Multi-State Alternate Assessment

2015–16 Technical Report

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

TABLE OF CONTENTS

CHAPTER 1 CURRENT YEAR UPDATES 1
1.1 MEASURED VALIDITY STATEMENT 1
CHAPTER 2 OVERVIEW OF MSAA 2
2.1 HISTORY OF MSAA 2
2.1.1 Core Beliefs 3
2.1.2 Stakeholders 3
2.2 PURPOSES AND USES OF MSAA 4
2.3 MSAA PARTICIPATION 5
CHAPTER 3 TEST CONTENT 6
3.1 HISTORY OF ALTERNATE ACHIEVEMENT STANDARDS AND CORE CONTENT CONNECTORS 6
3.1.1 The Learning Progression Frameworks 6
3.1.2 Core Content Connectors 7
3.2 ALIGNMENT AND LINKAGES 9
3.3 ASSESSMENT DESIGN 9
3.3.1 Operational Design 9
3.3.2 Operational Core Items and Embedded Field-Test Items 10
3.3.3 Item Design and Administration 12
3.3.4 Item Components 15
3.3.4.1 Selected-Response: Reading, Writing, Mathematics 15
3.3.4.2 Constructed-Response: Mathematics 15
3.3.4.3 Constructed-Response: Writing 16
3.3.4.4 Open-Response: Foundational Reading 16
3.4 CONTENT AND BLUEPRINTS 16
3.4.1 Mathematics 17
3.4.2 English Language Arts 18
CHAPTER 4 TEST DEVELOPMENT 19
4.1 GENERAL PHILOSOPHY AND ROLE OF TEST CONSTRUCTION SUBCOMMITTEE IN TEST DEVELOPMENT 19
CHAPTER 5 TRAINING AND ADMINISTRATION 21
5.1 TEST ADMINISTRATOR TRAINING 21
5.1.1 Training Modules 21
5.1.2 Test Administration Manual 22
5.1.2.1 MSAA Technical Support—Who to Call for Help 23
5.1.2.2 Overview of MSAA 23
5.1.2.3 Description of MSAA Test Sessions and Description of MSAA Item Types 23
5.1.2.4 Responsibilities of Test Administrators 23
5.1.2.5 Student Participation Criteria 24
5.1.2.6 Optimal Testing Conditions, Assessment Features, Test Accommodations, and Procedures for Assessing Students Who Are Blind, Deaf, or Deaf-Blind 25
5.1.2.7 Testing Integrity and Appropriate Test Practices 25

5.1.2.8 Administration of the Test ... 27
5.2 TEST COORDINATOR TRAINING ... 30
5.2.1 Training Modules ... 30
5.2.2 Test Administration Manual ... 31
5.2.2.1 MSAA Technical Support—Who to Call for Help ... 31
5.2.2.2 Responsibilities of Test Coordinators ... 31
5.2.2.3 Testing Integrity and Appropriate Test Practices ... 32
5.3 OPERATIONAL TEST ADMINISTRATION ... 33
5.3.1 Session Structure ... 33
5.3.2 Administration Support ... 33
5.3.3 MSAA Service Center ... 34
5.3.4 Additional Supports ... 34
5.3.5 Monitoring and Quality Control ... 35
5.3.6 Operational Test Survey Results ... 35
CHAPTER 6 SCORING ... 38
6.1 ITEM SCORING PROCESS ... 38
6.1.1 Scoring Processes and Rules for Selected-Response and Constructed-Response Items in Mathematics and Selected-Response Items in Reading and Writing ... 38
6.1.1.1 Overview of scoring process by item type ... 38
6.2 SCORING PROCESSES AND RULES FOR FIELD-TEST WRITING CONSTRUCTED-RESPONSE ITEMS ... 41
6.2.1 Overview of Scoring Process by Item Type ... 41
6.2.1.1 Constructed-response: Writing ... 41
6.2.1.2 Field-Tested Writing Constructed-response—Administrator training and monitoring ... 41
6.2.1.3 Field-Tested Writing Constructed-response: Scoring ... 42
6.2.2 Scorer Recruitment and Qualifications ... 43
6.2.3 Measured Progress Staff and Scoring Leadership ... 43
6.2.4 Qualification ... 45
6.2.5 Methodology for Scoring Field-Tested Writing Constructed-Response Items ... 47
6.2.6 Monitoring of Scoring Quality Control ... 50
6.2.6.1 Calibration Sets ... 50
6.2.6.2 Read-behind Scoring ... 50
6.2.6.3 Double-blind Scoring ... 51
6.2.7 Quality and Production Management Reports ... 52
6.2.8 Inter-Rater Reliability ... 53
CHAPTER 7 REPORTING ... 55
7.1 INTRODUCTION ... 55
7.2 DEVELOPMENT AND APPROVAL ... 55
7.3 PRIMARY REPORTS ... 56
7.3.1 Student Report ... 57
7.3.2 Student Roster ... 57
7.3.3 Summary Report ... 58
7.3.4 Quality Assurance ... 59
CHAPTER 8 CLASSICAL ITEM ANALYSIS ... 60
8.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES ... 60

8.2 DIFFERENTIAL ITEM FUNCTIONING 63
8.3 DIMENSIONALITY ANALYSIS 65
CHAPTER 9 ITEM RESPONSE THEORY SCALING AND EQUATING 69
9.1 ITEM RESPONSE THEORY 69
9.2 ITEM RESPONSE THEORY RESULTS 70
9.3 EQUATING 73
9.4 ACHIEVEMENT STANDARDS 73
9.5 REPORTED SCALED SCORES 75
CHAPTER 10 RELIABILITY 77
10.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT 78
10.2 SUBGROUP RELIABILITY 79
10.3 RELIABILITY OF PERFORMANCE-LEVEL CATEGORIZATION 79
10.3.1 Accuracy and Consistency 81
CHAPTER 11 VALIDITY 82
11.1 SCALE VALIDATION AND POST-EQUATED CHECK OF PRE-EQUATED TESTS 82
11.2 TEST SCORE VALIDATION EVIDENCE 83
REFERENCES 85
APPENDICES 87

APPENDIX A TEST BLUEPRINTS
APPENDIX B DECISION RULES
APPENDIX C MSAA 2016 GUIDE FOR SCORE REPORT INTERPRETATION
APPENDIX D ITEM LEVEL CLASSICAL STATISTICS
APPENDIX E DIFFERENTIAL ITEM FUNCTIONING RESULTS
APPENDIX F ITEM RESPONSE THEORY PARAMETERS
APPENDIX G TEST CHARACTERISTIC CURVES & TEST INFORMATION FUNCTIONS
APPENDIX H RAW TO SCALED SCORE LOOKUP TABLES
APPENDIX I SCORE DISTRIBUTIONS
APPENDIX J CLASSICAL RELIABILITY
APPENDIX K DECISION ACCURACY AND CONSISTENCY RESULTS
APPENDIX L ACCOMMODATION FREQUENCIES
APPENDIX M TECHNICAL ADVISORY COMMITTEE MEMBERS
APPENDIX N ITEM REVIEW AND BIAS AND SENSITIVITY REVIEW COMMITTEE MEMBERS
APPENDIX O PARTICIPATION RATES

CHAPTER 1 CURRENT YEAR UPDATES

The overall design of the Multi-State Alternate Assessment (MSAA) remained the same for 2015–16 as the National Center and State Collaborative (NCSC) 2014–15 operational assessment. New to the 2015–16 assessment were the addition of a small number of selected-response field-test items, the field-testing of two writing prompts each at differing levels of complexity for each grade level, and the use of only two forms per grade for each content area. Test documentation was updated to reflect changes in the *Test Administration Manual* (TAM), *MSAA System User Guide for Test Administrators*, *MSAA System User Guide for Test Coordinators*, *Directions for Test Administration* (DTA), and *MSAA 2016 Guide for Score Report Interpretation Guide*. Additional detailed information is available in Chapter 3. In addition, the MSAA participating states and entities differ slightly from those participating in the NCSC. MSAA comprises Arizona, Arkansas, Maryland, Maine, Montana, the Pacific Assessment Consortium (PAC-6: Guam and the Commonwealth of Northern Mariana Islands [CNMI]), Rhode Island, South Dakota, Tennessee, The U.S. Virgin Islands (USVI), and Washington D.C.

1.1 MEASURED VALIDITY STATEMENT

The 2015–16 report will describe several technical aspects of the MSAA in an effort to contribute to the accumulation of validity evidence to support MSAA score interpretations. Because the interpretations of test scores, not the test itself, are evaluated for validity, this report presents documentation to substantiate intended interpretations (AERA et al., 2014). Each section in this report contributes important information to MSAA: test development, test alignment, test administration, scoring, reliability, performance levels, and reporting.

Standards for Educational and Psychological Testing (AERA et al., 2014) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five general areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. Although each of these sources may speak to a different aspect of validity, they are not distinct types of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

CHAPTER 2 OVERVIEW OF MSAA

MSAA assesses English language arts (ELA; reading and writing) and mathematics at grades 3–8 and 11 and is aligned to the States' Content Standards and the MSAA Core Content Connectors. MSAA is a computer–based, on-demand 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 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 designed to be administered one-on-one, and may be delivered via an online or paper–pencil format. The needs of the student may also be addressed through other supports, such as assessment features built into the platform, and accommodations such as: use of assistive technology, paper versions of the items (state specific), having a scribe, communicating passages, items and response options in sign language.

Each content area consists of 30–40 items, which 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.

2.1 HISTORY OF MSAA

Work leading up to MSAA began in late 2010, when NCSC began development of the NCSC Alternate Assessments based on Alternate Achievement Standards (AA-AAS) for students with the most significant cognitive disabilities. This work culminated in the operationalized NCSC assessment in the spring of 2015. For additional information about the NCSC assessment, please refer to the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* or contact the MSAA states at MSAA@AZED.gov.

This assessment was developed through the research and development completed by NCSC and has been carried forward by the MSAA state partners. MSAA is currently being administered by 11 participating states and territories: Arizona, Arkansas, Maryland, Maine, Montana, the Pacific Assessment Consortium

(PAC-6: Guam and the Commonwealth of Northern Mariana Islands [CNMI]), Rhode Island, South Dakota, Tennessee, The U.S. Virgin Islands (USVI), and Washington D.C.

2.1.1 Core Beliefs

The core beliefs of MSAA began with NCSC, and were laid out in the prior planning and development of that assessment. 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. They required a blend of policy, educational, and technical solutions. Through policy discussions and in iterative research and design steps, the partners arrived at a shared philosophy and guiding principles that are reflected in the overall project resources. These project resources include the comprehensive system of curriculum, instruction, classroom assessment, and professional development, as well as in the operational assessment design.

MSAA partners, 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 to 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 on the assessment. The NCSC Theory of Action, available at www.ncscpartners.org/Media/Default/PDFs/Resources/NCSCBrief9.pdf, was developed to explain the bases for these resources and how they were intended to relate to one another, to college- and career-ready academic standards, and, ultimately, to the goals of having all students with significant cognitive disabilities leave high school ready to participate in college, careers, and their communities.

Practice-focused summaries of the foundational components reflected in the design of the NCSC assessment, known as the NCSC Brief series, are available to orient readers to the larger context of the comprehensive NCSC system of curriculum, instruction, assessment, and professional development. For additional information about the NCSC assessment, please refer to the *National Center and State Collaborative 2015 Operational Assessment Technical Manual* or contact the MSAA states at MSAA@AZED.gov.

2.1.2 Stakeholders

Many stakeholders are involved in the development of MSAA. State MSAA coordinators are key representatives from each partner state and together comprise the decision-making body for 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 partner states for each subcommittee's focal area.

The Manuals, User Guides, and Training Subcommittee that oversaw development of the *Test Administration Manual* (TAM), *MSAA System User Guide for Test Administrators, MSAA System User Guide for Coordinators*, *Procedures for Assessing Students Who Are Blind, Deaf, or Deaf- Blind: Additional Directions for Test Administration*, and the online training modules consisted of state coordinators from Arizona, Maine, Montana, Rhode Island, and South Dakota. The End of Test Survey Subcommittee that provided the content of the survey, determined relevant policies, and received the results after administration had representation from Arizona, Rhode Island, and South Dakota. Decisions and approvals related to the core item constructed sets and the front matter for the *Directions for Test Administration* (DTA) was addressed by the Test Construction and DTA Revisions Subcommittee, composed of state coordinators from Arizona, Maine, Maryland, Montana, Rhode Island, and South Dakota. Similarly, the Test Construction Writing Items Subcommittee reviewed and provided input and direction related to revisions made to the writing items. Finally, report revisions and decisions were the responsibility of the Reports Subcommittee, with representation from Arizona, Maine, Maryland, Rhode Island, South Dakota, Tennessee, and Washington, D.C.

2.2 Purposes and Uses of MSAA

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. MSAA is designed to measure academic content that is aligned to and derived from 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 possible. MSAA is administered in the areas of ELA and mathematics in grades 3–8 and 11.

MSAA was developed to ensure that all students with significant cognitive disabilities are able to participate in an assessment that is a measure of what they know and can do in relation to the grade-level State Content Standards. 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 to the grade-level State Content Standards.

MSAA's long-term goal is to ensure that students with the most significant cognitive disabilities achieve increasingly higher academic outcomes and leave high school capable of pursuing post-secondary options. A well-designed summative assessment alone is insufficient to achieve this goal.

MSAA is designed to meet the requirements of the Elementary and Secondary Education Act (ESEA) and Individuals with Disabilities Education Act (IDEA). These laws mandate that all students participate in assessments that measure student achievement on grade-level content standards.

2.3 MSAA PARTICIPATION

The criteria for student participation in 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:

Table 2-1. 2015–16 MSAA: 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 Common Core State Standards (CCSS).	Goals and instruction listed in the IEP for this student are linked to the enrolled grade-level State 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.

CHAPTER 3 TEST CONTENT

3.1 HISTORY OF ALTERNATE ACHIEVEMENT STANDARDS AND CORE CONTENT CONNECTORS

Designed specifically for students with significant cognitive disabilities, NCSC Alternate Assessment is a performance-based test that is aligned with grade-level State Content Standards for English language arts (ELA) and mathematics. The NCSC AA-AAS tests student performance in ELA and mathematics based on alternate achievement standards. The student's performance on the NCSC AA-AAS is reported by a scale score for each content area, as well as by a performance level. NCSC looked at the Learning Progression Frameworks (LPFs) together with the grade-level content expectations from the Common Core State Standards (CCSS) to identify and clarify the most salient grade-level, core academic content to guide instruction and assessment of students with the most significant cognitive disabilities from kindergarten through high school. This academic content is referred to as the Core Content Connectors (CCCs).

The NCSC state and center partners, which 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 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 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 CCCs to connect the LPFs to the CCSS took place and was led by NCSC with partner state involvement.

