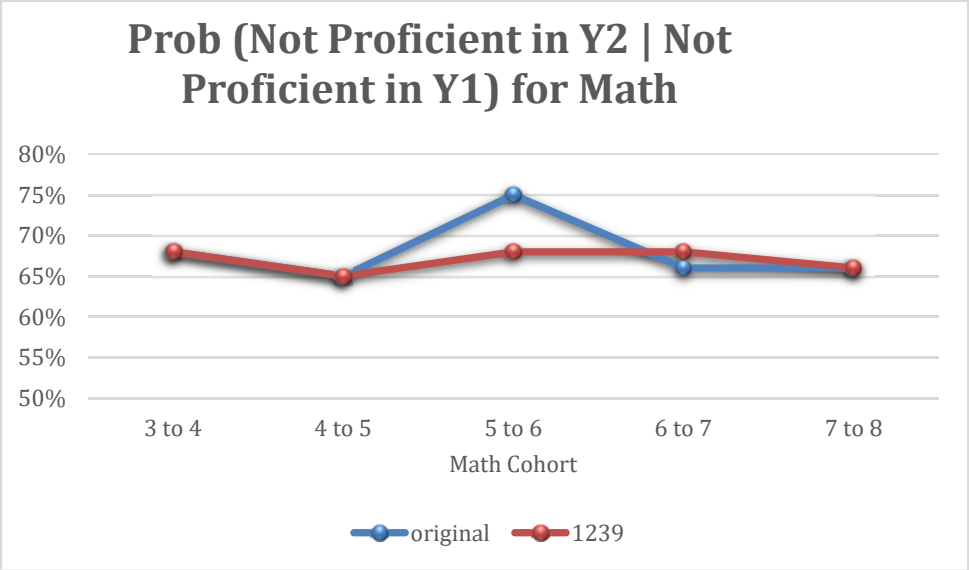


Figure 3. Probabilities for Mathematics—Not Proficient

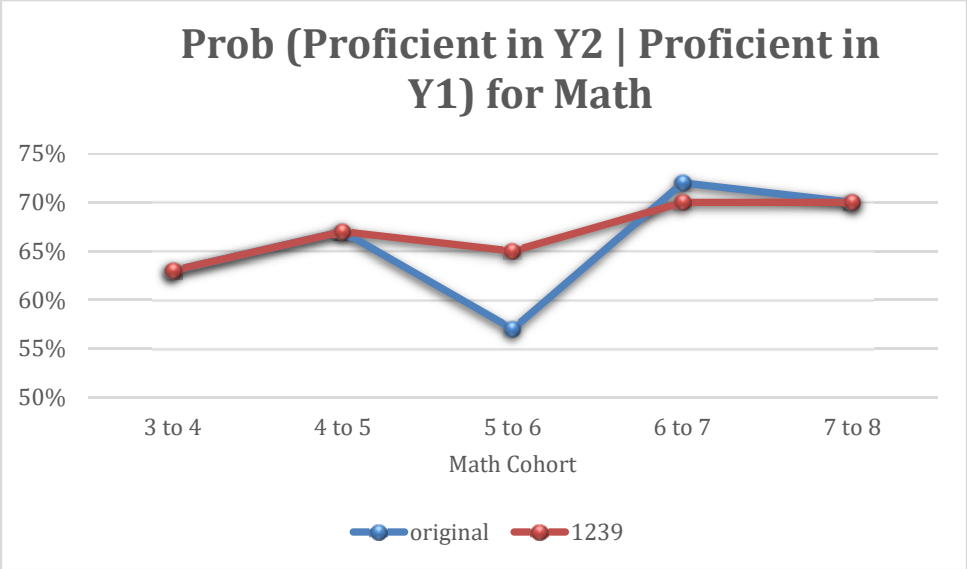


Figure 4. Probabilities for Mathematics—Proficient

If there is something unusual about grade 6, then we should see some evidence of that in the data for the grade 5 and grade 6 2016 students.

If the MAT06 proficient cut score is unreasonably high, while all the other grades are reasonable, we would expect:

> *(1) the probability of a non-proficient grade 5 student remaining non-proficient would be greater than for grades 3, 4, or 7; and*
>
> *(2) the probability of a proficient grade 5 student remaining proficient to be lower than for grades 3, 4, or 7.*
>
> *(3) proficient grade 6 students to have a lower probability of remaining non-proficient in Year 2, as compared to grades 3, 4, or 7 students; and*
>
> *(4) the probability of a proficient 4 grade 6 student remaining proficient to be higher than for grades 3, 4, or 7.*

First, let us look at the grade 5 2016 students.

The data above indicate that, in going from 2016 to 2017, grade 5 non-proficient students were more likely to stay non-proficient, as compared to grade 3, grade 4, or grade 7 students (75% compared to 68%, 65%, and 66%). Thus, the evidence supports the first hypothesis.

And grade 5 proficient 4 students were less likely to stay proficient than grades 3, 4, or 7 students (57% compared to 63%, 67%, and 70%). And the evidence supports the second hypothesis.

Next, let us look at the grade 6 2016 students.

The probability of a grade 6 non-proficient student remaining non-proficient is about the same compared to grades 3, 4, and 7 (66% compared to 68%, 65%, and 66%). The evidence does *not* support the third hypothesis.

The probability of a grade 6 proficient student remaining proficient is higher than for grades 3 and 4, but about the same as for grade 7 (72% compared to 63%, 67%, and 70%). This evidence does *not* support the third hypothesis.

Thus, there is mixed evidence that the grade 6 proficiency cut might be set too high. Only two of the four cases show consistent statistical evidence relative to the other grades. Indeed, the statistics for students moving from grade 6 to grade 7 seem no different from the statistics for students moving from grade 7 to grade 8, and only slightly different from the statistics for students moving from grade 4 to grade 5.

For the sake of completeness, we also calculated the probabilities based on a Cut 2 of 1239 for grade 6. As shown in the above graphs, this results in the other two 5-to-6 and 6-to-7 cohort probabilities falling into line with the probabilities for the other cohorts.

Overall, the statistical analysis provides at best only weak evidence that the mathematics grade 6 proficiency cut is more stringently set relative to the observed performance of the students as compared to other grades.

It is important to keep in mind that these analyses focus only on statistical indices. Given that the MSAA tests do not assume a vertical scale, the statistical arguments are strongest only when all four probabilities are in agreement with the hypothesis that the grade 6 performance level 2 cut score is overly stringent. In the case of mathematics, a substantive argument can be made that the apparent over-stringency may be due simply to the change in rigor in the standards in going from grade 5 to grade 6. This substantive argument is further supported by the fact that the grade 6-to-7 cohort has probabilities that are perfectly in line with the other cohorts, suggesting that the instruction in grade 6 has enabled the students to adjust by the end of grade 6.

Appendix CC—Articulation Displays and Slides for the May 4, 2018 Psychometric Subcommittee and TAC Meeting and Vertical Articulation Displays for 2017 ELA and Mathematics

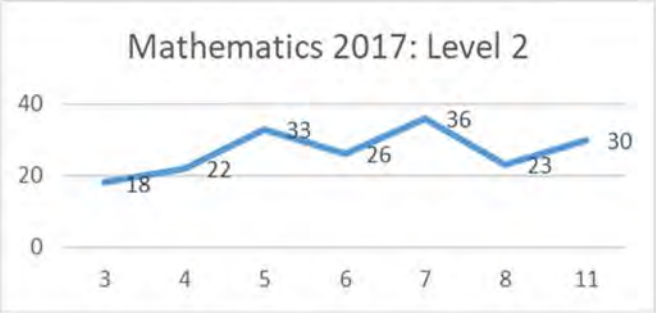
ELA 2017: Level 4
30
20
10
0
22
12
12
14
19
21
14
3
4
5
6
7
8
11

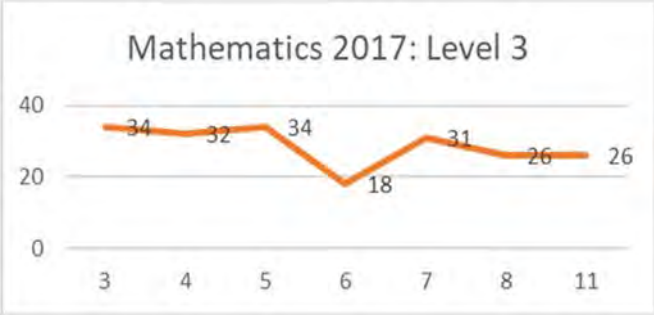
ELA 2017: Level 2
40
20
0
19
17
26
27
15
29
19
3
4
5
6
7
8
11

ELA 2017: Level 3
40
20
0
21
30
32
21
29
17
34
3
4
5
6
7
8
11

ELA 2017: Level 1
60
40
20
0
38
42
29
38
37
32
33
3
4
5
6
7
8
11

Mathematics 2017: Level 4
30
20
10
0
16
15
14
22
17
23
22
3
4
5
6
7
8
11


Mathematics 2017: Level 2
40
20
0
18
22
33
26
36
23
30
3
4
5
6
7
8
11


Mathematics 2017: Level 3
40
20
0
34
32
34
18
31
26
26
3
4
5
6
7
8
11

Mathematics 2017: Level 1
40
20
0
3
4
5
6
7
8
11

Appendix DD—Item-PLD Analysis

Alignment between Response Demands of Items Located in Proficiency Level 3 after Cut Score Adjustments

MSAA Vertical Articulation Content Based Rationales

ELA, Grade 6

Proposed Adjusted Cut = 1237

Item 115183A - Current Level 3 Cut = This item at the current cut aligns to the Level 3 PLD. The text is of moderate complexity and the item aligns to the KSAs of using domain specific words and phrases accurately.

Items that Move from Level 2 to Level 3 with the Articulated Cut Score

Item 121802A - This item aligns to the Level 3 PLD. The text is of moderate complexity and the item aligns to the KSAs of summarizing a literary text from beginning to end without including personal opinions.

Item 124242A - This item is aligned to the Level 2 PLD, due primarily to these types of items being written to a very brief, straightforward text. The item aligns to the KSAs of identifying the next event in a brief narrative.

ELA, Grade 8

Proposed Adjusted Cut = 1238

Item 121042A - Current Level 3 Cut - This item aligns to the Level 3 PLD. The text is of moderate complexity and the item aligns to the KSAs of using details to support an inference from an informational text.

Items that Move from Level 2 to Level 3 with the Articulated Cut Score

Item 124300A - This item aligns to the Level 2 PLD. The text is very short (one line). The item aligns to the KSAs of identifying an idea relevant to a claim.

Item 114879A - This item is aligned to the Level 3 PLD. The text is of higher complexity and the item aligns to the KSAs of using context to identify the meaning of grade-level words and phrases.

Mathematics, Grade 3

Proposed Adjusted Cut = 1242

Item 110959A - Current Level 3 Cut - This item is aligned to the lower end of the Level 3 PLD. The item is borderline between low and moderate task complexity, but aligns to the KSAs of transferring data from an organized list to a bar (picture) graph.

Items that Moves from Level 3 to Level 2 with the Articulated Cut Score

Item 112555A - This item is aligned to the upper end of the Level 2 PLD. The item is of moderate task complexity and aligns to the KSAs of identifying geometric figures that are divided into equal parts.

Item 110959A - Current Level 3 Cut - This item is aligned to the lower end of the Level 3 PLD. The item is borderline between low and moderate task complexity, but aligns to the KSAs of transferring data from an organized list to a bar (picture) graph (same analysis as above).

Item 111377A- This item is aligned to the Level 2 PLD. The item is of moderate text complexity and aligns to the KSAs of identifying geometric figures which are divided into equal parts.

Item 112551A- This item is aligned to the Level 2 PLD. The item is of low task complexity and is aligned to the KSAs of identifying a representation of the area of a rectangle.

Item 120682A- This item is aligned to the Level 2 PLD. The item is of moderate text complexity and aligns to the KSAs of identifying geometric figures which are divided into equal parts.

Note: The response demands associated with both item 110959A and item 112555A are very close. It appears both items are “borderline,” with alignment characteristics to both the Level 2 and Level 3 PLDs.

Mathematics, Grade 4

Proposed Adjusted Cut = 1239

Item 111123A - Current Level 3 Cut - This item is aligned to the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of computing the perimeter of rectangles.

Item that Moves from Level 2 to Level 3 with the Articulated Cut Score

Item 111685A - This item is aligned to the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of rounding numbers to the nearest 10, 100, or 1000.

Note: The scale score location for item 111685A that corresponds to adjusting the cut score by one scale score point is 1235.327, which is not over the threshold of 1239.

Mathematics, Grade 6

Proposed Adjusted Cut = 1239

Item 112663A - Current Level 3 Cut - This item aligns to the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of performing operations using up to three-digit numbers.

Item that Moves from Level 2 to Level 3 with the Articulated Cut Score

Item 111517A - This item is aligned to the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of solving word problems with expressions including variables.

Note: The scale score location for item 111517A that corresponds to adjusting the cut score by one scale score point is 1239.347, which is over the threshold of 1239.

Alignment between Response Demands of Items Located in Mathematics Performance Levels 2 and 4 after Cut Score Adjustments

MSAA Vertical Articulation Content Based Rationales

Mathematics, Grade 5

Level 2, Proposed Adjusted Cut = 1232

Item 112372A- Current Level 2 Cut - This item aligns to the Level 1 PLD. The item is of low task complexity and aligns to the KSAs of identifying values in the tenths place.

Item that Moves from Level 2 to Level 1 with the Articulated Cut Score

Item 111242A- This item aligns to the Level 1 PLD. The item is of low task complexity and aligns to the KSAs of identifying values in the tenths place.

Item that is at the Level 2 Articulated Cut Score

Item 113889A- This item aligns to the Level 2 PLD. The item is of low task complexity and aligns to the KSAs of identifying if the total will increase or decrease when combining sets.

Level 4, Proposed Adjusted Cut = 1253

Item 111299A- Current Level 4 Cut - This item aligns to the Level 4 PLD. The item is of high task complexity and aligns to the KSAs of making quantitative comparisons between data sets shown as line graphs.

Items that Move from Level 3 to Level 4 with the Articulated Cut Score

Item 120724A- This item aligns to the Level 4 PLD. The item is of high task complexity and aligns to the KSAs of locating a given point on a coordinate plane when given an ordered pair.

Item 113902A- This item is a “borderline” item aligning to the Level 3 and Level 4 PLDs. The item is of moderate to high task complexity and aligns to the KSAs of making quantitative comparisons between data sets shown as line graphs.

Mathematics, Grade 6

Level 2, Proposed Adjusted Cut = 1233

Item 120855A- Current Level 2 Cut - This item aligns to the Level 2 PLD. The item is of low task complexity and aligns to the KSAs of matching a given ratio to a model.

Items that Move from Level 1 to Level 2 with the Articulated Cut Score

Item 111508A- This item aligns to the Level 2 PLD. The item is of low task complexity and aligns to the KSAs of performing one-step operations with two decimal numbers.

Item 110990A- This item aligns to the Level 2 PLD. The item is of low to moderate task complexity and aligns to the KSAs of solving common problems presented in mathematical context using various mathematical terms and symbols.

Mathematics, Grade 7

Level 2, Proposed Adjusted Cut = 1234

Item 111749A- Current Level 2 Cut - This item is a “borderline” item aligning to the low end of Level 2 and high end of Level 1 PLDs. The item is of low task complexity and aligns to the KSAs about surface area.

Items that Move from Level 2 to Level 1 with the Articulated Cut Score

Item 111641A- This item is a “borderline” item aligning to the low end of Level 2 and high end of Level 1 PLDs. The item is of low task complexity and aligns to the KSAs about surface area.

Item 111744A- This item is a “borderline” item aligning to the low end of Level 2 and high end of Level 1 PLDs. The item is of low task complexity and aligns to the KSAs about surface area.

Item 112899A- This item aligns to the Level 1 PLD. The item is of low task complexity and aligns to the KSAs of making qualitative comparisons when presented with data in a graph or table.

Item that is at the Level 2 Articulated Cut Score

Item 112852A- This item is a “borderline” item aligning to the low end of Level 2 and high end of Level 1 PLDs. The item is of low task complexity and aligns to the KSAs of identifying representations of the area and circumference of a circle.

Mathematics, Grade 8

Level 4, Proposed Adjusted Cut = 1251

Item 111583A- Current Level 4 Cut - This item is a “borderline” item aligning to the Level 3 and Level 4 PLDs. The item is of moderate to high task complexity and aligns to the KSAs of identifying the relationship shown on a linear graph.

Items that Move from Level 4 to Level 3 with the Articulated Cut Score

Item 113937A- This item aligns to the Level 4 PLD. The item is of high task complexity and aligns to the KSAs of locating the approximate placement of an irrational number on a number line.

Item 112476A- This item aligns to the higher end of the Level 3 PLD. The item is of high task complexity and aligns to the KSAs of identifying congruent and similar figures.

Item 113957A- This item aligns to the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of identifying the relationship shown on a linear graph.

Item 112486A- This item aligns to the higher end of the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of computing the change in area of a figure when its dimensions change.

Item 111339A- This item aligns to the Level 3 PLD. The item is of moderate task complexity and aligns to the KSAs of solving for the volume of a cylinder.

Item that is at the Level 4 Articulated Cut Score

Item 113959A- This item aligns to Level 4 PLD. The item is of high task complexity and aligns to the KSAs of identifying the relationship shown on a linear graph.

Mathematics, Grade 11

Level 2, Proposed Adjusted Cut = 1235

Item 110915A - Current Level 2 Cut - This item aligns to the Level 1 PLD. The item is of low task complexity and aligns to the KSAs of using a table to match a unit conversion.

Item that is at the Level 2 Articulated Cut Score

Item 112924A- This item aligns to the Level 2 PLD. The item is of low task complexity and aligns to the KSAs of identifying the greatest or least value in a set of data shown on a number line.

Appendix B—Participants and Roles in the Standard Validation Process

ELA PLD Revision Committee Members

ELA PLD Revision-members from the subcommittees involved included: Psychometric Subcommittee, Item Development Subcommittee, and Scoring Subcommittee.

- Bethany Zimmerman (AZ)
- Lee Scott (AZ)
- Michael Craig (DC)
- Marsie Torchon (MD)
- Nancy Schmitt (MD)
- Sue Nay (ME)
- Yvonne Field (MT)
- Fasefulu Tigilau (PAC-6)
- Terese Crisostomo (PAC-6)
- Jan Martin (SD)
- Alexandria Baltimore-Hookfin (USVI)

Cut Score Review Meeting Attendees

Cut score review meeting attendees consisted of Psychometric Subcommittee members, State Content Specialists, Technical Advisory Committee members, and other State Partner attendees.

<u>Subcommittee Members</u>

- Bethany Zimmerman (AZ)
- Lee Scott (AZ)
- Marsie Torchon (MD)
- Sue Nay (ME)
- Yvonne Field (MT)
- Jan Martin (SD)

<u>ELA Content Specialist</u>

- Christy Mock-Stutz (MT)

<u>Technical Advisory Committee Members</u>

- Rachel Quenemoen
- Mike Russell

<u>Other State Partner Attendees</u>

- Hansley Mussotte (AZ)
- Cindy Sandner (AZ)

ELA PLD Review Meeting Attendees

Panelist List

Narrative Group – Grades 3-5 (Group 1)	
Name	**State**
Amy Cochran	DC
Michelle DeBlois	ME
Kesiah Frederick	SD
Nicole Greenplate	MD
Christine Hernandez	GU
Dana Lester	TN
Michelle Moen	SD
Sacha Richards	DC

Informational Group – Grades 6-8 (Group 2)	
Name	**State**
Janice Almoquera	GU
Bess Cropper	MD
Georgia Green	DC
Dedriene Rogers	TN
Sarah Stare	MD
Abby Trask	ME
Meredith Verrill	ME
Michelle Wood	AZ

Persuasive Group – Grade 11 (Group 3)	
Name	**State**
Allison Bennett	TN
Sandra Cookson	MD
Johanna Connell	ME
Helene Cruz	GU
Katie DiTullio	AZ
Carissa Hollinger	MD
Heather Saran	MD
Lesa Warrick	DC

State Partner Attendees

- Hansley Mussotte (AZ)
- Bethany Zimmerman (AZ)
- Marsie Torchon (MD)
- Sue Nay (ME)
- Yvonne Field (MT)
- Melissa Flor (SD)
- Megan Sellers (TN)

Appendix C—MSAA ELA Cutscore Review Meeting Agenda

DATE & TIME	Tues., 07/24/18 at 10:00 AM – 5:00 PM EST & Wed., 7/25/18 at 10:00 AM – 5:00 PM EST
LOCATION	
WebEx 7/24/18	Join Webex meeting Meeting number (access code): 634 621 794 Join by phone **1-866-469-3239** Call-in toll-free number (US/Canada) **1-650-429-3300** Call-in toll number (US/Canada) Toll-free calling restrictions Can't join the meeting?
WebEx 7/25/18	Join WebEx meeting Meeting number (access code): 637 980 464 Join by phone **1-866-469-3239** Call-in toll-free number (US/Canada) **1-650-429-3300** Call-in toll number (US/Canada) Toll-free calling restrictions Can't join the meeting?

PARTICIPANTS *("X" signifies attendance)*

STATE PARTNERS			MEASURED PROGRESS		TECHNICAL ADVISORY COMMITTEE	
	AZ	Hansley Mussotte		Steve Ferrara - **Facilitator (Day 1)**		Rachel Quenemoen
	AZ	Bethany Zimmerman		Stephen Murphy - **Facilitator (Day 2)**		Mike Russell
	AZ	Cindy Sandner		Kelly Ickes – **Note-Taker**		Joseph Martineau
	MD	Marsie Torchon		Lisa Jones-Kennedy (Day 1)		
	ME	Sue Nay		Jim Kroening (Day 2)		
	MT	Yvonne Field		Megan Bairstow		
	MT	Cindy Mock-Stutz		Chris Clough		
	SD	Jan Martin		Tina Fregeau		

Topic	Lead	Duration	Outcome
Welcome, Overview, Introductions	Steve	15 min	
MSAA ELA Design, Purpose, Goals, Attendee Role, Procedures Overview	Steve	45 min	Present info & address questions
Materials Review • ELA PLDs • Writing Prompt Level 2 and 3 DTAs • Writing Prompt Rubrics	Kelly	30 min	Present info & address questions
Cut Score Review Rounds • Grade 11 • Grade 8 • Grade 7 • Grade 6 • Grade 5 • Grade 4 • Grade 3	Steve/Stephen	Remainder of Day 1 and Day 2 • Day 1 • Day 1 • Day 1 • Day 2 • Day 2 • Day 2 • Day 2	Gather States Feedback
Adjourn (5:00PM EST)			

Appendix D—MSAA Cutscore Review Meeting

Cutscore Review Enhanced ELA Scale MSAA

MSAA Psychometric Subcommittee, Content Specialists, and Technical Advisory Committee

Measured Progress
July 24-25, 2018

measured progress

Overview

- Why we need a cutscore review
- Purpose and goal of the meeting
- Procedures overview
- Procedures and materials details

- Round 1
- Feedback and discussion
- Round 2
- Content based rationale
- Vertical articulation

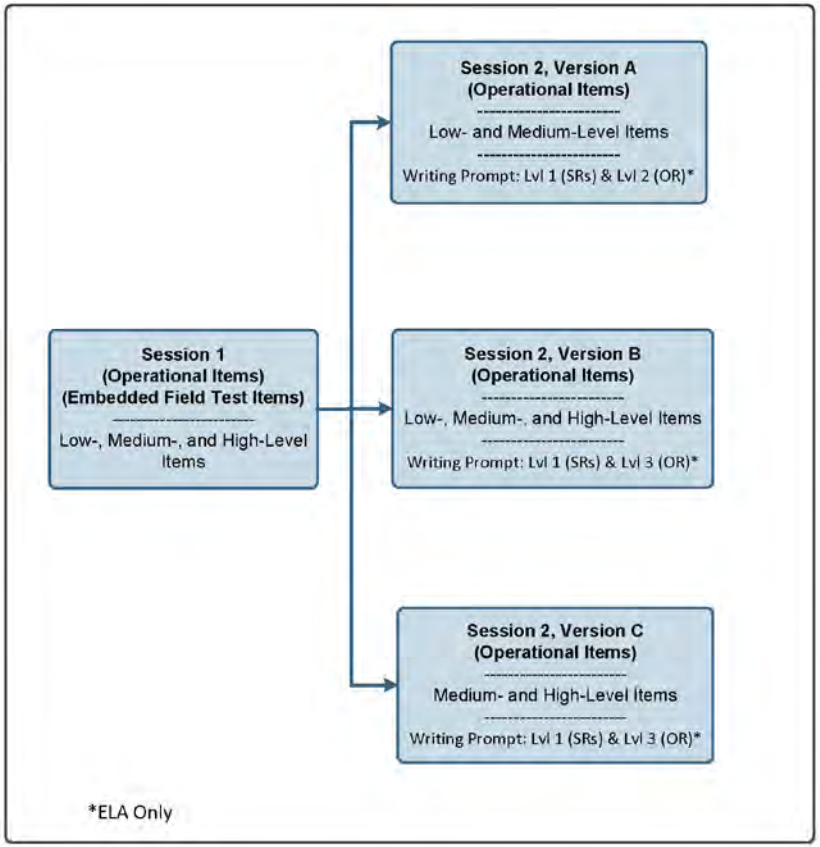
- Definitions of terms

Introductions

- MSAA Psychometric Subcommittee
- MSAA Content Specialists
- MSAA TAC
- Measured Progress

MSAA ELA design

- Multi-stage adaptive (MST)
- 2 stages, writing prompts in session 2
- IRT scaling
- Within grade scales
  - 1200-1290
  - PLs 1, 2, 3, and 4
- PL 3 cutscore = 1240 or thereabouts



MSAA ELA design (cont.)

- Scoring
  - Organization, Idea Development, Conventions
  - Each 0, 1, 2, 3
  - 0 = unrelated evidence
- Calibration
  - 0, 1, 2
    - (0→0; 1→1; 2→1; 3→2)
  - Two threshold locations: Prob .67 of 1 or higher, prob .67 of 2

Why we need a cutscore review

- Added writing prompt to existing ELA scale
- Enhanced scale
  - Same ELA scale, more information
- Enhancement is more information about ELA performance

Purpose of the cutscore review

- Determine if the current cutscores, established in 2015, are **appropriate** for the enhanced ELA scale →

- The original ELA scale contained only Reading, Writing, and Language selected response items
- The enhanced ELA scale also includes writing prompt scores

Goals

- Validate or consider adjusting cutscores that would improve alignment between rubric threshold locations and PLDs
- Ultimate goal
  - Enable valid interpretations of the enhanced ELA scale
  - In relation to the enhanced PLDs
- Consider vertical articulation

Your job, our job

- Follow the procedures and discussion rules
  - Psychometric Subcommittee and state content staff: Decide
  - TAC: Monitor, advise, support
  - Measured Progress: Facilitate, support
- Use your expertise to make judgments
- Psychometric Subcommittee make a group decision regarding the current cutscores and any possible adjustments
- Provide rationales

- Cutscore tables, description of the process, for review and approval in each state

Procedures overview

- Review locations of prompt rubric score thresholds[1]
  - On the enhanced scale
  - In relation to the current cutscores/PLDs
- Make **content based judgments** about the appropriateness of those relationships
- Write **content based rationales** for adjusting the current cutscores

[1] Actually, the **threshold values** for each rubric score: 0,1; 1,2

Procedures overview (cont.)

- Start at grade 11
- Review locations of prompt thresholds in PLs
  - Do the rubric demands and PLDs align?
- Two rounds per grade
  - Round 1: Work independently, discuss process
  - Review feedback, discuss
  - Round 2: Make joint recommendations and write group rationales
- Work downward to grade 3

Workshop materials

- Item maps: Impact Data Tool
- Directions for Test Administration (DTA)—prompts only
  - Display DTAs for other items as needed
- Scoring rubrics
- PLDs
- Content based rationale form
- Also, as needed: content standards

Let's practice the process

Review the prompt locations

- Item maps in Impact Data Tool
  - Items from 2018 operational, all three paths, ordered by difficulty, RP 67 locations
  - Location of score level thresholds in relation to the PLDs
  - Do they align?

- Model, independent practice, discussion

Review the PLDs, writing prompts, and rubrics

- Together

Review the PLDs (Grade 11)…

Grade 11 ELA Performance Level Descriptors

Level 1	Level 2	Level 3	Level 4
Low text complexity - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
In reading, he/she 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, he/she 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, he/she 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, he/she 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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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, he/she is able to: • identify information which 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, he/she 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, he/she 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, he/she is able to:** • use the writing process to create an argumentative product and demonstrate command of organization, idea development and/or conventions

MSAA 2017-18

Final Draft-PLD Updated 050318

Review the PLDs (Grade 11)…

AND in writing, he/she is able to:

- identify information which 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, he/she 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, he/she 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, he/she is able to:

- use the writing process to create an argumentative product and demonstrate command of organization, idea development and/or conventions

Grade 11 ELA Performance Level Descriptors

Level 1	Level 2	Level 3	Level 4
Low text complexity - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple, compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
In reading, he/she 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, he/she 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, he/she 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, he/she 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 - *Text with clear, complex ideas and relationships and simple, compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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, he/she is able to: • identify information which 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, he/she 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, he/she 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, he/she is able to:** • use the writing process to create an argumentative product and demonstrate command of organization, idea development and/or conventions

MSAA 2017-18 Final Draft-PLD Updated 050318

measured progress

Review the writing prompts…

- Electronic materials for participants
  - Writing Prompt Reference Materials

Grade 11 DTA Pages
Spring 2018

Writing Prompt
Level 2

Grade 11 DTA Pages
Spring 2018

Writing Prompt
Level 3

Review the Level 2 rubric (grade 11)…

msaa

Grade 11 Writing Prompt Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0
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)	☐ **no** evidence of organization
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	☐ **no** evidence of idea development
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 Ex: "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)	☐ **no** evidence of standard English conventions

measured progress

Review the Level 3 rubric (grade 11)…

msaa

Grade 11 Writing Prompt Rubric

Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0
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)	☐ **no** evidence of organization
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 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 ☐ 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	☐ **no** evidence of idea development
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 **majority** of thought units ☐ **one** complete sentence that expresses an idea **with** subject-verb agreement Ex: "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)	☐ **no** evidence of standard English conventions

Proposed guideline

- For prompt score level locations that appear misaligned with a PL
  - If within 1 SE, acceptable for now (now = 2018)
  - If outside 1 SE (out of bounds), consider more carefully
- In considering adjusting a cutscore to align a prompt score with a PL, consider the other R, W, L item locations that are impacted
- Could adjust cutscore so that the misaligned item is within 1 SE, not necessarily in the targeted PL

Ready to begin?

Round 1

- Review prompt score locations and PLDs
  - All 12 of them
- All seven grades
  - Start with grade 11
- Work independently, formulate your own initial recommendations
- Discuss
- Share insights, not persuasion

Rubric score threshold locations

PLO1	Locations	Locations	Locations	Locations
	Grade 11	Grade 8	Grade 7	Grade 6
PL 1	--	--	--	--
PL 2	C1	C1, C1, O1	O1	O1, C1
PL 3	O1, C1, O1, I1, C2	I1, O1, I1	I1, C1, C1, O1, I1, I2	I1, C1, I1, O1, I2, C2
PL 4	I1, C2, I2, O2, O2, I2	C2, C2, I2, I2, O2, O2	C2, C2, O2, I2, O2	C2, O2, I2, O2

	Grade 5	Grade 4	Grade 3
PL 1	--	--	--
PL 2	C1, O1	C1, O1	O1, C1
PL 3	I1, C1, C2, O1	C1, I1, C2, O1, I2	C1, I1, O1
PL 4	I1, O2, C2, I2, O2, I2	I1, O2, C2, I2, O2	C2, I1, I2, O2, C2, I2, O2

What to consider

1. Do the rubric descriptions for each score level align with the corresponding PLD?
2. What reasonably can be adjusted?
   - Without causing undue disruption to impact data and interpretation using the PLDs
3. Does the item-PLD alignment analysis support the adjusted cutscore?

Round 1

- Start with grade 11, all rubric locations in relation to PLDs
- Consider
  - Are score threshold locations aligned in relation to PLDs?
  - Using the Impact Data Tool
- Let's do this one together

Round 1 task

- Think about
  - Location of each score threshold location and its corresponding PLD
  - Rubric-PLD alignment/misalignment
- What recommendations do you want to make about each cutscore, given those relationships?
- What is your content based rationale for each recommendation?
- Think independently

Feedback and discussion

- Discussion
  - Share insights and understandings
  - Not persuasion
  - Initial recommendations and rationales
- Feedback
  - Locations summary
  - Impact data

Ready for round 2?

Round 2

- Review locations one last time
- Think and review independently one last time

- Discussion to achieve consensus recommendations
  - Write group content based rationales

Vertical articulation

- Goal
  - Ensure that percentages of students at/above PL 3 are reasonably similar…
  - If any of the PL 3 cut scores were adjusted
  - Articulation of PL 2 and PL 4 after 2018

Intentionally blank

Appropriate

- Aligned, consistent:
  - Response demands of each rubric score threshold
  - Knowledge and skill demands of the corresponding PLD
- Retain the current cutscores because these two elements are reasonably aligned
- Adjust a cut score to improve the alignment
  - Adjustments should be small (i.e., impact data, number of scale score points)
  - Must consider the other items that are in the range of adjustment
  - Content based rationales to support the adjustment

Content based judgments and rationales

- Basis for retaining or adjusting or retaining current cutscores
- Relationship between rubric thresholds and PLDs
- Make explicit references to both
  - What the item demands of students
  - What the PLD requires of students

Threshold values

- Location for each rubric score level **threshold**
  - 0,1 1,2

- Example: 0,1 threshold
  - Probability (.67) that students at this ELA scale location would achieve a score of 1 or higher rather than 0

It's all about student learning.

Period.

Thank you.

ferrara.steve@measuredprogress.org

+1 603-749-9102, ext. 7065

measured progress

Recommendation options

1. Retain current cutscores, write CB rationales to support
2. Adjust some cutscores, write CB rationales to support
3. Recommend not linking prompt to ELA scale in one or more grades in 2018; hold off until…
   a. Stability, research
4. Recommend implementing the linking in 2018 and retaining current cut scores
   a. Investigate why prompts are so difficult in 2018
   b. Conduct a CSR in 2019

- Can recommend studies to accompany recommendations
- Studies: Focus on appropriateness of expectations for 3rd and 4th graders—writing on line, prompts, rubrics

Focus cutscores

Grade	0,1	1,2	2,3	3,4
8	--	--	--	--
7	--	--	--	--
6	--	--	--	--
5	--	--	--	--
4	--	--	--	--
3	--	--	--	--

Appendix E—MSAA ELA CSR Summary

MSAA ELA Cut Score Review

Alignment of the 2018 PLDs and ELA Scale (with Writing Prompts Added)

Psychometric Subcommittee Recommendations

July 24, 2018

Purpose of this Document

The MSAA Psychometric Subcommittee, MSAA content specialists, the MSAA Technical Advisory Committee, and Measured Progress worked together to support valid interpretations of the 2018 ELA assessment and performance level descriptors, after linking writing prompt scores to the existing ELA scale. This document summarizes that process and results so that each MSAA partner state can review and consider approving the adjusted cut scores for reporting of 2018 and subsequent MSAA results.

Cut Score Review Goals, Process, and Results

Cut score reviews are designed to ensure that cut scores on test score scales are aligned appropriately with performance level descriptors (PLDs). This alignment is necessary to ensure valid interpretation of test scores, using the PLDs.

We linked the writing prompts to the 2018 ELA assessment, which originally included only selected response items in reading, writing, and language. Adding the writing prompt scores to the original scale enhances the information that the ELA assessment provides. We also updated the PLDs to include references to direct writing performance at each of performance levels 1, 2, 3, and 4. The 2018 PLDs and the ELA assessment do not change interpretation of ELA scores; rather, they add information about what students know and can do in ELA.

The cut score review process included the following steps:

- Prior to the cut score review, we reviewed the 2017 impact data (i.e., percentages of students in performance levels 3 and 4) and vertically articulated it, to smooth out differences in performance across grades of greater than 5 percent. We wrote content-based rationales to support small adjustments to the performance level 3 cut scores that achieved vertical articulation.
- The MSAA Psychometric Subcommittee and ELA content specialists from MSAA states reviewed the alignment between the location of writing prompt scores on the ELA scale and the PLDs and recommended either (a) retaining the existing cut scores, or (b) adjusting the cut scores to improve that alignment.
- They also wrote content-based rationales for adjusted cut scores.

The results of the cut score review meeting on July 24, 2018 are summarized in the table below. The content-based determinations that were made for the cut score adjustments outlined in the table below indicate that there is alignment between the items that are in performance level 3 and performance level 4 and PLD expectations for each performance level. Full details of the content-based rationales for the small number of cut score adjustments appear in a separate document.

Table 1. Cut Score Review Results: MSAA ELA

Grade	Performance Level	Original Cut Score [1]	2018 Impact Data (%)	Adjusted Cut Score	Resulting Impact Data (%)
11	PL 4	1255	20.3	--	--
	PL 3	1240	40.9	--	--
	PL 2	1236	15.0	--	--
	PL 1	--	23.8	--	--
8	PL 4	1250	21.4	--	--
	PL 3	1238	27.6	--	--
	PL 2	1230	27.8	--	--
	PL 1	--	23.2	--	--
7	PL 4	1255	21.2	--	--
	PL 3	1240	34.6	--	--
	PL 2	1236	16.0	--	--
	PL 1	--	28.2	--	--
6	PL 4	1253	15.9	1251	18.5
	PL 3	1237	39.9	--	37.3
	PL 2	1231	27.6	--	--
	PL 1	--	16.6	--	--
5	PL 4	1256	14.6	--	--
	PL 3	1240	34.6	--	--
	PL 2	1232	24.3	--	--
	PL 1	--	26.5	--	--
4	PL 4	1258	13.6	1259	10.6
	PL 3	1240	33.6	--	36.5
	PL 2	1234	18.4	--	--
	PL 1	--	34.4	--	--
3	PL 4	1251	24.7	1254	19.2
	PL 3	1240	26.0	--	31.5
	PL 2	1234	12.5	--	--
	PL 1	--	36.8	--	--

[1] Before the cut score review, after vertical articulation of the PL 3 cut scores on July 20, 2018. Each grade scale ranges from 1200 to 1290.

Appendix F—MSAA ELA CSR Adjusted Cut Scores CBRs

MSAA Cut Score Review

Writing Content Based Rationales for Adjusted Cut Scores

MSAA Psychometric Subcommittee

July 24, 2018

Grade	Performance Level	Original Cut Score [1]	2018 Impact Data (%)	Adjusted Cut Score	Resulting Impact Data (%)	Content Based Rationale
11	PL 4	1255	20.3	--	--	
	PL 3	1240	40.9	--	--	
	PL 2	1236	15.0	--	--	
	PL 1	--	23.8	--	--	
8	PL 4	1250	21.4	--	--	
	PL 3	1238	27.6	--	--	
	PL 2	1230	27.8	--	--	
	PL 1	--	23.2	--	--	
7	PL 4	1255	21.2	--	--	
	PL 3	1240	34.6	--	--	
	PL 2	1236	16.0	--	--	
	PL 1	--	28.2	--	--	
6	PL 4	1253	15.9	1251	18.5	Item 471934C (level 2 writing prompt), score level 2 on the reporting scale, is aligned to the borderline of the Level 4 and Level 3 PLD, and falls within the Level 4 PLD. The expectations for this grade shift from producing a narrative product in grades 3-5 to producing an explanatory product using the writing process. To obtain a score level 2 for Conventions the expectations are for the product to include more than one sentence with end punctuation and a minimum of one complete sentence with subject-verb agreement. This item is “borderline,” with alignment characteristics to both the expectations Level 4 and Level 3 PLDs.

continued

Grade	Performance Level	Original Cut Score [1]	2018 Impact Data (%)	Adjusted Cut Score	Resulting Impact Data (%)	Content Based Rationale
6	PL 3	1237	39.9	--	37.3	Item 471934I (level 2 writing prompt), score level 2 on the reporting scale, is aligned to the borderline of the Level 4 and Level 3 PLD, and falls within the Level 3 PLD. The expectations for this grade shift from producing a narrative product in grades 3-5 to producing an explanatory product using the writing process. To obtain a score level 2 for Idea Development the expectations are for the product to include a minimum of three activities, each with relevant details. This item is "borderline," with alignment characteristics to both the expectations Level 4 and Level 3 PLDs.
	PL 2	1231	27.6	--	--	
	PL 1	--	16.6	--	--	
5	PL 4	1256	14.6	--	--	
	PL 3	1240	34.6	--	--	
	PL 2	1232	24.3	--	--	
	PL 1	--	26.5	--	--	
4	PL 4	1258	13.6	1259	10.6	Item 126163I (level 3 writing prompt), score level 1 on the reporting scale, is more aligned to expectations at Level 3 PLD. Shifting the cut score brings this closer to the borderline between Level 4 PLD and Level 3 PLD and aligns the expectations more with a partial command of organization, idea development, and/or conventions as outlined in the PLD. Item 512069, which falls right at the adjusted cut, is aligned to the borderline of the Level 4 and Level 3 PLD, and falls within the Level 4 PLD. The text is of moderate complexity and the items align to the KSAs of using information presented visually, orally, or quantitatively to answer questions.
	PL 3	1240	33.6	--	36.5	Items 121580A and 122582A, are aligned to the Level 3 PLD. The cut score adjustment brings the items within Level 3 PLD. The text is of moderate complexity and the items align to the KSAs of explaining how information provided in charts, graphs, diagrams, or timelines contribute to an understanding of informational text.
	PL 2	1234	18.4	--	--	continued

Grade	Performance Level	Original Cut Score [1]	2018 Impact Data (%)	Adjusted Cut Score	Resulting Impact Data (%)	Content Based Rationale
4	PL 1	--	34.4			
3	PL 4	1251	24.7	1254	19.2	Item 125971I (level 3 writing prompt), score level 1 on the reporting scale, is more aligned to expectations at Level 3 PLD. Shifting the cut score brings this closer to the borderline between Level 4 PLD and Level 3 PLD and aligns the expectations more with a partial command of organization, idea development, and/or conventions as outlined in the PLD. Item 120914A, which falls right at the adjusted cut, is aligned to the Level 4 PLD. The text is of high complexity and the items align to the KSAs of determining the central idea and supporting details in literary text.
	PL 3	1240	26.0	--	31.5	Item 116202A, which was on the original cut, is aligned to the Level 3 PLD. The cut score adjustment brings the item within Level 3 PLD. The text is of moderate complexity and the items align to the KSAs of answering literal questions.
	PL 2	1234	12.5	--	--	
	PL 1	--	36.8	--	--	

[1] Before the cut score review, after vertical articulation of the PL 3 cut scores on July 20, 2018. Each grade scale ranges from 1200 to 1290.

Appendix G—MSAA Mathematics Vertical Articulation Results by Grade

MSAA Mathematics Vertical Articulation

Psychometric Subcommittee Recommendations

July 26, 2018

Purpose of this Document

The MSAA Psychometric Subcommittee, MSAA content specialists, the MSAA Technical Advisory Committee, and Measured Progress worked together to determine a process for conducting and completing the vertical articulation for mathematics cut scores. This document summarizes that process and results so that each MSAA partner state can review and consider approving the adjusted cut scores for reporting of 2018 and subsequent MSAA mathematics results.

Vertical Articulation Goals, Process, and Results

A Measured Progress team of Content Development, Psychometrics, and Program Management staff completed the vertical articulation process to review the current MSAA cut scores and 2016–2017 performance data (i.e., percentages of all MSAA students in performance levels 1, 2, 3, and 4) and suggest recommendations for changes to the mathematic cut scores. The results of this work was reviewed with the Psychometric Subcommittee and approval was given for the recommended adjustments detailed on page 2 in the table “Vertical Articulation Results: MSAA Mathematics”.

The vertical articulation process included the following steps:

- Review of the 2017 impact data (i.e., percentages of students in performance levels 3 and 4) and vertically articulated to smooth out differences in performance across grades of greater than 5 percent. Write content-based rationales to support the small adjustments to the performance level 3 cut scores that achieved the vertical articulation.
- Review of the 2017 impact data (i.e., percentage of students in performance levels 2 and 4) and vertically articulated to smooth out differences in performance across the grades of greater than 5 percent and doing so without altering the articulation of the impact data at performance levels 3 and 4. Write content based rationales to support the small adjustments to the performance level 2 and 4 cut scores that achieved vertical articulation.

The results of the vertical articulation process are summarized in the table below. In general, the content-based rationales that were made for the cut score adjustments outlined in the table below indicate that there is alignment between the items and PLD expectations for each performance level. Full details of the content-based rationales for the cut score adjustments appear in a separate document.

Table 1. Vertical Articulation Results: MSAA Mathematics

Grade	Performance Level	Original Cut Score [1]	2017 Impact Data by PL (%)	2017 Impact Data for PLs 3 & 4 (%)	Adjusted Cut Score	Resulting Impact Data (%)	2017 Impact Data for PLs 3 & 4 (%)	Adjustments made for Vertical Articulation
11	PL 4	1249	19.20	43.03	1250	17.92	--	Higher by 1
	PL 3	1240	23.83		--	25.11		
	PL 2	1234	26.81		1235	24.18		Higher by 1
	PL 1	--	30.16		--	32.79		
8	PL 4	1249	23.06	45.90	1251	18.78	--	Higher by 2
	PL 3	1240	22.84		--	27.12		
	PL 2	1234	23.41		--	--		
	PL 1	--	30.69		--	--		
7	PL 4	1254	17.67	47.00	--	--	--	
	PL 3	1240	29.33		--	--		
	PL 2	1232	32.36		1234	25.16		Higher by 2
	PL 1	--	20.64		--	27.84		
6	PL 4	1249	21.67	40.54	1251	17.53	46.50	Higher by 2
	PL 3	1240	18.87		1239	28.97		Lower by 1
	PL 2	1234	23.11		1233	21.49		Lower by 1
	PL 1	--	36.35		--	32.01		
5	PL 4	1255	13.60	46.54	1253	16.84	--	Lower by 2
	PL 3	1240	32.94		--	29.70		
	PL 2	1231	29.72		1232	24.08		Higher by 1
	PL 1	--	23.74		--	29.38		
4	PL 4	1251	17.40	41.01	--	17.40	44.13	
	PL 3	1240	23.61		1239	26.73		Lower by 1
	PL 2	1233	23.59		1232	24.83		Lower by 1
	PL 1	--	35.40		--	31.04		
3	PL 4	1254	19.31	49.46	--	19.31	44.66	
	PL 3	1240	30.15		1242	25.35		Higher by 2
	PL 2	1236	16.54		1235	22.55		Lower by 1
	PL 1	--	34.00		--	32.79		

[1] Original cut scores established in August 2015. Each grade scale ranges from 1200 to 1290.

The largest differences between each of the performance levels across grades are: PL 1 (4.95%), PL 2 (3.67%), PL 3 (4.59%), PL 4 (2.47%). The figure below (Figure 1) shows the articulation lines for the proposed sets of mathematics cut score adjustments for all grades as outlined in the table above.

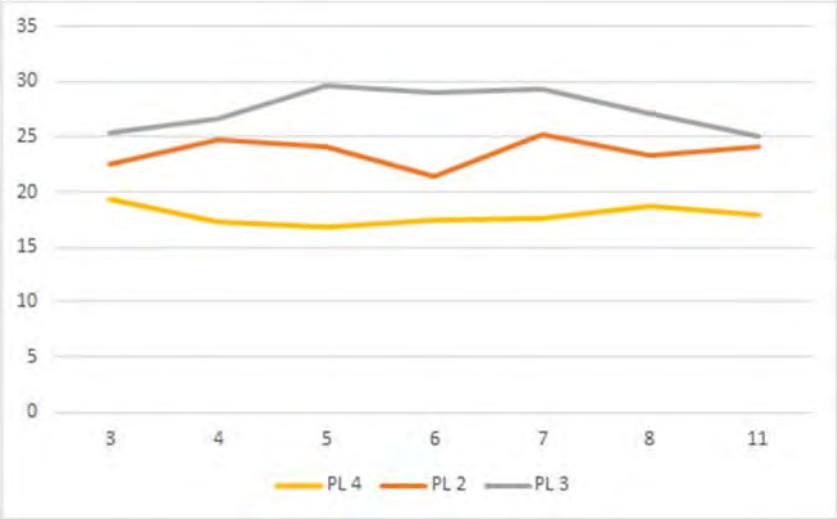


Figure 1. Mathematics vertical articulation lines for the all proposed adjustments at performance levels 2, 3, and 4.

The following figures show the percentage of students at each performance level across all grades for MSAA mathematics. The first one (Figure 2) shows the student percentages at the original cut scores. The second one (Figure 3) shows the student percentages at the adjusted cut scores.

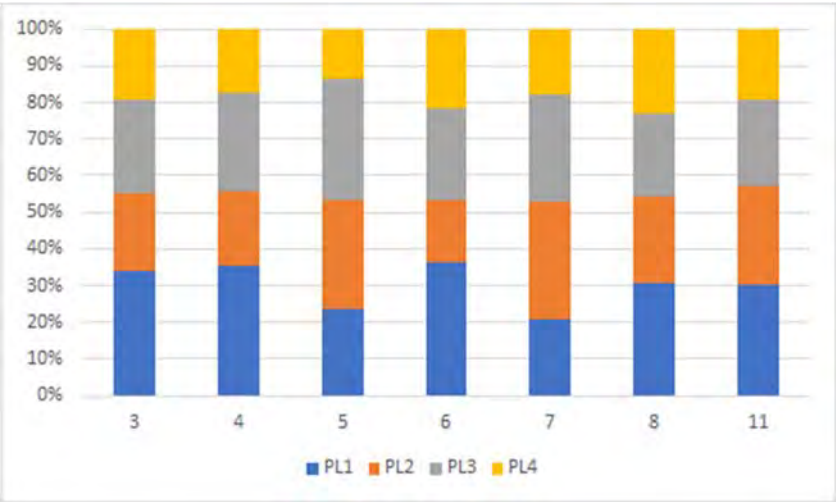


Figure 2. Percentage of students for each performance level based on original cut scores.

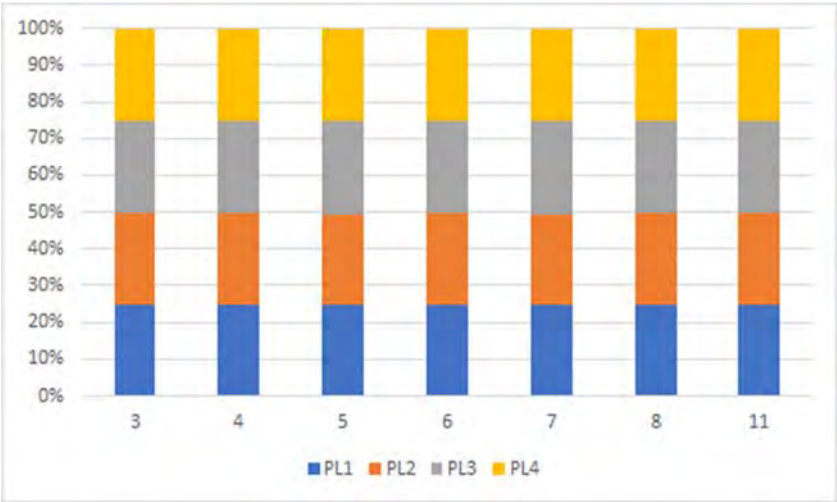


Figure 3. Percentage of students for each performance level based on proposed adjusted (vertically articulated) cut scores.

Appendix H—MSAA 2-18 ELA PLD Panelist Training Opening Presentation

Welcome.

MSAA English Language Arts Performance Level Descriptor Review Meeting

August 9, 2018

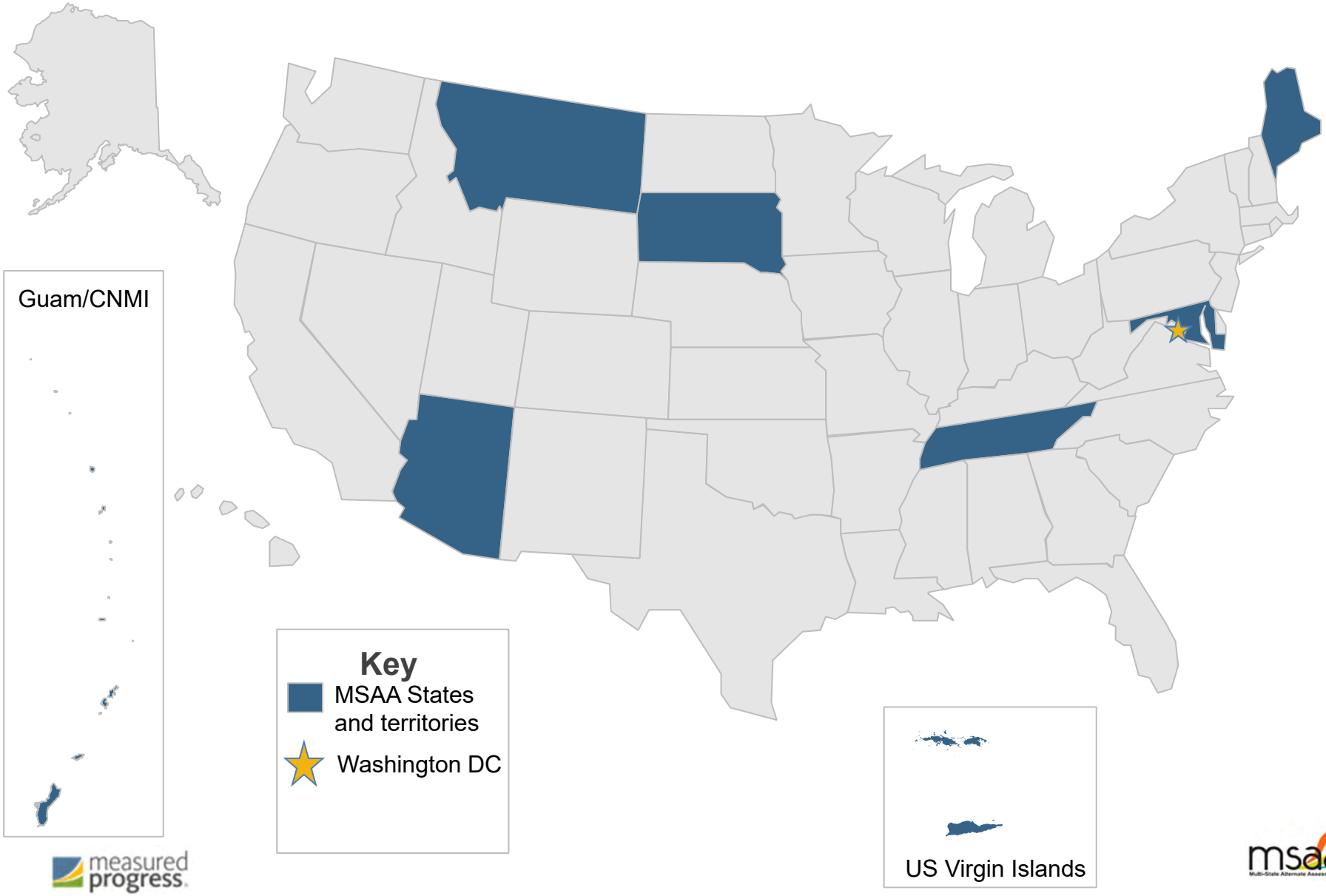
measured progress

msaa

Agenda

- MSAA Partner States
- Background and Purpose
- Overview of the Test
- MSAA Item Types
- Assessment Features & Accommodations
- Writing Prompt Overview
- Writing Prompt Considerations and Emphasis
- Writing Prompt Rubrics
- ELA PLD Overview
- Panelist Expectations
- Next Steps

MSAA Partner States



measured progress

msaa
Multi-State Alternate Assessment

MSAA Partner States

- **Arizona –** Bethany Zimmerman
- **Arizona –** Hansley Mussotte
- **DC –** Michael Craig
- **Guam/CNMI –** Mr. T
- **Maine –** Sue Nay
- **Maryland –** Marsie Torchon
- **Montana –** Yvonne Field
- **Tennessee –** Megan Sellers
- **South Dakota –** Melissa Flor

MSAA Learner Characteristics

Participation Criteria

Sample Participant

Background and Purpose

Participation

- Grade-level content aligned to State Content Standards

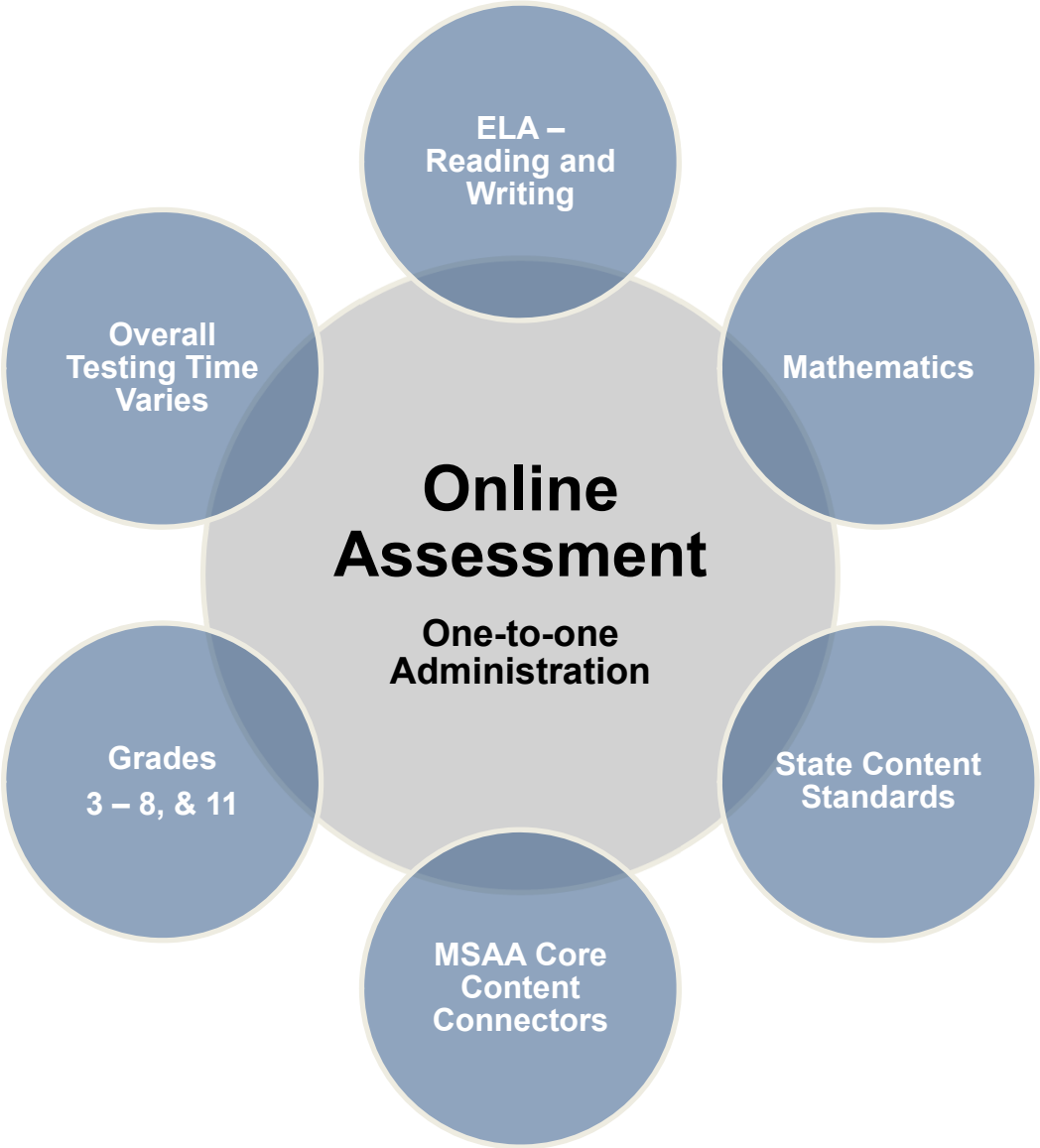
Achievement

- Higher academic outcomes
- Prepared for post-secondary options

Accountability

- Every Student Succeeds Act (**ESSA**)
- Individuals with Disabilities Education Act (**IDEA**)

Overview of the Test



measured progress

msaa
Multi-State Alternate Assessment

MSAA Item Types

ELA

- Selected-Response Items Reading and Writing
- Constructed-Response Items Writing
- Foundational Reading Items Grades 3-4

Math

- Selected-Response Items
- Constructed-Response Items

Assessment Features

- MSAA Assessment features
  - Answer Masking
  - Audio Player
  - Alternate Color Themes
  - Increase/Decrease Size of Text/Graphics
  - Increase Volume
  - Line Reader Tool
  - Read Aloud/Reread item directions, response options, passages

Accommodations

- MSAA Accommodations
  - Assistive Technology (AT) for viewing, responding, or interacting with test items
  - Paper version of items (downloaded from platform)
  - Scribe
  - Sign language

NOTE: The use of any physical prompting, including hand over hand, invalidates the results of the test for the student.

Writing Prompt Overview

- Writing Prompt: ELA requires students to produce a permanent product in response to a writing prompt.

- Each writing prompt DTA contains:
  - A standardized, scripted sequence of steps for the TA to follow
  - A graphic organizer for students to make notes and plan their essay
  - A template to write their essay before it is typed on the computer or uploaded into the system
  - A mentor text to present to the student as an example of a finished product (grades 3, 4, 5, 11 only)

Writing Prompt Overview

- Two different levels of constructed response writing prompts:
  - Level 2
  - Level 3
- Differentiated by the amount of support that is provided in terms of template and stimulus materials
- Each student only takes one constructed response writing prompt
- Students use their primary mode of communication to construct a writing product
- Standardized script and support materials are provided for each level

Writing Prompt Overview

- Each grade span assesses a different writing style. The writing prompts provide steps to guide students through the writing processes using stimulus materials:
  - Topic selection
  - Choosing characters/supporting details
  - Drafting
  - Revising
  - Editing
  - Producing final story or essay

Writing Prompt Considerations

- Traditional views of writing will need to be expanded in the creation of a permanent product.
- Important for students to make a connection to the writing by promoting personal relevance.
- Teachers prioritize steps of the writing process students participate initially. As mastery occurs, students may be able to participate in more of the steps.
- Extended time expectations may be needed for writing projects
- Writing for this population may require teachers to think anew to come up with strategies that will allow students to participate in the writing process (content and physical aspects).

Writing Prompt Emphasis

Across all grades students demonstrate:

- The ability to generate a permanent product to represent and/or organize ideas or thoughts so messages can be interpreted by someone else when the writer is not present;
- The ability to respond to a writing prompt to produce a Literary/Narrative, Informative/Explanatory, or Persuasive/Argument permanent product; and
- The ability to include grade-specific writing skills specific to a writing mode related to organization; language and vocabulary; idea development; and conventions.