3.1.1 The Learning Progression Frameworks

The LPFs presents a broad description of the essential content and general sequencing for student learning and skill development (Hess, 2010). It is the pathway that students typically take toward mastering skills for college and career readiness, as they move through the grades. The LPFs give us the educational logic to help move students with the most significant cognitive disabilities along with their peers in an

educationally sound way. Experts at NCSC looked at these learning targets together with the grade-level content expectations from the CCSS to identify and clarify the most salient grade-level, core academic content to guide instruction and assessment of students with the most significant cognitive disabilities from kindergarten through high school. This academic content is referred to as the CCCs. The CCCs identify the academic content designed to frame instruction and assessment while retaining the grade-level content focus of the CCSS and the learning targets of the LPFs. 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 to the next grade. They are the starting point for instruction, not necessarily everything an individual student can and should learn.

3.1.2 Core Content Connectors

According to NCSC, as outlined in the National Center and State Collaborative—General Supervision Enhancement Grant Project Request for Proposal RFP #2012-08-01, the CCCs are the academic content designed to frame the instruction and assessment of students with the most significant cognitive disabilities. The CCCs create a connection between the LPFs and CCSS for students with the most significant cognitive disabilities. The CCCs illustrate the necessary knowledge and skills students with the most significant cognitive disabilities need to reach the learning targets within the LPFs and the CCSS. 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 identify academic content in each subject area to guide the instruction for students in this population and for the alternate assessment. 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.

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 teachable and assessable segments of content. The connectors and corresponding *Curriculum Resource Guides* were developed to help 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.

Table 3-1. 2015–16 MSAA: Example of the Core Content Connectors Across Grades - Mathematics Strand: Geometry Big Idea: Shapes and Figures-Their Attributes, Properties, and Corresponding Parts

	Grades K-2	Grades 3-4	Grades 5-6	Grades 7-8	HS
Properties and attributes of shapes and figures and their corresponding parts	K.G.M1a1 Recognize two-dimensional shapes (e.g., circle, square, triangle, rectangle) regardless of orientation or size	3.GM.1h1 Identify shared attributes of shapes	5.GM.1a1 Recognize properties of simple plane figures	7.GM.1e Construct or draw plane figures using properties	H.GM.1e Make formal geometric constructions with a variety of tools and methods
	K.GM.1a2 Recognize two-dimensional shapes in environment regardless of orientation or size	4.GM.1h2 Classify two-dimensional shapes based on attributes (# of angles)	5.GM.1b1 Distinguish plane figures by their properties	8.GM.1g1 Recognize congruent and similar figures	H.GM.1b Use definitions to determine congruency and similarity of figures
	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				

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, H.GM.1b is based on an original progress indicator within the progression that stated, “Using congruence and similarity relationships to solve problems, including triangle congruence relationships.” The letter/number shows the grade level (in this case, high school), the next letters show the content (e.g., geometry), and the rest of the code relates to where the connector falls in the progression. 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 row, 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

As part of the assessment development process, NCSC conducted an alignment study in 2015. The purpose of the alignment study was to examine the relationship between the spring 2015 operational assessment items and academic grade-level content targets prioritized for the NCSC AA-AAS. NCSC used results from this study to provide content validity evidence for the NCSC summative assessment. The National Alternate Assessment Center (NAAC) developed the Links for Academic Learning (LAL; Flowers, Wakeman, Browder, & Karvonen, 2007), which established the criteria used in this alignment study. The criteria are:

1. The content is academic;
2. The focus of achievement maintains fidelity with the content of the original grade-level standards (content centrality) and when possible, the specified performance (performance centrality);
3. The Depth of Knowledge (DOK) differs from academic grade-level content targets, as described in the academic grade-level content targets, but maintains high expectations for students with the most significant cognitive disabilities;
4. There is some differentiation in achievement across grade levels or grade bands; and
5. The potential barriers to demonstrating what students know and can do are minimized in the assessment.

The *Alignment of NCSC English Language Arts and Mathematics Operational Items to Academic Grade-Level Content Targets* report is available in the 2014–15 NCSC Technical Report.

3.3 ASSESSMENT DESIGN

3.3.1 Operational Design

The operational MSAA assessment program was designed to produce valid and reliable mathematics and ELA (reading and writing) scores. The mathematics and reading portions of the test comprised primarily selected-response items. Some grade levels in mathematics included constructed-response items. Grade 3 and grade 4 ELA included foundational reading items. Writing comprised only selected-response items. The selected-response items for writing were from stand-alone writing items and a Tier 1 writing prompt.

The items utilized in the 15-16 assessment vary in complexity. There was a variety of tier levels used. Items were built as item families where each tier within the family addressed 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). Additional detailed information about the item development is available in section 3.3.3 of Chapter 3.

Within a particular grade and content area, two forms per grade were developed based on Form 1 of the 2014–15 administration for the core items with field-test items embedded (the 2014–15 Form 1 was the most accessible form built at each grade and content). This form followed guidelines informed by the NCSC blueprint. The inclusion of the field-test items created two different fixed forms with the core set of items being the same across the two forms.

3.3.2 Operational Core Items and Embedded Field-Test Items

The items on the Form 1 tests were reviewed by Psychometrics and Research for any validity and reliability concerns. The core items from the 2014–15 assessments from the remaining forms were also reviewed for validity and reliability. Any items within the Form 1 that were flagged for statistics, were then reviewed by Content, Design & Development (CDD) staff to determine if an appropriate replacement item was available from the remaining item pool of 2014–15 items. Flagged items were replaced whenever possible.

The mathematics tests consisted of 40–41 items across two testing sessions per grade, with 26 items replaced, as shown in Tables 3-2 and 3-3:

Table 3-2. 2015–16 MSAA: Mathematics Items

Grade	Total Items	Field-Test Items (total across both forms)
3	40	10
4	40	10
5	41	10
6	40	10
7	40	10
8	40	10
11	40	10

Table 3-3. 2015–16 MSAA: Mathematics Item Replacements

Grades	Number of Items Replaced
3	5 (0)
4	5 (6)
5	7 (10)
6	2 (3)
7	2 (2)
8	4 (0)
11	1 (1)

*The numbers in parentheses are the number of items flagged in 14/15 Form 1

The field-test items were selected from previously developed items that went through at least one of the pilot testing phases. During the original item development process, these items followed a typical development cycle, including reviews by NCSC and also Item Content and Bias and Sensitivity reviews by panelists. In making the selection of field-test items for the 2015–16 mathematics sessions, the statistics from the pilot data were reviewed and items with the least number of flags were set as priority to be selected first. The field-test items were selected based on the following criteria:

1. Specific tiers (Five items from Tier 2 and five items from either Tier 3 or Tier 4)
2. Items with no statistical flags or items with the fewest number of statistical flags were selected
3. Content issues such as clueing and variety in stimuli

The mathematics field-test items were used as they were presented on the assessment during the pilot testing.

The ELA tests consisted of a varying number of operational items and field-tested items, as shown in Table 3-4. It ranged from 38 to 48 items, depending on the grade and form. Because of the nature of ELA reading passage sets, the item pool did not support replacement passage sets in most instances. For this reason, ELA replacements were made only at grade 4. Five replacement items were used at grade 4, as shown in Table 3-5. Specifically, in grade 4, one Tier 3 informational passage set was replaced with another Tier 3 informational passage set.

Table 3-4. 2015–16 MSAA: ELA Items

Grade	Total Items (depending on form)	Selected-Response Field-Test Items (total across both forms)	Writing Prompt Field-Test Items (total across both forms)
3	47–48	9	2
4	47–48	11	2
5	38–39	11	2
6	40	12	2
7	42	12	2
8	42	12	2
11	38–39	9	2

Table 3-5. 2015–16 MSAA: ELA Item Replacements

Grades	Number of Items Replaced
3	N/A (1*)
4	5 (1)
5	N/A (2)
6	N/A (1)
7	N/A (0)
8	N/A (1)
11	N/A (0)

*The numbers in parentheses are the number of items flagged in 14/15 Form 1

The field-test items were selected from previously developed items that went through at least one of the pilot testing phases. During the original item development process, these items followed a typical development cycle, including reviews by NCSC and also Item Content and Bias and Sensitivity reviews by panelists. In making the selection of field-test items for the 2015–16 ELA sessions, the statistics from the pilot data were reviewed and items with the least number of flags were set as priority to be selected first. The field-test items were selected based on the following criteria:

- Tier level was 2, 3, or 4 (one Tier 2 passage set for one form and one Tier 3 or Tier 4 passage set for the other form).
- Passage topic or writing topic was unique to the form.
- Passage topic and genre provided greater variety across the forms.
- Passage and item content was engaging, accurate, and free of regional bias.

The reading field-test items were used as they were presented on the assessment during the pilot testing, with the exception of removal of some of the multipart items to keep the length of the sessions consistent with the previous session administrations. The Tier 3 writing prompt field-test items were slightly revised to match the changes that were incorporated in the Tier 2 prompts during previous pilot testing. These changes included a simplification of both the prompt and the teacher administration instructions.

All constructed tests, identifying the items to be replaced and their replacement items, as well as the field-test items, were posted on a secure FTP site for MSAA subcommittee review and approval. A webinar was held with the MSAA subcommittee to explain the constructed sets process and review the recommended replacement items and field-test items selected. The MSAA subcommittee then had an opportunity to review the constructed sets and provide input. The MSAA subcommittee approved all field-test item selections, all ELA constructed sets as initially presented, and requested one replacement to the constructed sets for mathematics in grade 6.

3.3.3 Item Design and Administration

The NCSC partnership designed the NCSC AA-AAS to capture student performance at different levels of skill 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. NCSC's intentional assessment development process addressed the targeted grade-level academic content linked to evidence-based curricular and instructional materials, and resulted in useful information for educators and families.

The NCSC content development processes addressed levels of cognitive and language complexity, specifically addressing the CCSS and the heterogeneous characteristics of the target student population. The assessment items vary systematically in complexity yet remain aligned to the focal knowledge, skill, and ability (FKSA) behind the CCCs. NCSC designed its AA-AAS to capture student performance through two specific item design features: (1) levels of content complexity, and (2) degrees and types of scaffolds and

supports. Items were built as item families where each tier within the family addressed both the content complexity and the degree of scaffolding and support provided with the items. The items were written intentionally to measure a range of academic abilities within the target population. The array of item characteristics to facilitate varying student needs include reminders, examples, and models. These are provided to focus the student on the task and elicit a response without guiding the student's response.

Overall Item Structure

Multiple item families were developed for each CCC. An item family is a cluster of items that are specific to one CCC and contain one item from each tier level, for a total of four items. Each tier provides variable features and supports that provide multiple entry points for a variety of students to demonstrate their FKSAs. All items in an item family assess grade-level academic concepts defined by either the FKSAs or Essential Understandings (EUs).

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 Item 3), and Tier 1 (least complex).

Guidelines for graduated complexity of the items within a task were developed and implemented in the initial design phase of the project. These guidelines were used to create items of graduated complexity that address the same FKSA but provide increased levels of support and/or decreased levels of complexity so that students with different levels of cognitive ability are able to access the content. In addition to the tier specific item content, all items include an introductory sentence and teacher directives.

Overall Item Types

Item types consistent with NCSC design pattern and task template guidelines were developed. The item types that have been developed include selected-response, multiple-part selected-response, constructed-response, and open-response.

Selected-response items are multiple-choice items where a student selects a response from three options (two options at Tier 1) and the answer is worth 0 or 1 point. 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) and the answer is worth 0 or 1 point. The overall cluster is 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. Multiple-part selected-response items exist in ELA, but do not exist in mathematics. A decision was made during the mathematics development process to address CCCs with multiple components by unique items. Therefore, a CCC that might ask a student to identify and solve an equation would have two items written that require the student to identify the correct equation for a word problem and two items written that require the student to solve an equation.

Constructed-response items require the student to interact in some way with response information to provide a response. In the case of the constructed-response mathematics items for NCSC, all items are worth 0 or 1 point because the items ultimately ask the student to show whether he or she understands a single concept or not; therefore, a 0/1 point score assignment is appropriate. There are no operational ELA constructed-response items.

The Reading Foundational items are each worth 1 point and are administered either as open-response or selected-response items. For open-response Reading Foundational items, students are requested to read aloud five words as each is individually presented. Each item/word is presented to the student in a standardized, scripted sequence of steps. Students with clear and consistent oral speech are administered the open-response Reading Foundational Items. Students using means of communication other than oral speech, such as Augmentative and Alternative Communication (AAC) devices, American Sign Language (ASL), or eye gaze, are administered the selected-response Reading Foundational Items.

The writing open-response items have been developed such that the student is required to compose a permanent product. The student response is evaluated against a grade- and tier-specific rubric. Writing open-response items are used at Tiers 3 and 2 only. The Tier 1 writing task for these item families is in the form of a set of related selected-response items.

Administration

Test administrators (TAs) could begin with either the mathematics test or the ELA test. Once a content-area test was started, TAs were required to complete that test before beginning the test in the other content area. Each content-area test consisted of a set of testing sessions. Students were administered the test sessions in order for a given content area. ELA consisted of four test sessions and mathematics consisted of two test sessions at each grade level, as shown in Tables 3-6 and 3-7.

Table 3-6. 2015–16 MSAA: ELA Test Sessions

MSAA ELA Test			
Session 1: Reading	*Session 2: Reading*	*Session 3: Writing*	*Session 4: Writing*
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 items	One constructed-response writing item (Field tested)
Open-response foundational reading items (grades 3 and 4 only)	Open-response foundational reading items (grades 3 and 4 only)		

Table 3-7. 2015–16 MSAA: Mathematics Test Sessions

MSAA Mathematics Test	
Mathematics Session 1	*Mathematics Session 2*
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, MATHEMATICS*

Selected-response items are presented to students in a standard format. All directions and materials needed for administering selected-response items are in the secure *Directions for Test Administration* (DTA) that accompanies each test form. Every item is presented in the following order:

- Item stimulus (which may include a passage, passage part, picture, graphic, or other illustration)
- Item question
- Response options presented in stacked, or vertical, formation

Students select a response from the options and may do so in a variety of ways (e.g., using the computer mouse, verbalizing, gesturing, using eye gaze or communication devices, using assistive technology). Students enter responses into the MSAA system. If the 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 constructed-response items, in selected grades for mathematics, require students to develop an answer instead of selecting an answer from response options. Constructed-response items are presented as novel tasks using materials and content presented in an on-demand test format. Each item is presented to the student in a standardized, scripted sequence of steps culminating in the TA scoring the student performance using the Mathematics Scoring Rubrics. The Mathematics Scoring Rubrics provide scoring standards that must be used to evaluate student responses. Directions and materials needed for administering mathematics constructed-response items are included in the secure DTA accompanying each mathematics test form. The TA enters the student constructed-response score into the MSAA system.

3.3.4.3 *CONSTRUCTED-RESPONSE: WRITING*

The constructed-response writing item requires students to produce a permanent product in response to a writing prompt. The student, or a scribe, will record the response to the writing prompt on either the response template that is in the online MSAA system or on the paper response template that is included in the writing DTA.

The constructed-response writing item is presented to the student by the TA in a standardized, scripted sequence of steps and includes directions to present grade- and prompt-specific writing stimulus materials that need to be printed and prepared. All writing stimulus materials, including the response template, are identified by a card number and are included in the writing DTA. If the student uses a paper version of this template to write a response, the TA will:

- Annotate or interpret the student's writing directly on the student's written product if the TA determines that a novel reader (i.e., a scorer) may not be able to interpret a component (e.g., inventive spelling, penmanship, or use of symbolic expressions) of the student's written product.
- Transcribe or type exactly the student's written response, including any annotations, into the MSAA system.