Writing Prompt Rubrics – Level 2

msaa

Grade 3 Writing Prompt Rubric

Level 2

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0
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	☐ **no** evidence of organization
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)	☐ **no** evidence of idea development
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 Ex: "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)	☐ **no** evidence of standard English conventions

measured progress

msaa

Writing Prompt Rubrics – Level 3

msaa

Grade 3 Writing Prompt Rubric

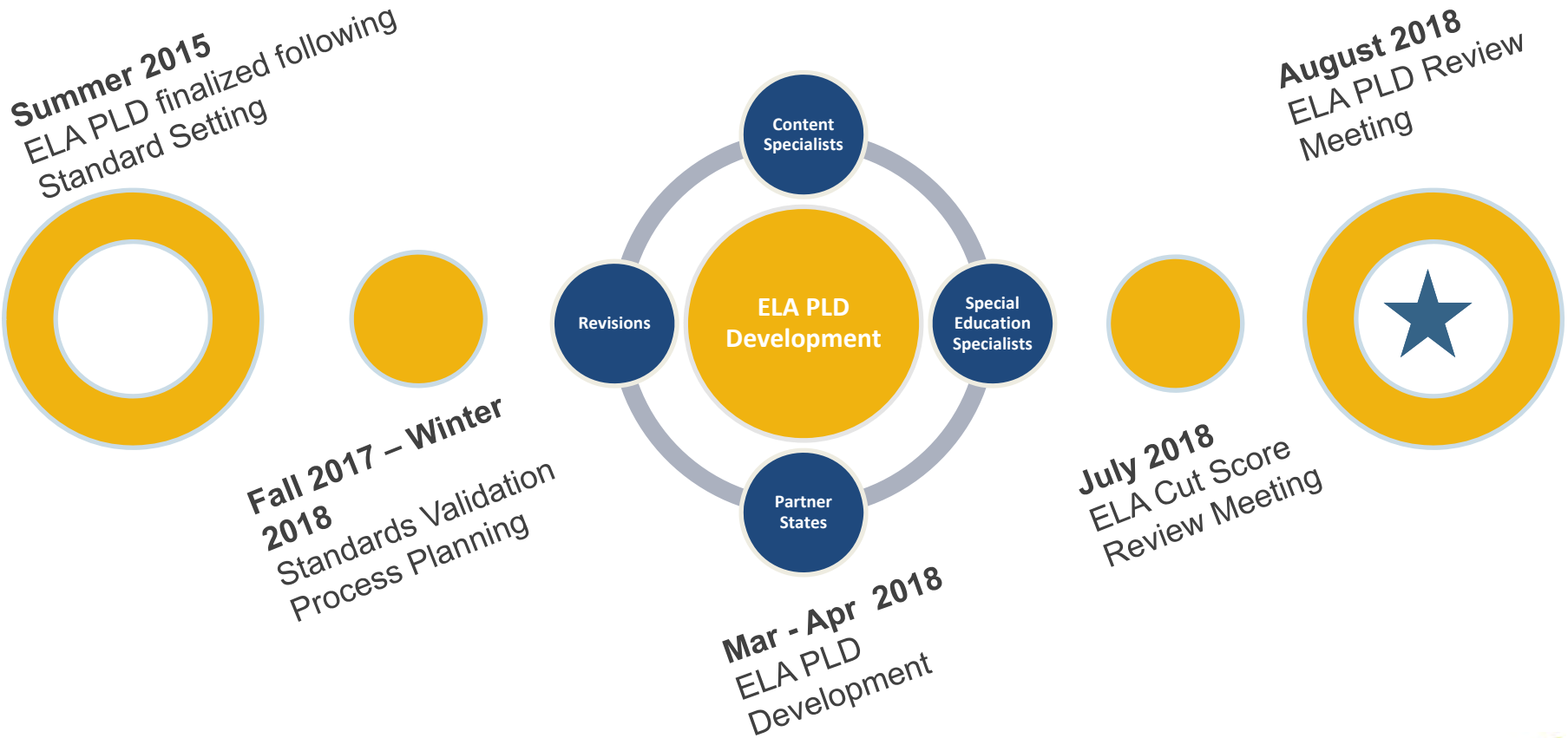
Level 3

Rubric Elements	Full Evidence 3	Partial Evidence 2	Limited Evidence 1	Unrelated Evidence 0
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) ☐ **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)	☐ **no** evidence of organization
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)	☐ **no** evidence of idea development
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 Ex: "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)	☐ **no** evidence of standard English conventions

measured progress

msaa

ELA PLD Development



measured progress

msaa
Multi-State Alternate Assessment

ELA PLD Overview (Draft)

Grade 3 ELA Performance Level Descriptors

Level 1	Level 2	Level 3	Level 4
Low text complexity - *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Low text complexity -** *Brief text with straightforward ideas and relationships; short, simple sentences.*	**Moderate text complexity -** *Text with clear, complex ideas and relationships and simple; compound sentences.*	**High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*
In reading, he/she 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, he/she 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, he/she 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, he/she 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 - *Text with clear, complex ideas and relationships and simple; compound sentences.*	**AND with High text complexity -** *Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.*	
	• 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, he/she 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, he/she is able to:** • identify grade level words	
AND in writing, he/she 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, he/she 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, he/she 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, he/she is able to:** • use the writing process to create a narrative product and demonstrate command of organization, idea development and/or conventions

MSAA 2017-18

Final Draft-PLD Updated 050318

measured progress

msaa
Multi-State Alternate Assessment

ELA PLD Overview (Draft)

AND in writing, he/she 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

Level 1

AND in writing, he/she 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

Level 2

AND in writing, he/she is able to:

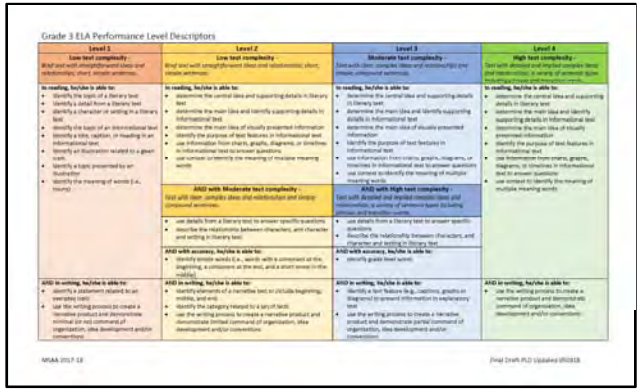
- identify a text feature (e.g., captions, graphs or diagrams) to present information in explanatory text
- use the writing process to create a narrative product and demonstrate partial command of organization, idea development and/or conventions

Level 3

AND in writing, he/she is able to:

- use the writing process to create a narrative product and demonstrate command of organization, idea development and/or conventions

Level 4


Grade 3 ELA Performance Level Descriptors

measured progress

msaa Multi-State Alternate Assessment

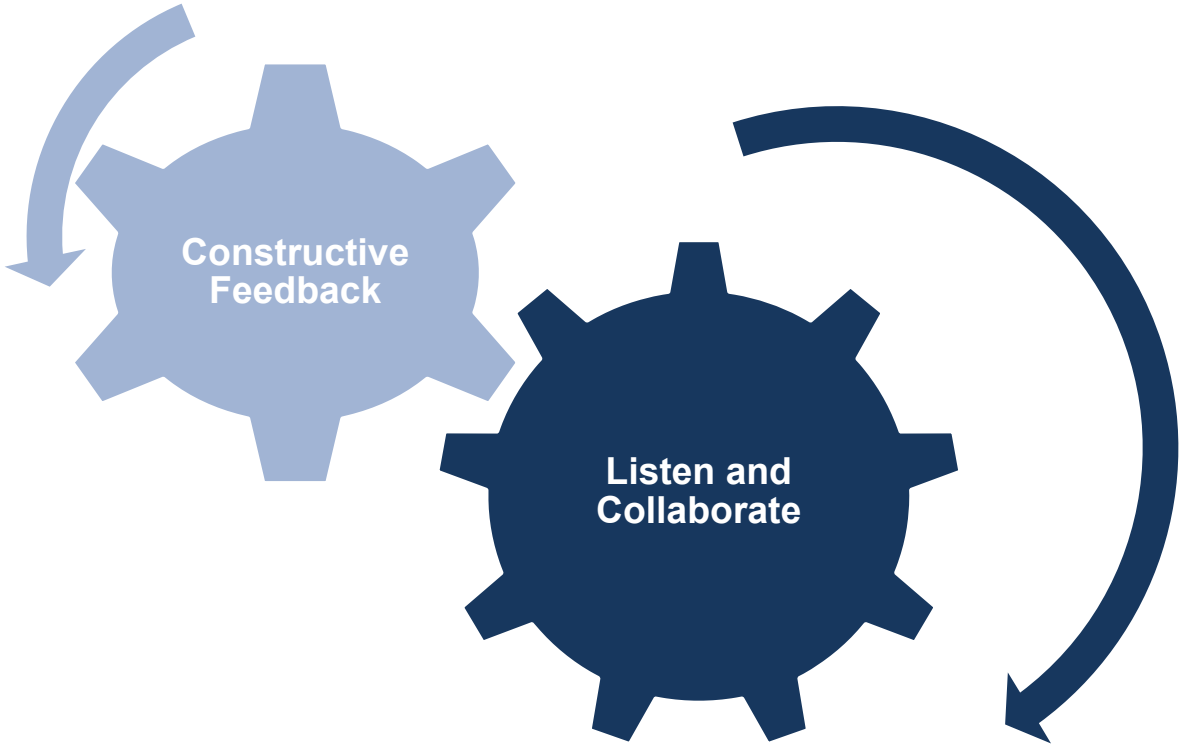
ELA PLD Review Expectations

- For each grade consider the:
  - Writing prompt definitions and emphasis
  - Scoring rubrics
  - Writing prompts
  - Student exemplars
- Answer the questions:
  - Does the writing prompt information contained in the PLD for each level accurately account for what the writing prompt is designed to measure?
  - Is the language clear and reflective of information that is understandable for administrators, teachers and parents?
  - How might the writing prompt measures and emphasis best be communicated to the field?
  - What avenues should be used to communicate information about the writing prompts to administrators, teachers and parents?

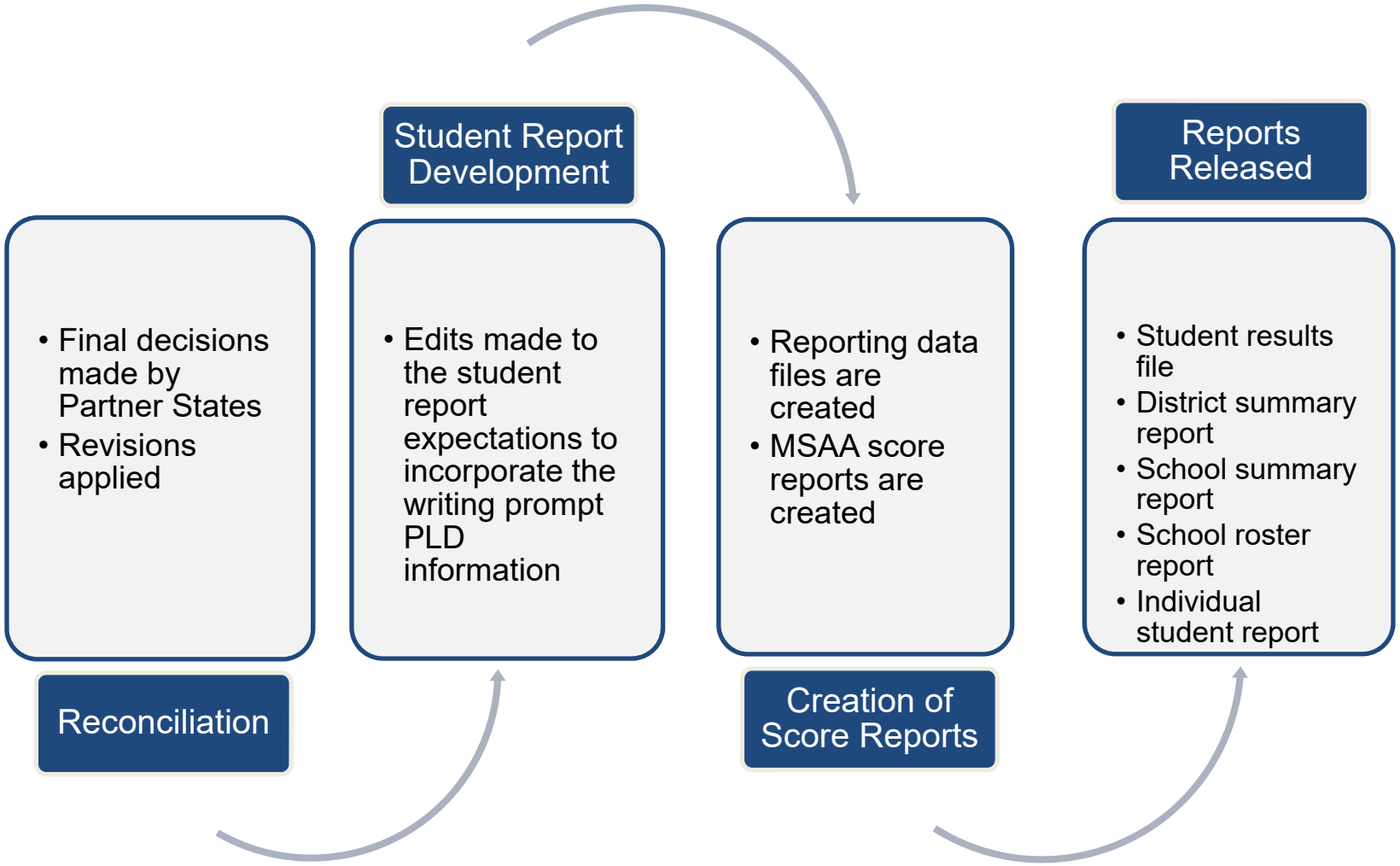
ELA PLD Review Process

1. For each grade, the facilitator will orient the group to the writing prompts, scoring rubrics and student samples
2. Reflect on these things individually, then discuss as a group for each grade
3. Reach consensus about any suggested edits to the PLDs
4. Provide suggestions on what information needs to be shared with the field and how best to go about providing that information

Expectations of All Panelists



Next Steps


Reconciliation
• Final decisions made by Partner States
• Revisions applied
Student Report Development
• Edits made to the student report expectations to incorporate the writing prompt PLD information
Creation of Score Reports
• Reporting data files are created
• MSAA score reports are created
Reports Released
• Student results file
• District summary report
• School summary report
• School roster report
• Individual student report

ELA PLD Review Groups

- Grades 3-5 (Narrative) Group
  - Room: Florida
- Grades 6-8 (Informational) Group
  - Room: Illinois
- Grade 11 (Persuasive) Group
  - Room: Utah

Thank you.

measured progress

msaa
Multi-State Alternate Assessment

Appendix I—MSAA ELA PLD Review Meeting Process Document

Measured Progress		
Role	**Name**	**Responsibility**
Oversight	Kelly Ickes	Handle morning registration of panelists Provide overview session training Float among the rooms during the ELA PLD Review process Assist panelists with accommodations Interact with facility staff
Facilitator	Megan Bairstow	Assist at the registration table Facilitate Grades 6-8 panel-Illinois room
Facilitator	Chris Clough	Assist at the registration table Facilitate Grade 11 panel-Utah room
Facilitator	Jim Kroening	Assist at the registration table Facilitate Grades 3-5 panel-Florida room

MSAA State Partners		
Role: Monitor process in each room; final reconciliation of PLD edits provided by the panelists		
Name	**State**	**Title**
Bethany Zimmerman	AZ	Director of Alternate Assessment
Hansley Mussotte	AZ	MSAA Project Coordinator
Michael Craig	DC	Assessment Specialist-Special Populations
Fasefulu Tigilau (Mr. T)	Guam/CNMI	CNMI PSS State Assessment Coordinator
Sue Nay	ME	Alternate Assessment & WIDA ACCESS Coordinator
Marsie Torchon	MD	Program Manager-Assessments for Students with Disabilities
Yvonne Field	MT	Assessment Coordinator
Megan Sellers	TN	Program Manager of Special Assessment
Melissa Flor	SD	

Facilitator Room Set-up

1. Number the Writing Prompt and Student Sample Booklets, Performance Level Descriptors (lower left corner), and ELA Blueprints (lower left corner)
2. Set out panel group meeting supplies (pens, highlighters, stickie-notes)
3. Have 2017-18 Test Administration Manual (TAM) and 2017 Guide for Score Report Interpretation available on materials table
    a. Purpose: meant as a reference doc should panelist conversation warrant review (may be valuable in discussions around Review questions 3 and 4).

Introductions

1. Ask each panelist to sign-in on the room sign-in sheet.
2. Welcome group, introduce yourself (name, affiliation, a little selected background information).
3. Have each participant introduce him/herself, including detail such as state, school, and background information, as well as noting whether they have administered the MSAA.
4. Ask each participant to sign a nondisclosure form. Do not proceed until a signed nondisclosure form has been collected from each participant. Mark on the Materials Tracking Sheet.

Review of Panelist Folder Materials

Overview: To help set the context for the meeting and the materials that will be used provide a brief review of what is in each panelist's folder.

- Left Side
    - Agenda
    - Non-Disclosure Form
    - Reimbursement Form
- Right Side
    - Opening PowerPoint
    - Writing Definition and Emphasis
    - Writing Prompt Rubric
    - Evaluation

Other ELA PLD Review Meeting Materials

Pass out the following ELA PLD Review Meeting Materials to panelists, marking the # on the Materials Tracking Sheet:

- Performance Level Descriptors (PLDs) (one grade at a time)
- Writing Prompt and Student Sample Booklet (one grade at a time)
- Note: Follow slip sheets in each booklet, do not worry about page numbering
- ELA Blueprints (one grade at a time)

Let panelists know that there is also a copy of the Test Administration Manual (1 per room) and Score Report Interpretation Guide (1 per room) available for review, if needed.

Review the Writing Prompts, Rubrics, and Student Samples

Overview: In order to establish an understanding of the writing prompts and for panelists to gain an understanding of what writing prompt evidence might look like from the students who take the test, each participant will review the Writing Prompt and Student Sample Booklet for a grade level. Panelists may wish to discuss or take issue with the writing prompts or scoring of the student samples. Tell them we will gladly take their feedback to MSAA, but that will not be the focus of discussion. This would be a strategy to help them move forward. (Panelists can write down feedback they have on stickie notes that can be provided to MSAA.) However, these are the actual writing prompts and scored samples of student writing evidence and these are what we have as reference in order to provide input on the ELA PLDs.

Activities

1. Introduce the Writing Prompts, Rubrics, and Student Samples:
2. Explain that there are two levels for the writing prompts: level 2 and level 3. Each student only takes one writing prompt.
3. Walk through the Directions for Test Administration (DTA) detail for each level writing prompt. Review corresponding writing prompt rubric for each level.
4. Have panelists review the Student Samples for each writing prompt level. These student samples came from the scoring anchor papers. These are just a sampling of scoring anchor papers spread across the scoring rubric ranges.
5. The purpose of the exercise is to help panelists establish a good understanding of the writing prompts administration and students experience and to gain an understanding of the writing prompt evidence that students create. Panelists may want to reference their Writing Prompt Considerations and Emphasis document when going through activities a-c to assist their understanding.
6. Remind panelists that they have access to the Test Administration Manual -- This manual outlines the following: participation guidelines, administration procedures, scoring directions, accommodations and assistive technology, security protocols, contact information, etc. One manual per room will be provided as a resource for participants. Remind panelists that this is available to refer to.

Discuss Performance Level Descriptors (PLDs) and Discuss the Four ELA PLD Review Questions

Overview: In order to establish an understanding of the expected performance of students on the writing prompt within each performance level, panelists will need to review the writing prompt information in the ELA PLDs and discuss.

Note: the ELA PLDs were established and finalized in 2015 with the standard setting process that was conducted. At that time performance expectations were established for the reading, writing stand-alone, and level 1 writing prompt items. The ELA PLDs were reviewed by the MSAA members of the Psychometric, Item Development and Scoring Subcommittees in Spring 2018. The writing prompt expectation for each performance level was added to the PLDs, no other expectation information was changed or edited. The Subcommittees referred to the Writing Prompt Considerations and Emphasis as part of the PLD update process. The revised PLDs were utilized during the Cut Score Review meeting to ensure there was alignment between the performance levels that the writing prompt score thresholds fell into and the expectations provided in the PLDs for the writing prompt.

Activities

1. Introduce the Performance Level Descriptors (PLDs):
2. Walk through the writing prompt detail that is in each performance level PLD (for a grade).
3. Have panelists discuss as a group and answer the following questions:
    (1) Does the open response writing prompt information contained in the PLD for each level accurately account for what the open response writing prompt is designed to measure?
    (2) Is the language clear and reflective of information that is understandable for administrators, teachers and parents/guardians?
    (3) How might the open response writing prompt measures and emphasis best be communicated to the field?
    (4) What avenues should be used to communicate information about the open response writing prompts to administrators, teachers and parents/guardians?

The purpose of the exercise is to have panelists provide suggestions to MSAA about possible edits to the ELA PLDs and suggestions on means of communicating the information to the field (including administrators teachers, and parents).

Remind panelists that they have access to the 2017 Guide for Score Report Interpretation -- This guide outlines the following: administration overview, score report descriptions, PLDs, ISR sample, writing prompt rubrics, parent overview, contact information, etc. One guide per room will be provided as a resource for participants. Remind panelists that this is available to refer to.

Facilitator collects feedback on each of the questions electronically-displays on screen.

Organization of Materials

Collect the following materials for each grade from panelists:

- Performance Level Descriptors
- Writing Prompt and Student Sample Booklets
- ELA Blueprints

Mark on the Materials Tracking Sheet when collected. Materials will be shipped back to Measured Progress for shredding.

Note: panelists may keep the following folder materials:

- Agenda
- Opening PowerPoint
- Writing Definition and Emphasis
- Writing Prompt Rubric
- Reimbursement Form

Complete Evaluation Form

Make sure panelists fill out the evaluation. Emphasize that their honest feedback is important. Mark on the Materials Tracking Sheet when collected.

Appendix J—ELA PLD Panelist Feedback

ELA PLD Panelist Feedback, August 9, 2018

Grades 3-5 - Narrative

1. Does the writing prompt information contained in the PLD for each level accurately account for what the writing prompt is designed to measure?

   - Level 4 – Panel agrees with change recommended by Grade 11 panel
     - Change “command” to “overall command”
   - Levels 3, 2, and 1 – Panel likes the wording as it is.
   - No other changes; comfortable with included PLD information

2. Is the language clear and reflective of information that is understandable for administrators, teachers, and parents?

   - Panelists agree: yes

3. How might the writing prompt measures and emphasis best be communicated to the field?

   - Critical: exemplar writing prompt and student evidence shared with the field in order to understand how writing is assessed and scored. This provides information about what the rubrics descriptions actually translate to in a student product.

4. What avenues should be used to communicate information about the writing prompts to administrators, teachers, and parents?

   - Information needs to be available through score report interpretation guide and websites.
   - Needs to be easily accessible to parents and other stakeholders.
   - Additional feedback was that printed reports about student performance are important for the field.

Grades 6-8 - Informational

1. Does the writing prompt information contained in the PLD for each level accurately account for what the writing prompt is designed to measure?

   - Level 4 – Panel agrees with change recommended by Grade 11 panel
     - Change “command” to “overall command”
   - Levels 3, 2, and 1 – Panel likes the wording as it is.
   - No other changes; comfortable with included PLD information
     - Discussion occurred: Do we need to define the describing words? What does it look like?
       - After talking it through as a group, decision was made that this was not needed. More valuable would be samples for the field (see question 3 below).

2. Is the language clear and reflective of information that is understandable for administrators, teachers, and parents?

   - Panelists agree: yes

3. How might the writing prompt measures and emphasis best be communicated to the field?

   - An exemplar writing prompt and student evidence needs to be shared with the field.
   - Writing prompt rubrics availability needs to be emphasized.
   - Writing prompt definition (considerations and emphasis) shared as part of the orientation training needs to be shared with the field.

4. What avenues should be used to communicate information about the writing prompts to administrators, teachers, and parents?

   - Through training, TAM, and state web pages.
   - Consider a MSAA specific web page for all to access, not just state level web pages.

Grades 11 - Persuasive

1. Does the writing prompt information contained in the PLD for each level accurately account for what the writing prompt is designed to measure?

   - Level 4 – Panel recommends changing “command” to “overall command.”
     - Refer to other discussion notes in the comments on PLD PDF.
   - Levels 3, 2, and 1 – Panel likes the wording as it is.
   - Panel feels there’s a good flow in the writing PLDs from Level 1 to Level 4.
   - Panel has a high level of confidence/consensus in their recommendations.

2. Is the language clear and reflective of information that is understandable for administrators, teachers, and parents?

   - Panelists agree: yes

3. How might the writing prompt measures and emphasis best be communicated to the field?

   - Additional information about the types of supports are available at each prompt level – especially helpful for non-test administrators (e.g., parents) and teachers developing IEP goals.
   - Additional ideas below.

4. What avenues should be used to communicate information about the writing prompts to administrators, teachers, and parents?

   - Face-to-face PD scenarios, webinar/PPT giving guidance to teachers (instructional, interacting with parents), or for parents directly.
   - Some schools run after-school meetings to involve and discuss with parents a variety of topics, including any assessment their child is taking (dinner, babysitters!)
   - Parent/Teacher conferences
   - Best Practice videos (new in 2018) were instrumental; changed the way some teachers approached the assessment.
   - Writing rubrics made widely available to teachers (other than interp guide; e.g., on DOE site).
   - Writing sample items (use with rubrics) available to teachers.
   - Item specs that explain standard by standard how each standard will be assessed, and connect these to the PLDs; also exemplars.
   - Parent materials indicating student performance compared to/where they will among other students.
   - Increase parent awareness; e.g.,
     - sample items via parent nights
     - parent-friendly overview video of the MSAA (greater overview, less investment required by parents than sample items; could be used during IEP meetings)

Add-On Notes: Observed Parent Responses to MSAA

- mixed responses, based largely around
  - student severity
  - parents' mindset (glad their student is doing something similar to gen ed assessments)
- increased awareness tends to limit fear and concern

Appendix K—MSAA ELA PLD Review Evaluation Forms

Grade(s): **Narrative Group: Grades 3-5**

	Not at all useful	Somewhat not useful	Neutral	Somewhat useful	Extremely useful
The session	☐	☐	1	2	5
The small group activities	☐	☐	☐	5	3
Becoming familiar with the assessment	☐	☐	☐	2	6
Discussions with other participants	☐	☐	☐	1	7

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I understood the goals of the ELA PLD Review meeting.	☐	☐	☐	2	6
The facilitator helped me understand the process.	☐	☐	☐	3	5
The materials contained the information needed to review the clarity of the ELA PLDs.	☐	☐	☐	4	5
I understood how to use the materials provided.	☐	☐	☐	2	4
The facilitator was able to provide answers to my questions.	☐	☐	☐	2	6
Sufficient time was allotted for training on the review process.	☐	☐	☐	2	6
Sufficient time was allotted to complete the review process.	☐	☐	☐	2	6

Comments/Suggestions:

- The 4 guiding questions were essential.
- I enjoyed discussing & collaborating with my colleagues; Great Process. Talking about areas we weren't clear in was extremely useful.
- Informative and well organized.
- There wasn't much communication provided beforehand about what to expect (especially related to meals & stipends). A schedule could have been provided beforehand.
- It would have been more helpful to have the opening sessions as a video to preview before our work. Since it was only one day, it would get me informed and give time to process everything before diving in (background and process). Just an opening speech for introductions and purpose would be great. Facilitators, food and accommodations were great. Thank you for the opportunity!
- The work was finished fairly quickly. I feel like the day was drawn out and work could have been completed in a ½ day. There was not enough communication prior to the meeting concerning schedule and work to be done

- Grade(s): **Informative Group: Grades 6-8**

	Not at all useful	Somewhat not useful	Neutral	Somewhat useful	Extremely useful
The session	☐	☐	☐	2	6
The small group activities	☐	☐	☐	☐	8
Becoming familiar with the assessment	☐	☐	☐	☐	8
Discussions with other participants	☐	☐	☐	☐	8

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I understood the goals of the ELA PLD Review meeting.	☐	☐	☐	2	6
The facilitator helped me understand the process.	☐	☐	☐	☐	8
The materials contained the information needed to review the clarity of the ELA PLDs.	☐	☐	☐	1	7
I understood how to use the materials provided.	☐	☐	☐	1	7
The facilitator was able to provide answers to my questions.	☐	☐	☐	☐	8
Sufficient time was allotted for training on the review process.	☐	☐	☐	☐	8
Sufficient time was allotted to complete the review process.	☐	☐	☐	☐	8

Comments/Suggestions:

- It was very helpful information. Thank you!
- Excellent job keeping group focused & providing feedback to keep us on track.
- Megan was great as always. I always learn something with each meeting I attend to take back to teachers and admins.
- Overall a good experience. It has been wonderful to watch the process as the years have progressed and see the test move to the next level.
- It was both helpful and distracting when additional State Leads joined the group. They shared helpful information at the end but checking email and side conversations were distracting. I wish they had either fully joined the activities or done their other work in a separate space.
- This session was very informative. I enjoyed the discussion and conversation on what is happening with the partner states. I was also very impressed with the samples provided. This gives hope that our students with SCD can succeed and produce the work. Thank You!!
- Attending this event was great. It allowed me to learn more about the MSAA. Looking at PLDs in relation to given prompts led to critical thinking and was a great spring board for rich discussion among the participants. Thank you! PS: The host/lead of the Grade6-8 group was so patient, understanding and knowledgeable. Thank you!

Grade(s): **Persuasive Group: Grades 11**

	Not at all useful	Somewhat not useful	Neutral	Somewhat useful	Extremely useful
The session	☐	☐	1	2	5
The small group activities	☐	☐	☐	☐	8
Becoming familiar with the assessment	☐	☐	1	2	5
Discussions with other participants	☐	☐	☐	☐	8

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I understood the goals of the ELA PLD Review meeting.	☐	☐	☐	1	7
The facilitator helped me understand the process.	☐	☐	☐	1	7
The materials contained the information needed to review the clarity of the ELA PLDs.	☐	☐	☐	1	7
I understood how to use the materials provided.	☐	☐	☐	1	7
The facilitator was able to provide answers to my questions.	☐	☐	☐	2	6
Sufficient time was allotted for training on the review process.	☐	☐	☐	2	6
Sufficient time was allotted to complete the review process.	☐	☐	☐	1	7

Comments/Suggestions:

- Very educational and helpful for me as a teacher!
- More review time for those less familiar with Test Administration.
- The booklet was difficult to navigate. Consider pagination & cut-outs. I would suggest providing the cut-out info as actual cut-outs for panelists.

APPENDIX N—RAW TO SCALED SCORE LOOK-UP TABLES

Table N-1. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 3

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
	0	1200	15.19	1	1200	18.3	1
	1	1200	12.04	1	1200	13.3	1
	2	1200	9.02	1	1201	8.9	1
	3	1204	6.87	1	1207	6.8	1
	4	1207	5.70	1	1211	5.7	1
	5	1210	4.96	1	1214	5.0	1
	6	1213	4.46	1	1217	4.6	1
	7	1215	4.10	1	1219	4.3	1
	8	1216	3.83	1	1221	4.1	1
	9	1218	3.62	1	1222	3.9	1
	10	1219	3.46	1	1224	3.8	1
	11	1221	3.34	1	1226	3.7	1
	12	1222	3.24	1	1227	3.6	1
	13	1223	3.17	1	1229	3.6	1
	14	1224	3.11	1	1230	3.6	1
	15	1225	3.08	1	1231	3.6	1
	16	1226	3.05	1	1233	3.6	1
	17	1227	3.04	1	1234	3.6	2
	18	1229	3.04	1	1236	3.7	2
	19	1230	3.04	1	1237	3.8	2
	20	1231	3.06	1	1239	3.9	2
	21	1232	3.09	1	1241	4.1	3
A	22	1233	3.13	1	1242	4.3	3
	23	1233	3.18	1	1244	4.5	3
	24	1235	3.24	2	1247	4.8	3
	25	1236	3.31	2	1249	5.2	3
	26	1237	3.39	2	1252	5.8	4
	27	1238	3.49	2	1255	6.6	4
	28	1240	3.60	3	1260	7.6	4
	29	1241	3.72	3	1265	9.4	4
	30	1242	3.86	3	1274	12.5	4
	31	1244	4.03	3	1289	20.1	4
	32	1245	4.21	3	1290	20.1	4
	33	1247	4.42	3	-	-	-
	34	1249	4.66	3	-	-	-
	35	1251	4.94	3	-	-	-
	36	1253	5.27	3	-	-	-
	37	1256	5.66	4	-	-	-
	38	1259	6.14	4	-	-	-
	39	1262	6.76	4	-	-	-
	40	1266	7.63	4	-	-	-
	41	1271	8.97	4	-	-	-
	42	1278	11.38	4	-	-	-
	43	1289	16.21	4	-	-	-
	44	1290	16.21	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	17.52	1	1200	19.4	1
	1	1200	13.83	1	1200	14.4	1
	2	1200	10.45	1	1201	10.0	1
	3	1204	7.85	1	1207	7.5	1
	4	1208	6.44	1	1211	6.3	1
	5	1212	5.56	1	1215	5.5	1
	6	1214	4.97	1	1217	5.0	1
	7	1217	4.54	1	1220	4.6	1
	8	1219	4.23	1	1222	4.3	1
	9	1220	4.00	1	1224	4.1	1
	10	1222	3.82	1	1226	4.0	1
	11	1224	3.69	1	1227	3.9	1
	12	1225	3.58	1	1229	3.8	1
	13	1226	3.50	1	1230	3.8	1
	14	1228	3.44	1	1232	3.7	1
	15	1229	3.40	1	1233	3.7	1
	16	1230	3.37	1	1235	3.8	2
	17	1231	3.35	1	1236	3.8	2
	18	1232	3.35	1	1238	3.9	2
	19	1233	3.36	1	1239	4.0	2
	20	1235	3.37	2	1241	4.1	3
	21	1236	3.40	2	1243	4.2	3
B	22	1237	3.44	2	1245	4.4	3
	23	1238	3.48	2	1247	4.7	3
	24	1240	3.54	3	1249	5.0	3
	25	1241	3.62	3	1252	5.5	4
	26	1242	3.70	3	1254	6.1	4
	27	1244	3.81	3	1258	6.9	4
	28	1245	3.93	3	1262	8.0	4
	29	1247	4.07	3	1268	9.8	4
	30	1248	4.23	3	1277	12.9	4
	31	1250	4.42	3	1289	18.7	4
	32	1252	4.64	3	1290	18.7	4
	33	1254	4.90	4	-	-	-
	34	1256	5.19	4	-	-	-
	35	1258	5.53	4	-	-	-
	36	1261	5.92	4	-	-	-
	37	1264	6.35	4	-	-	-
	38	1267	6.84	4	-	-	-
	39	1271	7.38	4	-	-	-
	40	1276	8.01	4	-	-	-
	41	1281	8.93	4	-	-	-
	42	1288	10.78	4	-	-	-
	43	1289	10.94	4	-	-	-
	44	1290	10.94	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	18.17	1	1200	19.4	1
	1	1200	13.94	1	1200	14.7	1
	2	1200	10.00	1	1200	10.3	1
	3	1206	7.52	1	1206	7.8	1
	4	1210	6.19	1	1211	6.5	1
	5	1213	5.37	1	1214	5.6	1
	6	1216	4.82	1	1217	5.1	1
	7	1218	4.43	1	1220	4.7	1
	8	1220	4.15	1	1222	4.4	1
	9	1222	3.93	1	1224	4.2	1
	10	1223	3.77	1	1225	4.1	1
	11	1225	3.64	1	1227	3.9	1
	12	1226	3.54	1	1229	3.9	1
	13	1227	3.46	1	1230	3.8	1
	14	1229	3.39	1	1232	3.8	1
	15	1230	3.34	1	1233	3.8	1
	16	1231	3.30	1	1235	3.8	2
	17	1232	3.27	1	1236	3.9	2
	18	1233	3.24	1	1238	3.9	2
	19	1235	3.23	2	1239	4.0	2
	20	1236	3.22	2	1241	4.2	3
	21	1237	3.22	2	1243	4.3	3
C	22	1238	3.23	2	1245	4.5	3
	23	1239	3.25	2	1247	4.8	3
	24	1240	3.28	3	1249	5.1	3
	25	1241	3.32	3	1252	3.9	4
	26	1243	3.37	3	1255	5.6	4
	27	1244	3.44	3	1258	6.2	4
	28	1245	3.53	3	1263	7.0	4
	29	1246	3.64	3	1268	8.1	4
	30	1248	3.78	3	1277	9.9	4
	31	1249	3.94	3	1289	13.1	4
	32	1251	4.14	3	1290	18.7	4
	33	1253	4.38	3	-	-	-
	34	1255	4.67	4	-	-	-
	35	1257	5.02	4	-	-	-
	36	1259	5.43	4	-	-	-
	37	1262	5.91	4	-	-	-
	38	1265	6.45	4	-	-	-
	39	1269	7.06	4	-	-	-
	40	1273	7.73	4	-	-	-
	41	1279	8.60	4	-	-	-
	42	1286	10.15	4	-	-	-
	43	1289	11.17	4	-	-	-
	44	1290	11.17	4	-	-	-

Table N-2. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 4

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	0	1200	15.31	1	1200	20.2	1
	1	1200	12.15	1	1200	15.6	1
	2	1200	9.16	1	1200	11.5	1
	3	1201	7.19	1	1202	8.6	1
	4	1205	6.07	1	1207	7.0	1
	5	1208	5.34	1	1211	6.0	1
	6	1211	4.82	1	1214	5.3	1
	7	1213	4.43	1	1216	4.8	1
	8	1215	4.13	1	1218	4.4	1
	9	1216	3.89	1	1220	4.1	1
	10	1218	3.69	1	1222	3.8	1
	11	1219	3.54	1	1224	3.7	1
	12	1221	3.41	1	1225	3.5	1
	13	1222	3.31	1	1227	3.4	1
	14	1223	3.22	1	1228	3.4	1
	15	1224	3.16	1	1230	3.4	1
	16	1225	3.11	1	1231	3.4	1
	17	1226	3.07	1	1232	3.4	1
	18	1227	3.05	1	1234	3.5	2
	19	1229	3.04	1	1235	3.7	2
	20	1230	3.04	1	1237	3.8	2
	21	1231	3.06	1	1239	4.0	2
	22	1232	3.09	1	1241	4.3	3
	23	1233	3.13	1	1243	4.6	3
	24	1234	3.18	2	1245	5.0	3
	25	1235	3.25	2	1248	5.6	3
	26	1236	3.33	2	1251	6.3	3
	27	1237	3.42	2	1255	7.2	3
	28	1238	3.53	2	1259	8.6	4
	29	1239	3.66	2	1266	10.6	4
	30	1241	3.80	3	1275	14.3	4
	31	1243	3.96	3	1288	20.7	4
	32	1244	4.14	3	1290	20.7	4
	33	1246	4.35	3	-	-	-
	34	1248	4.59	3	-	-	-
	35	1250	4.87	3	-	-	-
	36	1252	5.21	3	-	-	-
	37	1254	5.62	3	-	-	-
	38	1257	6.15	3	-	-	-
	39	1261	6.89	4	-	-	-
	40	1265	8.00	4	-	-	-
	41	1272	9.99	4	-	-	-
	42	1283	14.92	4	-	-	-
	43	1290	18.45	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
B	0	1200	16.74	1	1200	23.2	1
	1	1200	13.01	1	1200	17.4	1
	2	1200	9.55	1	1200	12.0	1
	3	1203	7.45	1	1204	8.6	1
	4	1207	6.26	1	1209	6.9	1
	5	1210	5.47	1	1213	5.8	1
	6	1213	4.91	1	1216	5.1	1
	7	1215	4.49	1	1219	4.6	1
	8	1217	4.17	1	1221	4.3	1
	9	1219	3.91	1	1223	4.0	1
	10	1220	3.71	1	1224	3.8	1
	11	1222	3.54	1	1226	3.7	1
	12	1223	3.41	1	1227	3.6	1
	13	1224	3.30	1	1229	3.5	1
	14	1226	3.22	1	1230	3.5	1
	15	1227	3.15	1	1232	3.4	1
	16	1228	3.10	1	1233	3.5	1
	17	1229	3.07	1	1235	3.5	2
	18	1230	3.05	1	1236	3.6	2
	19	1231	3.05	1	1238	3.7	2
	20	1232	3.05	1	1239	3.8	2
	21	1233	3.07	1	1241	4.0	3
	22	1234	3.11	2	1243	4.2	3
	23	1236	3.16	2	1245	4.6	3
	24	1237	3.22	2	1247	5.0	3
	25	1238	3.30	2	1250	5.5	3
	26	1239	3.40	2	1253	6.2	3
	27	1241	3.51	3	1256	7.2	3
	28	1242	3.65	3	1261	8.5	4
	29	1243	3.82	3	1267	10.6	4
	30	1245	4.02	3	1277	14.2	4
	31	1247	4.24	3	1288	19.7	4
	32	1248	4.50	3	1290	19.7	4
	33	1251	4.80	3	-	-	-
	34	1253	5.15	3	-	-	-
	35	1255	5.54	3	-	-	-
	36	1258	5.99	3	-	-	-
	37	1261	6.51	4	-	-	-
	38	1265	7.13	4	-	-	-
	39	1269	7.89	4	-	-	-
	40	1274	8.87	4	-	-	-
	41	1281	10.40	4	-	-	-
	42	1288	12.60	4	-	-	-
	43	1290	12.60	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	0	1200	18.28	1	1200	23.0	1
	1	1200	13.91	1	1200	17.6	1
	2	1200	9.97	1	1200	12.8	1
	3	1204	7.74	1	1203	9.3	1
	4	1208	6.49	1	1208	7.4	1
	5	1211	5.68	1	1212	6.3	1
	6	1214	5.10	1	1216	5.5	1
	7	1216	4.68	1	1218	5.0	1
	8	1218	4.35	1	1221	4.6	1
	9	1220	4.10	1	1223	4.3	1
	10	1222	3.90	1	1225	4.1	1
	11	1224	3.73	1	1226	4.0	1
	12	1225	3.60	1	1228	3.9	1
	13	1226	3.50	1	1230	3.8	1
	14	1228	3.42	1	1231	3.8	1
	15	1229	3.35	1	1233	3.8	1
	16	1230	3.30	1	1234	3.8	2
	17	1231	3.26	1	1236	3.9	2
	18	1233	3.24	1	1237	3.9	2
	19	1234	3.22	2	1239	4.1	2
	20	1235	3.22	2	1241	4.2	3
	21	1236	3.23	2	1243	4.4	3
	22	1237	3.24	2	1245	4.7	3
	23	1238	3.27	2	1247	5.0	3
	24	1239	3.32	2	1249	5.4	3
	25	1241	3.38	3	1252	5.9	3
	26	1242	3.45	3	1255	6.7	3
	27	1243	3.55	3	1259	7.7	4
	28	1245	3.67	3	1264	9.1	4
	29	1246	3.82	3	1271	11.4	4
	30	1248	3.99	3	1281	15.4	4
	31	1249	4.20	3	1288	18.6	4
	32	1251	4.44	3	1290	18.6	4
	33	1253	4.73	3	-	-	-
	34	1255	5.06	3	-	-	-
	35	1258	5.45	3	-	-	-
	36	1260	5.90	4	-	-	-
	37	1263	6.43	4	-	-	-
	38	1267	7.06	4	-	-	-
	39	1271	7.83	4	-	-	-
	40	1276	8.87	4	-	-	-
	41	1283	10.54	4	-	-	-
	42	1288	12.10	4	-	-	-
	43	1290	12.10	4	-	-	-

Table N-3. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 5

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
A	0	1200	18.32	1	1200	23.3	1
	1	1200	14.37	1	1200	15.6	1
	2	1200	10.77	1	1202	9.4	1
	3	1202	8.18	1	1208	7.0	1
	4	1206	6.76	1	1212	5.8	1
	5	1210	5.85	1	1215	5.0	1
	6	1212	5.22	1	1218	4.5	1
	7	1215	4.76	1	1220	4.1	1
	8	1217	4.41	1	1222	3.9	1
	9	1219	4.14	1	1224	3.7	1
	10	1220	3.93	1	1225	3.6	1
	11	1222	3.76	1	1227	3.5	1
	12	1223	3.62	1	1228	3.4	1
	13	1225	3.51	1	1229	3.4	1
	14	1226	3.43	1	1231	3.4	1
	15	1227	3.37	1	1232	3.4	2
	16	1228	3.33	1	1234	3.5	2
	17	1230	3.30	1	1235	3.6	2
	18	1231	3.29	1	1237	3.7	2
	19	1232	3.30	2	1238	3.8	2
	20	1233	3.32	2	1239	4.0	2
	21	1234	3.35	2	1242	4.3	3
	22	1235	3.40	2	1244	4.5	3
	23	1237	3.46	2	1246	4.9	3
	24	1238	3.54	2	1248	5.3	3
	25	1239	3.63	2	1251	5.8	3
	26	1240	3.73	3	1254	6.5	3
	27	1242	3.85	3	1258	7.4	4
	28	1243	3.98	3	1263	8.7	4
	29	1245	4.13	3	1270	10.8	4
	30	1246	4.31	3	1280	14.8	4
	31	1248	4.50	3	1290	20.5	4
	32	1250	4.73	3	1290	20.5	4
	33	1252	4.99	3	-	-	-
	34	1254	5.29	3	-	-	-
	35	1257	5.66	4	-	-	-
	36	1260	6.13	4	-	-	-
	37	1263	6.76	4	-	-	-
	38	1267	7.64	4	-	-	-
	39	1272	8.99	4	-	-	-
	40	1278	11.27	4	-	-	-
	41	1289	15.70	4	-	-	-
	42	1290	16.83	4	-	-	-
	43	1290	16.83	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
B	0	1200	21.90	1	1200	23.4	1
	1	1200	16.69	1	1200	16.1	1
	2	1200	12.05	1	1201	10.0	1
	3	1203	8.79	1	1208	7.5	1
	4	1208	7.08	1	1212	6.1	1
	5	1212	6.02	1	1216	5.3	1
	6	1215	5.30	1	1218	4.8	1
	7	1217	4.79	1	1220	4.4	1
	8	1219	4.41	1	1222	4.2	1
	9	1221	4.12	1	1224	4.0	1
	10	1223	3.90	1	1226	3.9	1
	11	1224	3.73	1	1228	3.8	1
	12	1226	3.61	1	1229	3.8	1
	13	1227	3.51	1	1231	3.8	1
	14	1229	3.45	1	1232	3.8	2
	15	1230	3.40	1	1234	3.8	2
	16	1232	3.38	2	1235	3.9	2
	17	1233	3.38	2	1237	4.0	2
	18	1234	3.39	2	1239	4.1	2
	19	1235	3.42	2	1240	4.2	3
	20	1236	3.47	2	1242	4.4	3
	21	1238	3.53	2	1244	4.6	3
	22	1239	3.60	2	1246	4.9	3
	23	1240	3.68	3	1248	5.2	3
	24	1242	3.78	3	1251	5.5	3
	25	1243	3.89	3	1254	6.0	3
	26	1245	4.01	3	1257	6.7	4
	27	1246	4.15	3	1261	7.5	4
	28	1248	4.30	3	1266	8.8	4
	29	1250	4.47	3	1272	10.9	4
	30	1252	4.67	3	1282	14.7	4
	31	1254	4.89	3	1290	19.1	4
	32	1255	5.14	3	1290	19.1	4
	33	1258	5.42	4	-	-	-
	34	1261	5.74	4	-	-	-
	35	1263	6.10	4	-	-	-
	36	1267	6.52	4	-	-	-
	37	1270	7.03	4	-	-	-
	38	1274	7.76	4	-	-	-
	39	1280	9.01	4	-	-	-
	40	1287	11.57	4	-	-	-
	41	1290	13.52	4	-	-	-
	42	1290	13.52	4	-	-	-
	43	1290	13.52	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	0	1200	21.89	1	1200	24.7	1
	1	1200	16.90	1	1200	17.2	1
	2	1200	12.54	1	1200	10.8	1
	3	1202	9.14	1	1207	8.0	1
	4	1207	7.37	1	1212	6.5	1
	5	1211	6.28	1	1215	5.6	1
	6	1214	5.54	1	1218	5.1	1
	7	1217	5.02	1	1221	4.7	1
	8	1219	4.63	1	1223	4.4	1
	9	1221	4.33	1	1225	4.2	1
	10	1223	4.10	1	1226	4.1	1
	11	1225	3.93	1	1228	4.0	1
	12	1226	3.79	1	1230	3.9	1
	13	1228	3.69	1	1232	3.9	2
	14	1229	3.62	1	1233	3.9	2
	15	1230	3.57	1	1234	3.9	2
	16	1232	3.55	2	1236	4.0	2
	17	1233	3.54	2	1238	4.1	2
	18	1234	3.55	2	1239	4.2	2
	19	1236	3.57	2	1241	4.4	3
	20	1237	3.60	2	1243	4.5	3
	21	1238	3.65	2	1245	4.7	3
	22	1239	3.72	2	1247	5.0	3
	23	1241	3.79	3	1250	5.3	3
	24	1242	3.88	3	1252	5.8	3
	25	1244	3.98	3	1255	6.3	3
	26	1245	4.09	3	1259	7.0	4
	27	1247	4.22	3	1263	7.9	4
	28	1249	4.37	3	1268	9.4	4
	29	1250	4.54	3	1275	11.7	4
	30	1252	4.74	3	1286	16.2	4
	31	1254	4.96	3	1290	18.7	4
	32	1257	5.22	4	1290	18.7	4
	33	1259	5.50	4	-	-	-
	34	1261	5.83	4	-	-	-
	35	1264	6.20	4	-	-	-
	36	1268	6.62	4	-	-	-
	37	1271	7.16	4	-	-	-
	38	1276	7.96	4	-	-	-
	39	1281	9.33	4	-	-	-
	40	1289	12.17	4	-	-	-
	41	1290	13.35	4	-	-	-
	42	1290	13.35	4	-	-	-
	43	1290	13.35	4	-	-	-

Table N-4. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 6

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
	0	1200	26.79	1	1200	24.8	1
	1	1200	13.45	1	1200	11.4	1
	2	1204	8.04	1	1205	6.9	1
	3	1209	6.06	1	1209	5.2	1
	4	1212	5.02	1	1212	4.3	1
	5	1215	4.37	1	1214	3.8	1
	6	1217	3.92	1	1216	3.5	1
	7	1219	3.58	1	1218	3.2	1
	8	1220	3.33	1	1219	3.1	1
	9	1222	3.13	1	1221	2.9	1
	10	1223	2.97	1	1222	2.9	1
	11	1224	2.85	1	1223	2.8	1
	12	1225	2.74	1	1224	2.8	1
	13	1226	2.66	1	1225	2.8	1
	14	1227	2.60	1	1226	2.8	1
	15	1228	2.55	1	1228	2.8	1
	16	1229	2.51	1	1229	2.9	1
	17	1230	2.48	1	1230	2.9	1
	18	1231	2.47	2	1231	3.0	2
	19	1232	2.47	2	1232	3.2	2
	20	1232	2.47	2	1234	3.3	2
	21	1233	2.49	2	1235	3.5	2
A	22	1234	2.51	2	1237	3.7	2
	23	1235	2.55	2	1239	4.0	2
	24	1236	2.59	2	1241	4.4	3
	25	1237	2.65	3	1243	4.8	3
	26	1238	2.71	3	1246	5.3	3
	27	1239	2.79	3	1249	6.0	3
	28	1240	2.88	3	1253	7.0	4
	29	1241	2.98	3	1258	8.4	4
	30	1242	3.10	3	1265	10.8	4
	31	1243	3.24	3	1277	16.3	4
	32	1244	3.40	3	1290	22.4	4
	33	1246	3.58	3	-	-	-
	34	1247	3.79	3	-	-	-
	35	1249	4.05	3	-	-	-
	36	1251	4.35	4	-	-	-
	37	1253	4.74	4	-	-	-
	38	1255	5.24	4	-	-	-
	39	1258	5.92	4	-	-	-
	40	1262	6.90	4	-	-	-
	41	1267	8.52	4	-	-	-
	42	1275	11.73	4	-	-	-
	43	1287	18.30	4	-	-	-
	44	1290	18.30	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	27.86	1	1200	27.0	1
	1	1200	14.35	1	1200	17.2	1
	2	1204	8.67	1	1201	9.6	1
	3	1210	6.56	1	1207	6.7	1
	4	1214	5.43	1	1211	5.3	1
	5	1216	4.72	1	1214	4.5	1
	6	1219	4.24	1	1216	4.0	1
	7	1221	3.88	1	1218	3.6	1
	8	1222	3.62	1	1220	3.4	1
	9	1224	3.40	1	1222	3.3	1
	10	1225	3.24	1	1223	3.2	1
	11	1226	3.10	1	1224	3.1	1
	12	1227	2.99	1	1226	3.1	1
	13	1228	2.90	1	1227	3.2	1
	14	1229	2.83	1	1228	3.2	1
	15	1231	2.78	2	1230	3.3	1
	16	1231	2.74	2	1231	3.4	2
	17	1232	2.71	2	1233	3.5	2
	18	1233	2.69	2	1234	3.7	2
	19	1234	2.69	2	1236	3.9	2
	20	1235	2.69	2	1238	4.1	2
	21	1236	2.71	2	1239	4.4	2
B	22	1237	2.74	3	1242	4.7	3
	23	1238	2.77	3	1244	5.0	3
	24	1239	2.82	3	1246	5.5	3
	25	1240	2.88	3	1249	6.0	3
	26	1241	2.96	3	1253	6.7	4
	27	1242	3.05	3	1257	7.7	4
	28	1243	3.15	3	1262	9.0	4
	29	1244	3.27	3	1269	11.2	4
	30	1245	3.41	3	1279	15.1	4
	31	1247	3.57	3	1287	18.9	4
	32	1248	3.77	3	1290	18.9	4
	33	1250	4.00	3	-	-	-
	34	1252	4.27	4	-	-	-
	35	1254	4.59	4	-	-	-
	36	1256	4.99	4	-	-	-
	37	1258	5.47	4	-	-	-
	38	1262	6.07	4	-	-	-
	39	1265	6.83	4	-	-	-
	40	1270	7.80	4	-	-	-
	41	1276	9.22	4	-	-	-
	42	1284	12.02	4	-	-	-
	43	1287	13.22	4	-	-	-
	44	1290	13.22	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	28.51	1	1200	27.6	1
	1	1200	14.36	1	1200	17.3	1
	2	1205	8.66	1	1201	9.6	1
	3	1210	6.56	1	1208	6.7	1
	4	1214	5.44	1	1212	5.2	1
	5	1217	4.74	1	1215	4.4	1
	6	1219	4.26	1	1217	3.9	1
	7	1221	3.91	1	1219	3.6	1
	8	1223	3.64	1	1220	3.4	1
	9	1224	3.43	1	1222	3.3	1
	10	1226	3.27	1	1223	3.2	1
	11	1227	3.13	1	1225	3.2	1
	12	1228	3.03	1	1226	3.2	1
	13	1229	2.94	1	1228	3.2	1
	14	1231	2.87	2	1229	3.3	1
	15	1231	2.82	2	1231	3.4	2
	16	1232	2.78	2	1232	3.5	2
	17	1233	2.76	2	1233	3.6	2
	18	1234	2.75	2	1235	3.8	2
	19	1235	2.74	2	1237	4.0	2
	20	1236	2.75	2	1239	4.3	2
	21	1237	2.77	3	1241	4.6	3
C	22	1238	2.80	3	1243	4.9	3
	23	1239	2.84	3	1245	5.3	3
	24	1240	2.90	3	1248	5.8	3
	25	1241	2.96	3	1251	6.5	3
	26	1242	3.03	3	1255	7.2	4
	27	1243	3.12	3	1259	8.3	4
	28	1244	3.22	3	1264	9.7	4
	29	1245	3.33	3	1272	11.9	4
	30	1246	3.46	3	1282	15.5	4
	31	1248	3.62	3	1287	17.6	4
	32	1249	3.80	3	1290	17.6	4
	33	1251	4.01	4	-	-	-
	34	1253	4.26	4	-	-	-
	35	1255	4.56	4	-	-	-
	36	1257	4.92	4	-	-	-
	37	1259	5.37	4	-	-	-
	38	1262	5.93	4	-	-	-
	39	1266	6.64	4	-	-	-
	40	1270	7.60	4	-	-	-
	41	1276	9.03	4	-	-	-
	42	1285	11.85	4	-	-	-
	43	1287	12.99	4	-	-	-
	44	1290	12.99	4	-	-	-

Table N-5. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 7

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
	0	1200	23.80	1	1200	23.9	1
	1	1200	14.85	1	1200	15.0	1
	2	1204	8.31	1	1205	8.7	1
	3	1210	6.08	1	1211	6.3	1
	4	1213	4.93	1	1215	5.0	1
	5	1216	4.22	1	1218	4.3	1
	6	1218	3.74	1	1220	3.8	1
	7	1220	3.40	1	1222	3.5	1
	8	1221	3.15	1	1223	3.3	1
	9	1222	2.96	1	1225	3.1	1
	10	1223	2.81	1	1226	3.0	1
	11	1225	2.69	1	1227	2.9	1
	12	1226	2.60	1	1228	2.8	1
	13	1227	2.53	1	1230	2.8	1
	14	1227	2.47	1	1231	2.8	1
	15	1228	2.43	1	1232	2.8	1
	16	1229	2.40	1	1233	2.9	1
	17	1230	2.39	1	1234	2.9	1
	18	1231	2.38	1	1236	3.0	2
	19	1232	2.39	1	1237	3.1	2
	20	1232	2.40	1	1238	3.3	2
	21	1233	2.43	1	1239	3.5	2
A	22	1234	2.46	1	1242	3.7	3
	23	1235	2.51	1	1243	4.0	3
	24	1236	2.56	2	1245	4.4	3
	25	1237	2.63	2	1248	4.9	3
	26	1238	2.70	2	1251	5.5	3
	27	1239	2.79	2	1254	6.3	3
	28	1239	2.89	2	1258	7.5	4
	29	1241	3.00	3	1264	9.2	4
	30	1242	3.13	3	1272	12.1	4
	31	1243	3.28	3	1285	18.4	4
	32	1245	3.46	3	1290	21.8	4
	33	1246	3.66	3	-	-	-
	34	1248	3.89	3	-	-	-
	35	1249	4.16	3	-	-	-
	36	1251	4.49	3	-	-	-
	37	1254	4.89	3	-	-	-
	38	1256	5.39	4	-	-	-
	39	1259	6.04	4	-	-	-
	40	1263	6.96	4	-	-	-
	41	1268	8.45	4	-	-	-
	42	1276	11.58	4	-	-	-
	43	1290	21.20	4	-	-	-
	44	1290	21.20	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	23.69	1	1200	23.2	1
	1	1200	17.02	1	1200	16.6	1
	2	1200	11.09	1	1202	10.9	1
	3	1207	8.03	1	1209	8.0	1
	4	1211	6.48	1	1214	6.4	1
	5	1215	5.53	1	1217	5.4	1
	6	1217	4.90	1	1220	4.7	1
	7	1220	4.45	1	1222	4.3	1
	8	1222	4.11	1	1224	4.0	1
	9	1223	3.86	1	1226	3.7	1
	10	1225	3.66	1	1228	3.6	1
	11	1226	3.50	1	1229	3.5	1
	12	1228	3.38	1	1231	3.4	1
	13	1229	3.28	1	1232	3.4	1
	14	1230	3.20	1	1233	3.4	1
	15	1231	3.15	1	1235	3.4	1
	16	1232	3.10	1	1236	3.5	2
	17	1233	3.08	1	1238	3.6	2
	18	1234	3.06	1	1239	3.7	2
	19	1236	3.06	2	1241	3.9	3
	20	1237	3.07	2	1243	4.1	3
	21	1238	3.09	2	1245	4.4	3
B	22	1239	3.12	2	1247	4.7	3
	23	1239	3.17	2	1249	5.0	3
	24	1241	3.22	3	1251	5.5	3
	25	1242	3.29	3	1255	6.0	4
	26	1243	3.38	3	1257	6.7	4
	27	1244	3.48	3	1261	7.5	4
	28	1246	3.60	3	1266	8.7	4
	29	1247	3.74	3	1272	10.3	4
	30	1249	3.90	3	1280	13.0	4
	31	1250	4.10	3	1290	17.9	4
	32	1252	4.32	3	1290	17.9	4
	33	1254	4.58	3	-	-	-
	34	1256	4.89	4	-	-	-
	35	1258	5.23	4	-	-	-
	36	1261	5.63	4	-	-	-
	37	1264	6.06	4	-	-	-
	38	1267	6.51	4	-	-	-
	39	1271	6.96	4	-	-	-
	40	1276	7.45	4	-	-	-
	41	1281	8.31	4	-	-	-
	42	1289	10.62	4	-	-	-
	43	1290	11.97	4	-	-	-
	44	1290	11.97	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	24.60	1	1200	23.0	1
	1	1200	17.15	1	1200	17.1	1
	2	1201	10.74	1	1200	11.8	1
	3	1208	7.77	1	1208	8.8	1
	4	1212	6.28	1	1213	7.0	1
	5	1215	5.38	1	1216	5.9	1
	6	1218	4.78	1	1219	5.1	1
	7	1220	4.35	1	1222	4.5	1
	8	1222	4.03	1	1224	4.2	1
	9	1224	3.79	1	1226	3.9	1
	10	1225	3.60	1	1228	3.7	1
	11	1227	3.45	1	1229	3.6	1
	12	1228	3.34	1	1231	3.5	1
	13	1229	3.24	1	1232	3.5	1
	14	1230	3.17	1	1234	3.5	1
	15	1232	3.12	1	1236	3.5	2
	16	1233	3.08	1	1237	3.6	2
	17	1234	3.06	1	1238	3.7	2
	18	1235	3.05	1	1239	3.9	2
	19	1236	3.05	2	1242	4.0	3
	20	1237	3.06	2	1243	4.3	3
	21	1238	3.08	2	1245	4.5	3
C	22	1239	3.12	2	1247	4.8	3
	23	1240	3.16	3	1250	5.2	3
	24	1241	3.22	3	1252	5.7	3
	25	1242	3.29	3	1255	6.3	4
	26	1244	3.38	3	1259	7.0	4
	27	1245	3.48	3	1263	7.9	4
	28	1246	3.60	3	1267	9.2	4
	29	1248	3.74	3	1274	11.0	4
	30	1249	3.90	3	1283	14.0	4
	31	1251	4.08	3	1290	17.6	4
	32	1252	4.30	3	1290	17.6	4
	33	1255	4.55	4	-	-	-
	34	1256	4.84	4	-	-	-
	35	1259	5.17	4	-	-	-
	36	1261	5.55	4	-	-	-
	37	1264	5.96	4	-	-	-
	38	1267	6.39	4	-	-	-
	39	1271	6.84	4	-	-	-
	40	1276	7.35	4	-	-	-
	41	1281	8.21	4	-	-	-
	42	1289	10.47	4	-	-	-
	43	1290	11.87	4	-	-	-
	44	1290	11.87	4	-	-	-

Table N-6. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 8

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
	0	1200	21.99	1	1200	23.0	1
	1	1200	14.90	1	1200	15.4	1
	2	1200	8.93	1	1200	9.4	1
	3	1205	6.56	1	1206	6.9	1
	4	1208	5.31	1	1210	5.6	1
	5	1211	4.53	1	1213	4.7	1
	6	1213	4.01	1	1216	4.2	1
	7	1215	3.63	1	1218	3.8	1
	8	1217	3.35	1	1219	3.5	1
	9	1218	3.14	1	1221	3.3	1
	10	1219	2.98	1	1222	3.1	1
	11	1221	2.85	1	1224	3.0	1
	12	1222	2.76	1	1225	3.0	1
	13	1223	2.68	1	1226	3.0	1
	14	1224	2.63	1	1227	3.0	1
	15	1225	2.59	1	1229	3.0	1
	16	1226	2.57	1	1229	3.1	1
	17	1226	2.56	1	1231	3.1	2
	18	1227	2.56	1	1233	3.2	2
	19	1228	2.58	1	1234	3.4	2
	20	1229	2.60	1	1235	3.5	2
	21	1230	2.63	2	1237	3.7	2
A	22	1231	2.67	2	1239	3.9	2
	23	1232	2.72	2	1241	4.2	3
	24	1233	2.78	2	1243	4.5	3
	25	1234	2.85	2	1245	4.9	3
	26	1235	2.93	2	1247	5.4	3
	27	1236	3.02	2	1251	6.1	4
	28	1237	3.12	2	1255	7.1	4
	29	1238	3.24	3	1260	8.8	4
	30	1239	3.37	3	1268	11.7	4
	31	1241	3.52	3	1282	18.8	4
	32	1242	3.70	3	1290	23.9	4
	33	1244	3.91	3	-	-	-
	34	1245	4.15	3	-	-	-
	35	1247	4.44	3	-	-	-
	36	1249	4.79	3	-	-	-
	37	1252	5.23	4	-	-	-
	38	1254	5.78	4	-	-	-
	39	1258	6.48	4	-	-	-
	40	1262	7.43	4	-	-	-
	41	1267	8.77	4	-	-	-
	42	1274	10.98	4	-	-	-
	43	1287	16.39	4	-	-	-
	44	1290	18.21	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	22.20	1	1200	22.8	1
	1	1200	13.89	1	1200	15.6	1
	2	1201	8.36	1	1200	9.8	1
	3	1206	6.24	1	1206	7.3	1
	4	1210	5.10	1	1210	5.9	1
	5	1212	4.40	1	1213	5.0	1
	6	1215	3.92	1	1216	4.4	1
	7	1216	3.57	1	1218	4.0	1
	8	1218	3.32	1	1220	3.7	1
	9	1219	3.13	1	1222	3.5	1
	10	1221	2.98	1	1223	3.4	1
	11	1222	2.87	1	1225	3.3	1
	12	1223	2.78	1	1226	3.3	1
	13	1224	2.72	1	1228	3.2	1
	14	1225	2.67	1	1229	3.3	1
	15	1226	2.64	1	1230	3.3	2
	16	1227	2.63	1	1232	3.4	2
	17	1228	2.62	1	1233	3.4	2
	18	1229	2.62	1	1234	3.6	2
	19	1229	2.64	1	1236	3.7	2
	20	1230	2.66	2	1238	3.8	2
	21	1231	2.69	2	1239	4.0	2
B	22	1232	2.73	2	1241	4.2	3
	23	1233	2.78	2	1243	4.5	3
	24	1234	2.84	2	1245	4.8	3
	25	1235	2.90	2	1248	5.3	3
	26	1236	2.98	2	1251	5.8	4
	27	1238	3.07	3	1254	6.6	4
	28	1239	3.17	3	1258	7.7	4
	29	1240	3.29	3	1264	9.4	4
	30	1241	3.43	3	1272	12.3	4
	31	1243	3.58	3	1286	19.2	4
	32	1244	3.76	3	1290	21.4	4
	33	1246	3.97	3	-	-	-
	34	1248	4.20	3	-	-	-
	35	1250	4.45	4	-	-	-
	36	1252	4.71	4	-	-	-
	37	1254	4.98	4	-	-	-
	38	1257	5.23	4	-	-	-
	39	1260	5.49	4	-	-	-
	40	1263	5.89	4	-	-	-
	41	1268	6.70	4	-	-	-
	42	1274	8.69	4	-	-	-
	43	1285	15.27	4	-	-	-
	44	1290	18.89	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	22.57	1	1200	23.0	1
	1	1200	14.69	1	1200	16.3	1
	2	1200	8.70	1	1200	10.7	1
	3	1206	6.41	1	1205	8.0	1
	4	1210	5.21	1	1210	6.4	1
	5	1212	4.46	1	1213	5.4	1
	6	1214	3.96	1	1216	4.7	1
	7	1216	3.61	1	1218	4.2	1
	8	1218	3.35	1	1220	3.9	1
	9	1219	3.15	1	1222	3.7	1
	10	1221	3.00	1	1224	3.5	1
	11	1222	2.89	1	1225	3.4	1
	12	1223	2.80	1	1227	3.4	1
	13	1224	2.74	1	1228	3.4	1
	14	1225	2.69	1	1229	3.4	1
	15	1226	2.66	1	1231	3.4	2
	16	1227	2.65	1	1233	3.5	2
	17	1228	2.64	1	1234	3.6	2
	18	1229	2.65	1	1236	3.7	2
	19	1229	2.67	1	1237	3.9	2
	20	1231	2.69	2	1239	4.1	2
	21	1232	2.72	2	1241	4.3	3
C	22	1233	2.77	2	1243	4.5	3
	23	1234	2.82	2	1245	4.9	3
	24	1235	2.88	2	1247	5.3	3
	25	1236	2.95	2	1250	5.7	4
	26	1237	3.03	2	1253	6.4	4
	27	1238	3.12	3	1257	7.3	4
	28	1239	3.23	3	1261	8.5	4
	29	1240	3.35	3	1267	10.4	4
	30	1242	3.49	3	1276	13.6	4
	31	1243	3.64	3	1290	19.8	4
	32	1245	3.82	3	1290	19.8	4
	33	1246	4.02	3	-	-	-
	34	1248	4.24	3	-	-	-
	35	1250	4.48	4	-	-	-
	36	1252	4.73	4	-	-	-
	37	1255	4.97	4	-	-	-
	38	1258	5.21	4	-	-	-
	39	1261	5.48	4	-	-	-
	40	1264	5.91	4	-	-	-
	41	1268	6.78	4	-	-	-
	42	1274	8.90	4	-	-	-
	43	1286	15.69	4	-	-	-
	44	1290	18.50	4	-	-	-