Note: If the TA determines that a student can enter an online response using a keyboard, a printed response template does not need to be used during the administration.

Information related to preparation and the standard administration of the constructed-response writing item is included in the DTA.

3.3.4.4 *OPEN-RESPONSE: FOUNDATIONAL READING*

Open-response foundational reading items are included in the reading test in grades 3 and 4 only. The items are word identification tasks. Students identify three to five words as each item is presented. The TA enters the student's scores into the MSAA system.

3.4 CONTENT AND BLUEPRINTS

The NCSC test blueprint was designed to be consistent with the NCSC Theory of Action, the evidence-centered design undertaken to develop the summative assessment, and best practices in educational measurement. Tables 3-8 and 3-9 show the broad targets developed to guide the item development process and to inform test construction. They provide general guidance for identifying areas of emphasis in the development of the mathematics and ELA test forms. The blueprint tables in Appendix A 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. The items for each form of the test in each grade and content were revisited following the operational assessment window to balance both the content requirements of the blueprints and the psychometric characteristics of the items. The core set of items on each form was established from this balance.

Table 3-8. 2015–16 MSAA: Guidelines for Distribution of Mathematics Content by Grade Level

Math Content Category	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	Gr 11
Operations and Algebraic Thinking	30%	30%	10%				
Number and Operations Base Ten	20%	10%	40%				
Number and Operations Fractions	20%	30%	20%				
Measurement and Data	20%	20%	20%				
Geometry	10%	10%	10%	10%	20%	30%	10%
Ratio and Proportions				30%	40%		
Expressions and Equations				20%	10%	20%	
The Number System				30%	20%	10%	
Statistics and Probability				10%	10%	20%	20%
Functions						20%	
Algebra and Functions							50%
Number and Quantity							20%

Table 3-9. 2015–16 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	30%	30%	30%	20%	20%	20%	15%
Reading Informational	25%	25%	30%	40%	40%	40%	45%
Reading Vocabulary	9%	9%	10%	10%	10%	10%	10%
Reading Foundation	6%	6%					
Writing	30%	30%	30%	30%	30%	30%	30%

3.4.1 Mathematics

Mathematics items are aligned to prioritized CCCs, which are in turn connected to the CCSS and the LPFs. Mathematical knowledge is assessed across the CCCs through selected-response items and constructed-response items. Constructed-response items were present at grades 3, 4, 5, 8, and 11 only. The need for constructed-response items was determined by the FKSA associated with a given CCC. Students might construct a graph, solve a problem, or complete a table in a constructed-response item. Constructed-response items were scored dichotomously, or “correct/incorrect,” only.

In some cases, the selected FKSAs were best addressed by separating the skill into two parts. Therefore, two unique items are necessary to fully address a single content standard. For example, the CCC 8.DPS.1h1 asks students to both graph bivariate data using scatter plots and identify possible associations between the variables. Items were developed to address both parts of the standard. Tables in Appendix A identify which CCCs require two item versions.

In addition, in mathematics, there were items identified as not allowing the use of calculators in responding to the item. These items tended 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

ELA items in reading and writing are aligned to prioritized CCCs, which are in turn connected to the CCSS and 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.

The project determined that all reading comprehension assessment items be presented in a selected-response format. Thus, to measure more complex reading skills, some selected-response items were built as a set of two or three sequenced items ("multipart"), which, when combined, serve to measure the breadth of one prioritized content standard. In other words, in some instances the FKSAs aligned to a specific CCC are designed to have two or three selected-response items associated with them.

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.

The three CCCs prioritized for writing at each grade level consist of one CCC assessed by a constructed-response item and two assessed by selected-response items. The constructed-response writing items 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 clarify meaning and the use of standard English conventions. The constructed-response writing items were considered field-test items and did not count toward the student's score. Each state was given the choice of reporting information on the field-test constructed-response writing item to the student's teacher/school.

CHAPTER 4 TEST DEVELOPMENT

4.1 General Philosophy and Role of Test Construction Subcommittee in Test Development

As noted previously, 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. MSAA assesses English language arts (ELA; reading and writing) and mathematics at grades 3–8 and 11 and is aligned to the State Content Standards and the MSAA Core Content Connectors (CCCs). MSAA is a computer-based, on-demand 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 items on the 2015–16 administration were from the previous NCSC 2014–15 administration. The 2014–15 item pool was used given the late start date of the contract. The utilization of the 2014-15 item pool affected the specific criteria for replacement and field-test items for each content area. As described in Chapter 3, the items selected as field-test items were developed by NCSC and have been transitioned to MSAA. These previously developed items were written in a manner that supports the assessment's design. The item development process was an iterative one, which allowed multiple opportunities for review of the items by various stakeholders including NCSC, the state partners at that time, and external passage bias and item content review participants. This multistage development and review process provided ample opportunity to evaluate items for their accessibility, appropriateness, and adherence to the principles of Universal Design. In this way, accessibility emerges 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 significant cognitive disabilities.

For the 2015–16 administration, a small number of previously developed selected-response field-test items were selected and reviewed by the MSAA Test Construction subcommittee. The selected-response items were not edited with the exception of the removal of some items within certain ELA passage sets so that the number of items for each passage set did not exceed six items per passage set. This was done in an effort to keep the length of the test similar to the 2014–15 administration. The subcommittee was in agreement with the selection of the field-test items for both ELA and mathematics. In addition to the selected-response items, there was field-testing of a Tier 2 and a Tier 3 writing prompt at each grade level. The Tier 2 writing prompts

were used as previously developed. The Tier 3 writing prompts were slightly modified to simplify both the prompt and the teacher administration instructions to make them more similar to the Tier 2 writing prompt. The MSAA Test Construction subcommittee reviewed and approved the Tier 2 and Tier 3 writing prompts for each grade level.

CHAPTER 5 TRAINING AND ADMINISTRATION

5.1 TEST ADMINISTRATOR TRAINING

MSAA adhered to the premise from the testing standards (AERA et al., 2014) that a key consideration in developing test administration procedures and manuals is that the test administration should be fair to all examinees. As MSAA was a computer-administered test, the administration procedures were consistent with the hardware and software requirements of the test specifications. MSAA required completion of training by all test administrators (TAs) to support standardized test processes and procedures. MSAA provided ancillary testing materials outlining specific practices and policies including (a) the *Test Administration Manual* (TAM); (b) *MSAA Online Test Administration Training*; (c) *MSAA System User Guide for Test Administrators*; (d) *MSAA System User Guide for Test Coordinators,* and (d) grade- and content-specific *Directions for Test Administration* (DTA). TAs received 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 testing integrity practices for technology-based assessments.

5.1.1 Training Modules

The online training modules were made available prior to the beginning of the testing window and throughout the testing window. They were customized to address the specific responsibilities of the TA and to provide important information from the three documents TAs were required to use: (1) the TAM, (2) DTA, and (3) *MSAA System User Guide for Test Administrators*. MSAA developed 13 training modules for TAs (see Table 5-1).

Table 5-1. 2015–16 MSAA: Modules for Test Administrators

Module 1: Training Requirements and Responsibilities of Test Administrators
Module 2: Overview of the MSAA Test and Testing Integrity
Module 3: Optimal Testing Conditions and Assessment Features
Module 4: Test Accommodations and Procedures for Assessing Students Who Are Blind, Deaf, or Deaf-Blind: Additional Directions for Test Administration
Module 5: Navigate the MSAA System
Module 6: Before Test: Complete Demographics, LCI, and Accommodations
Module 7: Student Response Check

continued

Module 8: Student Experience in the MSAA System
Module 9: Mathematics DTA – Administer the Test
Module 10: ELA DTA: Reading – Administer the Test
Module 11: ELA DTA: Writing – Administer the Test
Module 12: Upload Evidence for ELA Constructed Response Writing Item
Module 13: Submitting or Closing a Test, Accommodations- After Test, and End of Test Survey

All online training recordings were accessed by TAs through the MSAA system. It was a requirement that the online training modules be viewed in sequence, and one module had to be viewed before the link to the subsequent module would become accessible. Once a module was accessed, that module would be marked as complete in the MSAA system and the link to the next module in the sequence would become available. Once all 13 modules were marked as complete, an end-of-training final quiz became available to TAs within the MSAA system.

Additionally, TAs were instructed to become familiar with the online system by accessing the sample items supplied within the system. MSAA utilized the same set of sample assessment items developed by content and measurement experts for teachers, administrators, and policymakers for the NCSC assessment. The sample items did not address all assessed content at each grade level and were not representative of every item type. Rather, the sample items provided a preview of the array of items and illustrated multiple item features supporting ways in which students with a wide range of learner characteristics interact with the assessment process.

There were quiz questions pertaining to information from the module at the end of each online training module for TAs and test coordinators (TCs). The quiz questions were included as a review of the content and to prepare TAs for the end-of-training final quiz. The end-of training final quiz was accessed via the MSAA system following completion of all online training modules. TAs were required to take the end-of-training final quiz, which covered content across all modules, and had to obtain a score of 80% or higher to be provided access to secure test administration materials. If TAs did not fulfill this certification requirement, they were not allowed access to the secure test materials. The TAs were notified within the MSAA system if they had passed or not passed the end-of-training final quiz. TAs were allowed multiple attempts to obtain a score of 80% or higher on the end-of-training final.

5.1.2 Test Administration Manual

The MSAA TAM provided an overview of and the guidelines for planning and managing MSAA administration for district and school personnel. Additionally, the TAM defined the role and responsibilities of the TAs involved in the administration of MSAA. Although the TAM included information pertinent to both TAs and TCs, areas of particular relevance to TAs are described below.

5.1.2.1 *MSAA TECHNICAL SUPPORT—WHO TO CALL FOR HELP*

Intentionally positioned early in the TAM, this section instructs TAs when and how to obtain answers pertaining to the MSAA system and test administration procedures. Directions were included that specify when to contact the State MSAA Coordinator or the MSAA Service Center.

5.1.2.2 *OVERVIEW OF MSAA*

This section includes background and purpose of the test, information about content areas covered, test delivery method, testing window, assessment features, accommodations, how to pause and resume administration of the test, and how to access the documents that are needed for test administration.

5.1.2.3 *DESCRIPTION OF MSAA TEST SESSIONS AND DESCRIPTION OF MSAA ITEM TYPES*

This section previews the testing sessions for each content area of the test and describes item types used in the test, including selected-response: reading, writing, mathematics; constructed-response: mathematics; constructed-response: writing; and open-response: foundational reading.

5.1.2.4 *RESPONSIBILITIES OF TEST ADMINISTRATORS*

This section explains that whoever serves as a TA must be a certified and licensed educator familiar with the student, typically the student's teacher, who has completed the required MSAA Test Administration Training, the end-of-module quizzes, and the end-of-training final quiz with at least an 80% accuracy score. Alternatively, if a student's teacher has a long-term substitute who is a certified and licensed educator, has completed the required MSAA Test Administration Training and end-of- module quizzes, and attained at least an 80% accuracy score on the end-of-training final quiz, then the long-term substitute may administer the test. Relevant state-specific criteria are also provided where applicable.

The before-, during-, and after-test administration responsibilities for this role are also detailed in this section, as described below in Table 5-2.

Table 5-2. 2015–16 MSAA: Test Administration Responsibilities

BEFORE TEST ADMINISTRATION
Sign and submit state-specific test security and confidentiality forms.
Complete MSAA Test Administration Training and attain at least an 80% score on end-of-training final quiz.
Complete Student Demographics and Learner Characteristics Inventory (LCI).
Complete Accommodations: Before Test, for each student being tested.
Complete the Student Response Check, as needed for each student.

continued

BEFORE TEST ADMINISTRATION
Review the TAM section on Optimal Testing Conditions, Assessment Features, and Test Accommodations and make appropriate arrangements for students; provide a printed version of the test if a student needs this accommodation (TAs were referred to TAM Appendix A-State-Specific Information for contact and resource links pertaining to state-specific information).
Ensure the student is taking the appropriate grade-level test.
• Download a hard copy of the DTA, or use a separate computer to access the DTA for the test. • Read the DTA; prepare and organize materials and print the reference sheets as specified in the DTA.
Review and complete the sample items with the student.
DURING TEST ADMINISTRATION
Implement the DTA as written and complete test administration by May 13, 2016.
Provide accommodations.
Report technology concerns to the TC or MSAA Service Center.
Maintain test security by ensuring all test materials are in a secure and locked location when not testing.
Report inappropriate test practices to the TC.
AFTER TEST ADMINISTRATION
Report any inappropriate test practices and suspected irregularities to the TC, according to state policy.
Give all printed copies of the test, DTAs, scoring rubrics, student login information, scratch paper, student work, etc., to the TC for secure shredding as well as all materials that were ordered.
Submit the tests and complete the Accommodations: After Test and the End of Test Survey. TAs should complete **one** End of Test Survey using the link for any one student. This feedback is important to us and will provide useful information for future planning.

5.1.2.5 *STUDENT PARTICIPATION CRITERIA*

This section specifies the criteria, which reflect the pervasive nature of a significant cognitive disability, and descriptors used by each student's individualized education plan (IEP) team to determine eligibility for participation in MSAA. Three criteria are named in particular: (1) The student has a significant cognitive disability; (2) the student is learning content linked to (derived from) the States' Content Standards; and (3) the student requires extensive direct individualized instruction and substantial supports to achieve measurable gains in the grade- and age-appropriate curriculum.

5.1.2.6 *OPTIMAL TESTING CONDITIONS, ASSESSMENT FEATURES, TEST ACCOMMODATIONS, AND PROCEDURES FOR ASSESSING STUDENTS WHO ARE BLIND, DEAF, OR DEAF-BLIND*

This section identifies various ways the TA must provide each student an appropriate testing environment during every testing session. This includes optimal testing conditions, confirmation that the student has an observable way to communicate his or her response to the items by completing the Student Response Check (SRC), use of appropriate assessment features itemized in the TAM, identifying IEP-defined accommodations that align with MSAA accommodations policies, administering the test in a familiar and distraction-free setting, and developing a schedule to administer the test during the best time of day for the student.

This section provides information pertaining to MSAA accommodations, including assistive technology, paper versions of the item(s), scribing, and sign language. The TAM explicitly states that physical prompting, including hand-over-hand, is not permitted and is considered to be an inappropriate test practice and a test irregularity

Finally, TAs are directed to the secure document *Procedures for Assessing Students Who Are Blind, Deaf, or Deaf- Blind: Additional Directions for Test Administration*, which provides additional tasks to complete before, during, and after the assessment; strategies, with definitions and examples, that may be used by the TA, as appropriate to enhance access to MSAA; and the DTA, which must be used to administer open-response foundational reading items to students in grades 3 and 4 who are blind, deaf, or deaf-blind.

5.1.2.7 *TESTING INTEGRITY AND APPROPRIATE TEST PRACTICES*

This section describes MSAA policies related to testing integrity and appropriate and inappropriate test practices. The importance of test security and the practices required for appropriate handling of secure test materials is explained, including

- maintaining all printed test materials in a secure, locked location
- refraining from duplicating, reproducing, or sharing items or other secure test materials
- giving all printed test item(s) or other printed material to the TC as prescribed in state-specific policy, after the test session is completed, for secure shredding
- deleting any test materials, items, or information from the computer and/or any assistive technology used by the student after testing is complete.

TAs are required to ensure that all aspects of the test are maintained in a secure manner. TAs are informed that items are for the exclusive use of testing and are not to be used for instruction and are not to be shared, e-mailed, copied, or distributed in any manner. To do so is considered a test irregularity and a violation of test security.

TAs are instructed to prepare a secure testing environment as students are administered the test individually, one-to-one, most likely in their classroom or a similar environment familiar to the student. A secure testing environment is explained to include but not be limited to

- administering the test only through the password-protected testing environment, the MSAA system
- restricting student access to resources that are explicitly identified in the DTA
- viewing of test items only by the student taking the test and the certified, licensed, and trained TA administering the test
- removing electronic devices and photography technology that could jeopardize test content in the test-taking environment (with the exception of the webcam used only to capture the student writing response)
- ensuring a quiet test-taking environment, void of talking or other distractions, and one that does not permit other students hearing the responses to the test items of the student being tested.

To underscore the importance of appropriate test practices, this section provides specific examples of inappropriate and prohibited test practices, including but not limited to

- changing the wording of test directions, items/questions, response options, or any text as it is written in the DTA
- using any physical prompting, including hand-over-hand; providing students a preview of the test at any time
- providing answers to students in advance of or during test administration; providing students clues or supports not indicated in the DTA
- manipulation of testing materials in a way that hints at a correct or incorrect answer (e.g., reducing the number of answer options)
- changing a student's answer
- reminding the student of previously used materials or experiences related to concepts in an item
- teaching test content immediately before the test or the administration of an item
- leaving any test materials unattended or in a nonsecure setting, including but not limited to DTAs, test items, materials related to test items, and scoring rubrics
- leaving the MSAA system unattended while logged in to the test or the DTA
- administering the test by a staff member who has not completed the online training modules and passed the final quiz.

This section emphasizes that each person participating in the state assessment program is directly responsible for immediately reporting any violation or suspected violation of test security or confidentiality.

TAs and other staff are required to notify their school or district TC if they witness or become aware of an inappropriate test practice or suspect one has occurred.

5.1.2.8 *ADMINISTRATION OF THE TEST*

Step-by-step directions are given for TAs to execute before and after testing, such as accessing their MSAA system account, supplying student demographic information, completing the Learner Characteristics Inventory (LCI), providing the Accommodations: Before Test and Accommodations: After Test information, conducting the SRC and interpreting its results, submitting completed tests, closing student tests, and completing one End of Test Survey.

Appendices to the TAM provided TAs with additional information. Appendix A-State-Specific Information provide links and contact information to access state-specific policies related to MSAA.

Appendix B-MSAA Scribe Accommodation: Protocol for Reading, Mathematics, and Writing addresses the scribe accommodation, where a scribe enters in the MSAA system the student-indicated answer to a selected-response item, and for the constructed-response writing item records the student's response to the writing prompt on the response templates in the MSAA system. For selected-response items for English language arts (ELA; reading and writing) and mathematics, the scribe protocol is defined as follows:

- The scribe may not question or correct student choices, alert students to errors or mistakes, guide the student to a correct answer, or otherwise influence a student's answer or answer choice in any way.
- The student must be tested in a setting that does not permit his or her responses to test items to be heard by other students.
- The scribe will comply with student requests for use of all available and allowable Assessment Features on the MSAA test system (e.g., when to turn a feature on or off, when to change size of graphic).
- A TA who is a scribe may answer procedural questions asked by the student (e.g., test directions, navigation within the test environment). However, a scribe who is a district employee working under the direction of a qualified, trained TA who is administering the item, must refer such questions to the TA.
- For computer-based administrations, the scribe must enter student responses directly into the MSAA system.
- The scribe may ask the student to repeat a response.
- The scribe allows the student to indicate when he or she wants to move to the next test item.
- The scribe must provide an opportunity for the student to review and modify what the scribe has recorded.
- After testing, the scribe collects scratch paper, graphic organizers, other ancillary materials, and login information and gives it to the school TC for secure shredding. Neither the scribe nor the TA may keep any testing materials after testing is complete.

- NOTE: If the scribe indicated the student responses on a paper-version of the test, the scribe will transcribe (type exactly) the student's responses into the MSAA system.

For constructed-response writing items, the scribe protocol was described as follows:

- The scribe types exactly what the student communicates directly on the response template in the MSAA system or writes exactly what the student communicates on a paper-version of the response template.
- The scribe correctly spells all words (spelling is not scored).
- The scribe does not capitalize words or punctuate text unless indicated by the student.
- The scribe allows the student to edit for punctuation, capitalization, or other edits as described in the DTA.
- The scribe makes student-requested changes, even if incorrect.
- The scribe orally confirms meaning of homonyms and commonly confused homophones (e.g., than and then; to, two, and too; there, their, and they're).
- After testing, the scribe collects scratch paper, graphic organizers, other ancillary materials, and login information and gives it to the TC for secure shredding. Scribes and/or teachers may not keep any testing materials after testing is complete.
- NOTE: If the scribe wrote the student response on a paper-version of the response template, the scribe will transcribe (type exactly) this written response into the MSAA system.

Appendix B of the TAM also described the qualifications and preparation necessary for TAs and district employees who function in the role of scribe. For TAs, a scribe must be a state-certified educator employed by the district; complete all training for the MSAA test; be familiar to the student; and be familiar with all the accommodations in the student's IEP. For district employees, a scribe must sign and submit state test security agreements according to state policy; be familiar to the student; review Appendix B-MSAA Scribe Accommodation: Protocol for Reading, Mathematics, and Writing; and scribe under the direction of a qualified, trained TA who is administering the item.

Appendix C-MSAA Augmentative and Alternative Communication Guidelines for Constructed-Response Writing explained that the TA must record the student's response(s) for constructed-response writing exactly as the student indicates using the student's existing Augmentative and Alternative Communication (AAC) system or device. It was explained that a student may complete the constructed-response writing following the methods provided in the TAM, and that when a method currently used by the student instructionally conflicts with the provided testing protocol for using AAC, the protocol should guide the method used during the assessment. The TAM defined the AAC methods as follows:

- Student should use the communication mode/system with which the student is the most competent and which provides the most accessibility to producing a writing product.

- Allow the student to select the word/symbol/picture/phrase in the communication mode/system in the same manner as used in instruction (e.g., direct select, indirect such as scanning, eye gaze).
- Allow the student to access words/symbols/pictures/phrases within their communication mode/system in the same manner as in instruction (e.g., subject specific boards, multiple levels by categories). Visual examples were provided in the TAM for this method.

Finally, Appendix C instructed TAs in the AAC protocol for completing constructed-response writing, which emphasized that when administering the constructed-response writing, TAs must adhere to the AAC protocol to ensure that the student's response is generated in a manner that allows for accurate measurement of the student's writing ability. Specific instruction included the following:

- Words/symbols/pictures/phrases that the student typically uses during instruction to communicate can be provided and should be words/pictures/symbols/phrases that are familiar to the student (e.g., events, descriptive words).
- Introduce vocabulary related to the prompt, but do not practice the prompt or teach the vocabulary in the context of the prompt. For example, if the prompt refers to supporting a claim related to "solar energy," it is appropriate to define and describe "solar energy" and its uses in order to familiarize the student with the related symbol(s) using the AAC device. However, it is not appropriate to practice writing a persuasive essay using "solar energy" as the context.
- Any content represented in the grade-specific stimulus materials can be added to the student's AAC device (e.g., list of temporal words, problem/solution cards, words from mentor text or sample essay) to support student responding. Ensure the words/pictures/symbols/phrases used from the stimulus materials are familiar or can readily be understood.
- A constructed-response answer cannot be the result of a series of dichotomous choices of words, phrases, or sentences selected by the TA. An example of a series of dichotomous choices that would not be allowed is: The teacher asks, "Do you want to say that the girl was tall or short?" The student chooses tall. The teacher then asks, "Do you want to say the girl ran or swam?" The student chooses swam. The teacher asks, "Do you want to say the girl swam in the lake or in the pool?" The student chooses the pool.
- A constructed-response answer can be the result of the student completing a process directed by the TA using a series of two categories to communicate his or her word/picture/symbol/phrase preference. For example, a series of dichotomous choices that is allowable is: The teacher asks, "Do you want People-Thing words or Action words?" The student selects People-Thing words and the teacher then gives the choice of People or Thing words. The student chooses People words. The teacher then presents a series of choices of People words to allow the student to select the preferred person from those provided on the board. (As stated above, this should not result in a series of dichotomous choices of words, phrases, or sentences selected by the TA.)
- Words/symbols/pictures/phrases cannot be arranged by the TA on a student's communication board so that any selection would be correct. An exception to this would be if the student requests or selects a specific category level or board that has all words that

could be used in a response (e.g., the student selects or requests the board filled with adjectives and all would apply to the writing piece).

- Refer to the guidance regarding the placement and presentation of words, pictures, checklists, graphic organizers, and/or templates in the DTA.

Where appropriate, the TAM refers TAs to the *MSAA System User Guide for Test Administrators*, which outlines using the system to accomplish the tasks for which TAs are responsible. User guides provide step-by-step instruction with MSAA system screenshots to facilitate use of the system.

5.2 TEST COORDINATOR TRAINING

MSAA requires completion of training by all TCs to support standardized test processes and procedures. MSAA provides ancillary testing materials outlining specific practices and policies including (a) the TAM; (b) *MSAA Online Test Administration Training*; (c) *MSAA System User Guide for Test Administrators*; (d) *MSAA System User Guide for Test Coordinators,* and (d) grade- and content-specific DTA. TCs 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 testing integrity practices for technology-based assessments.

5.2.1 Training Modules

In addition to the training modules for TAs described above, online modules specific to the role of TCs are made available prior to the beginning of the test window and throughout the testing window. These training modules are customized to address the specific responsibilities of the TC and to provide important information from the documents TCs are required to use: (1) the TAM and (2) *MSAA System User Guide for Test Coordinators*. MSAA developed four training modules for TCs (see Table 5-3).

Table 5-3. 2015–16 MSAA: Modules for Test Coordinators

Module 1: Responsibilities of Test Coordinators
Module 2: Overview of the MSAA Test and Testing Integrity
Module 3: Navigate the MSAA System
Module 4: Create Users and Organizations

All online training recordings are accessed by TCs through the MSAA system. It is a requirement that the online training modules be viewed in sequence, and one module has to be viewed before the link to the subsequent module will become accessible. Once a module is accessed, that module is marked as complete in the MSAA system and the link to the next module in the sequence becomes available. TCs are required to complete the online training for TCs but are not required to take the end-of-training final quiz. At the end of

each online training module for TCs are quiz questions pertaining to information from the module. The quiz questions are included as a review of the content.

5.2.2 Test Administration Manual

In addition to the instruction for TAs detailed above, the MSAA TAM also provides directions specific to the role and responsibilities of TCs involved in the administration of MSAA. To support TAs before, during, and after administration, it is imperative that TCs understand the TA roles and responsibilities as well as their own. This section describes aspects of the TAM of particular relevance to TCs.

5.2.2.1 *MSAA TECHNICAL SUPPORT—WHO TO CALL FOR HELP*

This section instructs TCs when and how to obtain answers pertaining to the MSAA system and test administration procedures. Directions are included that specify when to contact the State MSAA Coordinator or the MSAA Service Center.

5.2.2.2 *RESPONSIBILITIES OF TEST COORDINATORS*

This section specifies the training required for TCs and essential TC tasks that are required before, during, and after administration of the test. TCs are required to complete the MSAA test administration training for test coordinators, described in the previous section of this report, and the MSAA test administration training that focuses on information in the TAM and *MSAA System User Guide for Test Coordinators*. This training helps ensure that the TCs are knowledgeable about their responsibilities, testing integrity, and appropriate test practices for the test. TCs are not required to take an end-of-training final quiz. Relevant state-specific criteria are also provided where applicable.

The before-, during-, and after-test administration responsibilities for this role are also detailed in this section, as described below in Table 5-4.

Table 5-4. 2015–16 MSAA: Test Administration Responsibilities

BEFORE TEST ADMINISTRATION
Sign and submit state-specific test security and confidentiality forms.
Complete MSAA Test Administration Training for TCs. (Not required to take End-of-Training final quiz)
Ensure TAs and TCs have received the required training and can access the online MSAA system.
Communicate all information to TAs about the Test received from the State MSAA Coordinator.
Ensure technology capacity is met. Work with district/school IT personnel to ensure that the online MSAA System is accessible and functioning on every computer that is used for testing.
Support TAs to develop a testing schedule so that all tests will be submitted within the administration window.

continued

DURING TEST ADMINISTRATION
Monitor to ensure implementation of appropriate test practices and appropriate student participation so that test administration is completed by May 13, 2016.
Ensure students and TAs have the materials and resources needed to administer the Test.
Maintain test security by ensuring all test materials are in a secure and locked location when not testing.
DURING TEST ADMINISTRATION
TCs report all inappropriate test practices to the State MSAA Coordinator immediately. Refer to Appendix A-State-Specific Information.
AFTER TEST ADMINISTRATION
Investigate and report any inappropriate test practices and suspected irregularities to the State MSAA Coordinator. Refer to Appendix A-State-Specific Information.
Securely shred all printed copies of the test, DTAs, scoring rubrics, and student work (e.g., writing materials). Return all ordered materials using the Return Envelope with UPS label provided in the materials shipment. Do not return materials that were downloaded and printed locally.

5.2.2.3 *TESTING INTEGRITY AND APPROPRIATE TEST PRACTICES*

This section describes MSAA policies related to testing integrity and appropriate and inappropriate test practices. The importance of test security and the practices required for appropriate handling of secure test materials was explained, including maintaining all printed test materials in a secure, locked location; refraining from duplicating, reproducing, or sharing items or other secure test materials; receipt of all printed test item(s) or other printed material by the TC as prescribed in state-specific policy, after the test session is completed, for secure shredding; and deleting any test materials, items, or information from the computer. TCs were required to ensure that all aspects of the test are maintained in a secure manner. TCs were informed that items are for the exclusive use of testing and are not to be used for instruction and are not to be shared, e-mailed, copied, or distributed in any manner. To do so is considered a test irregularity and a violation of test security.

TCs were instructed in the secure handling of test materials, required preparation of a secure testing environment, and appropriate and inappropriate testing practices, as described above for TAs. This equipped TCs to respond to reports of inappropriate practices in the field. All TCs were instructed to follow their state procedures regarding reporting (referring to TAM Appendix A-State-Specific Information). District TCs were required to report any incidents involving alleged or suspected violations that fall under the category of a serious irregularity to the State MSAA Coordinator. State professional codes of ethics and state law provide the guidelines for determining the consequences for any inappropriate test practices

Where appropriate, the TAM referred TCs to the *MSAA System User Guide for Test Coordinators*, which outlined using the system to accomplish the tasks for which TCs were responsible. User guides provided step-by-step instruction with MSAA system screenshots to facilitate use of the system.