Table N-7. 2017–18 MSAA: Raw to Scaled Score Look-up Table — ELA Grade 11

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
	0	1200	19.16	1	1200	32.1	1
	1	1203	10.16	1	1200	20.8	1
	2	1211	6.45	1	1206	11.0	1
	3	1215	5.00	1	1214	6.9	1
	4	1217	4.20	1	1218	5.2	1
	5	1219	3.68	1	1221	4.3	1
	6	1221	3.32	1	1223	3.7	1
	7	1223	3.05	1	1225	3.3	1
	8	1224	2.85	1	1227	3.1	1
	9	1225	2.69	1	1228	2.9	1
	10	1226	2.57	1	1229	2.7	1
	11	1227	2.47	1	1231	2.6	1
	12	1228	2.39	1	1232	2.6	1
	13	1229	2.32	1	1233	2.6	1
	14	1230	2.27	1	1234	2.6	1
	15	1230	2.24	1	1235	2.6	1
	16	1231	2.21	1	1236	2.6	2
	17	1232	2.19	1	1237	2.7	2
	18	1233	2.18	1	1238	2.8	2
	19	1233	2.18	1	1239	2.9	2
	20	1234	2.19	1	1241	3.0	3
	21	1235	2.20	1	1242	3.2	3
A	22	1236	2.23	2	1244	3.4	3
	23	1236	2.25	2	1245	3.7	3
	24	1237	2.29	2	1247	4.1	3
	25	1238	2.34	2	1250	4.6	3
	26	1239	2.39	2	1253	5.4	3
	27	1240	2.45	3	1257	6.7	4
	28	1240	2.53	3	1262	8.7	4
	29	1241	2.61	3	1270	12.4	4
	30	1242	2.71	3	1285	20.1	4
	31	1243	2.83	3	1290	22.7	4
	32	1244	2.97	3	1290	22.7	4
	33	1246	3.13	3	-	-	-
	34	1247	3.33	3	-	-	-
	35	1248	3.56	3	-	-	-
	36	1250	3.85	3	-	-	-
	37	1252	4.21	3	-	-	-
	38	1254	4.65	3	-	-	-
	39	1257	5.23	4	-	-	-
	40	1260	5.98	4	-	-	-
	41	1264	7.05	4	-	-	-
	42	1270	8.86	4	-	-	-
	43	1281	13.57	4	-	-	-
	44	1290	20.30	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	19.42	1	1200	32.1	1
	1	1202	10.77	1	1200	20.6	1
	2	1210	6.77	1	1206	10.8	1
	3	1214	5.23	1	1214	7.0	1
	4	1217	4.39	1	1218	5.4	1
	5	1220	3.85	1	1221	4.5	1
	6	1221	3.47	1	1223	4.0	1
	7	1223	3.19	1	1225	3.6	1
	8	1224	2.99	1	1227	3.4	1
	9	1225	2.82	1	1228	3.2	1
	10	1227	2.69	1	1230	3.1	1
	11	1228	2.59	1	1231	3.0	1
	12	1229	2.51	1	1232	2.9	1
	13	1229	2.45	1	1234	2.8	1
	14	1230	2.40	1	1235	2.8	1
	15	1231	2.36	1	1236	2.8	2
	16	1232	2.34	1	1237	2.9	2
	17	1233	2.32	1	1238	2.9	2
	18	1234	2.32	1	1239	3.0	2
	19	1234	2.32	1	1241	3.1	3
	20	1235	2.33	1	1242	3.2	3
	21	1236	2.35	2	1243	3.4	3
B	22	1237	2.37	2	1245	3.6	3
	23	1238	2.41	2	1247	3.9	3
	24	1238	2.45	2	1249	4.3	3
	25	1239	2.51	2	1251	4.9	3
	26	1240	2.57	3	1254	5.7	3
	27	1241	2.64	3	1258	6.8	4
	28	1242	2.73	3	1264	8.7	4
	29	1243	2.84	3	1272	12.2	4
	30	1244	2.96	3	1286	19.5	4
	31	1245	3.10	3	1290	21.7	4
	32	1247	3.28	3	1290	21.7	4
	33	1248	3.49	3	-	-	-
	34	1250	3.74	3	-	-	-
	35	1251	4.04	3	-	-	-
	36	1253	4.42	3	-	-	-
	37	1256	4.88	4	-	-	-
	38	1259	5.46	4	-	-	-
	39	1262	6.18	4	-	-	-
	40	1266	7.05	4	-	-	-
	41	1272	8.09	4	-	-	-
	42	1279	9.53	4	-	-	-
	43	1289	13.19	4	-	-	-
	44	1290	13.63	4	-	-	-

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
	0	1200	20.44	1	1200	33.9	1
	1	1202	11.58	1	1200	20.1	1
	2	1210	7.22	1	1208	10.0	1
	3	1215	5.59	1	1215	6.7	1
	4	1218	4.71	1	1219	5.3	1
	5	1220	4.15	1	1222	4.5	1
	6	1222	3.76	1	1224	4.0	1
	7	1224	3.48	1	1226	3.6	1
	8	1225	3.26	1	1228	3.4	1
	9	1227	3.09	1	1230	3.2	1
	10	1228	2.95	1	1231	3.1	1
	11	1229	2.84	1	1232	3.0	1
	12	1230	2.75	1	1234	3.0	1
	13	1231	2.68	1	1235	3.0	1
	14	1232	2.62	1	1236	3.0	2
	15	1233	2.58	1	1237	3.0	2
	16	1234	2.54	1	1239	3.0	2
	17	1235	2.51	1	1239	3.1	2
	18	1236	2.49	2	1241	3.2	3
	19	1237	2.48	2	1243	3.4	3
	20	1237	2.48	2	1244	3.6	3
	21	1238	2.49	2	1246	3.8	3
C	22	1239	2.51	2	1248	4.1	3
	23	1240	2.54	3	1250	4.5	3
	24	1241	2.58	3	1252	5.0	3
	25	1242	2.63	3	1255	5.7	4
	26	1243	2.70	3	1259	6.6	4
	27	1244	2.78	3	1263	7.8	4
	28	1245	2.89	3	1269	9.7	4
	29	1246	3.00	3	1278	12.9	4
	30	1247	3.15	3	1290	18.5	4
	31	1248	3.31	3	1290	18.5	4
	32	1250	3.50	3	1290	18.5	4
	33	1251	3.73	3	-	-	-
	34	1253	3.99	3	-	-	-
	35	1255	4.30	4	-	-	-
	36	1257	4.66	4	-	-	-
	37	1260	5.08	4	-	-	-
	38	1262	5.59	4	-	-	-
	39	1266	6.20	4	-	-	-
	40	1270	6.93	4	-	-	-
	41	1275	7.88	4	-	-	-
	42	1281	9.43	4	-	-	-
	43	1290	12.59	4	-	-	-
	44	1290	12.59	4	-	-	-

Table N-8. 2017–18 MSAA: Raw to Scaled Score Look-up Table — Mathematics Grade 3

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
A	0	1200	30.67	1	1200	22.9	1
	1	1200	21.61	1	1200	16.7	1
	2	1200	13.38	1	1200	11.5	1
	3	1209	9.10	1	1207	8.9	1
	4	1215	6.98	1	1212	7.5	1
	5	1219	5.75	1	1216	6.6	1
	6	1222	4.96	1	1219	5.9	1
	7	1224	4.42	1	1222	5.5	1
	8	1226	4.04	1	1225	5.1	1
	9	1228	3.75	1	1227	4.8	1
	10	1230	3.54	1	1229	4.6	1
	11	1231	3.39	1	1231	4.4	1
	12	1232	3.27	1	1232	4.3	1
	13	1234	3.19	1	1234	4.1	1
	14	1235	3.13	2	1236	4.0	2
	15	1236	3.11	2	1237	4.0	2
	16	1237	3.10	2	1239	3.9	2
	17	1239	3.12	2	1240	3.9	3
	18	1240	3.16	2	1242	3.9	3
	19	1241	3.23	2	1243	4.0	3
	20	1243	3.32	3	1245	4.1	3
	21	1244	3.44	3	1247	4.2	3
	22	1245	3.59	3	1248	4.4	3
	23	1247	3.77	3	1250	4.6	3
	24	1249	3.99	3	1252	4.8	3
	25	1250	4.26	3	1254	5.1	4
	26	1252	4.58	3	1257	5.4	4
	27	1255	4.97	4	1259	5.8	4
	28	1257	5.44	4	1262	6.3	4
	29	1260	6.02	4	1265	6.9	4
	30	1263	6.78	4	1269	7.7	4
	31	1268	7.82	4	1274	8.8	4
	32	1273	9.40	4	1280	10.4	4
	33	1282	12.27	4	1288	13.1	4
	34	1290	19.45	4	1290	16.2	4
	35	1290	19.45	4	1290	16.2	4
B	0	1200	28.88	1	1200	25.4	1
	1	1200	18.69	1	1200	17.5	1
	2	1204	11.31	1	1203	11.4	1
	3	1212	8.37	1	1210	8.8	1
	4	1217	6.78	1	1215	7.4	1
	5	1220	5.79	1	1219	6.5	1
	6	1223	5.12	1	1222	5.8	1
	7	1226	4.63	1	1225	5.3	1
	8	1228	4.27	1	1227	4.9	1
	9	1230	3.99	1	1229	4.6	1

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
B	10	1231	3.78	1	1231	4.4	1
	11	1233	3.62	1	1233	4.2	1
	12	1234	3.49	1	1235	4.0	1
	13	1236	3.41	2	1237	3.8	2
	14	1237	3.34	2	1238	3.7	2
	15	1238	3.31	2	1239	3.6	2
	16	1240	3.30	2	1241	3.6	3
	17	1241	3.32	2	1242	3.6	3
	18	1242	3.36	3	1244	3.6	3
	19	1244	3.43	3	1245	3.7	3
	20	1245	3.53	3	1247	3.8	3
	21	1247	3.67	3	1248	3.9	3
	22	1248	3.85	3	1250	4.1	3
	23	1250	4.07	3	1252	4.3	3
	24	1252	4.34	3	1253	4.6	3
	25	1254	4.66	4	1256	4.9	4
	26	1256	5.05	4	1258	5.2	4
	27	1258	5.50	4	1260	5.6	4
	28	1261	6.05	4	1263	6.1	4
	29	1265	6.73	4	1266	6.6	4
	30	1268	7.60	4	1270	7.4	4
	31	1273	8.79	4	1274	8.4	4
	32	1280	10.59	4	1280	9.9	4
	33	1289	13.85	4	1288	12.7	4
	34	1290	16.89	4	1290	16.1	4
	35	1290	16.89	4	1290	16.1	4
C	0	1200	28.01	1	1200	25.6	1
	1	1200	15.87	1	1200	17.7	1
	2	1208	9.98	1	1203	11.6	1
	3	1215	7.55	1	1210	9.0	1
	4	1219	6.21	1	1215	7.6	1
	5	1223	5.36	1	1219	6.6	1
	6	1225	4.78	1	1222	5.9	1
	7	1228	4.36	1	1225	5.3	1
	8	1229	4.04	1	1228	4.8	1
	9	1231	3.79	1	1230	4.4	1
	10	1233	3.59	1	1232	4.1	1
	11	1234	3.44	1	1234	3.9	1
	12	1236	3.32	2	1236	3.7	2
	13	1237	3.24	2	1237	3.6	2
	14	1238	3.18	2	1238	3.6	2
	15	1240	3.14	2	1239	3.6	2
	16	1241	3.13	2	1241	3.6	3
	17	1242	3.14	3	1243	3.6	3
	18	1243	3.18	3	1244	3.7	3
	19	1245	3.24	3	1246	3.8	3
	20	1246	3.33	3	1247	3.9	3
	21	1247	3.44	3	1249	4.1	3

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	22	1249	3.59	3	1251	4.3	3
	23	1250	3.77	3	1253	4.6	3
	24	1252	3.99	3	1255	4.9	4
	25	1253	4.26	3	1257	5.3	4
	26	1256	4.58	4	1259	5.7	4
	27	1258	4.96	4	1262	6.1	4
	28	1260	5.42	4	1265	6.6	4
	29	1263	5.99	4	1268	7.2	4
	30	1267	6.72	4	1273	8.1	4
	31	1271	7.69	4	1277	9.2	4
	32	1276	9.12	4	1284	11.0	4
	33	1283	11.51	4	1290	14.2	4
	34	1290	16.87	4	1290	15.3	4
	35	1290	16.87	4	1290	15.3	4

Table N-9. 2017–18 MSAA: Raw to Scaled Score Look-up Table — Mathematics Grade 4

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	0	1200	31.98	1	1200	25.2	1
	1	1200	22.34	1	1200	19.9	1
	2	1200	14.08	1	1200	15.5	1
	3	1207	9.83	1	1200	11.7	1
	4	1213	7.60	1	1205	9.6	1
	5	1217	6.26	1	1210	8.2	1
	6	1221	5.38	1	1214	7.3	1
	7	1224	4.77	1	1218	6.6	1
	8	1226	4.32	1	1221	6.1	1
	9	1228	3.98	1	1224	5.6	1
	10	1230	3.73	1	1226	5.3	1
	11	1231	3.53	1	1228	5.0	1
	12	1233	3.39	2	1230	4.8	1
	13	1234	3.28	2	1232	4.6	1
	14	1235	3.21	2	1234	4.5	2
	15	1237	3.17	2	1236	4.4	2
	16	1238	3.15	2	1238	4.4	2
	17	1239	3.16	3	1239	4.3	2
	18	1241	3.20	3	1241	4.3	3
	19	1242	3.27	3	1243	4.4	3
	20	1243	3.37	3	1244	4.4	3
	21	1245	3.50	3	1246	4.5	3
	22	1246	3.67	3	1248	4.6	3
	23	1248	3.89	3	1250	4.8	3
	24	1250	4.18	3	1252	5.0	4
	25	1252	4.56	4	1254	5.2	4

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	26	1254	5.05	4	1257	5.6	4
	27	1257	5.70	4	1259	6.0	4
	28	1261	6.60	4	1262	6.6	4
	29	1265	7.86	4	1266	7.3	4
	30	1271	9.74	4	1270	8.4	4
	31	1279	12.74	4	1276	10.0	4
	32	1290	18.30	4	1283	12.4	4
	33	1290	18.41	4	1290	15.8	4
	34	1290	18.41	4	1290	15.8	4
	35	-	-	-	1290	15.8	4
B	0	1200	29.34	1	1200	27.8	1
	1	1200	21.22	1	1200	19.7	1
	2	1200	14.19	1	1200	13.1	1
	3	1204	10.27	1	1205	10.0	1
	4	1211	8.15	1	1211	8.3	1
	5	1215	6.85	1	1215	7.2	1
	6	1219	5.97	1	1219	6.5	1
	7	1222	5.34	1	1222	5.9	1
	8	1224	4.87	1	1225	5.5	1
	9	1227	4.51	1	1227	5.1	1
	10	1228	4.24	1	1229	4.9	1
	11	1230	4.03	1	1232	4.7	1
	12	1232	3.87	2	1233	4.5	2
	13	1234	3.75	2	1235	4.4	2
	14	1235	3.67	2	1237	4.3	2
	15	1237	3.62	2	1239	4.2	2
	16	1238	3.60	2	1240	4.2	3
	17	1239	3.61	3	1242	4.1	3
	18	1241	3.65	3	1244	4.1	3
	19	1242	3.71	3	1245	4.2	3
	20	1244	3.81	3	1247	4.2	3
	21	1245	3.93	3	1249	4.3	3
	22	1247	4.10	3	1250	4.4	3
	23	1249	4.30	3	1252	4.5	4
	24	1251	4.56	4	1254	4.8	4
	25	1253	4.89	4	1256	5.1	4
	26	1255	5.30	4	1259	5.4	4
	27	1258	5.82	4	1262	5.9	4
	28	1261	6.50	4	1265	6.6	4
	29	1265	7.42	4	1268	7.5	4
	30	1270	8.73	4	1273	8.6	4
	31	1276	10.71	4	1279	10.3	4
	32	1285	13.98	4	1286	12.7	4
	33	1290	17.15	4	1290	14.9	4
	34	1290	17.15	4	1290	14.9	4
	35	1290	17.15	4	1290	14.9	4

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	0	1200	28.20	1	1200	26.1	1
	1	1200	18.14	1	1200	19.5	1
	2	1202	11.63	1	1200	13.8	1
	3	1210	8.97	1	1204	10.7	1
	4	1215	7.52	1	1210	9.0	1
	5	1219	6.59	1	1215	7.9	1
	6	1222	5.95	1	1219	7.2	1
	7	1225	5.48	1	1222	6.6	1
	8	1228	5.10	1	1225	6.2	1
	9	1230	4.79	1	1228	5.9	1
	10	1232	4.53	2	1230	5.6	1
	11	1234	4.29	2	1232	5.4	1
	12	1236	4.07	2	1235	5.2	2
	13	1237	3.89	2	1237	5.0	2
	14	1239	3.72	3	1239	4.9	2
	15	1240	3.59	3	1241	4.8	3
	16	1242	3.48	3	1243	4.7	3
	17	1243	3.41	3	1244	4.7	3
	18	1245	3.38	3	1246	4.6	3
	19	1246	3.38	3	1248	4.7	3
	20	1247	3.43	3	1250	4.7	3
	21	1249	3.52	3	1252	4.9	4
	22	1250	3.65	3	1254	5.1	4
	23	1252	3.83	4	1256	5.3	4
	24	1253	4.05	4	1259	5.7	4
	25	1255	4.31	4	1261	6.1	4
	26	1257	4.63	4	1264	6.7	4
	27	1259	4.99	4	1268	7.4	4
	28	1262	5.42	4	1272	8.2	4
	29	1264	5.95	4	1276	9.1	4
	30	1268	6.63	4	1281	10.3	4
	31	1272	7.56	4	1287	11.8	4
	32	1277	8.99	4	1290	13.1	4
	33	1284	11.58	4	1290	13.1	4
	34	1290	15.05	4	1290	13.1	4
	35	1290	15.05	4	1290	13.1	4

Table N-10. 2017–18 MSAA: Raw to Scaled Score Look-up Table — Mathematics Grade 5

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
A	0	1200	39.66	1	1200	27.3	1
	1	1200	29.52	1	1200	19.7	1
	2	1200	20.61	1	1200	13.3	1
	3	1200	12.98	1	1206	9.9	1
	4	1207	9.32	1	1212	8.1	1
	5	1213	7.42	1	1217	6.9	1
	6	1217	6.28	1	1220	6.1	1
	7	1220	5.54	1	1223	5.6	1
	8	1223	5.04	1	1226	5.2	1
	9	1225	4.69	1	1228	4.9	1
	10	1227	4.44	1	1230	4.7	1
	11	1229	4.26	1	1232	4.6	2
	12	1231	4.15	1	1234	4.4	2
	13	1233	4.07	2	1236	4.4	2
	14	1235	4.04	2	1238	4.3	2
	15	1236	4.03	2	1239	4.3	2
	16	1238	4.05	2	1241	4.2	3
	17	1239	4.09	2	1243	4.3	3
	18	1241	4.15	3	1244	4.3	3
	19	1243	4.23	3	1246	4.3	3
	20	1245	4.33	3	1248	4.4	3
	21	1247	4.46	3	1250	4.5	3
	22	1248	4.60	3	1252	4.7	3
	23	1250	4.78	3	1254	4.9	3
	24	1253	4.99	4	1256	5.1	4
	25	1255	5.25	4	1258	5.4	4
	26	1257	5.57	4	1261	5.8	4
	27	1260	5.99	4	1264	6.3	4
	28	1263	6.54	4	1267	7.0	4
	29	1267	7.32	4	1271	7.8	4
	30	1272	8.46	4	1275	8.9	4
	31	1278	10.28	4	1281	10.3	4
	32	1287	13.59	4	1288	12.4	4
	33	1290	16.75	4	1290	14.2	4
	34	1290	16.75	4	1290	14.2	4
	35	1290	16.75	4	1290	14.2	4
B	0	1200	33.68	1	1200	29.3	1
	1	1200	22.93	1	1200	21.1	1
	2	1200	14.69	1	1200	14.2	1
	3	1206	10.69	1	1207	10.6	1
	4	1213	8.71	1	1213	8.5	1
	5	1218	7.51	1	1218	7.3	1
	6	1222	6.71	1	1221	6.4	1
	7	1225	6.14	1	1224	5.8	1
	8	1228	5.72	1	1227	5.4	1
	9	1230	5.39	1	1230	5.1	1

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
B	10	1232	5.14	2	1232	4.9	2
	11	1235	4.95	2	1234	4.7	2
	12	1237	4.79	2	1236	4.6	2
	13	1238	4.68	2	1238	4.5	2
	14	1240	4.59	3	1239	4.4	2
	15	1242	4.53	3	1241	4.4	3
	16	1244	4.49	3	1243	4.4	3
	17	1246	4.48	3	1245	4.4	3
	18	1247	4.50	3	1247	4.5	3
	19	1249	4.54	3	1248	4.5	3
	20	1251	4.60	3	1250	4.6	3
	21	1253	4.70	4	1252	4.8	3
	22	1255	4.83	4	1254	4.9	3
	23	1257	5.01	4	1256	5.2	4
	24	1259	5.23	4	1258	5.4	4
	25	1261	5.51	4	1261	5.8	4
	26	1264	5.86	4	1264	6.2	4
	27	1266	6.32	4	1267	6.8	4
	28	1270	6.91	4	1270	7.4	4
	29	1273	7.70	4	1274	8.3	4
	30	1278	8.79	4	1279	9.5	4
	31	1284	10.39	4	1285	11.1	4
	32	1290	13.04	4	1290	13.4	4
	33	1290	13.72	4	1290	13.6	4
	34	1290	13.72	4	1290	13.6	4
	35	1290	13.72	4	1290	13.6	4
C	0	1200	29.92	1	1200	28.7	1
	1	1200	17.80	1	1200	21.0	1
	2	1205	11.36	1	1200	14.5	1
	3	1213	8.83	1	1206	10.9	1
	4	1218	7.44	1	1212	8.9	1
	5	1222	6.55	1	1217	7.6	1
	6	1225	5.93	1	1221	6.7	1
	7	1228	5.47	1	1224	6.1	1
	8	1230	5.12	1	1227	5.7	1
	9	1232	4.84	2	1229	5.4	1
	10	1234	4.62	2	1232	5.2	2
	11	1236	4.45	2	1234	5.0	2
	12	1238	4.30	2	1236	4.9	2
	13	1239	4.19	2	1238	4.8	2
	14	1241	4.11	3	1240	4.8	3
	15	1243	4.05	3	1242	4.7	3
	16	1245	4.01	3	1244	4.7	3
	17	1246	4.00	3	1246	4.7	3
	18	1248	4.00	3	1248	4.7	3
	19	1249	4.03	3	1250	4.8	3
	20	1251	4.07	3	1252	4.9	3
	21	1252	4.14	3	1254	5.0	3

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	22	1254	4.24	4	1256	5.2	4
	23	1256	4.36	4	1258	5.4	4
	24	1257	4.51	4	1260	5.7	4
	25	1259	4.71	4	1263	6.1	4
	26	1261	4.96	4	1266	6.6	4
	27	1264	5.27	4	1269	7.2	4
	28	1266	5.67	4	1273	8.0	4
	29	1269	6.19	4	1277	9.1	4
	30	1273	6.90	4	1283	10.4	4
	31	1277	7.91	4	1289	12.1	4
	32	1282	9.47	4	1290	13.4	4
	33	1290	12.26	4	1290	13.4	4
	34	1290	13.70	4	1290	13.4	4
	35	1290	13.70	4	1290	13.4	4

Table N-11. 2017–18 MSAA: Raw to Scaled Score Look-up Table — Mathematics Grade 6

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	0	1200	23.83	1	1200	21.0	1
	1	1200	16.02	1	1200	14.2	1
	2	1201	9.97	1	1201	9.5	1
	3	1208	7.55	1	1207	7.5	1
	4	1212	6.25	1	1211	6.4	1
	5	1215	5.44	1	1215	5.7	1
	6	1218	4.88	1	1217	5.2	1
	7	1220	4.48	1	1220	4.8	1
	8	1222	4.19	1	1222	4.5	1
	9	1224	3.96	1	1224	4.3	1
	10	1226	3.79	1	1225	4.1	1
	11	1227	3.66	1	1227	4.0	1
	12	1228	3.57	1	1229	3.8	1
	13	1230	3.50	1	1230	3.8	1
	14	1231	3.46	1	1231	3.7	1
	15	1232	3.44	1	1233	3.7	1
	16	1234	3.44	2	1234	3.6	2
	17	1235	3.45	2	1235	3.6	2
	18	1236	3.49	2	1237	3.7	2
	19	1238	3.54	2	1238	3.7	2
	20	1239	3.61	3	1239	3.7	2
	21	1240	3.70	3	1241	3.8	3
	22	1242	3.81	3	1242	3.9	3
	23	1243	3.94	3	1244	4.1	3
	24	1245	4.10	3	1246	4.2	3
	25	1247	4.29	3	1247	4.4	3

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	26	1249	4.52	3	1249	4.7	4
	27	1251	4.80	4	1252	5.0	4
	28	1253	5.16	4	1254	5.4	4
	29	1256	5.62	4	1257	5.9	4
	30	1259	6.23	4	1260	6.5	4
	31	1262	7.10	4	1264	7.4	4
	32	1267	8.45	4	1269	8.9	4
	33	1274	10.91	4	1276	11.4	4
	34	1287	17.30	4	1290	17.7	4
	35	1290	20.61	4	1290	19.4	4
B	0	1200	26.41	1	1200	24.0	1
	1	1200	13.62	1	1200	13.2	1
	2	1206	8.69	1	1206	8.7	1
	3	1212	6.82	1	1211	6.9	1
	4	1216	5.80	1	1215	5.9	1
	5	1219	5.16	1	1218	5.3	1
	6	1221	4.72	1	1220	4.8	1
	7	1223	4.39	1	1223	4.5	1
	8	1225	4.15	1	1225	4.2	1
	9	1227	3.96	1	1226	4.0	1
	10	1229	3.81	1	1228	3.8	1
	11	1230	3.69	1	1229	3.7	1
	12	1232	3.60	1	1231	3.6	1
	13	1233	3.53	2	1232	3.6	1
	14	1234	3.48	2	1234	3.5	2
	15	1236	3.44	2	1235	3.5	2
	16	1237	3.42	2	1236	3.5	2
	17	1238	3.42	2	1237	3.5	2
	18	1239	3.42	3	1239	3.5	2
	19	1241	3.44	3	1239	3.5	2
	20	1242	3.48	3	1241	3.6	3
	21	1243	3.53	3	1243	3.7	3
	22	1245	3.61	3	1244	3.8	3
	23	1246	3.70	3	1246	3.9	3
	24	1248	3.82	3	1247	4.0	3
	25	1249	3.97	3	1249	4.2	4
	26	1251	4.16	4	1251	4.5	4
	27	1253	4.41	4	1253	4.8	4
	28	1255	4.72	4	1255	5.1	4
	29	1257	5.15	4	1258	5.6	4
	30	1260	5.73	4	1261	6.2	4
	31	1264	6.58	4	1264	7.1	4
	32	1268	7.95	4	1269	8.5	4
	33	1275	10.50	4	1276	10.9	4
	34	1288	17.23	4	1289	17.1	4
	35	1290	19.89	4	1290	19.1	4

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	0	1200	28.17	1	1200	24.7	1
	1	1203	11.19	1	1200	14.9	1
	2	1211	7.61	1	1204	9.6	1
	3	1216	6.15	1	1210	7.5	1
	4	1219	5.33	1	1214	6.4	1
	5	1222	4.79	1	1217	5.6	1
	6	1224	4.41	1	1220	5.1	1
	7	1226	4.13	1	1222	4.8	1
	8	1228	3.91	1	1224	4.5	1
	9	1230	3.73	1	1226	4.3	1
	10	1232	3.59	1	1228	4.1	1
	11	1233	3.47	2	1230	4.0	1
	12	1234	3.38	2	1231	3.9	1
	13	1236	3.30	2	1233	3.8	1
	14	1237	3.24	2	1234	3.7	2
	15	1238	3.19	2	1236	3.7	2
	16	1239	3.16	3	1237	3.7	2
	17	1241	3.14	3	1238	3.7	2
	18	1242	3.13	3	1239	3.7	2
	19	1243	3.14	3	1241	3.7	3
	20	1244	3.16	3	1243	3.8	3
	21	1245	3.19	3	1244	3.9	3
	22	1247	3.24	3	1246	4.0	3
	23	1248	3.31	3	1247	4.2	3
	24	1249	3.41	3	1249	4.4	4
	25	1251	3.53	4	1251	4.6	4
	26	1252	3.68	4	1253	4.9	4
	27	1254	3.88	4	1255	5.2	4
	28	1256	4.15	4	1258	5.7	4
	29	1258	4.50	4	1261	6.2	4
	30	1260	4.98	4	1264	7.0	4
	31	1263	5.68	4	1268	8.0	4
	32	1267	6.78	4	1274	9.7	4
	33	1273	8.79	4	1283	12.8	4
	34	1283	13.89	4	1290	17.5	4
	35	1290	20.13	4	1290	17.5	4

Table N-12. 2017–18 MSAA: Raw to Scaled Score Look-up Table — Mathematics Grade 7

Path	Raw Score	2018 Scaled Score	2018 Standard Error	2018 Performance Level	2017 Scaled Score	2017 Standard Error	2017 Performance Level
A	0	1200	23.97	1	1200	22.3	1
	1	1200	15.86	1	1200	15.8	1
	2	1203	9.79	1	1202	10.5	1
	3	1209	7.44	1	1209	8.1	1
	4	1214	6.19	1	1215	6.8	1
	5	1217	5.43	1	1219	6.0	1
	6	1220	4.92	1	1222	5.4	1
	7	1222	4.56	1	1224	5.0	1
	8	1224	4.30	1	1227	4.7	1
	9	1226	4.12	1	1229	4.5	1
	10	1227	3.98	1	1231	4.3	1
	11	1229	3.88	1	1233	4.2	2
	12	1231	3.81	1	1235	4.1	2
	13	1232	3.77	1	1236	4.0	2
	14	1233	3.74	1	1238	4.0	2
	15	1235	3.73	2	1239	3.9	2
	16	1236	3.74	2	1241	3.9	3
	17	1238	3.76	2	1243	3.9	3
	18	1239	3.80	2	1245	4.0	3
	19	1241	3.85	3	1247	4.0	3
	20	1242	3.92	3	1248	4.1	3
	21	1244	4.01	3	1250	4.2	3
	22	1245	4.12	3	1252	4.3	3
	23	1247	4.25	3	1254	4.4	4
	24	1249	4.42	3	1256	4.6	4
	25	1250	4.62	3	1259	4.8	4
	26	1252	4.88	3	1261	5.1	4
	27	1255	5.19	4	1264	5.4	4
	28	1257	5.60	4	1268	5.8	4
	29	1260	6.13	4	1272	6.4	4
	30	1264	6.86	4	1277	7.1	4
	31	1268	7.91	4	1284	8.1	4
	32	1274	9.60	4	1290	9.6	4
	33	1282	12.78	4	1290	12.4	4
	34	1290	19.26	4	1290	17.6	4
	35	1290	19.26	4	-	-	-
B	0	1200	27.45	1	1200	17.6	1
	1	1200	17.41	1	1200	24.5	1
	2	1204	10.54	1	1202	16.9	1
	3	1211	8.00	1	1210	11.0	1
	4	1216	6.66	1	1214	8.4	1
	5	1219	5.83	1	1218	7.1	1
	6	1222	5.27	1	1221	6.2	1
	7	1225	4.87	1	1224	5.6	1
	8	1227	4.57	1	1226	5.2	1
	9	1229	4.34	1	1228	4.8	1

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
B	10	1231	4.17	1	1230	4.6	1
	11	1232	4.03	1	1232	4.3	2
	12	1234	3.93	2	1233	4.2	2
	13	1235	3.85	2	1235	4.1	2
	14	1237	3.79	2	1236	4.0	2
	15	1238	3.75	2	1238	4.0	2
	16	1239	3.73	2	1239	4.0	2
	17	1241	3.73	3	1241	4.0	3
	18	1242	3.74	3	1243	4.1	3
	19	1244	3.77	3	1244	4.1	3
	20	1245	3.81	3	1246	4.2	3
	21	1246	3.87	3	1247	4.3	3
	22	1248	3.96	3	1249	4.4	3
	23	1250	4.07	3	1251	4.6	3
	24	1251	4.21	3	1253	4.8	3
	25	1253	4.38	3	1255	5.0	4
	26	1255	4.61	4	1257	5.3	4
	27	1257	4.90	4	1259	5.6	4
	28	1259	5.27	4	1262	6.1	4
	29	1262	5.79	4	1265	6.6	4
	30	1265	6.51	4	1269	7.4	4
	31	1270	7.61	4	1273	8.5	4
	32	1275	9.48	4	1279	10.2	4
	33	1284	13.42	4	1288	13.3	4
	34	1290	19.34	4	1290	16.4	4
	35	1290	19.34	4	1290	16.4	4
C	0	1200	27.38	1	1200	25.4	1
	1	1201	13.17	1	1200	17.8	1
	2	1211	8.76	1	1202	11.7	1
	3	1217	6.94	1	1209	8.9	1
	4	1221	5.92	1	1215	7.4	1
	5	1224	5.26	1	1219	6.5	1
	6	1226	4.80	1	1222	5.8	1
	7	1229	4.45	1	1224	5.4	1
	8	1230	4.18	1	1227	5.1	1
	9	1232	3.96	1	1229	4.8	1
	10	1234	3.77	2	1231	4.6	1
	11	1235	3.62	2	1233	4.5	2
	12	1237	3.50	2	1235	4.4	2
	13	1238	3.40	2	1236	4.3	2
	14	1239	3.32	2	1238	4.3	2
	15	1241	3.26	3	1239	4.3	2
	16	1242	3.23	3	1241	4.3	3
	17	1243	3.21	3	1243	4.3	3
	18	1244	3.21	3	1245	4.4	3
	19	1245	3.23	3	1247	4.5	3
	20	1247	3.27	3	1248	4.6	3
	21	1248	3.32	3	1250	4.7	3

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	22	1249	3.40	3	1252	4.9	3
	23	1251	3.49	3	1254	5.1	4
	24	1252	3.61	3	1256	5.3	4
	25	1254	3.75	4	1259	5.7	4
	26	1255	3.93	4	1261	6.1	4
	27	1257	4.16	4	1264	6.6	4
	28	1259	4.44	4	1268	7.3	4
	29	1261	4.80	4	1272	8.3	4
	30	1263	5.29	4	1277	9.7	4
	31	1266	5.97	4	1284	11.9	4
	32	1270	7.00	4	1290	15.7	4
	33	1276	8.79	4	1290	15.7	4
	34	1286	13.06	4	1290	15.7	4
	35	1290	18.91	4	-	-	-

Table N-13. Mathematics Grade 8

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	0	1200	36.30	1	1200	24.1	1
	1	1200	21.11	1	1200	16.9	1
	2	1203	10.98	1	1200	11.2	1
	3	1210	7.68	1	1207	8.6	1
	4	1215	6.11	1	1212	7.1	1
	5	1218	5.19	1	1216	6.2	1
	6	1221	4.60	1	1219	5.5	1
	7	1223	4.20	1	1222	5.1	1
	8	1225	3.92	1	1224	4.7	1
	9	1227	3.72	1	1226	4.5	1
	10	1228	3.58	1	1228	4.3	1
	11	1230	3.49	1	1229	4.2	1
	12	1231	3.43	1	1231	4.1	1
	13	1232	3.40	1	1233	4.0	1
	14	1233	3.39	1	1234	4.0	2
	15	1235	3.40	2	1236	4.0	2
	16	1236	3.43	2	1237	4.0	2
	17	1238	3.47	2	1239	4.0	2
	18	1239	3.52	2	1240	4.0	3
	19	1240	3.58	3	1242	4.0	3
	20	1242	3.65	3	1243	4.1	3
	21	1243	3.73	3	1245	4.2	3
	22	1245	3.83	3	1246	4.3	3
	23	1246	3.95	3	1248	4.4	3
	24	1248	4.09	3	1250	4.5	4
	25	1250	4.25	3	1252	4.7	4
	26	1251	4.46	4	1254	4.9	4

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	27	1253	4.72	4	1256	5.2	4
	28	1256	5.04	4	1258	5.6	4
	29	1258	5.48	4	1261	6.1	4
	30	1261	6.08	4	1264	6.7	4
	31	1265	6.96	4	1268	7.6	4
	32	1270	8.40	4	1274	9.1	4
	33	1277	11.24	4	1281	11.7	4
	34	1290	20.63	4	1290	18.1	4
	35	1290	21.44	4	1290	18.1	4
B	0	1200	32.16	1	1200	26.4	1
	1	1200	15.54	1	1200	15.9	1
	2	1207	9.42	1	1204	10.1	1
	3	1214	7.26	1	1211	7.9	1
	4	1218	6.11	1	1215	6.7	1
	5	1221	5.39	1	1219	6.0	1
	6	1224	4.89	1	1221	5.4	1
	7	1226	4.53	1	1224	5.1	1
	8	1228	4.25	1	1226	4.8	1
	9	1230	4.04	1	1228	4.6	1
	10	1231	3.88	1	1230	4.4	1
	11	1233	3.75	1	1232	4.3	1
	12	1234	3.66	2	1233	4.2	1
	13	1236	3.58	2	1235	4.1	2
	14	1237	3.53	2	1236	4.0	2
	15	1238	3.49	2	1238	4.0	2
	16	1239	3.48	2	1239	4.0	2
	17	1241	3.47	3	1241	4.0	3
	18	1242	3.49	3	1242	4.0	3
	19	1243	3.51	3	1244	4.0	3
	20	1245	3.55	3	1245	4.1	3
	21	1246	3.61	3	1247	4.2	3
	22	1247	3.68	3	1249	4.2	4
	23	1249	3.77	3	1250	4.4	4
	24	1250	3.89	3	1252	4.5	4
	25	1252	4.04	4	1253	4.7	4
	26	1253	4.22	4	1255	4.9	4
	27	1255	4.44	4	1258	5.2	4
	28	1257	4.74	4	1260	5.5	4
	29	1260	5.12	4	1263	6.0	4
	30	1262	5.64	4	1266	6.6	4
	31	1266	6.39	4	1270	7.5	4
	32	1270	7.58	4	1275	8.9	4
	33	1276	9.81	4	1282	11.4	4
	34	1289	16.35	4	1290	17.3	4
	35	1290	20.29	4	1290	17.3	4

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	0	1200	27.99	1	1200	27.0	1
	1	1200	13.60	1	1200	17.8	1
	2	1209	8.88	1	1202	11.2	1
	3	1215	6.98	1	1209	8.6	1
	4	1219	5.92	1	1214	7.2	1
	5	1222	5.25	1	1218	6.3	1
	6	1224	4.78	1	1221	5.7	1
	7	1226	4.43	1	1223	5.3	1
	8	1228	4.17	1	1226	5.0	1
	9	1230	3.97	1	1228	4.8	1
	10	1232	3.81	1	1230	4.6	1
	11	1233	3.69	1	1232	4.4	1
	12	1235	3.59	2	1233	4.3	1
	13	1236	3.52	2	1235	4.2	2
	14	1237	3.46	2	1237	4.2	2
	15	1239	3.42	2	1238	4.2	2
	16	1240	3.40	3	1239	4.1	2
	17	1241	3.39	3	1241	4.2	3
	18	1242	3.40	3	1243	4.2	3
	19	1244	3.42	3	1244	4.2	3
	20	1245	3.46	3	1246	4.3	3
	21	1246	3.51	3	1248	4.4	3
	22	1248	3.58	3	1249	4.5	4
	23	1249	3.67	3	1251	4.6	4
	24	1251	3.78	4	1253	4.8	4
	25	1252	3.93	4	1255	5.0	4
	26	1254	4.11	4	1257	5.3	4
	27	1256	4.34	4	1260	5.7	4
	28	1258	4.62	4	1262	6.1	4
	29	1260	5.00	4	1266	6.7	4
	30	1262	5.50	4	1269	7.6	4
	31	1266	6.19	4	1274	8.8	4
	32	1270	7.25	4	1281	10.8	4
	33	1275	9.09	4	1290	14.5	4
	34	1285	13.50	4	1290	16.4	4
	35	1290	19.31	4	1290	16.4	4

Table N-14. 2017–18 MSAA: Raw to Scaled Score Look-up Table — Mathematics Grade 11

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
A	0	1200	31.31	1	1200	25.2	1
	1	1200	15.59	1	1200	13.6	1
	2	1209	9.66	1	1207	9.0	1
	3	1216	7.21	1	1213	7.1	1
	4	1220	5.80	1	1217	6.1	1
	5	1223	4.89	1	1220	5.4	1
	6	1226	4.26	1	1222	5.0	1
	7	1228	3.80	1	1225	4.6	1
	8	1229	3.47	1	1227	4.4	1
	9	1231	3.23	1	1228	4.2	1
	10	1232	3.06	1	1230	4.0	1
	11	1234	2.95	1	1232	3.9	1
	12	1235	2.87	2	1233	3.8	1
	13	1236	2.83	2	1235	3.8	2
	14	1237	2.81	2	1236	3.8	2
	15	1238	2.81	2	1237	3.7	2
	16	1239	2.83	2	1239	3.7	2
	17	1240	2.86	3	1240	3.7	3
	18	1242	2.90	3	1242	3.8	3
	19	1243	2.96	3	1243	3.8	3
	20	1244	3.04	3	1244	3.8	3
	21	1245	3.14	3	1246	3.9	3
	22	1246	3.26	3	1247	4.0	3
	23	1248	3.41	3	1249	4.1	4
	24	1249	3.59	3	1251	4.3	4
	25	1251	3.82	4	1252	4.4	4
	26	1253	4.11	4	1254	4.6	4
	27	1255	4.48	4	1256	4.9	4
	28	1257	4.96	4	1259	5.2	4
	29	1260	5.59	4	1261	5.6	4
	30	1263	6.47	4	1264	6.2	4
	31	1268	7.75	4	1268	7.0	4
	32	1274	9.80	4	1272	8.2	4
	33	1283	13.71	4	1279	10.3	4
	34	1290	19.57	4	1290	15.6	4
	35	1290	19.57	4	1290	17.9	4
B	0	1200	37.59	1	1200	28.9	1
	1	1201	15.54	1	1200	14.6	1
	2	1212	9.29	1	1208	9.4	1
	3	1218	6.96	1	1214	7.3	1
	4	1223	5.70	1	1218	6.2	1
	5	1226	4.91	1	1222	5.6	1
	6	1228	4.38	1	1224	5.1	1
	7	1230	3.99	1	1226	4.7	1
	8	1232	3.71	1	1228	4.5	1
	9	1233	3.50	1	1230	4.3	1

continued

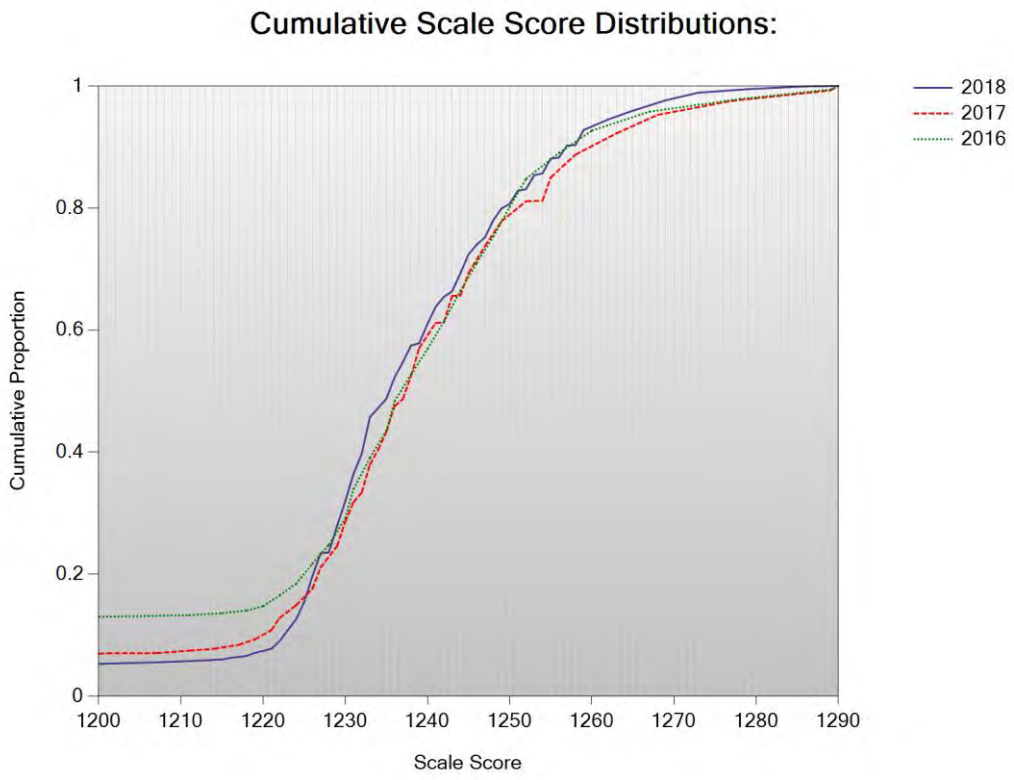
Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
B	10	1235	3.33	2	1232	4.1	1
	11	1236	3.21	2	1234	4.0	2
	12	1237	3.11	2	1235	3.9	2
	13	1239	3.04	2	1237	3.9	2
	14	1240	3.00	3	1238	3.8	2
	15	1241	2.97	3	1239	3.8	2
	16	1242	2.96	3	1241	3.8	3
	17	1243	2.96	3	1242	3.8	3
	18	1244	2.98	3	1244	3.8	3
	19	1245	3.02	3	1245	3.9	3
	20	1247	3.08	3	1247	4.0	3
	21	1248	3.15	3	1248	4.0	3
	22	1249	3.25	3	1250	4.1	4
	23	1250	3.37	4	1251	4.3	4
	24	1252	3.52	4	1253	4.4	4
	25	1253	3.72	4	1255	4.6	4
	26	1255	3.96	4	1257	4.8	4
	27	1257	4.28	4	1259	5.1	4
	28	1259	4.69	4	1261	5.4	4
	29	1262	5.24	4	1264	5.9	4
	30	1265	6.04	4	1267	6.5	4
	31	1269	7.24	4	1271	7.4	4
	32	1275	9.30	4	1276	8.7	4
	33	1284	13.55	4	1283	11.2	4
	34	1290	19.50	4	1290	16.5	4
	35	1290	19.50	4	1290	16.5	4
C	0	1200	36.86	1	1200	29.3	1
	1	1208	12.28	1	1200	14.9	1
	2	1217	7.93	1	1208	9.5	1
	3	1222	6.12	1	1214	7.5	1
	4	1226	5.11	1	1219	6.4	1
	5	1228	4.46	1	1222	5.7	1
	6	1230	4.01	1	1225	5.2	1
	7	1232	3.67	1	1227	4.9	1
	8	1234	3.42	1	1229	4.6	1
	9	1235	3.22	2	1231	4.4	1
	10	1237	3.06	2	1233	4.3	1
	11	1238	2.93	2	1234	4.2	2
	12	1239	2.83	2	1236	4.1	2
	13	1240	2.75	3	1237	4.0	2
	14	1241	2.69	3	1239	4.0	2
	15	1242	2.65	3	1240	3.9	3
	16	1243	2.62	3	1242	3.9	3
	17	1244	2.61	3	1243	3.9	3
	18	1245	2.61	3	1245	4.0	3
	19	1246	2.64	3	1246	4.0	3
	20	1247	2.67	3	1248	4.1	3
	21	1248	2.73	3	1249	4.2	4

continued

Path	Raw Score	2018			2017		
		Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
C	22	1249	2.80	3	1251	4.3	4
	23	1250	2.89	4	1253	4.4	4
	24	1252	3.01	4	1254	4.6	4
	25	1253	3.15	4	1256	4.8	4
	26	1254	3.32	4	1258	5.1	4
	27	1256	3.55	4	1261	5.4	4
	28	1258	3.83	4	1263	5.8	4
	29	1260	4.20	4	1266	6.4	4
	30	1262	4.71	4	1270	7.1	4
	31	1265	5.45	4	1274	8.1	4
	32	1269	6.64	4	1280	9.7	4
	33	1274	8.83	4	1288	12.5	4
	34	1285	14.43	4	1290	15.5	4
	35	1290	20.75	4	1290	15.5	4

APPENDIX O—SCORE DISTRIBUTIONS

Figure O-1. 2017–18 MSAA: Cumulative Score Distribution
Top: ELA Grade 3 Bottom: ELA Grade 4



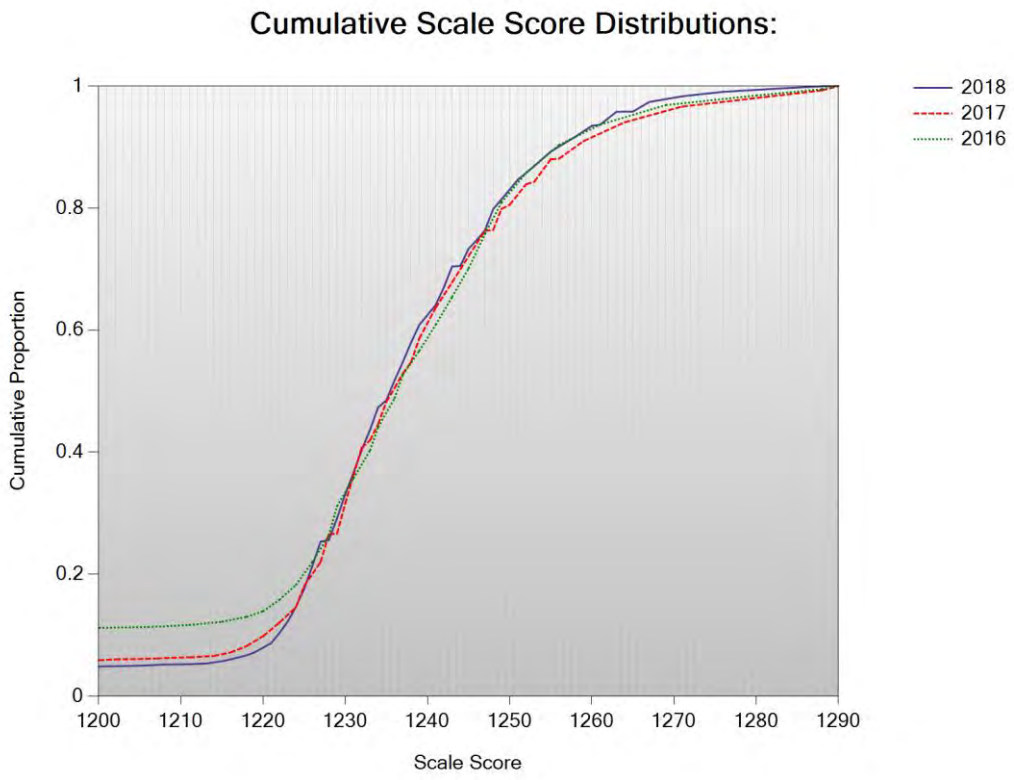
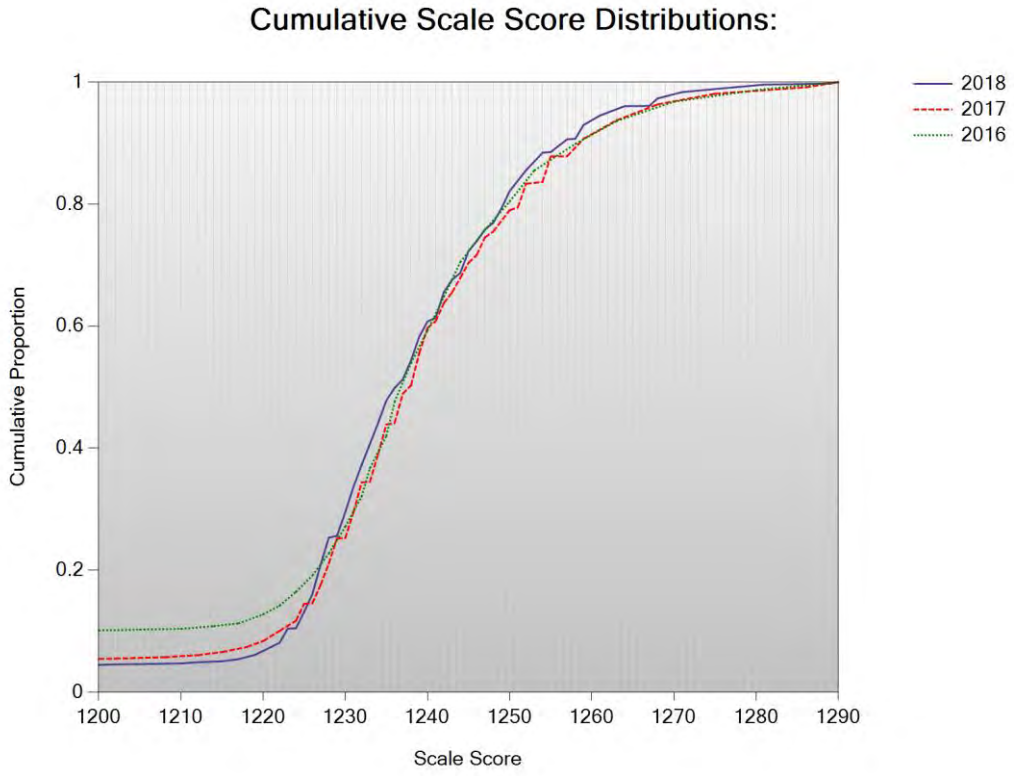


Figure O-2. 2017–18 MSAA: Cumulative Score Distribution
Top: ELA Grade 5 Bottom: ELA Grade 6



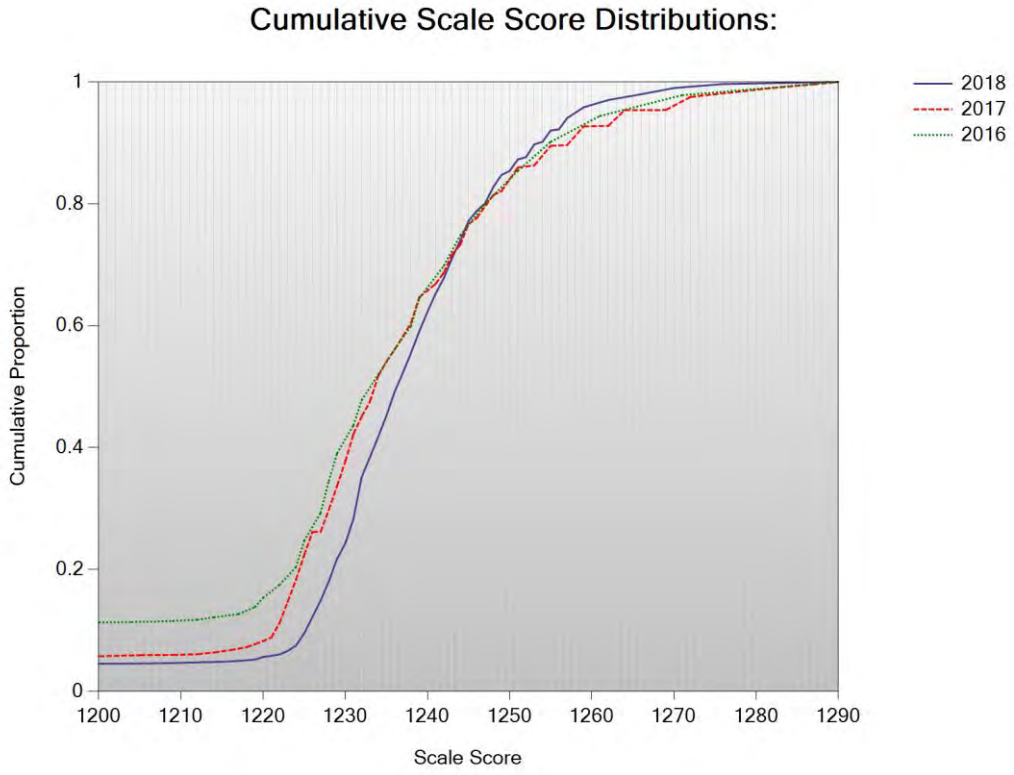
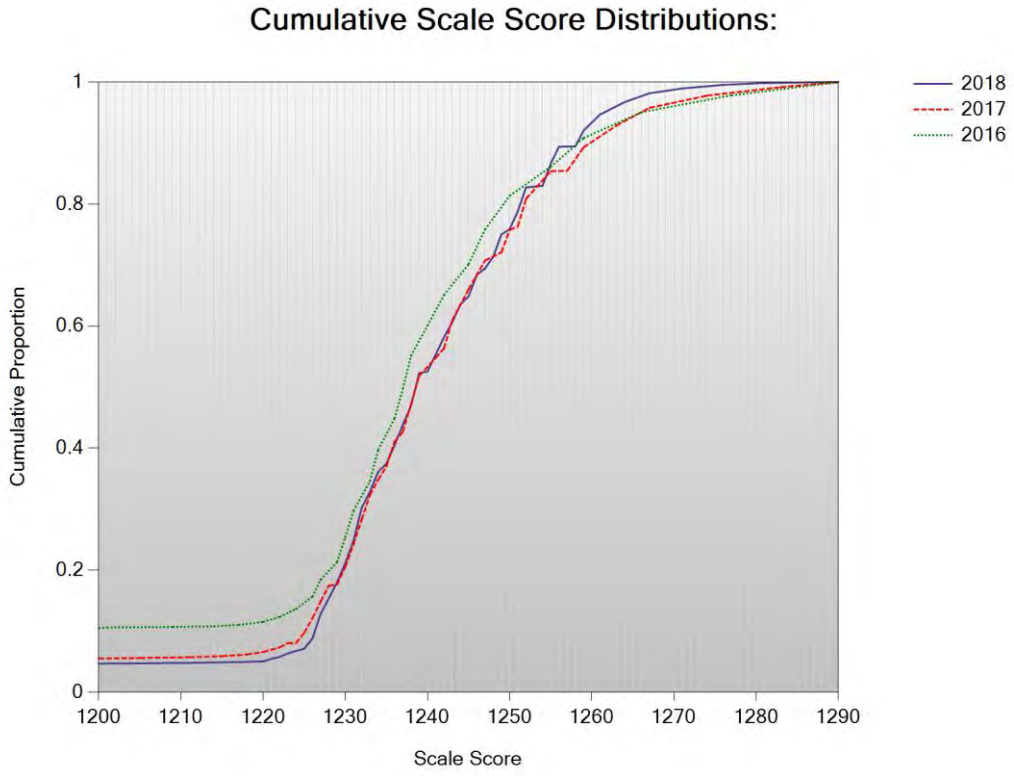


Figure O-3. 2017–18 MSAA: Cumulative Score Distribution
Top: ELA Grade 7 Bottom: ELA Grade 8



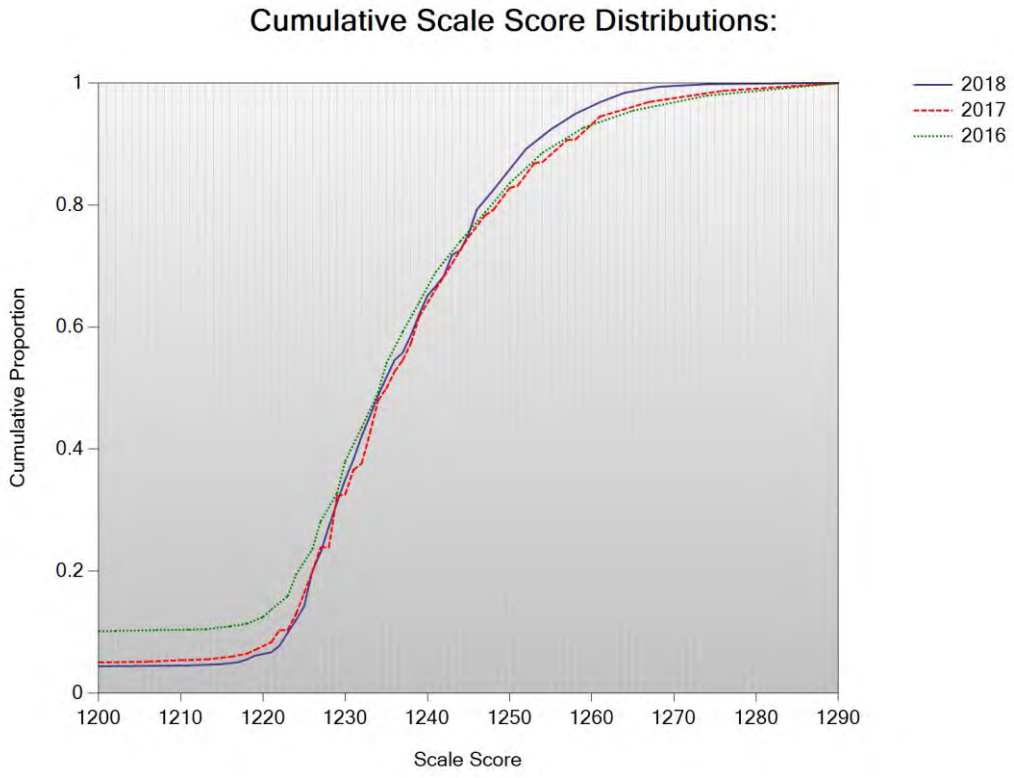


Figure O-4. 2017–18 MSAA: Cumulative Score Distribution
ELA Grade 11

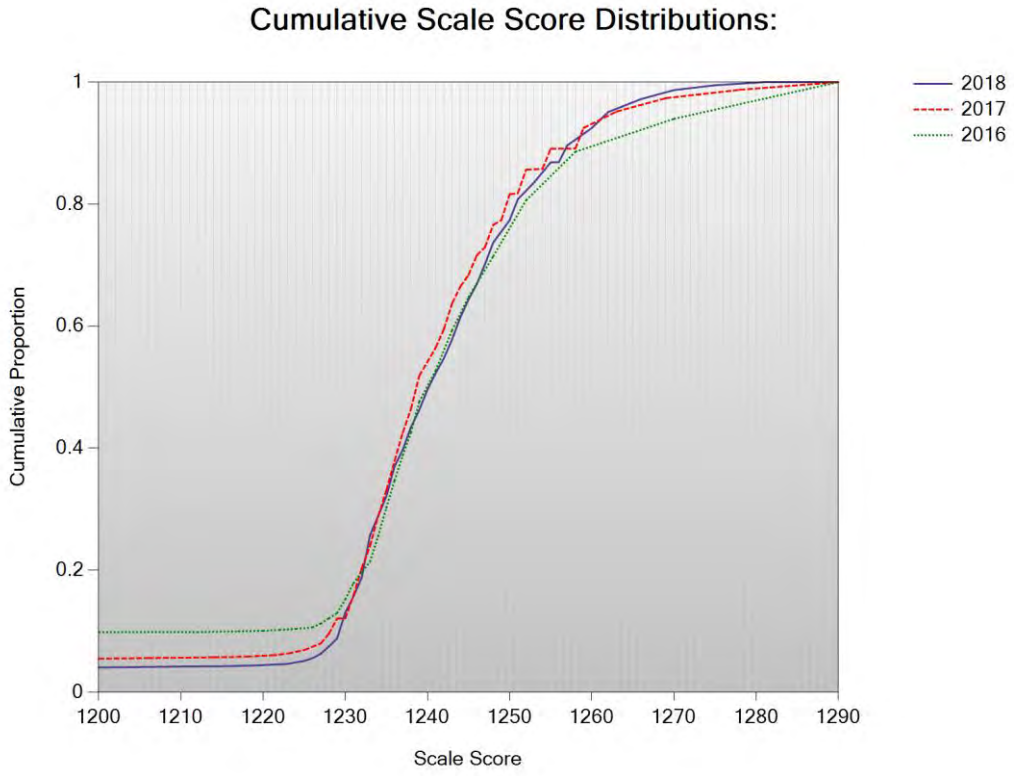
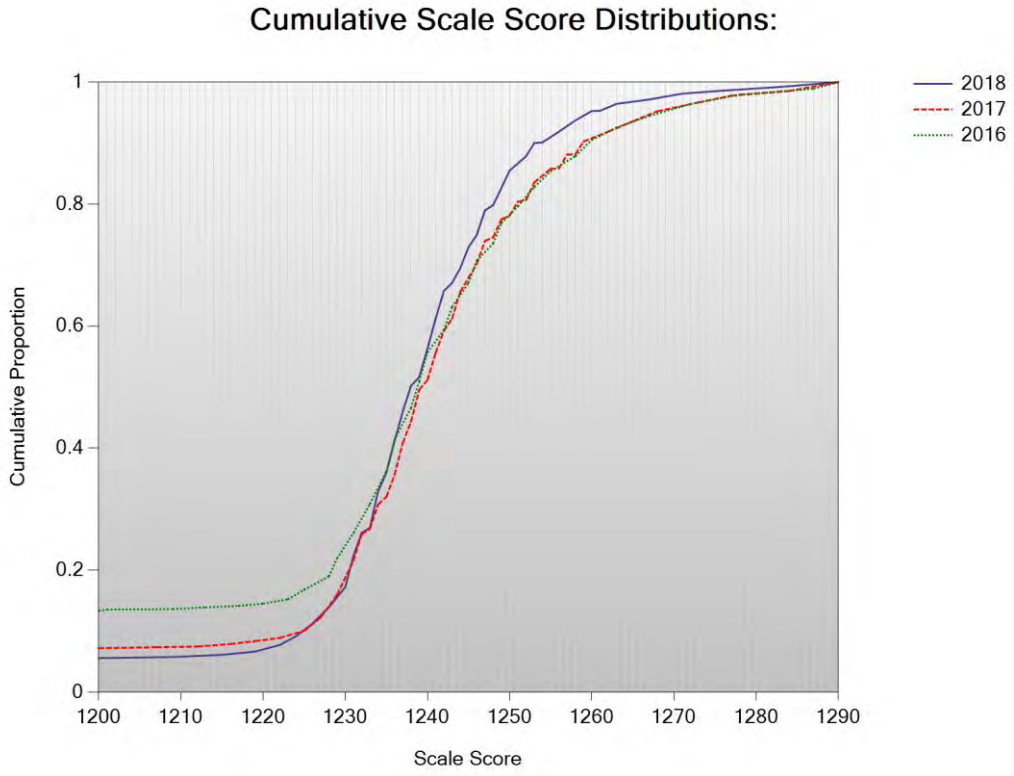


Figure O-5. 2017–18 MSAA: Cumulative Score Distribution
Top: Mathematics Grade 3 Bottom: Mathematics Grade 4