5.3 Operational Test Administration

The test administration window was March 30 to May 13, 2016. The tests were delivered for the online administration using the MSAA system, following the MSAA test design requiring test administration in six separate sessions (four for ELA and two for mathematics).

MSAA was not a timed test. Testing time varied for each student with testing paused and resumed, based on student needs. If a student exhibited frustration, lack of engagement, refusal to participate, or became sick during the administration of MSAA, TAs were directed to pause the testing and take a break—which may have been a few minutes to a few days, depending on the student's needs. MSAA protocols allowed the TA to pause and resume the administration of the test as often as necessary during the test window, based on a student's needs.

5.3.1 Session Structure

TAs could begin with either the mathematics test or the ELA test. Once a content-area test was started, TAs were required to complete that test before beginning the test in the other content area. Each content area test consisted of a set of testing sessions. Students were administered the test sessions in order for a given content area. ELA consisted of four test sessions (see Table 5-5) and mathematics consisted of two test sessions (see Table 5-6) at each grade level.

Table 5-5. 2015–16 MSAA: ELA Test Sessions

Session 1: Reading	Session 2: Reading	Session 3: Writing	Session 4: Writing
Literary and informational reading passages and associated **selected-response reading** items **Open-response foundational reading** items (grades 3 and 4 only)	Literary and informational reading passages and associated **selected-response reading** items **Open-response foundational reading** items (grades 3 and 4 only)	**Selected-response writing** items	One **constructed-response writing** item

Table 5-6. 2015–16 MSAA: Mathematics Test Sessions

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

5.3.2 Administration Support

The TAM directed TAs and school and district TCs to contact State MSAA Coordinators for assistance with general questions about the MSAA system or to obtain state-specific information. Appendix

A-State-Specific Information in the TAM listed contact information for each State MSAA Coordinator and the link to state-specific policies related to MSAA.

5.3.3 MSAA Service Center

To provide additional support to schools before, during, and after testing, the test administration vendor provided technical support through and operated the MSAA Service Center. The MSAA Service Center provided a centralized location that those involved in test administration could call, using a toll-free number, to ask specific questions or report problems they may be experiencing. MSAA Service Center operators were responsible for receiving, responding to, and tracking calls, then routing issues to the appropriate person(s) for resolution. The MSAA Service Center was available for extended hours throughout the test window (from 7:00 a.m. to 8:00 p.m. EST, Monday through Friday) because the test was administered in multiple time zones.

The TAM directed TAs and TCs to contact the MSAA Service Center with questions pertaining to the MSAA system and test administration procedures. The MSAA Service Center responded to questions and requests from the field via phone and e-mail throughout the administration window, with extended hours during key administrative activities such as registration. The MSAA Service Center's toll-free support number and e-mail address were promoted to the field through the MSAA system and related communications.

Support was provided in a tiered manner, where Tier 1 support represented direct support to the caller by MSAA Service Center representatives, Tier 2 support represented promoted support by the program management team for items such as policy questions, and Tier 3 support represented technical requests that were escalated to the technology vendor for attention. Wherever possible, callers were directed to the appropriate section of the TAM, *MSAA System User Guide for Test Administrators, or MSAA System User Guide for Test Coordinators*, available to users within the MSAA system.

All activity was tracked in the MSAA Service Center log and included in weekly status reports that were provided to MSAA partner states. These reports summarized activity, ticket requests, call analysis data (e.g., call duration, hold time), and per-grade/content and per-state test status summaries throughout the administration window.

5.3.4 Additional Supports

In addition to the MSAA Service Center, the test administration vendor program management team periodically provided direct phone and e-mail support to the state coordinators. In cases where logistical or procedural support was needed, program management worked with State MSAA Coordinators to resolve questions or issues. In cases with policy or consortium-wide implications, however, program management referred the state lead to the partner states.

A messaging system in the MSAA system was implemented to notify users of important information during the administration window. Upon logging in to the system, a message appeared at the top of the screen that notified users of system information and upcoming system activities, such as known issues and scheduled system maintenance, as well as courtesy messages.

5.3.5 Monitoring and Quality Control

To ensure that proper testing procedures and appropriate test practices were maintained throughout administration, numerous measures were taken both to communicate participants' responsibilities and to monitor the appropriateness, accuracy, and completion of key procedures and tasks. The TAM explained to TAs and TCs that each person participating in the assessment program was directly responsible for immediately reporting any violation or suspected violation of test security or confidentiality by notifying the school or district TC. TCs were then instructed to follow state procedures regarding reporting the issue or suspected issue; however, district TCs were informed that they must report to the State MSAA Coordinator any incidents involving alleged or suspected violations that would be considered a serious irregularity. The TAM further explained that the consequences for inappropriate test practices would be determined by their 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. As described in the Session Structure section of this chapter, tests were administered in item groupings referred to as test sessions. Immediately prior to completion of a test session, the online system displayed an End of Session screen that notified the TA of the number of questions left unanswered in that session, allowing the TA to ensure that no student responses were inadvertently omitted prior to leaving the session. The TA could navigate back to an item in the session and continue working or pause the test at that item in order to return to that session at a later time. However, the TA was notified that the system would effectively "lock" the session once it was completed by clicking Save & Exit on the End of Session screen. This prevented them from returning to the earlier session, which ensured that items in a subsequent session could not be used to inform responses to items in the previous session.

Throughout the administration window the test administration vendor monitored and provided weekly updates to partner states on the test statuses across MSAA states and trends identified in support calls. This provided 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 responsiveness throughout the assessment process contributed to a proper and valid administration of MSAA.

5.3.6 Operational Test Survey Results

An End of Test Survey (EOTS) was developed to learn from the experience of each TA administering MSAA. TAs were instructed to complete one EOTS after submitting or closing one of his or her students'

content area test. Specific directions for completing the EOTS were provided in the User Guide for Test Administrators. The survey questions focused on several themes:

- Challenges experienced while using and providing accommodations
- Challenges experienced while using the embedded supports and materials
- Instructional time spent on the State Content Standards
- Teacher attitudes and priorities when developing instruction for students
- Technical challenges with the online system
- Students' ability to communicate and access the test

In addition to identifying issues that were unknown to MSAA, the results of the EOTS also highlighted several issues that MSAA was already addressing prior to reviewing these data. In that way, the EOTS data provided confirmation that MSAA was developing plans to address several known issues. These issues include:

- Addressing what is and is not a Consistent Observable Response
- Clarifying directions and stimulus materials for the writing prompts
- Resolving issues surrounding test administrators accidentally being locked out
- Students accidentally submitting tests

The survey data also identified several issues with little strain on timelines to prepare for the 2016-17 administration. These included:

- Improving the online training modules to address allowable supports to students in a clearer, more explicit manner
- Restricting the application of the Early Stopping Rule to test coordinators, rather than test administrators, and what a Consistent Observable Response is and is not
- Simplify directions for converting writing responses from a PDF into a JPG or PNG file format
- Making it clearer to users when they are about to submit a test for scoring

There were also issues that the survey uncovered that will take thoughtful, long-range planning to resolve. One of these issues is how students and teachers scroll to see the entire test item in the online platform. Currently, the item display is such that the whole item cannot be seen on the screen. This has proven difficult to fix as it requires code changes and changes to the APIP for each item. Another issue raised by the teachers is the lack of familiarity and relatability with the contexts and scenarios used in the writing prompts and other items. MSAA will focus on developing test items and writing prompts that contain contexts and scenarios that are more relatable to students in this population but this will have to be a sustained goal stretched over several testing cycles.

Several questions on the survey addressed teachers' attitudes and philosophies regarding teaching students with significant cognitive disabilities. MSAA has several challenges to address in this area. Specifically, the need for more instructional materials to illustrate how students in this population can learn rigorous academic content and for the test to become stage-adaptive in a way that teachers can trust. The 2016-17 administration will be the first year of the stage adaptive design and MSAA anticipates that this design will help to alleviate the concerns that many teachers have about the test being too difficult or "too high" for their students.

The EOTS data also show that teachers are struggling with providing instruction in several academic areas. These include fractions and data and statistics in mathematics, and writing an argument and reading informational texts in ELA. Given that education for students in this population has traditionally centered on life and functional skills, the heavier focus on academics is something that teachers may not feel adequately prepared for. The EOTS data show several academic areas in which teachers had difficulty instructing their students. In order to effectively help teachers, MSAA will need more information on why teachers are having difficulties teaching the areas identified above.

CHAPTER 6 SCORING

6.1 ITEM SCORING PROCESS

MSAA was completed through an online administration. Students responded to a variety of item types, including selected-response items, constructed-response items and open-response items. In addition, all students responded to a writing constructed-response item. The writing constructed-response item was an embedded field test item. For the writing constructed-response item, the student responses were entered directly into the MSAA system and/or uploaded into the system through scanning or webcam. The selected-response items were scored according to the answer keys provided in each test package. The Foundational reading open-response items and mathematics constructed-response items were scored as a correct or incorrect student response and this was entered by the Test administrator (TA).

The writing constructed-response items were scored by Measured Progress. Additional detailed information about the scoring of the writing constructed-response items is available in Section 6.2.

6.1.1 Scoring Processes and Rules for Selected-Response and Constructed-Response Items in Mathematics and Selected-Response Items in Reading and Writing

6.1.1.1 *OVERVIEW OF SCORING PROCESS BY ITEM TYPE*

Selected-Response: Reading, Mathematics, and Writing

Selected-response items (multiple-choice) were presented to students in a standard format. All directions and materials needed for administering selected-response items were provided in the secure *Directions for Test Administration* (DTA) that accompanied each test form. The TAs received training in the administration and scoring of selected-response reading and mathematics items in online training module 9 (Mathematics DTA—Administer the Test) and module 10 (English language arts [ELA] DTA: Reading—Administer the Test). In module 11 (ELA DTA: Writing—Administer the Test) TAs were given instruction to refer back to module 10 and follow the same protocols outlined for the selected-response reading items with the selected-response writing items. The DTA provided the full items, including the teacher scripts, to be read aloud to the student and any direction to the teacher related to the item and item setup, such as what to point to in the item as the script was read to the student. Every item was presented in the following order:

- Item stimulus (which may include a passage, passage part, picture, graphic, or other illustration)

- Item question
- Answer options presented in stacked, or vertical, formation

Students selected a response from the options in a variety of ways (e.g., using the computer mouse, verbalizing, gesturing, using eye gaze or communication devices, assistive technology). Many students entered responses directly into the MSAA system. If the student had the scribe accommodation, the scribe entered the student-selected response on behalf of the student.

Constructed-Response: Mathematics Completion

The constructed-response items, in selected grades for mathematics, required students to develop an answer instead of selecting an answer from response options. Constructed-response items were presented as novel tasks using materials and content presented in an on-demand test format. Each item was presented to the student in a standardized, scripted sequence of steps culminating in the TA scoring the student performance using the Mathematics Scoring Rubrics provided for the item. The Mathematics Scoring Rubrics provided scoring standards that were used to evaluate student responses. TAs received training in the administration and scoring of constructed-response mathematics items in online training module 9 (Mathematics DTA—Administer the Test). Directions and materials needed for administering mathematics constructed-response items were included in the secure DTA that accompanied each mathematics test form. The TA entered the student constructed-response score into the MSAA system as either correct or incorrect.

Open-Response: Foundational Reading

Open-response foundational reading items were included in the reading test in grades 3 and 4 only. The items were word identification tasks. Students identified either three or five words, depending on the tier level of the items presented. TAs received training in the administration and scoring of open-response foundational reading items in online training module 10 (ELA DTA: Reading—Administer the Test). Directions and materials needed for administering reading open-response items were included in the secure DTA that accompanied each ELA grades 3 and 4 test form. The TA entered the student's scores into the MSAA system.

Students with clear and consistent oral speech were administered the open-response foundational reading items. Students using communication other than oral speech, such as Augmentative and Alternative Communication (AAC) devices, American Sign Language (ASL), Braille, or eye gaze were administered selected-response foundational reading items included in the reading test.

For scoring purposes the Tier 1 foundational reading items required the student to identify all three words correctly to earn one score point for the foundational item set. The Tiers 2, 3, and 4 foundational reading items required the student to identify four or five of the five words correctly to earn one score point for the foundational item set.

Overview of Scoring Process within the Assessment System

The MSAA system provided automated machine scoring for all item types, aside from extended/constructed-response writing items, which required human scoring. The system also allowed for teacher entry of student responses to be used for paper-based test delivery. The MSAA system automatically scored question types that were machine-scorable and where scoring data had been provided. At the completion of the operational test, all test data were extracted from the system and were then compiled to generate full result sets for each student's tests.

The selected-response items were scored according to the answer keys provided in each test package. All item responses were exported from the system and provided to the Measured Progress Data and Reporting Services (DRS) Department. DRS then applied the scoring rules. Items were scored as correct or incorrect, with the majority of them contributing a score of 1 or 0 to the content area raw score. The selected-response items for the Tier 1 writing multiple-part selected-response suite were treated as a set for scoring purposes. Each item set consisted of 4, 5, or 6 items. In all cases, a student's score on the item set ranged from 0 to 2 points. If the student got none of the items correct, he or she earned 0 points; if the student got one or two items correct, he or she earned one point; and if the student got three or more of the items correct, he or she earned two points.

Administrator/Scorer Training

All TAs were required to participate in administration training modules and pass a final quiz in order to be certified to administer MSAA. The training included modules for each of the content-area DTAs (mathematics, reading, and writing). The modules reviewed the parameters for the administration and scoring of each item type, as well as how to enter the student responses into the MSAA system (mathematics and reading only).

During the test administration, TAs used the content-area DTAs to administer each item. The DTAs included the teacher scripting and directions related to any item setup, providing directions for the teacher to follow during administration. For the mathematics constructed-response items, the DTA included any templates required by the items, the directions related to how to present the items to the student, and the rubrics used to score the items.

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

During the administration window TAs were able to call or e-mail the MSAA Service Center with any questions related to the administration of test items and submission of the student responses within the MSAA system.

6.2 Scoring Processes and Rules for Field-Test Writing Constructed-Response Items

6.2.1 Overview of Scoring Process by Item Type

6.2.1.1 *CONSTRUCTED-RESPONSE: WRITING*

The writing constructed-response items were field-tested and required students to produce a permanent product in response to a writing prompt. The student, or a qualified scribe, recorded the response to the writing prompt on either the response template that was in the MSAA system or the paper response template that was included in the writing DTA. The MSAA grades 3–8 and grade 11 writing assessments included Tier 2 and Tier 3 constructed-response items. In Tier 2 a student generated a response using a response template, which included prepopulated sentence starters. Tier 3 items did not contain sentence starters.

The writing constructed-response item was presented to the student by the TA in a standardized, scripted sequence of steps and included directions to present grade- and prompt-specific writing stimulus materials that needed to be printed and prepared. All writing stimulus materials, including the response template, were identified by a card number and were included in the writing DTA. If the student used a paper version of the template to write a response, the following parameters were followed:

- The TA was instructed to annotate or interpret the student's writing directly on the student's written product if he or she determined that a novel reader (i.e., a scorer) may not be able to interpret a component (e.g., inventive spelling, penmanship, or use of symbolic expressions) of the student's written product.
- The TA was instructed to transcribe or type exactly the student's written response, including any annotations, into the MSAA system.

6.2.1.2 *FIELD-TESTED WRITING CONSTRUCTED-RESPONSE—ADMINISTRATOR TRAINING AND MONITORING*

All TAs were required to participate in administration training modules and pass a final quiz in order to be certified to administer the MSAA assessment. The training included module 11 (ELA DTA: Writing—Administer the Test), which reviewed the parameters for the administration of the constructed-response item, as well as how to enter the student responses into the MSAA system.

During the test administration, TAs used the DTAs to administer each writing constructed-response item. The DTAs included the teacher scripting and directions related to any item setup, providing directions for the teacher to follow during administration.

Further direction was provided to TAs on the entering of item responses in the MSAA system through the *MSAA System User Guide for Test Administrators*. The guide outlined the use of the system, including how to enter student responses and submit the writing session of the test. TAs were 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 were entered) were then extracted from the online system and provided to Measured Progress for human scoring.

6.2.1.3 *FIELD-TESTED WRITING CONSTRUCTED-RESPONSE: SCORING*

Benchmarking Meetings and Identification of Scoring Materials

The benchmarking for the constructed-response writing field-test items began on June 10, 2016. Scoring experts (Scoring Supervisors and Scoring Team Leaders [STLs], defined below) reviewed student responses in *iScore* and identified item-specific anchor and practice sets. The anchor sets consisted of 8–12 responses representing a full range of scores across traits. Each practice set consisted of five responses representing trends or points to be clarified during training. In addition, 10 extra responses were included as possible replacements or supplements to the anchor or practice sets. The proposed materials were reviewed by a Scoring Content Specialist. On June 21 and June 22 materials were posted to a secure FTP site for state partner representatives to preview.

The benchmarking meetings for MSAA began on June 27. Participants included members of Client Services' Special Education Department (SPED) and Scoring Services at Measured Progress as well as the following five partner state representatives:

- Arizona - Audra Ahumada
- Maine – Sue Nay
- Maryland – Marsie Torchon
- Rhode Island – Heather Heineke
- South Dakota – Christina Booth

The purpose of the meeting was to apply the rubric to proposed sets and come to a consensus on the final materials that would serve as score point exemplars during the training and scoring of each item. There were a total of 14 field-test items to review, with a Tier 2 and a Tier 3 item appearing in each grade. One group was assigned to review the Tier 2 and Tier 3 items for grades 8 and 11. Another group initially began with a review of the grade 6 items for Tier 2 and Tier 3. After grade 6 was complete, the group divided to focus on either Tier 2 or Tier 3 items for grades 3, 4, 5, and 7. At the end of each day, the teams met to report issues and share observations regarding the responses and the item-specific rubrics with the entire group. This process moved along at a faster pace than initially anticipated, and the benchmarking for all 14 items was completed in three days. The final scores for the anchor and practice sets were recorded and representatives from the partner states acknowledged their consensus on the signoff document for each item.

Following the benchmarking meetings, two qualification sets were identified for each item. These sets were assembled by members of Scoring Leadership who had been in attendance at the benchmarking meetings. Each qualification set consisted of 10 responses and scores were based on anchor responses and decisions made during the benchmarking meetings.

6.2.2 Scorer Recruitment and Qualifications

The MSAA scorers were a diverse group of individuals with a broad range of backgrounds, including teachers, business professionals, graduate students, and retired educators. They were primarily obtained through Kelly Services, a temporary employment agency. All scorers assigned to work on the MSAA writing assessment held the minimum of a four-year college degree, which included ELA or writing coursework. Approximately 76% of the leadership and scorer group assigned to MSAA had previous experience in scoring alternate assessments. All scorers signed a nondisclosure/confidentiality agreement.

Table 6-1 summarizes the qualifications of the 2016 MSAA scoring leadership and scorers.

Table 6-1. 2015–16 MSAA: Qualifications of Scoring Leadership and Scorers

Scoring Responsibility	Educational Credentials			
	Doctorate	*Master's*	*Bachelor's*	*Total*
Scoring Leadership[1]		2	8	10
Scorers	2	14	25	41

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

6.2.3 Measured Progress Staff and Scoring Leadership

The MSAA field test for writing constructed-response items were scored in Dover, New Hampshire, between July 13 and July 26, 2016. The following staff members were involved in scoring:

- 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 and had overall responsibility for MSAA scoring-related activities.
- ELA Group Manager for Scoring: Responsible for managing scoring-related activities, 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 coordinated 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.

Scoring Supervisors assigned to train the STLs and scorers thoroughly reviewed the decisions and materials that resulted from the benchmarking meetings in preparation for training. One Scoring Supervisor was assigned to Tier 2 items across grades and another Scoring Supervisor was assigned to all Tier 3 items. Leadership training took place on July 11 and July 12. Scoring Supervisors trained the STLs who were required to meet the minimum accuracy standard of 80% exact and 90% exact plus adjacent agreement on all items. This process was applied to each trait individually across qualification sets 1 and 2. The STLs were also present during scorer training, which further reinforced their understanding of the rubric and training materials.

Scoring Supervisors conducted training on each writing constructed-response field-test item before scorers were allowed access to student responses. Scorers were divided into two groups. One group focused on Tier 2 items and the other on Tier 3 items. Training sessions for scorers were facilitated by a Scoring Supervisor and conducted in the following manner:

- Scorer training for MSAA commenced 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 addressed the security, confidentiality, and proprietary nature of testing, scoring materials, and procedures.
- Training for each item focused on the three traits of the MSAA analytic rubrics for writing and how the scoring for each trait would be applied to student work.
- The training for each item included pertinent information on the testing instructions and item stimuli.
- Scorers reviewed actual responses with an item-specific anchor set, averaging 10 responses representing a range of scores across traits.
- Scorers were instructed to refer back to the anchor set frequently during scoring.
- Anchor exemplars were presented in a predetermined order and consisted of responses that were typical, rather than unusual or uncommon, solid, rather than controversial or borderline, and true.
- The Scoring Supervisor announced the anchor response score and explained the scoring rationale, allowing scorers to internalize typical characteristics of each score point.

- Supplementary training materials, averaging six responses, contained practice responses representing all score points across traits, when possible, and often contained responses that were 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 contained nonscorable codes.
- During the review of practice responses, the Scoring Supervisor often focused review efforts on the lines between adjacent score points or clarification of other scoring issues that are traditionally difficult for scorers to internalize.
- Scorers independently read and scored each practice response, and the Scoring Supervisor discussed the actual score and explained the scoring rationale for each response.
- A Q&A segment addressed any remaining questions from scorers and provided clarification.

6.2.4 Qualification

Following the training for each item, scorers were required to complete a qualification set to determine eligibility to score student work. There were two qualification sets consisting of 10 responses each. The responses, which represented a range of score points, were randomly distributed to scorers through *iScore*.

Scorers had two opportunities to qualify for scoring each item. If scorers attained a score match of at least 80% exact and 90% exact & adjacent agreement on all traits for the first qualification set, they were permitted to score live student responses. If scorers were unable to attain a score match of at least 80% exact and 90% exact & adjacent agreement for all traits on the first qualification set, the Scoring Supervisor conducted a retraining by discussing the qualifying responses in terms of the score point descriptions and the original anchor set. Following this training, scorers were assigned qualification set 2. For example, a scorer who qualified on the first and third trait in qualification set 1 received the retraining referenced above. However, they would only be required to qualify on trait 2 in qualification set 2. When the disaggregated data indicated that a qualified scorer had demonstrated a weakness in a particular trait, that qualified scorer received additional training prior to start of scoring.

Scorers who failed to achieve the minimum levels of agreement were not allowed to score live student responses. When scorers demonstrated a level of understanding and the ability to apply feedback, Scoring Leadership chose to include the scorer in future trainings.

Once the first field-test writing constructed-response item was completely scored, the training process was repeated for the next item. This continued until all field-test writing constructed-response items were scored.

Table 6-2 summarizes the qualifications rates for MSAA. Rates of success during qualification varied and this is visible in Table 6-2. 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. For instance, in

grades 6 and 7 all scorers passed the first qualification set for WRCC001 but this was not the case for all items.

Table 6-2. 2015–16 MSAA: Qualification Summary

Grade 3	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified WRCC001	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified WRCC002
Total Passed	8	6	14	12	9	21
Total Failed	10	4		11	2	
Grade 4	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified WRCC001	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified WRCC002
Total Passed	9	5	14	10	8	18
Total Failed	8	3		11	3	
Grade 5	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified WRCC001	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified WRCC002
Total Passed	11	5	16	17	0	17
Total Failed	6	1		2	2	
Grade 6	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified WRCC001	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified WRCC002
Total Passed	17	0	17	18	1	19
Total Failed	0	0		1	0	
Grade 7	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified
Total Passed	16	0	16	10	0	10
Total Failed	0	0		10	10	
Grade 8	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified WRCC001	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified WRCC002
Total Passed	9	3	12	4	5	9
Total Failed	4	1		8	3	
Grade 11	W R C C 0 0 1 Qual 1	W R C C 0 0 1 Qual 2	Scorers Qualified WRCC001	W R C C 0 0 2 Qual 1	W R C C 0 0 2 Qual 2	Scorers Qualified WRCC002
Total Passed	14	6	20	7	1	8
Total Failed	6	0		6	5	

Note: For identification purposes in iScore, Tier 2 items were designated as WRCC001 across all grades and Tier 3 items were designated WRCC002.

6.2.5 Methodology for Scoring Field-Tested Writing Constructed-Response Items

Student responses to the writing constructed-response items and any uploaded material were exported from the platform and imported to the Measured Progress *iScore* system. Through *iScore*, qualified scorers read and evaluated student responses and submitted scores electronically. The processes by which images were logged in, scanned, and uploaded into *iScore* provided anonymity to individual students and ensured random distribution of all responses during scoring.

All student responses were scored from either uploaded evidence or computer-generated text, defined as student work directly entered into the MSAA system. TAs were directed to capture an image of the Student Response Template pages when the student did not enter a response directly into the system. The template contained the student's original product and the TA uploaded the captured image into the MSAA system. The system allowed a TA to either use a webcam to take a snapshot of the student's paper or scan it with the school's scanner and upload to the system.

For Tier 2 items, when both uploaded and computer-generated text were available, the uploaded evidence was scored first and the computer-generated text was used for clarification and confirmation of the uploaded student writing evidence. When there was only uploaded writing evidence but no computer-generated text to provide clarification and confirmation, then the uploaded writing evidence was scored. When there was only computer-generated text but no uploaded writing evidence, the computer-generated text was scored. Because Tier 3 items had an emphasis on revision, the computer-generated text was determined to be the final product and was scored first. Any uploaded evidence was considered to be a draft and was used to provide clarification if needed. When there was only uploaded writing evidence but no computer-generated text, then the uploaded writing evidence was scored.

The following processes were in place during the scoring of the MSAA field test writing constructed-response items:

- The *iScore* system forced scorers to review all available pages before allowing a score to be submitted.
- All scoring was "blind" (i.e., no student names were visible to scorers, only booklet numbers within *iScore* linked student responses).
- Measured Progress maintained security during scoring by using a highly secure; server-to-server interface to ensure that access to all student response images was limited to only those who were scoring or working for Measured Progress in a scoring management capacity.
- During scoring, *iScore* enabled a constant measuring and monitoring of scorers for scoring accuracy and consistency. Each scorer's reading rate and total number of scored responses were also monitored.
- Scorers were required to maintain an acceptable scoring accuracy rate (80% exact/90% adjacent agreement) on a daily basis as measured across read-behinds, double-blinds, and daily calibration sets.

- Scorers who repeatedly fell below standard were retrained and/or dismissed from scoring that item.
- Scoring rules were in place to determine the final score of record, or when a final score was to be provided by Scoring Leadership (see tables 6-5 through 6-8).

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

Table 6-3. 2015–16 MSAA: Student Responses per Grade

Grade	Item	Number of Student Responses
3	WRCC001	1,599
3	WRCC002	1,627
Total		3,226
4	WRCC001	1,758
4	WRCC002	1,662
Total		3,420
5	WRCC001	1,749
5	WRCC002	1,806
Total		3,555
6	WRCC001	1,769
6	WRCC002	1,811
Total		3,580
7	WRCC001	1,909
7	WRCC002	1,644
Total		3,553
8	WRCC001	1,726
8	WRCC002	1,659
Total		3,385
11	WRCC001	1,251
11	WRCC002	1,149
Total		2,400

Scoring Rules

All field-test writing constructed-response items were scored against a three-trait analytic rubric. The scoring scale options of 0, 1, 2, and 3 were applied to each trait. When a response did not conform to score point parameters, scorers could designate the response as one of the following (see Table 6-4):

- **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.

- **Escalate:** The response requires clarification or adjudication by Scoring Leadership.
- **No Score:** The response requires designation by Scoring Leadership.

Table 6-4. 2015–16 MSAA: Scoring Resolution Process

Designation	Resolution Process
Blank	Responses scored Blank were sent to another scorer for a second read. Responses scored Blank twice were converted to zeros (“0”) for reporting purposes. Any discrepancies were resolved by the Scoring Leadership.
Unreadable	Those responses judged unreadable were forwarded to a special queue within *iScore* to be reviewed by a Scoring Supervisor who resolved the student score. (If the response remained unreadable after review, the Scoring Supervisor assigned a score of “0.”)
Non-English	Responses written in a language other than English were marked Non-English and converted to zeros (“0”) for reporting purposes.
Escalation	Responses that required additional clarification or adjudication were escalated to Scoring Leadership. This included responses where it appeared that more than one students’ work had been uploaded to the response. Responses where the uploaded evidence was a mismatch to the typed response were escalated. These responses were forwarded to Scoring Leadership for response appraisal and scoring. Responses that legitimately responded to another item were 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 were scored based on its merits and then forwarded for review. If further attention was warranted, our Client Services team would notify the appropriate MSAA partner state.
No Score	Responses that were determined to be nonscorable were resolved by the Measured Progress leadership team and in consultation with MSAA partner state representatives, if necessary.

Scorers also had the option of flagging a response as an “Alert” paper requiring immediate review and possible immediate action by Scoring Leadership and MSAA representatives. One student response was deemed a possible “Alert” during the scoring of MSAA and was forwarded to the appropriate partner state.

“Alert” responses could include but were not limited to 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
- Neglect

6.2.6 Monitoring of Scoring Quality Control

Measured Progress scorers assigned to the MSAA project that met or exceeded the minimum standard for qualification were allowed access to score live student responses. Scorers were continuously monitored to ensure that scoring was accurate and consistent. Read-behind and double-blind statistics were reviewed daily. Calibration sets were 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 paid close attention to the disaggregated read-behind, double-blind, and calibration statistics. Scorers in need of additional clarification on applying scores to specific traits were coached by Scoring Leadership. This continuous training allowed scorers an opportunity to resolve issues, ask questions, reiterate scoring guidelines, and establish parameters for atypical student responses. Scorers who demonstrated inaccurate or inconsistent scoring were retrained and allowed to resume scoring under increased supervision. Scoring Leadership removed scorers who continued to fall below accuracy standards. On any day that a scorer fell below accuracy standards the work was voided and rescored by other qualified scorers. During MSAA scoring, the work of three scorers was voided. There were two voids in grade 7 and one void in grade 4.