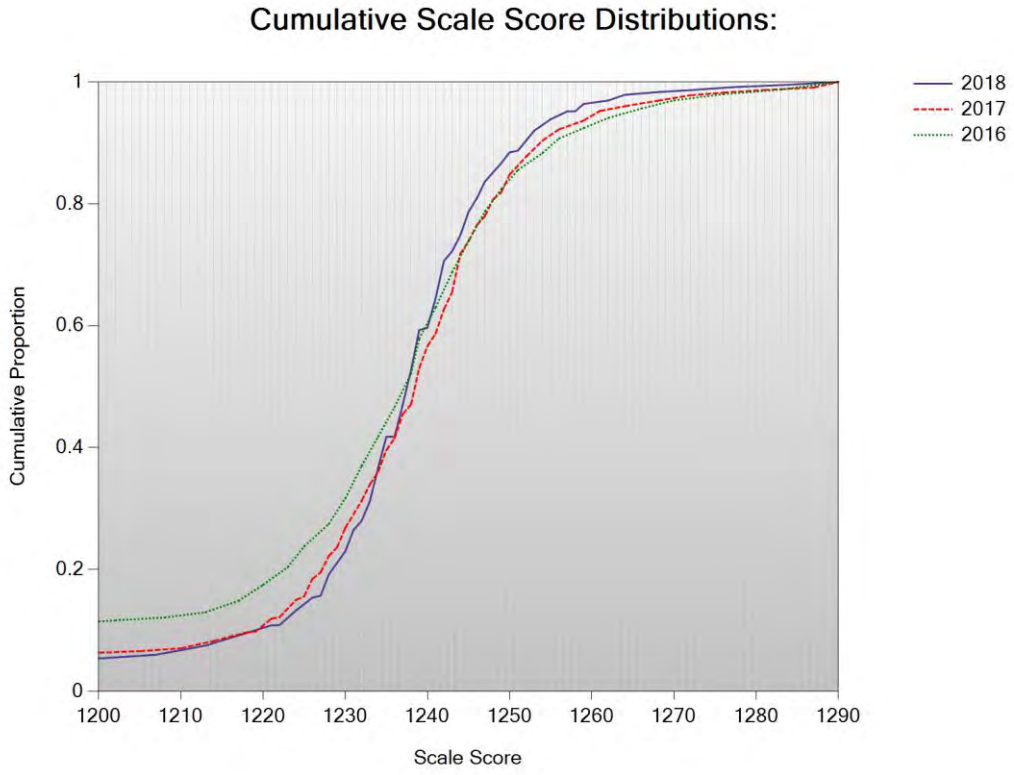
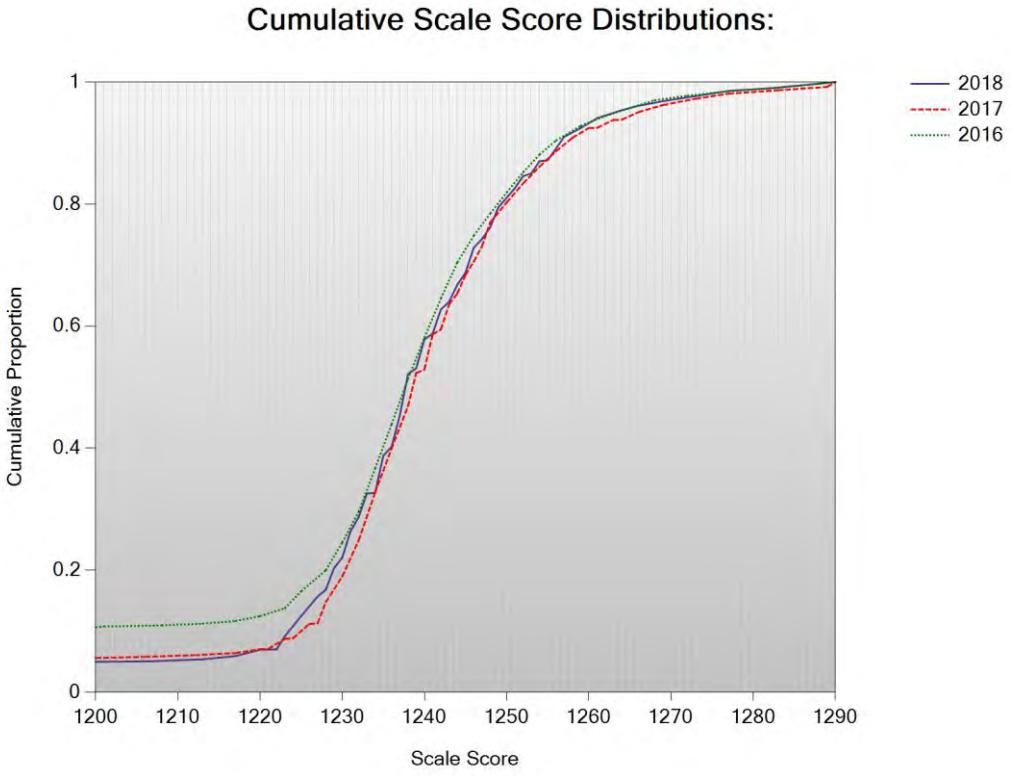


Figure O-6. 2017–18 MSAA: Cumulative Score Distribution
Top: Mathematics Grade 5 Bottom: Mathematics Grade 6



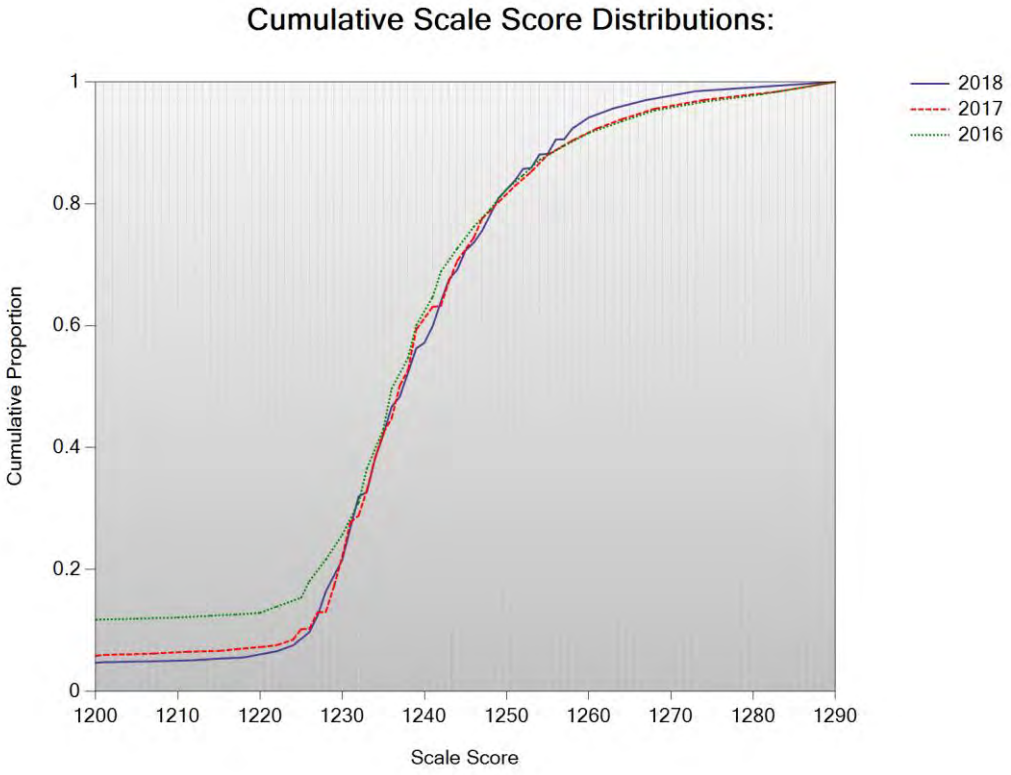
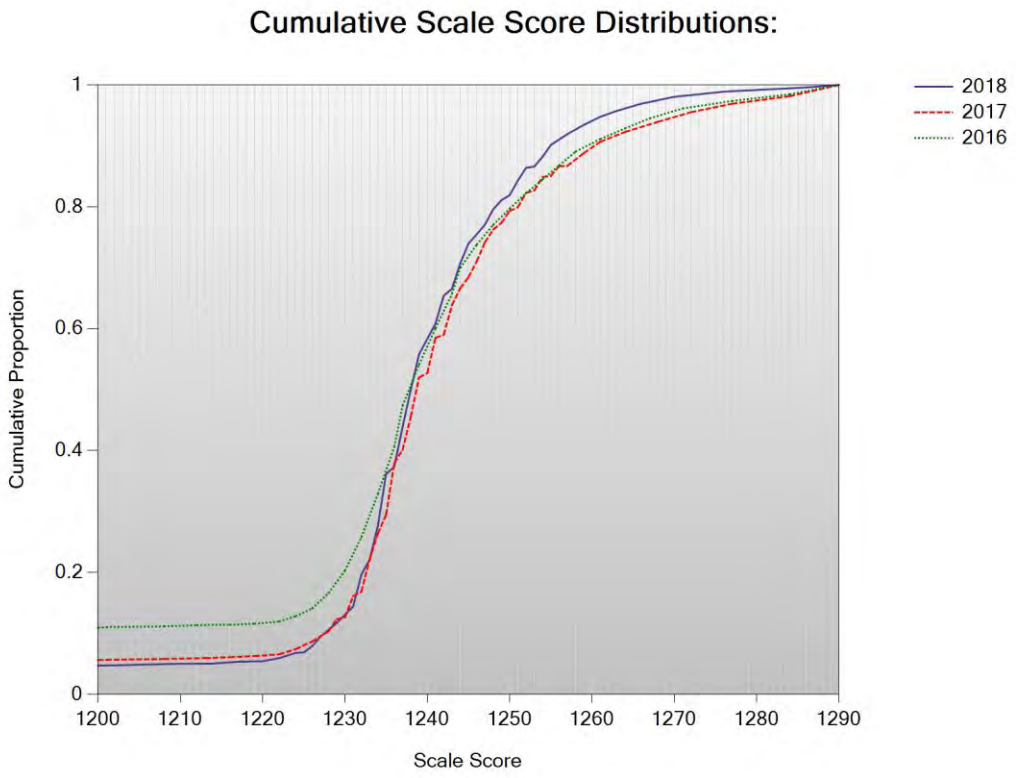


Figure O-7. 2017–18 MSAA: Cumulative Score Distribution
Top: Mathematics Grade 7 Bottom: Mathematics Grade 8



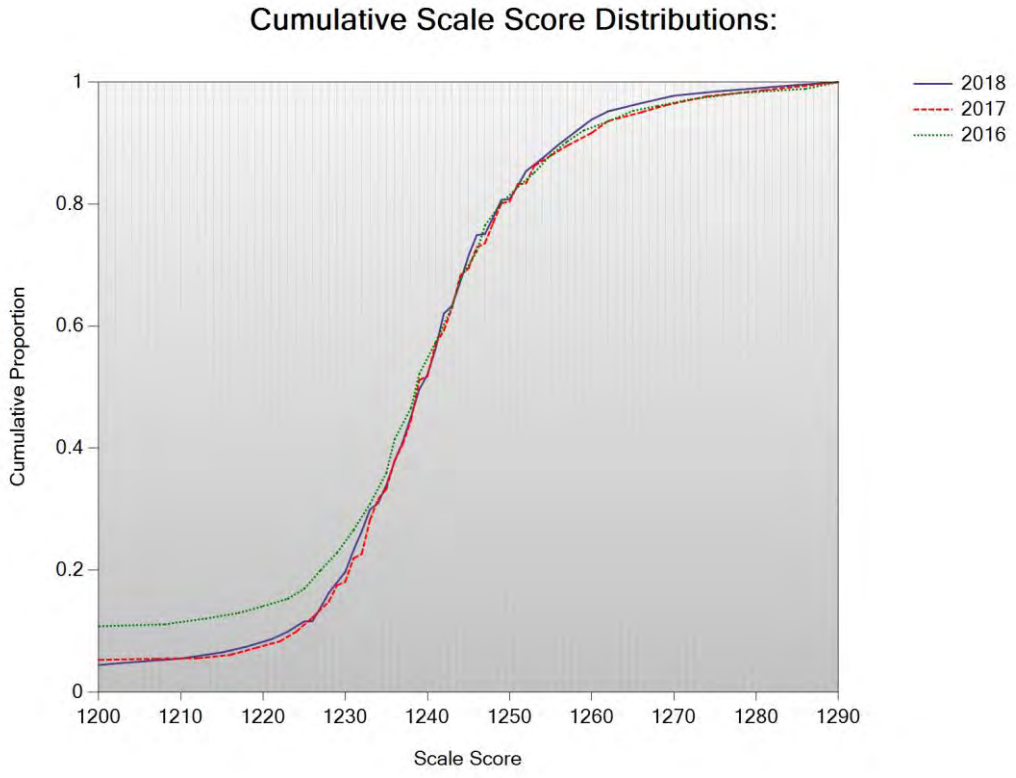
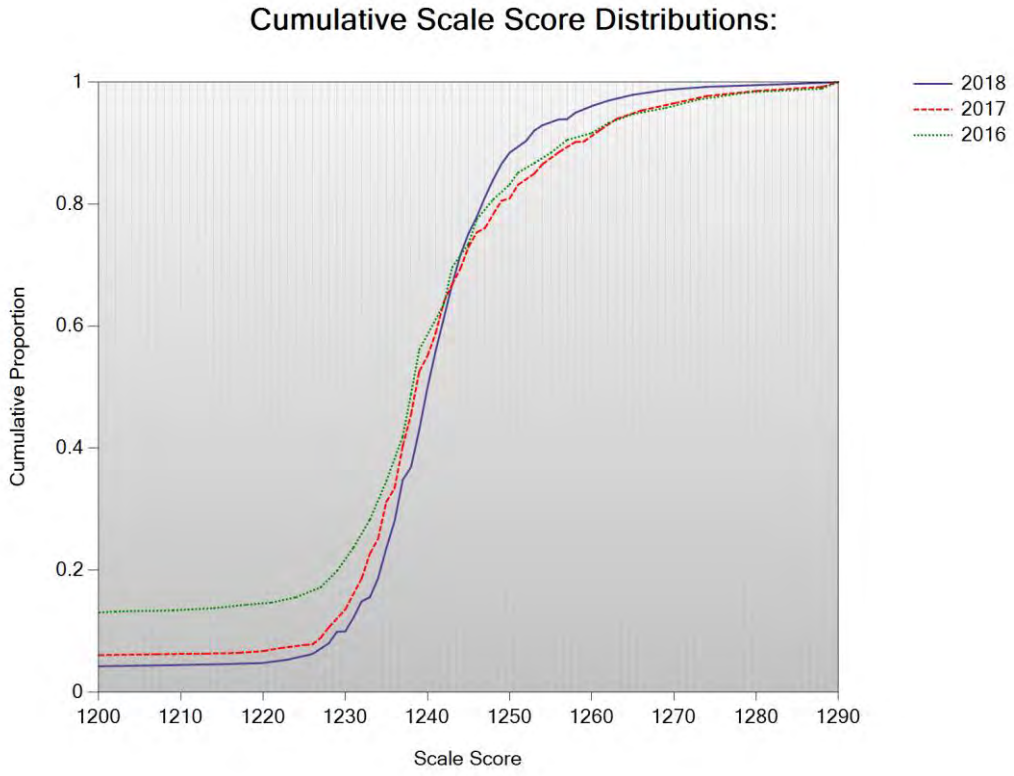


Figure O-8. 2017–18 MSAA: Cumulative Score Distribution Mathematics Grade 11



APPENDIX P—CLASSICAL RELIABILITY

Table P-1. 2016–17 MSAA: Reliability: Subgroup-ELA Grade 3

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	2990	1200	1289	1235.39	17.44	0.88	6.11
Female	1024	1200	1286	1235.06	17.32	0.87	6.18
Male	1957	1200	1289	1235.56	17.51	0.88	6.08
Gender Undefined	9	1200	1255	1237.89	16.39	0.86	6.21
Hispanic or Latino	665	1200	1279	1234.5	17.04	0.87	6.18
American Indian or Alaska Native	192	1200	1273	1233.83	18.11	0.87	6.62
Asian	90	1200	1265	1232.89	15.79	0.86	5.97
Black or African American	759	1200	1289	1235.83	16.67	0.88	5.79
Native Hawaiian or Pacific Islander	29	1200	1262	1235.41	12.79	0.87	4.55
White (non-Hispanic)	1899	1200	1286	1235.72	17.74	0.88	6.18
Two or More Races (non-Hispanic)	636	1200	1279	1235.29	17.53	0.87	6.23
No Primary race/Ethnicity Undefined	44	1200	1279	1237.25	16.69	0.90	5.16
Currently receiving LEP services	117	1200	1273	1236.27	16.13	0.89	5.46
Not receiving LEP services	1240	1200	1289	1236.5	18.92	0.88	6.57
LEP: All Other Students	1633	1200	1286	1234.49	16.28	0.87	5.79
Economically Disadvantaged Students	998	1200	1289	1236.64	17.47	0.89	5.9
Non-economically Disadvantaged Students	876	1200	1286	1234.51	17.21	0.87	6.31
SES: All Other Students	1116	1200	1286	1234.98	17.54	0.88	6.14
Migrant	2	1225	1251	1238	18.38	0.96	3.63
Non-migrant	1224	1200	1286	1236.65	19.06	0.88	6.61
Undefined Migrant Status	1764	1200	1289	1234.52	16.17	0.87	5.75
Augmentative Communication	657	1200	1273	1227.92	13.94	0.79	6.36
No Augmentative Communication	2243	1200	1289	1238.7	16.59	0.89	5.48
Undefined Augmentative Communications	90	1200	1286	1207.47	-	-	-
Hearing Loss	107	1200	1273	1221.98	-	-	-
Within Normal Limits	2797	1200	1289	1236.89	16.4	0.89	5.51
Undefined Hearing Loss	86	1200	1259	1203.58	-	-	-
Visual Impairment	212	1200	1262	1218.75	-	-	-
Within Normal Limits	2689	1200	1289	1237.7	15.85	0.89	5.22
Undefined Visual Impairment	89	1200	1257	1205.44	-	-	-
Sensory Stimuli Response	433	1200	1269	1216.9	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2482	1200	1289	1239.69	14.45	0.9	4.56
Undefined Receptive Language	75	1200	1200	1200	-	-	-
Special School	302	1200	1273	1227.35	14.95	0.81	6.53
Regular School Self-contained	1899	1200	1289	1235.06	16.46	0.88	5.81
Regular School Resource Room	397	1200	1286	1241.83	15.07	0.9	4.85
Regular School Primarily Self-contained	229	1200	1286	1246.05	14.77	0.89	4.86
Regular School General Education	88	1200	1286	1243.59	17.28	0.89	5.6
Undefined Classroom Setting	75	1200	1200	1200	-	-	-
Communicates Primarily Through Cries	382	1200	1262	1215.99	-	-	-
Uses Intentional Communication	725	1200	1289	1230.86	12.88	0.84	5.14
Uses Symbolic Language	1808	1200	1286	1242.78	13.85	0.9	4.42
Undefined Expressive Communication	75	1200	1200	1200	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-2. 2016–17 MSAA: Reliability: Subgroup-ELA Grade 4

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3304	1200	1290	1235.25	17.11	0.87	6.09
Female	1175	1200	1288	1235.21	17.38	0.87	6.19
Male	2125	1200	1290	1235.26	16.98	0.87	6.04
Gender Undefined	4	1232	1243	1240	5.35	0.53	3.67
Hispanic or Latino	767	1200	1283	1233.44	15.97	0.86	6.04
American Indian or Alaska Native	217	1200	1283	1233.6	17.98	0.87	6.53
Asian	131	1200	1271	1232.12	12.98	0.83	5.42
Black or African American	799	1200	1290	1235.95	16.69	0.88	5.83
Native Hawaiian or Pacific Islander	30	1200	1249	1228.83	15.76	0.8	6.99
White (non-Hispanic)	2106	1200	1288	1235.6	17.41	0.88	6.15
Two or More Races (non-Hispanic)	752	1200	1288	1233.96	16.09	0.86	5.97
No Primary race/Ethnicity Undefined	45	1200	1271	1233.42	15.8	0.86	5.86
Currently receiving LEP services	146	1200	1283	1235.38	15.11	0.86	5.62
Not receiving LEP services	1363	1200	1290	1236.85	19.18	0.88	6.66
LEP: All Other Students	1795	1200	1288	1234.02	15.43	0.87	5.66
Economically Disadvantaged Students	1138	1200	1290	1235.96	17.87	0.88	6.16
Non-economically Disadvantaged Students	966	1200	1288	1235.64	16.99	0.87	6.13
SES: All Other Students	1200	1200	1288	1234.26	16.43	0.87	5.99
Migrant	0	-	-	-	-	-	-
Non-migrant	1356	1200	1290	1236.71	19.13	0.88	6.68
Undefined Migrant Status	1948	1200	1288	1234.23	15.48	0.87	5.65
Augmentative Communication	674	1200	1283	1227.74	13.14	0.79	5.96
No Augmentative Communication	2539	1200	1290	1238.31	16.48	0.88	5.61
Undefined Augmentative Communications	91	1200	1267	1205.49	-	-	-
Hearing Loss	135	1200	1251	1220.27	-	-	-
Within Normal Limits	3081	1200	1288	1236.77	16.11	0.88	5.5
Undefined Hearing Loss	88	1200	1290	1204.85	-	-	-
Visual Impairment	203	1200	1267	1220.62	-	-	-
Within Normal Limits	3015	1200	1290	1237.15	15.73	0.89	5.3
Undefined Visual Impairment	86	1200	1261	1203.22	-	-	-
Sensory Stimuli Response	399	1200	1251	1215.84	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2824	1200	1290	1238.97	14.48	0.9	4.67
Undefined Receptive Language	81	1200	1261	1201.15	-	-	-
Special School	332	1200	1271	1225.92	14.29	0.79	6.6
Regular School Self-contained	2029	1200	1290	1234.88	16.15	0.87	5.78
Regular School Resource Room	523	1200	1288	1242.54	15.26	0.89	5.06
Regular School Primarily Self-contained	233	1200	1283	1243.26	12.79	0.87	4.56
Regular School General Education	106	1200	1288	1244.04	17.47	0.89	5.84
Undefined Classroom Setting	81	1200	1261	1201.15	-	-	-
Communicates Primarily Through Cries	347	1200	1253	1214.57	-	-	-
Uses Intentional Communication	704	1200	1288	1230.28	11.86	0.85	4.65
Uses Symbolic Language	2172	1200	1290	1241.43	14.37	0.89	4.77
Undefined Expressive Communication	81	1200	1261	1201.15	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-3. MSAA: Reliability: Subgroup-ELA Grade 5

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3376	1200	1290	1236.02	16.89	0.84	6.8
Female	1146	1200	1290	1235.9	16.97	0.83	6.9
Male	2224	1200	1290	1236.12	16.81	0.84	6.73
Gender Undefined	6	1200	1259	1222.83	-	-	-
Hispanic or Latino	758	1200	1289	1234.84	15.97	0.82	6.74
American Indian or Alaska Native	182	1200	1276	1234.96	14.4	0.82	6.05
Asian	100	1200	1271	1234.25	12.86	0.8	5.8
Black or African American	868	1200	1290	1237.24	16.41	0.85	6.36
Native Hawaiian or Pacific Islander	25	1222	1254	1235.16	11.39	0.88	3.87
White (non-Hispanic)	2160	1200	1290	1235.72	17.47	0.83	7.11
Two or More Races (non-Hispanic)	707	1200	1289	1234.4	15.85	0.81	6.82
No Primary race/Ethnicity Undefined	31	1200	1271	1235.26	11.91	0.83	4.98
Currently receiving LEP services	150	1200	1281	1237.21	15.26	0.86	5.65
Not receiving LEP services	1446	1200	1290	1237.01	18.74	0.84	7.41
LEP: All Other Students	1780	1200	1289	1235.13	15.31	0.83	6.36
Economically Disadvantaged Students	1165	1200	1290	1237.32	17.28	0.85	6.69
Non-economically Disadvantaged Students	1050	1200	1290	1235.35	16.89	0.83	6.99
SES: All Other Students	1161	1200	1289	1235.33	16.42	0.83	6.74
Migrant	1	1226	1226	1226	-	-	3.4
Non-migrant	1413	1200	1290	1237.06	18.68	0.84	7.37
Undefined Migrant Status	1962	1200	1289	1235.28	15.43	0.83	6.36
Augmentative Communication	613	1200	1271	1228.89	12.49	0.72	6.58
No Augmentative Communication	2660	1200	1290	1238.83	16.26	0.86	6.16
Undefined Augmentative Communications	103	1200	1254	1206.1	-	-	-
Hearing Loss	114	1200	1276	1224.96	-	-	-
Within Normal Limits	3162	1200	1290	1237.38	15.81	0.86	6.02
Undefined Hearing Loss	100	1200	1254	1205.63	-	-	-
Visual Impairment	190	1200	1276	1221.65	-	-	-
Within Normal Limits	3084	1200	1290	1237.91	15.4	0.86	5.74
Undefined Visual Impairment	102	1200	1254	1205.76	-	-	-
Sensory Stimuli Response	379	1200	1281	1217.85	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2911	1200	1290	1239.44	14.29	0.87	5.07
Undefined Receptive Language	86	1200	1254	1200.63	-	-	-
Special School	350	1200	1271	1228.48	15.29	0.76	7.52
Regular School Self-contained	2044	1200	1290	1235.73	16.02	0.84	6.42
Regular School Resource Room	545	1200	1290	1242.71	14.09	0.87	5
Regular School Primarily Self-contained	245	1200	1281	1243.81	14.27	0.87	5.19
Regular School General Education	104	1200	1276	1243.72	13.87	0.87	5.08
Undefined Classroom Setting	88	1200	1254	1200.61	-	-	-
Communicates Primarily Through Cries	311	1200	1276	1215.68	-	-	-
Uses Intentional Communication	697	1200	1281	1230.53	12.97	0.78	6.08
Uses Symbolic Language	2282	1200	1290	1241.81	13.72	0.88	4.79
Undefined Expressive Communication	86	1200	1254	1200.63	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-4. 2016–17 MSAA: Reliability: Subgroup-ELA Grade 6

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3543	1200	1290	1235.58	15.51	0.68	8.71
Female	1242	1200	1290	1235.37	16	0.67	9.25
Male	2292	1200	1290	1235.68	15.24	0.7	8.41
Gender Undefined	9	1200	1266	1238.44	17.4	0.7	9.57
Hispanic or Latino	755	1200	1285	1234.69	14.79	0.67	8.45
American Indian or Alaska Native	198	1200	1270	1234.57	16.17	0.63	9.85
Asian	126	1200	1276	1235.53	12.91	0.73	6.65
Black or African American	894	1200	1287	1235.54	15.05	0.69	8.44
Native Hawaiian or Pacific Islander	29	1200	1259	1232.41	16.33	0.6	10.39
White (non-Hispanic)	2254	1200	1290	1235.88	15.94	0.69	8.86
Two or More Races (non-Hispanic)	732	1200	1285	1235.26	15.4	0.68	8.65
No Primary race/Ethnicity Undefined	45	1200	1262	1236.93	15.6	0.69	8.65
Currently receiving LEP services	128	1200	1285	1237.98	13.83	0.73	7.14
Not receiving LEP services	1501	1200	1290	1235.45	17.01	0.66	9.87
LEP: All Other Students	1914	1200	1290	1235.52	14.34	0.7	7.79
Economically Disadvantaged Students	1147	1200	1285	1236.67	15.43	0.7	8.5
Non-economically Disadvantaged Students	1190	1200	1290	1234.82	15.36	0.67	8.82
SES: All Other Students	1206	1200	1290	1235.29	15.7	0.69	8.8
Migrant	1	1240	1240	1240	NA	NA	2.9
Non-migrant	1494	1200	1290	1235.76	16.96	0.67	9.72
Undefined Migrant Status	2048	1200	1290	1235.45	14.37	0.7	7.89
Augmentative Communication	619	1200	1270	1228.09	12.97	0.41	9.92
No Augmentative Communication	2833	1200	1290	1238.15	14.49	0.74	7.34
Undefined Augmentative Communications	91	1200	1270	1206.4	-	-	-
Hearing Loss	113	1200	1255	1222.7	-	-	-
Within Normal Limits	3336	1200	1290	1236.85	14.43	0.73	7.44
Undefined Hearing Loss	94	1200	1249	1205.78	-	-	-
Visual Impairment	208	1200	1270	1222.56	-	-	-
Within Normal Limits	3226	1200	1290	1237.28	14.03	0.75	7
Undefined Visual Impairment	109	1200	1255	1210.17	-	-	-
Sensory Stimuli Response	403	1200	1257	1216.6	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	3062	1200	1290	1238.97	12.48	0.82	5.29
Undefined Receptive Language	78	1200	1238	1200.49	-	-	-
Special School	375	1200	1270	1227.43	14.26	0.42	10.84
Regular School Self-contained	2259	1200	1290	1235.58	14.33	0.69	7.98
Regular School Resource Room	515	1200	1287	1241.74	13.07	0.82	5.56
Regular School Primarily Self-contained	218	1200	1285	1243.7	13.61	0.84	5.46
Regular School General Education	96	1217	1290	1245.09	13.5	0.9	4.23
Undefined Classroom Setting	80	1200	1238	1200.47	-	-	-
Communicates Primarily Through Cries	354	1200	1270	1215.95	-	-	-
Uses Intentional Communication	753	1200	1285	1230.84	11.93	0.6	7.53
Uses Symbolic Language	2358	1200	1290	1241.2	11.86	0.85	4.65
Undefined Expressive Communication	78	1200	1238	1200.49	NA	NA	NA

Note: Values are calculated only for subgroups with 100 or more students.

Table P-5. 2016–17 MSAA: Reliability: Subgroup-ELA Grade 7

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3613	1200	1289	1237.6	17.04	0.76	8.34
Female	1264	1200	1289	1237.78	17.46	0.75	8.66
Male	2340	1200	1289	1237.51	16.84	0.76	8.18
Gender Undefined	9	1226	1252	1237.33	8.6	0.86	3.16
Hispanic or Latino	761	1200	1289	1236.25	16.74	0.75	8.37
American Indian or Alaska Native	202	1200	1289	1236.63	18.15	0.76	8.93
Asian	113	1200	1264	1235.58	13.19	0.74	6.72
Black or African American	875	1200	1289	1239.04	15.89	0.78	7.38
Native Hawaiian or Pacific Islander	31	1200	1271	1232.52	11.95	0.81	5.27
White (non-Hispanic)	2328	1200	1289	1237.35	17.59	0.75	8.72
Two or More Races (non-Hispanic)	738	1200	1289	1236.42	16.83	0.76	8.3
No Primary race/Ethnicity Undefined	56	1200	1264	1238.3	15.32	0.78	7.23
Currently receiving LEP services	130	1200	1271	1240.89	13.06	0.82	5.51
Not receiving LEP services	1454	1200	1289	1238.07	18.36	0.76	9.04
LEP: All Other Students	2029	1200	1289	1237.05	16.24	0.76	7.96
Economically Disadvantaged Students	1183	1200	1281	1238.78	17.49	0.77	8.36
Non-economically Disadvantaged Students	1201	1200	1289	1237.6	16.21	0.76	7.98
SES: All Other Students	1229	1200	1289	1236.46	17.32	0.75	8.66
Migrant	1	1244	1244	1244	-	-	3.5
Non-migrant	1421	1200	1289	1238.48	18.32	0.76	8.94
Undefined Migrant Status	2191	1200	1289	1237.03	16.14	0.76	7.93
Augmentative Communication	586	1200	1267	1228.83	13.38	0.55	8.95
No Augmentative Communication	2910	1200	1289	1240.64	15.79	0.8	7.1
Undefined Augmentative Communications	117	1200	1281	1205.79	-	-	-
Hearing Loss	118	1200	1261	1221.73	-	-	-
Within Normal Limits	3382	1200	1289	1239.24	15.58	0.8	7.02
Undefined Hearing Loss	113	1200	1267	1205.24	-	-	-
Visual Impairment	209	1200	1271	1218.56	-	-	-
Within Normal Limits	3270	1200	1289	1239.98	14.87	0.82	6.4
Undefined Visual Impairment	134	1200	1281	1209.14	-	-	-
Sensory Stimuli Response	374	1200	1264	1216.23	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	3139	1200	1289	1241.34	13.7	0.84	5.44
Undefined Receptive Language	100	1200	1200	1200	-	-	-
Special School	401	1200	1289	1228.52	16.12	0.57	10.52
Regular School Self-contained	2238	1200	1289	1237.52	15.63	0.77	7.48
Regular School Resource Room	550	1200	1289	1245.21	13.77	0.83	5.62
Regular School Primarily Self-contained	231	1200	1281	1248.08	13.02	0.85	5.08
Regular School General Education	93	1220	1281	1247.97	12.23	0.88	4.24
Undefined Classroom Setting	100	1200	1200	1200	-	-	-
Communicates Primarily Through Cries	323	1200	1264	1214.2	-	-	-
Uses Intentional Communication	696	1200	1276	1232.55	12.52	0.7	6.87
Uses Symbolic Language	2494	1200	1289	1243.55	13.06	0.86	4.96
Undefined Expressive Communication	100	1200	1200	1200	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-6. 2016–17 MSAA: Reliability: Subgroup-ELA Grade 8

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3632	1200	1286	1234.37	15.69	0.77	7.48
Female	1315	1200	1286	1234.26	16.43	0.77	7.83
Male	2306	1200	1286	1234.46	15.24	0.77	7.24
Gender Undefined	11	1200	1246	1225.64	-	-	-
Hispanic or Latino	733	1200	1274	1233.1	15.28	0.75	7.62
American Indian or Alaska Native	181	1200	1274	1233.77	15.75	0.74	8.04
Asian	99	1200	1258	1231.67	13.49	0.76	6.65
Black or African American	944	1200	1286	1234.16	14.92	0.77	7.12
Native Hawaiian or Pacific Islander	25	1200	1268	1230.96	14.79	0.78	7
White (non-Hispanic)	2300	1200	1286	1234.72	16.19	0.78	7.66
Two or More Races (non-Hispanic)	716	1200	1274	1233.16	15.64	0.75	7.75
No Primary race/Ethnicity Undefined	49	1200	1261	1236.18	12.47	0.86	4.66
Currently receiving LEP services	101	1200	1268	1236.31	13.69	0.78	6.36
Not receiving LEP services	1549	1200	1286	1234.47	16.99	0.76	8.23
LEP: All Other Students	1982	1200	1286	1234.19	14.71	0.78	6.89
Economically Disadvantaged Students	1186	1200	1286	1235.58	15.51	0.79	7.13
Non-economically Disadvantaged Students	1225	1200	1286	1233.7	16.05	0.76	7.8
SES: All Other Students	1221	1200	1286	1233.86	15.46	0.77	7.48
Migrant	0	-	-	-	-	-	-
Non-migrant	1504	1200	1286	1234.64	16.82	0.76	8.17
Undefined Migrant Status	2128	1200	1286	1234.17	14.85	0.78	6.95
Augmentative Communication	534	1200	1261	1225.43	11.6	0.54	7.91
No Augmentative Communication	2970	1200	1286	1237.22	14.52	0.81	6.28
Undefined Augmentative Communications	128	1200	1255	1205.45	-	-	-
Hearing Loss	122	1200	1274	1223.98	-	-	-
Within Normal Limits	3384	1200	1286	1235.84	14.43	0.81	6.29
Undefined Hearing Loss	126	1200	1268	1204.9	-	-	-
Visual Impairment	214	1200	1274	1220.56	-	-	-
Within Normal Limits	3288	1200	1286	1236.43	13.91	0.83	5.82
Undefined Visual Impairment	130	1200	1252	1205.01	-	-	-
Sensory Stimuli Response	332	1200	1248	1213.95	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	3192	1200	1286	1237.65	12.85	0.86	4.84
Undefined Receptive Language	108	1200	1200	1200	-	-	-
Special School	404	1200	1274	1227.21	14.56	0.64	8.79
Regular School Self-contained	2284	1200	1286	1234.15	14.42	0.79	6.66
Regular School Resource Room	525	1200	1286	1241.98	12.3	0.85	4.76
Regular School Primarily Self-contained	219	1200	1274	1244.23	12.12	0.85	4.63
Regular School General Education	93	1200	1286	1244.11	12.83	0.85	4.91
Undefined Classroom Setting	107	1200	1200	1200	-	-	-
Communicates Primarily Through Cries	278	1200	1246	1213.32	-	-	-
Uses Intentional Communication	719	1200	1268	1227.92	12.38	0.68	6.98
Uses Symbolic Language	2526	1200	1286	1240	12.07	0.87	4.42
Undefined Expressive Communication	109	1200	1200	1200	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-7. 2016–17 MSAA: Reliability: Subgroup-ELA Grade 11