Throughout the scoring of the field-tested writing constructed-response items for MSAA, read-behind scoring, double-blind scoring, and calibration sets were used as quality-control measures. MSAA partner state representatives, along with the Measured Progress SPED and Scoring teams, monitored reports daily.

6.2.6.1 *CALIBRATION SETS*

Calibration sets commenced on the second day and each subsequent day of scoring each item. Scoring Leadership selected the responses used for daily calibration. For calibration, scorers were assigned five prescored responses representing a variety of scores and selected from recently scored responses. Scorers who correctly assigned at least 12 out of 15 exact scores on the calibration set began scoring for the day. Scorers who failed to meet the standard were retrained by discussing the calibration responses in terms of the rubric and the anchor set. Scoring Leadership determined if scorers who were retrained after the calibration set should be given access to live student responses to begin scoring. Retrained scorers who did return to scoring student work continued to be closely monitored. Over the course of scoring MSAA only 12 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.

6.2.6.2 *READ-BEHIND SCORING*

Read-behinds provide a crucial tool in verifying scorer accuracy. The STL completed read-behinds on individual scorers throughout the shift. The STL’s evaluation of each response was performed with no

knowledge of the scores assigned across traits. The scores were only available to the STLs after they also scored the response. If there was a difference in scores, either adjacent or discrepant (more than one score point difference), the STL score became the score of record. If the scores were discrepant, or if there were a significant number of adjacent scores between the scorer and the STL, the STL discussed the rationale with the scorer.

The average number of read-behinds for each scorer was 5–10 reads a day, but this varied depending on the accuracy of each scorer. Read-behinds provided an immediate means of identifying scorers in need of further clarification on how to effectively apply the scoring rubrics to student responses. If scorers fell consistently below the 80% exact and 90% exact & adjacent threshold, Scoring Leadership had the prerogative to void their scores for the day and/or stop them from scoring the item. Read-behinds were also performed on the STLs. Scoring Leadership monitored scoring accuracy and consistency by reviewing the read-behinds performed by the STLs as well as reading behind them whenever possible.

6.2.6.3 *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 were 100% double-blind scored.

In addition to monitoring inter-rater agreement rates, double-blind scoring also allowed Scoring Leadership to resolve arbitrations when two scorers' double-blind scores did not agree across any of the three traits. If there was not exact agreement, *iScore* automatically placed the response into an arbitration queue. Scoring Leadership, with no prior knowledge of the scores assigned, evaluated the response and their score became the score of record. The double-blind statistics provided an overview of agreement rate among the entire pool of scorers and assisted in identifying any need of group 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 6-5, 6-6, and 6-7 provide examples of how the final score of record may be determined through resolutions.

Table 6-5. 2015–16 MSAA: Examples of Scoring Resolutions: Read-Behind Scoring
Read-Behind Scoring [1]
(Trait 1-Trait 2-Trait 3)

Scorer Score	Leadership Score	Score
4-4-4	4-4-4	4-4-4
4-3-3	3-3-3	3-3-3
4-3-3	2-2-2	2-2-2

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

Table 6-6. 2015–16 MSAA: Examples of Scoring Resolutions: Double-Blind Scoring
Double-Blind Scoring[1]
(Trait 1-Trait 2-Trait 3)

Scorer #1	Scorer #2	Leadership Resolution	Final
4-4-4	4-4-3	4-4-4	4-4-4
4-3-3	2-2-2	3-3-2	3-3-2
2-1-1	1-1-1	2-2-1	2-2-1
2-2-2	4-4-4	3-3-3	3-3-3

[1]All adjacent or discrepant scores were resolved in arbitration and in these cases the leadership score became the final score of record.

Table 6-7. 2015–16 MSAA: Examples of Scoring Resolutions: Edit Scoring
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 received more than one read-behind and the scores supplied by the STLs did not agree, a resolution score was needed. In these cases, the Scoring Supervisor provided a final score during the post-scoring edit process.

6.2.7 Quality and Production Management Reports

Reports generated through *iScore* were essential during the scoring of MSAA. Reports provided real-time statistics to be reviewed by the Measured Progress Scoring team and MSAA partner state representatives to closely monitor scoring and to ensure:

- Overall accuracy, consistency, and reliability of scoring (group level) was maintained.
- Scorer data (individual level) were monitored in real-time to allow early scorer intervention when necessary.
- Individual traits in need of further clarification were identified.

- Scoring schedules were upheld.

The following reports, shown in Table 6-8, provided the comprehensive tools and statistical information needed to execute quality control and manage production.

Table 6-8. 2015–16 MSAA: Scoring Quality Control and Production Management

Report	Description
Read-Behind Disaggregated Summary	The Read-Behind Disaggregated Summary report showed the total number of read-behind responses read by both the scorer and the STL and noted the number and percentage of exact, adjacent, and discrepant scores across each trait.
Double-Blind Disaggregated Summary	The Double-Blind Disaggregated Summary report showed the total number of double-blind responses read by a scorer and noted the number and percentage of exact, adjacent, and discrepant scores across each trait.
Compilation Report	The Compilation Report showed 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 listed the total number of student responses loaded into *iScore*. This report included how many reads had been completed to date and how many reads remained.

6.2.8 Inter-Rater Reliability

Kappa statistics (kappa coefficient) measure the agreement between 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 6-9 below shows that agreement between raters fell into the substantial agreement or almost perfect agreement range for most of the field-test writing constructed-response across grades. The only exception was in the Tier 3 item for grade 7, where Trait 1 and Trait 2 were in the moderate agreement range.

Table 6-9. 2015–16 MSAA: Kappa Agreement—Field-Tested Writing Constructed-Response Items

	Organization Trait 1	Idea Development Trait 2	Conventions Trait 3
Grade 3			
WRCC001	0.67	0.68	0.84
WRCC002	0.68	0.71	0.80
Grade 4			
WRCC001	0.73	0.71	0.88
WRCC002	0.73	0.66	0.84
Grade 5			
WRCC001	0.75	0.74	0.89
WRCC002	0.64	0.69	0.80

continued

	Organization Trait 1	Idea Development Trait 2	Conventions Trait 3
Grade 6			
WRCC001	0.73	0.64	0.86
WRCC002	0.77	0.73	0.84
Grade 7			
WRCC001	0.77	0.70	0.84
WRCC002	0.51	0.55	0.88
Grade 8			
WRCC001	0.79	0.72	0.88
WRCC002	0.70	0.63	0.83
Grade 11			
WRCC001	0.94	0.80	0.83
WRCC002	0.76	0.68	0.84

Agreement Translation:
< 0 Less than 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 INTRODUCTION

To ensure that reported results for MSAA were accurate relative to collected data, a decision rules document delineating processing rules was prepared and approved by all participating states prior to processing of the results. The decision rules and included participation status structure provided the framework for the reporting requirements, which were defined for each unique report and similarly approved by all participating states prior to reporting.

7.2 DEVELOPMENT AND APPROVAL

The decision rules document was developed by Measured Progress in collaboration with the MSAA steering committee. The decision rules document contains the hierarchy by which the participation statuses were assigned for each individual test incorporating data elements collected by the test platform and directly from the states. The reporting requirements and corresponding report design templates were developed by Measured Progress with the guidance of the MSAA Reporting Committee. Both documents underwent iterative review processes that included draft reviews by the appropriate committee, incorporation of edits, draft reviews by all participating states, and committee review and integration of feedback, until final revisions were approved by all participating states. The approved decision rules are in Appendix B.

To develop the report design templates, Measured Progress worked with the MSAA Reporting Committee to identify modifications to the templates used for NCSC that would ensure the data elements, layout, and report text were meaningful for reporting the MSAA results. Once finalized, the results of this collaborative process were presented to participating state coordinators for final approval.

Measured Progress worked with the MSAA Reports Committee on the *MSAA 2016 Guide for Score Report Interpretation* beginning in May and collaborated to determine the layout and information that would be most helpful to teachers, administrators, and district and school staff, as they reviewed reports and discussed reports with parents or guardians. The guide included an overview of MSAA, student participation criteria, score reporting overview, and samples of the various types of reports available to schools and districts. Guidelines were provided to inform the interpretation and utilization of MSAA scores. The guide also included explanations for all special reporting codes and messages, as well as performance-level scale score ranges. States were permitted to remove codes not used in their state. Measured Progress revised the base document through an iterative process with the MSAA state partners. The Parents/Guardians letter was included as Appendix A and Performance-Level Descriptors for English language arts (ELA) and

Mathematics as Appendix B. The final, approved document was delivered to the MSAA state partners for state-specific revisions and distribution (see Appendix C).

7.3 PRIMARY REPORTS

Measured Progress, in collaboration with the MSAA Reporting Committee created the following primary reports for MSAA:

- Student Reports
- School Roster Reports
- School, District, and State Summary Reports

These confidential reports, along with student results data files, were posted online via the MSAA Assessment System's secure data and reporting portal, with access controlled by user-permissioned accounts, as illustrated in Table 7-1:

Table 7-1. 2015–16 MSAA: Matrix by Users by Report

	State TC	District TC	School TC
Student Reports	Yes	Yes	Yes
School Roster Reports	Yes	Yes	Yes
School Summary Reports	Yes	Yes	Yes
District Summary Reports	Yes	Yes	No
State Summary Reports	Yes	No	No

As determined by MSAA state partners, only test coordinators (TCs) were granted access to the online reports. For the purposes of the assessment system, State MSAA Coordinators were regarded as state TCs. As such, they were able to add new district and school TCs to the online system and to block users no longer in the TC role from accessing the system. Reports were generated for each school, district, or state that had results, as defined by the MSAA decision rules and reporting requirements.

The primary results reported were the student's scaled score and performance-level classification for mathematics and ELA. The performance-level classifications, with cuts determined through the standard setting process, were reported under the generic titles of: Level 1, Level 2, Level 3, and Level 4, with Level 1 as the lowest level, and Level 4 as the highest attainable performance level.

The average scale score and percent of students in each performance level were summarized by school, district, and state on both the roster and summary reports. This allowed for the comparison of individual student performance in relation to the state, as well as for comparison of school and district results against the overall state results.

7.3.1 Student Report

The student report was 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 contained results for both content areas and was developed for parents and guardians of students who participated in MSAA. Reports were organized by school and posted via the secure-access portal for permissioned users to download, print, and disseminate to parents and guardians as appropriate. Each report contained the student name, test grade, and school on the front and back of the report. The back page also included the state student ID for additional confirmation of the student's identification. Additionally, some states chose to print and distribute paper versions of these reports to districts/schools for distribution to students' parents/guardians. Sample student reports are provided in Appendix C.

The front page of the report contained a brief overview of MSAA, including examples of some of the built-in supports available during testing, and highlighted the compatibility of the assessment with various modes of communication. The front also contained a short overview of the results included on the back page, as well as a link to where more information could be accessed online. Parents and guardians were encouraged to communicate with their child's teacher regarding their specific mode of communication and performance.

The back page of the report contained the scaled 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 encapsulated 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.

For students unable to show an observable mode of communication, the lowest scaled score was assigned and displayed along with the Level 1 performance level. This was annotated, and in place of the Level 1 performance-level descriptor, the following text was 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.*

In the event that a student received a student report but did not receive results for one of the two content areas, results for the missing content area are replaced with text encouraging the parent or guardian to contact the child's teacher or school for more information.

7.3.2 Student Roster

The student roster was organized at the school level and provided a by-grade list of all students enrolled in MSAA, with a snapshot of their participation status and results for both content areas. The number of tested students, the average scaled score, and the percent of students by performance level were summarized for the school, district, and state at the top of the roster. Roster reporting requirements identified

which of the participation status codes were included on the roster and which of the participation status codes were included in each calculation.

The summary information at the top of the student roster supported 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 state partners did not suppress information when the number of students participating was small. This practice placed a burden on users to understand the data in the context of small numbers and to use all information provided to understand the results, as explained in the *MSAA 2016 Guide for Score Report Interpretation*.

Student results were listed below the summary section and were identified by name and state student identification number. For each content area the following student-level elements were reported:

- Participation Status
- Scaled Score
- Performance Level
- Comparison to the State Average

It is intended that these data points are to be used in conjunction with the *MSAA 2016 Guide for Score Report Interpretation* (see Appendix C).

7.3.3 Summary Report

Summary reports were 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 was defined by the decision rules and summary report requirements. The following information was summarized by grade and content area and displayed for the school, district, and state, based on the level of the report:

- Number of students enrolled
- Number of valid student tests
- Number of students that did not test
- Average scaled score
- Number and percent of students at each performance level

This summary provided 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 2016 Guide for Score Report Interpretation* (see Appendix C).

7.3.4 Quality Assurance

Proprietary quality-assurance measures at Measured Progress were embedded throughout the entire process of analysis and reporting. The data processors and data analysts who worked on the project implemented quality-control checks of their respective computer programs. Moreover, when data were handed off to different functions within the Data and Reporting Services (DRS) Department, the sending function verified that the data were accurate prior to handoff. Additionally, when a function received a data set, the first step was to verify the data for accuracy.

A second level of quality-assurance measure was parallel processing. One data analyst was responsible for writing all programs required to populate the student and aggregate reporting tables for the administration. Each reporting table was assigned to another data analyst on staff who used the decision rules to independently program the reporting table. The production and quality-assurance tables were compared, and only once there was 100% agreement were the tables released for report generation.

The third aspect of quality control at Measured Progress involved the procedures implemented by the quality-assurance group to check the accuracy of reported data. Using a sample of schools and districts, the quality-assurance group verified that reported information is correct.

The second set of samples included districts or schools that had unique reporting situations as indicated by decision rules. This set was necessary to check that each rule was applied correctly. The third set included districts and schools identified by the client for its review and approval before reports were produced for distribution.

The quality-assurance group used a checklist to implement its procedures. Once the checklist was completed, it underwent an internal parallel verification and then sample reports were circulated for psychometric checks and program management review. Samples of the final reports were then sent for client review and signoff. Simultaneously, Arizona ran successful independent confirmations of the results contained in their state data file. Once signoff was received from all states, the final reports were uploaded into the MSAA Assessment System reporting portal.

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 are conducted to ensure that MSAA English language arts (ELA) and mathematics items meet 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 2016.

8.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

All items are evaluated in terms of item difficulty according to standard classical test theory practices. 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 one-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.

Items that are answered correctly by almost all students provide little information about differences in student abilities, but do indicate knowledge or skills that have been mastered by most students. Similarly, items that are correctly answered by very few students provide little information about differences in student abilities, but may indicate knowledge or skills that have not yet been mastered by most students. In general, to provide the best measurement, difficulty indices should range from near-chance performance of 0.25 (for four-option selected-response items) to 0.90, with the majority of items generally falling between approximately 0.4 and 0.6 for the mathematics grades and between about 0.4 and 0.8 for ELA. However, on a standards-referenced assessment, it may be appropriate to include some items with very low or very high item difficulty values to ensure sufficient content coverage.