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3404	1200	1281	1238.3	17.13	0.83	7.13
Female	1222	1200	1281	1238.52	17.79	0.82	7.44
Male	2174	1200	1281	1238.18	16.77	0.83	6.95
Gender Undefined	8	1230	1266	1239.88	12.31	0.93	3.21
Hispanic or Latino	594	1200	1281	1236.27	15.78	0.81	6.91
American Indian or Alaska Native	154	1200	1281	1236.34	17.08	0.82	7.33
Asian	110	1200	1275	1237.35	14.78	0.82	6.25
Black or African American	855	1200	1281	1238.99	15.96	0.83	6.51
Native Hawaiian or Pacific Islander	23	1200	1260	1228.78	18.15	0.73	9.36
White (non-Hispanic)	2195	1200	1281	1238.31	17.52	0.83	7.3
Two or More Races (non-Hispanic)	567	1200	1281	1236.66	15.66	0.81	6.77
No Primary race/Ethnicity Undefined	37	1200	1275	1243.43	16.9	0.85	6.58
Currently receiving LEP services	82	1200	1275	1239.13	14.82	0.85	5.72
Not receiving LEP services	1446	1200	1281	1239.23	17.96	0.83	7.45
LEP: All Other Students	1876	1200	1281	1237.56	16.53	0.82	6.92
Economically Disadvantaged Students	1024	1200	1281	1240.52	17.69	0.84	7.05
Non-economically Disadvantaged Students	1293	1200	1281	1238.29	16.88	0.82	7.11
SES: All Other Students	1087	1200	1281	1236.23	16.64	0.81	7.22
Migrant	0	-	-	-	-	-	-
Non-migrant	1433	1200	1281	1239.27	17.9	0.83	7.42
Undefined Migrant Status	1971	1200	1281	1237.6	16.53	0.83	6.9
Augmentative Communication	497	1200	1281	1230.07	14.42	0.72	7.69
No Augmentative Communication	2755	1200	1281	1241.7	15.09	0.85	5.79
Undefined Augmentative Communications	152	1200	1248	1203.64	-	-	-
Hearing Loss	128	1200	1275	1225.47	-	-	-
Within Normal Limits	3131	1200	1281	1240.54	15.02	0.85	5.79
Undefined Hearing Loss	145	1200	1239	1201.38	-	-	-
Visual Impairment	191	1200	1270	1223.74	-	-	-
Within Normal Limits	3053	1200	1281	1240.95	14.58	0.86	5.52
Undefined Visual Impairment	160	1200	1275	1205.27	-	-	-
Sensory Stimuli Response	303	1200	1270	1217.29	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2962	1200	1281	1242.24	13.39	0.87	4.8
Undefined Receptive Language	139	1200	1234	1200.24	-	-	-
Special School	501	1200	1281	1232.6	15.99	0.77	7.6
Regular School Self-contained	2077	1200	1281	1240.03	15.08	0.84	5.97
Regular School Resource Room	488	1200	1281	1244.02	14.89	0.87	5.43
Regular School Primarily Self-contained	162	1200	1281	1247.45	12.33	0.87	4.45
Regular School General Education	36	1200	1281	1247.08	15.1	0.88	5.21
Undefined Classroom Setting	140	1200	1234	1200.24	-	-	-
Communicates Primarily Through Cries	264	1200	1257	1216.69	-	-	-
Uses Intentional Communication	541	1200	1275	1233.46	13.53	0.79	6.21
Uses Symbolic Language	2456	1200	1281	1243.91	12.82	0.87	4.57
Undefined Expressive Communication	143	1200	1234	1200.24	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-8. MSAA: Reliability: Subgroup-Mathematics Grade 3 Path A

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	2990	1200	1290	1236.6	16.93	0.59	10.86
Female	1024	1200	1290	1235.7	16.38	0.55	11.01
Male	1957	1200	1290	1237.07	17.2	0.61	10.79
Gender Undefined	9	1200	1260	1237.33	16.55	0.57	10.8
Hispanic or Latino	665	1200	1290	1236.76	17.42	0.6	10.97
American Indian or Alaska Native	192	1200	1290	1234.98	18.72	0.57	12.3
Asian	90	1200	1276	1234.47	16.91	0.51	11.78
Black or African American	759	1200	1290	1237.16	16.25	0.61	10.21
Native Hawaiian or Pacific Islander	29	1200	1256	1236.55	11	0.63	6.71
White (non-Hispanic)	1899	1200	1290	1236.73	17.02	0.59	10.92
Two or More Races (non-Hispanic)	636	1200	1290	1236.98	17.12	0.6	10.86
No Primary race/Ethnicity Undefined	44	1200	1263	1237.59	12.72	0.65	7.49
Currently receiving LEP services	117	1200	1283	1237.59	16.63	0.65	9.85
Not receiving LEP services	1240	1200	1290	1236.97	18.09	0.58	11.72
LEP: All Other Students	1633	1200	1290	1236.25	16.02	0.59	10.24
Economically Disadvantaged Students	998	1200	1290	1237.26	16.89	0.62	10.42
Non-economically Disadvantaged Students	876	1200	1290	1235.77	16.61	0.54	11.24
SES: All Other Students	1116	1200	1290	1236.67	17.2	0.59	10.95
Migrant	2	1237	1240	1238.5	-	-	-
Non-migrant	1224	1200	1290	1236.95	18.17	0.58	11.81
Undefined Migrant Status	1764	1200	1290	1236.36	16.02	0.6	10.16
Augmentative Communication	657	1200	1290	1230.88	15.01	0.36	12
No Augmentative Communication	2243	1200	1290	1239.45	15.78	0.67	9.08
Undefined Augmentative Communications	90	1200	1290	1207.34	-	-	-
Hearing Loss	107	1200	1260	1223.7	-	-	-
Within Normal Limits	2797	1200	1290	1238.1	15.7	0.65	9.34
Undefined Hearing Loss	86	1200	1260	1203.83	-	-	-
Visual Impairment	212	1200	1276	1219.88	-	-	-
Within Normal Limits	2689	1200	1290	1238.91	15.02	0.68	8.54
Undefined Visual Impairment	89	1200	1271	1206.79	-	-	-
Sensory Stimuli Response	433	1200	1258	1218.6	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2482	1200	1290	1240.85	13.31	0.76	6.46
Undefined Receptive Language	75	1200	1200	1200	-	-	-
Special School	302	1200	1271	1230.09	15.54	0.34	12.62
Regular School Self-contained	1899	1200	1290	1236.3	15.79	0.58	10.2
Regular School Resource Room	397	1200	1290	1243.07	14.72	0.76	7.2
Regular School Primarily Self-contained	229	1200	1290	1245.99	14.09	0.82	6.01
Regular School General Education	88	1200	1290	1243.01	17.2	0.71	9.3
Undefined Classroom Setting	75	1200	1200	1200	-	-	-
Communicates Primarily Through Cries	382	1200	1267	1218.02	-	-	-
Uses Intentional Communication	725	1200	1290	1233.92	13.82	0.55	9.32
Uses Symbolic Language	1808	1200	1290	1243.12	12.53	0.81	5.41
Undefined Expressive Communication	75	1200	1200	1200	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-9. 2016–17 MSAA: Reliability: Subgroup- Mathematics Grade 4

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3304	1200	1290	1235.16	16.13	0.56	10.69
Female	1175	1200	1290	1234.35	15.59	0.51	10.87
Male	2125	1200	1290	1235.61	16.42	0.58	10.6
Gender Undefined	4	1217	1243	1235.25	12.39	0.88	4.27
Hispanic or Latino	767	1200	1290	1234.23	15.43	0.52	10.67
American Indian or Alaska Native	217	1200	1290	1233.46	17.33	0.55	11.68
Asian	131	1200	1284	1233.8	13.69	0.52	9.47
Black or African American	799	1200	1290	1236.47	16.42	0.63	10
Native Hawaiian or Pacific Islander	30	1200	1253	1231.33	16.79	0.35	13.57
White (non-Hispanic)	2106	1200	1290	1235.16	15.99	0.55	10.77
Two or More Races (non-Hispanic)	752	1200	1290	1234.44	15.16	0.53	10.35
No Primary race/Ethnicity Undefined	45	1200	1257	1233.16	14.96	0.51	10.45
Currently receiving LEP services	146	1200	1268	1235.91	15.11	0.6	9.5
Not receiving LEP services	1363	1200	1290	1235.9	17.61	0.56	11.66
LEP: All Other Students	1795	1200	1290	1234.54	14.99	0.56	9.98
Economically Disadvantaged Students	1138	1200	1290	1236.02	17.03	0.61	10.63
Non-economically Disadvantaged Students	966	1200	1290	1235.06	15.84	0.54	10.79
SES: All Other Students	1200	1200	1290	1234.43	15.46	0.52	10.66
Migrant	0	-	-	-	-	-	-
Non-migrant	1356	1200	1290	1235.72	17.52	0.55	11.72
Undefined Migrant Status	1948	1200	1290	1234.78	15.09	0.57	9.91
Augmentative Communication	674	1200	1277	1230.42	15.47	0.43	11.68
No Augmentative Communication	2539	1200	1290	1237.49	15	0.64	8.97
Undefined Augmentative Communications	91	1200	1253	1205.26	-	-	-
Hearing Loss	135	1200	1253	1222.36	-	-	-
Within Normal Limits	3081	1200	1290	1236.6	15.03	0.63	9.08
Undefined Hearing Loss	88	1200	1277	1204.61	-	-	-
Visual Impairment	203	1200	1277	1221.87	-	-	-
Within Normal Limits	3015	1200	1290	1236.98	14.57	0.66	8.45
Undefined Visual Impairment	86	1200	1252	1202.95	-	-	-
Sensory Stimuli Response	399	1200	1255	1217.76	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2824	1200	1290	1238.6	13.21	0.76	6.51
Undefined Receptive Language	81	1200	1246	1201.07	-	-	-
Special School	332	1200	1262	1228.89	16.19	0.36	12.95
Regular School Self-contained	2029	1200	1290	1235.12	15.45	0.58	10.05
Regular School Resource Room	523	1200	1290	1240.07	13.08	0.77	6.23
Regular School Primarily Self-contained	233	1200	1284	1241.83	12.12	0.75	6.04
Regular School General Education	106	1200	1290	1242.72	15.56	0.72	8.27
Undefined Classroom Setting	81	1200	1246	1201.07	-	-	-
Communicates Primarily Through Cries	347	1200	1255	1216.72	-	-	-
Uses Intentional Communication	704	1200	1290	1232.96	13.91	0.62	8.52
Uses Symbolic Language	2172	1200	1290	1240.09	12.74	0.77	6.1
Undefined Expressive Communication	81	1200	1246	1201.07	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-10. 2016–17 MSAA: Reliability: Subgroup- Mathematics Grade 5

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3376	1200	1290	1237.29	17.46	0.44	13.06
Female	1146	1200	1290	1236.9	17.45	0.41	13.38
Male	2224	1200	1290	1237.53	17.43	0.46	12.84
Gender Undefined	6	1200	1261	1222.67	-	-	-
Hispanic or Latino	758	1200	1290	1236.27	16.76	0.35	13.5
American Indian or Alaska Native	182	1200	1282	1237.62	16.01	0.45	11.89
Asian	100	1200	1273	1235.68	12.87	0.31	10.68
Black or African American	868	1200	1290	1239.33	17.21	0.54	11.68
Native Hawaiian or Pacific Islander	25	1223	1269	1240.04	11.07	0.83	4.54
White (non-Hispanic)	2160	1200	1290	1236.54	17.81	0.4	13.81
Two or More Races (non-Hispanic)	707	1200	1290	1236.25	16.46	0.33	13.42
No Primary race/Ethnicity Undefined	31	1200	1273	1238.26	12.68	0.55	8.47
Currently receiving LEP services	150	1200	1290	1238.14	15.31	0.46	11.21
Not receiving LEP services	1446	1200	1290	1237.96	19.21	0.45	14.31
LEP: All Other Students	1780	1200	1290	1236.67	16.07	0.43	12.12
Economically Disadvantaged Students	1165	1200	1290	1238.62	18.22	0.51	12.82
Non-economically Disadvantaged Students	1050	1200	1290	1236.48	17.25	0.4	13.4
SES: All Other Students	1161	1200	1290	1236.68	16.8	0.4	13
Migrant	1	1230	1230	1230	-	-	5.4
Non-migrant	1413	1200	1290	1237.99	19.2	0.45	14.26
Undefined Migrant Status	1962	1200	1290	1236.79	16.09	0.43	12.13
Augmentative Communication	613	1200	1290	1232.06	14.4	0.18	13.06
No Augmentative Communication	2660	1200	1290	1239.66	16.8	0.55	11.26
Undefined Augmentative Communications	103	1200	1266	1207.12	-	-	-
Hearing Loss	114	1200	1277	1226.14	-	-	-
Within Normal Limits	3162	1200	1290	1238.68	16.31	0.54	11.02
Undefined Hearing Loss	100	1200	1254	1205.96	-	-	-
Visual Impairment	190	1200	1264	1222.84	-	-	-
Within Normal Limits	3084	1200	1290	1239.2	15.99	0.58	10.35
Undefined Visual Impairment	102	1200	1259	1206.27	-	-	-
Sensory Stimuli Response	379	1200	1290	1221.09	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2911	1200	1290	1240.48	14.77	0.68	8.3
Undefined Receptive Language	86	1200	1246	1200.53	-	-	-
Special School	350	1200	1290	1231.21	17.22	0.17	15.66
Regular School Self-contained	2044	1200	1290	1237.22	16.77	0.48	12.14
Regular School Resource Room	545	1200	1290	1243.21	14.33	0.71	7.69
Regular School Primarily Self-contained	245	1200	1290	1243.35	13.78	0.7	7.51
Regular School General Education	104	1200	1290	1244.79	15.64	0.77	7.44
Undefined Classroom Setting	88	1200	1246	1200.52	-	-	-
Communicates Primarily Through Cries	311	1200	1277	1218.86	-	-	-
Uses Intentional Communication	697	1200	1290	1234.19	14.97	0.33	12.28
Uses Symbolic Language	2282	1200	1290	1242.13	14.34	0.76	7.04
Undefined Expressive Communication	86	1200	1246	1200.53	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-11. 2016–17 MSAA: Reliability: Subgroup-Mathematics Grade 6

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3543	1200	1290	1236.93	16.9	0.75	8.42
Female	1242	1200	1290	1236.3	16.9	0.73	8.74
Male	2292	1200	1290	1237.27	16.9	0.76	8.24
Gender Undefined	9	1200	1256	1238	15.94	0.71	8.65
Hispanic or Latino	755	1200	1290	1236.52	16.1	0.75	8.07
American Indian or Alaska Native	198	1200	1290	1235.12	17.94	0.71	9.71
Asian	126	1200	1283	1240.25	16.67	0.83	6.96
Black or African American	894	1200	1290	1237.17	16.55	0.76	8.11
Native Hawaiian or Pacific Islander	29	1200	1283	1234.34	20.52	0.72	10.82
White (non-Hispanic)	2254	1200	1290	1236.93	17.05	0.75	8.54
Two or More Races (non-Hispanic)	732	1200	1290	1236.77	16.44	0.75	8.23
No Primary race/Ethnicity Undefined	45	1200	1273	1238.36	15	0.77	7.21
Currently receiving LEP services	128	1200	1283	1240.79	16.21	0.79	7.36
Not receiving LEP services	1501	1200	1290	1236.7	18.29	0.74	9.41
LEP: All Other Students	1914	1200	1290	1236.86	15.75	0.77	7.63
Economically Disadvantaged Students	1147	1200	1290	1238.24	16.85	0.77	8.13
Non-economically Disadvantaged Students	1190	1200	1290	1235.96	16.95	0.73	8.74
SES: All Other Students	1206	1200	1290	1236.66	16.84	0.75	8.38
Migrant	1	1236	1236	1236	-	-	3.5
Non-migrant	1494	1200	1290	1236.98	18.26	0.74	9.32
Undefined Migrant Status	2048	1200	1290	1236.9	15.85	0.76	7.7
Augmentative Communication	619	1200	1283	1229.24	15.36	0.61	9.64
No Augmentative Communication	2833	1200	1290	1239.61	15.75	0.79	7.27
Undefined Augmentative Communications	91	1200	1260	1206.07	-	-	-
Hearing Loss	113	1200	1283	1224.91	-	-	-
Within Normal Limits	3336	1200	1290	1238.21	15.82	0.78	7.42
Undefined Hearing Loss	94	1200	1263	1205.96	-	-	-
Visual Impairment	208	1200	1290	1221.67	-	-	-
Within Normal Limits	3226	1200	1290	1238.79	15.37	0.79	7.05
Undefined Visual Impairment	109	1200	1283	1211.06	-	-	-
Sensory Stimuli Response	403	1200	1283	1217.35	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	3062	1200	1290	1240.44	13.9	0.83	5.76
Undefined Receptive Language	78	1200	1237	1200.47	-	-	-
Special School	375	1200	1290	1228.26	16.09	0.58	10.4
Regular School Self-contained	2259	1200	1290	1237.25	15.9	0.76	7.8
Regular School Resource Room	515	1200	1290	1242.57	13.68	0.83	5.64
Regular School Primarily Self-contained	218	1200	1290	1243.94	13.68	0.83	5.58
Regular School General Education	96	1200	1290	1247.77	18.58	0.84	7.52
Undefined Classroom Setting	80	1200	1237	1200.46	-	-	-
Communicates Primarily Through Cries	354	1200	1283	1216.84	-	-	-
Uses Intentional Communication	753	1200	1290	1232.37	14.18	0.69	7.83
Uses Symbolic Language	2358	1200	1290	1242.61	13.16	0.84	5.2
Undefined Expressive Communication	78	1200	1237	1200.47	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-12. 2016–17 MSAA: Reliability: Subgroup- Mathematics Grade 7

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3613	1200	1290	1237.38	16.67	0.73	8.71
Female	1264	1200	1290	1236.85	16.64	0.71	8.9
Male	2340	1200	1290	1237.66	16.7	0.73	8.62
Gender Undefined	9	1226	1249	1239.22	7.51	0.76	3.71
Hispanic or Latino	761	1200	1286	1236.8	16.45	0.72	8.72
American Indian or Alaska Native	202	1200	1290	1236	17.57	0.7	9.61
Asian	113	1200	1286	1237.38	15.35	0.75	7.61
Black or African American	875	1200	1290	1238.51	15.48	0.74	7.82
Native Hawaiian or Pacific Islander	31	1200	1246	1232.39	10.53	0.54	7.15
White (non-Hispanic)	2328	1200	1290	1237.2	17.17	0.73	9
Two or More Races (non-Hispanic)	738	1200	1290	1236.84	16.57	0.72	8.78
No Primary race/Ethnicity Undefined	56	1200	1266	1237.04	14.78	0.7	8.1
Currently receiving LEP services	130	1200	1286	1241.92	12.34	0.8	5.48
Not receiving LEP services	1454	1200	1290	1237.4	17.94	0.73	9.36
LEP: All Other Students	2029	1200	1290	1237.08	15.91	0.72	8.39
Economically Disadvantaged Students	1183	1200	1290	1238.12	16.93	0.74	8.66
Non-economically Disadvantaged Students	1201	1200	1290	1237.65	16.13	0.73	8.4
SES: All Other Students	1229	1200	1290	1236.41	16.89	0.71	9.06
Migrant	1	1235	1235	1235	-	-	3.8
Non-migrant	1421	1200	1290	1237.75	17.91	0.73	9.28
Undefined Migrant Status	2191	1200	1290	1237.15	15.81	0.72	8.33
Augmentative Communication	586	1200	1270	1230.87	14.17	0.56	9.45
No Augmentative Communication	2910	1200	1290	1239.95	15.51	0.77	7.51
Undefined Augmentative Communications	117	1200	1263	1206.28	-	-	-
Hearing Loss	118	1200	1261	1222.28	-	-	-
Within Normal Limits	3382	1200	1290	1238.99	15.13	0.76	7.41
Undefined Hearing Loss	113	1200	1259	1204.98	-	-	-
Visual Impairment	209	1200	1266	1218.43	-	-	-
Within Normal Limits	3270	1200	1290	1239.74	14.45	0.78	6.83
Undefined Visual Impairment	134	1200	1276	1209.34	-	-	-
Sensory Stimuli Response	374	1200	1255	1217.37	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	3139	1200	1290	1240.96	13.31	0.8	5.94
Undefined Receptive Language	100	1200	1200	1200	-	-	-
Special School	401	1200	1290	1230.07	17.27	0.59	11.06
Regular School Self-contained	2238	1200	1290	1237.44	15.1	0.73	7.87
Regular School Resource Room	550	1200	1290	1243.29	13.13	0.8	5.83
Regular School Primarily Self-contained	231	1200	1290	1247.12	14.01	0.84	5.52
Regular School General Education	93	1200	1290	1248.67	14.8	0.85	5.81
Undefined Classroom Setting	100	1200	1200	1200	-	-	-
Communicates Primarily Through Cries	323	1200	1259	1215.15	-	-	-
Uses Intentional Communication	696	1200	1290	1234.19	12.78	0.66	7.42
Uses Symbolic Language	2494	1200	1290	1242.65	12.89	0.82	5.47
Undefined Expressive Communication	100	1200	1200	1200	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-13. 2016–17 MSAA: Reliability: Subgroup- Mathematics Grade 8

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3632	1200	1290	1237.32	17.37	0.52	12.04
Female	1315	1200	1290	1236.66	18.04	0.51	12.66
Male	2306	1200	1290	1237.73	16.93	0.53	11.63
Gender Undefined	11	1200	1270	1230.09	-	-	-
Hispanic or Latino	733	1200	1290	1236.93	17.08	0.5	12.12
American Indian or Alaska Native	181	1200	1270	1236.32	16.53	0.42	12.61
Asian	99	1200	1285	1235.12	18.49	0.55	12.39
Black or African American	944	1200	1290	1237.65	16.74	0.55	11.24
Native Hawaiian or Pacific Islander	25	1200	1258	1233	18.31	0.33	14.98
White (non-Hispanic)	2300	1200	1290	1237.34	17.55	0.51	12.23
Two or More Races (non-Hispanic)	716	1200	1290	1236.94	17.13	0.51	12.03
No Primary race/Ethnicity Undefined	49	1200	1285	1240.33	14.18	0.78	6.72
Currently receiving LEP services	101	1200	1290	1240.97	15.19	0.58	9.82
Not receiving LEP services	1549	1200	1290	1236.89	18.59	0.48	13.35
LEP: All Other Students	1982	1200	1290	1237.46	16.44	0.55	11.03
Economically Disadvantaged Students	1186	1200	1290	1239.04	17.52	0.57	11.42
Non-economically Disadvantaged Students	1225	1200	1290	1236.18	17.54	0.48	12.68
SES: All Other Students	1221	1200	1290	1236.79	16.92	0.5	11.97
Migrant	0	-	-	-	-	-	-
Non-migrant	1504	1200	1290	1236.96	18.22	0.48	13.16
Undefined Migrant Status	2128	1200	1290	1237.58	16.74	0.55	11.19
Augmentative Communication	534	1200	1270	1229.83	15.12	0.17	13.79
No Augmentative Communication	2970	1200	1290	1239.99	16.09	0.63	9.76
Undefined Augmentative Communications	128	1200	1270	1206.56	-	-	-
Hearing Loss	122	1200	1270	1227.03	-	-	-
Within Normal Limits	3384	1200	1290	1238.85	16.01	0.61	10.01
Undefined Hearing Loss	126	1200	1290	1206.09	-	-	-
Visual Impairment	214	1200	1275	1220.59	-	-	-
Within Normal Limits	3288	1200	1290	1239.65	15.44	0.65	9.09
Undefined Visual Impairment	130	1200	1285	1205.98	-	-	-
Sensory Stimuli Response	332	1200	1258	1216.27	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	3192	1200	1290	1240.77	14.34	0.74	7.29
Undefined Receptive Language	108	1200	1200	1200	-	-	-
Special School	404	1200	1285	1230.08	16.74	0.23	14.72
Regular School Self-contained	2284	1200	1290	1237.31	16.14	0.56	10.76
Regular School Resource Room	525	1200	1290	1244.33	13.39	0.77	6.46
Regular School Primarily Self-contained	219	1200	1290	1246.98	14.05	0.79	6.5
Regular School General Education	93	1218	1290	1249.66	14.64	0.86	5.52
Undefined Classroom Setting	107	1200	1200	1200	-	-	-
Communicates Primarily Through Cries	278	1200	1258	1216.07	-	-	-
Uses Intentional Communication	719	1200	1285	1232.54	15.17	0.41	11.69
Uses Symbolic Language	2526	1200	1290	1242.63	13.75	0.78	6.39
Undefined Expressive Communication	109	1200	1200	1200	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

Table P-14. 2016–17 MSAA: Reliability: Subgroup-Mathematics Grade 11

Description	*Number of Students*	*Raw Score*				*IRT Reliability*	*Standard Error*
		Maximum	*Minimum*	*Mean*	*Standard Deviation*		
All Students	3404	1200	1290	1237.02	16.17	0.52	11.24
Female	1222	1200	1290	1236.69	16.82	0.51	11.76
Male	2174	1200	1290	1237.19	15.81	0.52	10.95
Gender Undefined	8	1232	1252	1240.12	6.77	0.79	3.09
Hispanic or Latino	594	1200	1290	1235.67	15.22	0.46	11.15
American Indian or Alaska Native	154	1200	1274	1234.82	16.28	0.44	12.21
Asian	110	1200	1285	1237.61	15.23	0.56	10.05
Black or African American	855	1200	1290	1238.24	15.42	0.57	10.14
Native Hawaiian or Pacific Islander	23	1200	1285	1229.65	-	-	-
White (non-Hispanic)	2195	1200	1290	1236.76	16.34	0.51	11.48
Two or More Races (non-Hispanic)	567	1200	1290	1235.85	15.11	0.48	10.89
No Primary race/Ethnicity Undefined	37	1200	1274	1240.41	15.54	0.62	9.53
Currently receiving LEP services	82	1200	1269	1238.95	13.02	0.59	8.32
Not receiving LEP services	1446	1200	1290	1237.36	16.45	0.51	11.55
LEP: All Other Students	1876	1200	1290	1236.67	16.07	0.52	11.11
Economically Disadvantaged Students	1024	1200	1290	1239.04	16.74	0.58	10.86
Non-economically Disadvantaged Students	1293	1200	1290	1237.08	15.78	0.51	11.09
SES: All Other Students	1087	1200	1285	1235.03	15.85	0.45	11.76
Migrant	0	-	-	-	-	-	-
Non-migrant	1433	1200	1290	1237.38	16.34	0.51	11.47
Undefined Migrant Status	1971	1200	1290	1236.75	16.04	0.52	11.07
Augmentative Communication	497	1200	1285	1231.95	16.08	0.34	13.02
No Augmentative Communication	2755	1200	1290	1239.76	13.98	0.61	8.74
Undefined Augmentative Communications	152	1200	1248	1203.79	-	-	-
Hearing Loss	128	1200	1285	1225.91	-	-	-
Within Normal Limits	3131	1200	1290	1239.11	14.02	0.59	8.92
Undefined Hearing Loss	145	1200	1243	1201.52	-	-	-
Visual Impairment	191	1200	1285	1223.9	-	-	-
Within Normal Limits	3053	1200	1290	1239.49	13.59	0.61	8.46
Undefined Visual Impairment	160	1200	1260	1205.4	-	-	-
Sensory Stimuli Response	303	1200	1285	1218.69	-	-	-

continued

Description	Number of Students	Raw Score				IRT Reliability	Standard Error
		Maximum	Minimum	Mean	Standard Deviation		
Follow Directions	2962	1200	1290	1240.61	12.34	0.67	7.12
Undefined Receptive Language	139	1200	1243	1200.31	-	-	-
Special School	501	1200	1285	1233.05	15.97	0.39	12.46
Regular School Self-contained	2077	1200	1290	1238.7	14.14	0.57	9.27
Regular School Resource Room	488	1200	1290	1241.88	13.83	0.68	7.87
Regular School Primarily Self-contained	162	1200	1274	1243.55	10.37	0.78	4.82
Regular School General Education	36	1200	1285	1242.33	19.04	0.71	10.21
Undefined Classroom Setting	140	1200	1243	1200.31	-	-	-
Communicates Primarily Through Cries	264	1200	1248	1217.56	-	-	-
Uses Intentional Communication	541	1200	1285	1234.8	13.11	0.43	9.93
Uses Symbolic Language	2456	1200	1290	1241.73	12.02	0.7	6.55
Undefined Expressive Communication	143	1200	1243	1200.3	-	-	-

Note: Values are calculated only for subgroups with 100 or more students.

.

APPENDIX Q—DECISION ACCURACY AND CONSISTENCY RESULTS

Table Q-1. 2017–18 MSAA: Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Performance Level

Content Area	*Grade*	*Overall*	*Kappa*	*Conditional on Level*			
				Level 1	*Level 2*	*Level 3*	*Level 4*
ELA	3	0.8(0.73)	0.62	0.87(0.85)	0.56(0.42)	0.77(0.7)	0.85(0.79)
	4	0.8(0.73)	0.62	0.9(0.84)	0.59(0.47)	0.82(0.76)	0.82(0.72)
	5	0.78(0.7)	0.58	0.83(0.77)	0.7(0.57)	0.79(0.72)	0.82(0.74)
	6	0.79(0.71)	0.59	0.8(0.72)	0.69(0.58)	0.83(0.75)	0.84(0.78)
	7	0.79(0.71)	0.6	0.89(0.83)	0.54(0.42)	0.81(0.72)	0.79(0.74)
	8	0.78(0.7)	0.6	0.87(0.79)	0.7(0.59)	0.75(0.66)	0.84(0.78)
	11	0.81(0.74)	0.63	0.87(0.81)	0.56(0.43)	0.85(0.79)	0.81(0.77)
Mathematics	3	0.76(0.67)	0.55	0.83(0.75)	0.64(0.53)	0.77(0.68)	0.87(0.76)
	4	0.74(0.65)	0.51	0.8(0.7)	0.61(0.51)	0.77(0.68)	0.86(0.75)
	5	0.73(0.63)	0.51	0.79(0.69)	0.6(0.49)	0.71(0.63)	0.88(0.78)
	6	0.76(0.67)	0.56	0.78(0.72)	0.57(0.45)	0.79(0.68)	0.85(0.81)
	7	0.73(0.64)	0.5	0.77(0.64)	0.56(0.47)	0.79(0.68)	0.86(0.79)
	8	0.76(0.68)	0.56	0.84(0.77)	0.54(0.43)	0.77(0.66)	0.85(0.8)
	11	0.74(0.65)	0.5	0.81(0.62)	0.55(0.45)	0.76(0.7)	0.9(0.78)

Table Q-2. 2017–18 MSAA: Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint

Content Area	Grade	*Level 1/Level 2*				*Level 2/Level 3*				*Level 3/Level 4*			
		Accuracy	*Consistency*	*False*		*Accuracy*	*Consistency*	*False*		*Accuracy*	*Consistency*	*False*	
				Positive	*Negative*			*Positive*	*Negative*			*Positive*	*Negative*
ELA	3	0.92	0.89	0.05	0.03	0.93	0.90	0.03	0.04	0.94	0.92	0.03	0.03
	4	0.92	0.89	0.03	0.04	0.93	0.89	0.03	0.04	0.95	0.94	0.03	0.02
	5	0.92	0.88	0.04	0.04	0.92	0.88	0.04	0.05	0.95	0.92	0.03	0.03
	6	0.93	0.91	0.03	0.03	0.92	0.88	0.04	0.04	0.94	0.92	0.03	0.03
	7	0.93	0.90	0.03	0.04	0.93	0.90	0.04	0.03	0.93	0.90	0.03	0.04
	8	0.92	0.89	0.03	0.05	0.93	0.89	0.03	0.05	0.93	0.91	0.03	0.03
	11	0.94	0.91	0.03	0.03	0.93	0.91	0.03	0.04	0.94	0.91	0.03	0.04
Mathematics	3	0.91	0.87	0.05	0.05	0.90	0.86	0.05	0.04	0.95	0.93	0.03	0.02
	4	0.92	0.88	0.04	0.05	0.88	0.83	0.06	0.06	0.95	0.92	0.03	0.02
	5	0.90	0.86	0.05	0.05	0.88	0.84	0.06	0.06	0.94	0.91	0.04	0.02
	6	0.90	0.87	0.06	0.04	0.91	0.87	0.05	0.04	0.94	0.91	0.03	0.04
	7	0.89	0.85	0.04	0.07	0.88	0.83	0.07	0.05	0.95	0.93	0.03	0.02
	8	0.91	0.88	0.04	0.04	0.90	0.86	0.06	0.04	0.94	0.92	0.03	0.03
	11	0.92	0.88	0.02	0.06	0.87	0.82	0.06	0.07	0.94	0.92	0.04	0.02

Note: Due to the small sample size, students in Levels 3 and 4 were collapsed for purposes of the decision accuracy and consistency analysis.