A desirable characteristic of an item is for higher-ability students to perform better on the item than lower-ability students do. The correlation between student performance on a single item and total test score is

a commonly used measure of this characteristic of the item. 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 test. For selected-response items, the corresponding statistic is commonly referred to as a point-biserial correlation. The theoretical range of these statistics is –1.0 to 1.0, with a typical observed range from 0.2 to 0.7.

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 and discrimination values shown in the table are within typically observed ranges. One item in the grade 5 ELA test displayed a slightly negative discrimination statistic, but it was not significantly different from 0. Discrimination values near 0 indicate that getting the item correct or incorrect is not indicative of high or low performance on the test as a whole. This item was included in the operational test form to ensure content representativeness.

Table 8-1. 2015–16 MSAA: Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade[1]

Content Area	Grade	Number of Items	*p*-value				Discrimination			
			Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
ELA	3	42	0.34	0.86	0.58	0.14	0.02	0.57	0.39	0.13
	4	41	0.44	0.83	0.60	0.12	0.04	0.60	0.38	0.15
	5	32	0.32	0.85	0.59	0.15	-0.05	0.48	0.31	0.12
	6	33	0.33	0.90	0.63	0.12	0.09	0.49	0.35	0.10
	7	33	0.35	0.86	0.62	0.13	0.13	0.54	0.34	0.10
	8	35	0.35	0.86	0.63	0.13	0.13	0.46	0.32	0.09
	11	32	0.30	0.89	0.71	0.15	0.06	0.51	0.37	0.12
Mathematics	3	35	0.31	0.75	0.53	0.12	0.08	0.58	0.34	0.11
	4	33	0.22	0.79	0.50	0.14	0.10	0.48	0.31	0.11
	5	35	0.24	0.79	0.48	0.15	0.06	0.45	0.25	0.10
	6	35	0.30	0.78	0.54	0.13	0.07	0.47	0.31	0.10
	7	34	0.38	0.87	0.53	0.13	0.12	0.42	0.30	0.07
	8	35	0.33	0.79	0.51	0.11	0.06	0.46	0.32	0.09
	11	34	0.34	0.80	0.50	0.12	0.08	0.49	0.30	0.10

[1]Note: For ELA, the number of items does not equal the number of raw score points.

The individual item statistics can be found in Appendix D. Note that the classical statistics should be interpreted with caution because the items are primarily two- or three-option selected-response items. Because the items were developed to correspond to different tiers, the item statistics have been summarized by tier (Tables 8-2 and 8-3). Although the Tier 1 items tend to be easier than items from the other tiers, the relative difference is much greater when comparing the Tier 1 items to the other tiers than it is among Tiers 2, 3, and 4.

Table 8-2. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Summary by Grade, Content, and Tier- ELA

Content Area	Grade	Tier	Number of Items	*p*-value				Discrimination			
				Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
ELA	3	1	11	0.59	0.86	0.72	0.10	0.17	0.45	0.37	0.08
		2	12	0.34	0.76	0.53	0.13	0.11	0.57	0.42	0.14
		3	18	0.34	0.74	0.53	0.12	0.02	0.54	0.38	0.15
		4	1	0.57	0.57	0.57		0.35	0.35	0.35	
	4	1	11	0.64	0.83	0.76	0.06	0.20	0.53	0.39	0.10
		2	11	0.47	0.68	0.55	0.06	0.15	0.60	0.44	0.15
		3	18	0.44	0.65	0.54	0.07	0.04	0.59	0.34	0.17
		4	1	0.45	0.45	0.45		0.29	0.29	0.29	
	5	1	10	0.63	0.85	0.76	0.07	0.31	0.44	0.38	0.04
		2	7	0.34	0.70	0.52	0.13	-0.05	0.48	0.28	0.19
		3	14	0.32	0.72	0.52	0.11	0.12	0.45	0.29	0.11
		4	1	0.47	0.47	0.47		0.21	0.21	0.21	
	6	1	12	0.60	0.90	0.73	0.09	0.16	0.49	0.37	0.09
		2	7	0.46	0.74	0.62	0.09	0.17	0.48	0.38	0.10
		3	13	0.33	0.65	0.55	0.09	0.09	0.49	0.33	0.11
		4	1	0.52	0.52	0.52		0.25	0.25	0.25	
	7	1	13	0.49	0.86	0.74	0.10	0.15	0.43	0.35	0.07
		2	7	0.40	0.69	0.58	0.10	0.18	0.48	0.36	0.13
		3	12	0.35	0.65	0.53	0.08	0.13	0.54	0.33	0.12
		4	1	0.51	0.51	0.51		0.18	0.18	0.18	
	8	1	15	0.56	0.86	0.73	0.10	0.13	0.45	0.34	0.10
		2	7	0.56	0.72	0.63	0.07	0.29	0.46	0.38	0.07
		3	12	0.35	0.66	0.53	0.10	0.16	0.41	0.29	0.08
		4	1	0.42	0.42	0.42		0.17	0.17	0.17	
	11	1	19	0.58	0.89	0.79	0.08	0.15	0.50	0.39	0.08
		2	6	0.63	0.75	0.71	0.05	0.42	0.51	0.47	0.03
		3	6	0.30	0.68	0.48	0.15	0.06	0.47	0.23	0.18
		4	1	0.48	0.48	0.48		0.33	0.33	0.33	

Table 8-3. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Summary by Grade, Content, and Tier-Mathematics

Content Area	Grade	Tier	Number of Items	*p*-value				Discrimination			
				Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
Mathematics	3	1	7	0.60	0.75	0.70	0.05	0.23	0.35	0.29	0.04
		2	14	0.38	0.71	0.52	0.09	0.24	0.55	0.39	0.09
		3	10	0.31	0.60	0.47	0.09	0.20	0.58	0.35	0.12
		4	4	0.37	0.59	0.44	0.10	0.08	0.37	0.26	0.14
	4	1	8	0.49	0.79	0.66	0.11	0.15	0.33	0.22	0.06
		2	11	0.28	0.68	0.47	0.10	0.10	0.45	0.34	0.12
		3	12	0.23	0.58	0.44	0.10	0.14	0.48	0.34	0.10
		4	2	0.22	0.51	0.37	0.21	0.20	0.32	0.26	0.08
	5	1	8	0.60	0.79	0.70	0.07	0.21	0.37	0.29	0.06
		2	11	0.34	0.55	0.47	0.07	0.10	0.39	0.24	0.08
		3	12	0.24	0.56	0.40	0.10	0.06	0.42	0.25	0.12
		4	4	0.25	0.50	0.34	0.11	0.08	0.45	0.21	0.17

continued

Content Area	Grade	Tier	Number of Items	*p*-value				Discrimination			
				Min	*Max*	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*
Mathematics	6	1	7	0.59	0.78	0.71	0.07	0.22	0.34	0.27	0.05
		2	12	0.32	0.66	0.54	0.09	0.14	0.47	0.36	0.09
		3	12	0.34	0.58	0.49	0.07	0.07	0.44	0.30	0.11
		4	4	0.30	0.43	0.36	0.06	0.08	0.30	0.22	0.10
	7	1	7	0.62	0.87	0.73	0.09	0.16	0.29	0.23	0.05
		2	15	0.38	0.71	0.51	0.10	0.23	0.42	0.34	0.05
		3	10	0.40	0.50	0.44	0.03	0.12	0.36	0.28	0.08
		4	2	0.45	0.52	0.49	0.05	0.33	0.33	0.33	0.00
	8	1	6	0.60	0.79	0.69	0.07	0.17	0.39	0.29	0.08
		2	12	0.40	0.68	0.49	0.09	0.24	0.43	0.36	0.06
		3	13	0.33	0.55	0.45	0.07	0.06	0.46	0.30	0.12
		4	4	0.38	0.47	0.43	0.05	0.25	0.38	0.32	0.06
	11	1	7	0.51	0.80	0.70	0.11	0.12	0.32	0.23	0.07
		2	11	0.34	0.60	0.45	0.07	0.14	0.49	0.34	0.11
		3	14	0.37	0.52	0.45	0.05	0.14	0.43	0.31	0.09
		4	2	0.37	0.46	0.42	0.06	0.08	0.30	0.19	0.16

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 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 2015–16 MSAA, six subgroup comparisons were evaluated for DIF:

- Male vs. female
- White vs. Black

- White vs. Hispanic
- White vs. American Indian
- Not low socioeconomic status (SES) vs. low SES.
- Not Limited English Proficiency (LEP) vs. LEP (including current, exited 1 year, and exited 2 year).

The DIF statistics were calculated based only on the members of the subgroup in question in the computations; values are calculated only for subgroups with 100 or more students. The tables in Appendix E 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 E-1 and E-2 in Appendix E). Dorans and Holland further stated that items with values between –0.10 and –0.05 and between 0.05 and 0.10 (i.e., "low" DIF) should be inspected to ensure that no possible effect is overlooked, and that items with values outside the –0.10 to 0.10 range (i.e., "high" DIF) are more unusual and should be examined very carefully.

The number of items with a "high" DIF index for each tier are shown in Tables 8-4 and 8-5. Since an item can exhibit DIF for multiple comparisons, the item was counted once if any of the comparisons showed "high" DIF.

Table 8-4. 2015–16 MSAA: Number of Items with "High" DIF by Tier—ELA

Grade	Tier 1	Tier 2	Tier 3	Tier 4
3	0 (11)	0 (12)	1 (18)	0 (1)
4	0 (11)	0 (11)	3 (18)	1 (1)
5	0 (10)	1 (7)	2 (14)	0 (1)
6	1 (12)	1 (7)	1 (13)	0 (1)
7	1 (13)	0 (7)	0 (12)	0 (1)
8	2 (15)	0 (7)	1 (12)	0 (1)
11	0 (19)	0 (6)	0 (6)	0 (1)

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

Table 8-5. 2015–16 MSAA: Number of Items with "High" DIF by Tier—Mathematics

Grade	Tier 1	Tier 2	Tier 3	Tier 4
3	0 (7)	0 (14)	1 (10)	0 (4)
4	1 (8)	1 (11)	0 (12)	0 (2)
5	0 (8)	3 (11)	2 (12)	0 (4)
6	0 (7)	1 (12)	3 (12)	1 (4)
7	0 (7)	0 (15)	1 (10)	0 (2)
8	1 (6)	1 (12)	2 (13)	1 (4)
11	0 (7)	0 (11)	0 (14)	0 (2)

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

In addition to the values seen in Appendix E (Tables E-1 and E-2), Tables 8-4 and 8-5 also show that only few items were classified as "high" DIF for each grade and 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 2015–16 MSAA test forms.

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 2015–16 MSAA core items for ELA and mathematics are reported below. (Note: Only core items were analyzed since they are used for score reporting. Also, the MSAA participating states and entities differ slightly from those participating in the NCSC.)

The dimensionality analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999). Both of these 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).

DIMTEST and DETECT were applied to each grade on the 2015–16 MSAA ELA and mathematics tests. The data for each form were split into a training sample and a cross-validation sample. Every form had at least 1,921 student examinees, so every training sample and cross-validation sample had at least 960 students. DIMTEST was then applied to every form. DETECT was applied to each data set for which the DIMTEST null hypothesis was rejected to estimate the effect size of the multidimensionality.

Even though the sample sizes were not large for the MSAA test forms, the DIMTEST null hypothesis was rejected at a significance level of 0.01 for every data set, suggesting that the violations of local independence were sizeable. Thus, it was important to use DETECT to estimate the effect size of the violations of local independence found by DIMTEST. Table 8-6 displays the multidimensional effect size estimates from DETECT. (Note: Since the 2015–16 MSAA operational form used 2014–15 NCSC Form 1 as a basis, the 2014–15 multidimensional effect size presented in the table was calculated using Form 1 only. Also, the multidimensionality effect sizes among the 2014–15 NCSC four operational forms were quite similar. For more detailed information about 2014–15 NCSC multidimensional effect sizes, the reader is referred to the *National Center and State Collaborative 2015 Operational Assessment Technical Manual*.)

Table 8-6. 2015–16 MSAA: Multidimensional Effect Sizes by Content Area and Grade

Content Area	Grade	Multidimensionality Effect Size	
		2014–15	*2015–16*
ELA	3	1.58	1.39
	4	1.33	1.27
	5	0.86	0.72
	6	0.80	1.07
	7	0.89	0.98
	8	0.95	0.96
	11	1.08	0.78
	Average	1.07	1.02
Mathematics	3	1.36	1.09
	4	1.19	1.10
	5	1.64	1.12
	6	0.88	1.18
	7	1.01	1.22
	8	0.66	0.92
	11	0.67	0.71
	Average	1.06	1.05

All the DETECT values for 2015–16 indicated moderate to strong and very strong multidimensionality. Given the unusually large DETECT indices, it was important to identify the source(s) of the violations of local independence. 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, such as item type (e.g., foundational items and Tier 1 writing prompt selected-response items), item position, cognitive load (e.g., tier level), and assorted other item content considerations. From our investigation we found no evidence that any of these characteristics were related to the DETECT clusters. However, we did find a strong and consistent, though unusual, pattern related to the clusters—the placement of the correct-response key option was a very strong indicator of the cluster membership of nearly every item. As an example, consider the grade 3 mathematics test. This test form had 35 items, and the DETECT analysis reported a three-cluster solution. The first cluster contained 13 items, the second had 11 items, and the third had 11 items. The first cluster included all 11 items for which “A” was the correct response option, and the remaining two items in that cluster were open-response items. The second cluster contained all 11 items for which the last response option (“B” for two-option items and “C” for three-option items) was the correct response option. The third cluster contained all 11 items for which the middle response option (“B” for three-option items) was the correct response option. All the test forms for all the grades for both mathematics and ELA showed similar clustering. In addition to this type of clustering, one other type of clustering was observed. For grades 3 and 4 on the ELA tests, the foundational items formed their own separate clusters.

These dimensionality analysis results indicate two types of violations of local independence: one having to do with how some student scores are related to the placement of the correct response options and

the other having to do with the construct underlying the foundational items. In general, it is important that violations of local independence be understood, monitored, and controlled on tests. The violations of local independence having to do with the foundational items are well controlled by the combination of strict test specifications in regard to content, psychometric characteristics, and scoring so that their influence on student scores can be kept consistent from form to form and year to year. The violations of local independence that are related to the ordering of the correct-response option in selected-response items is a new phenomenon that will require close study and some changes to either test administration practices or the test specifications, or both, to ensure that it is either reduced to a minimal level or is constrained in a consistent and equitable manner with respect to student scores.

CHAPTER 9 ITEM RESPONSE THEORY SCALING AND EQUATING

This chapter describes the procedures used to calibrate, equate, and scale the MSAA tests. During the course of 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 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) of getting a dichotomous item correct. 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 (Hambleton & Swaminathan, 1985; Hambleton & van der Linden, 1997). The process of determining the specific mathematical relationship between θ and p is called item calibration. After items are calibrated, they are defined by a set of parameters that specify a nonlinear, monotonically increasing relationship between θ and p. 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.

For the 2015–16 MSAA tests, the two-parameter logistic (2PL) model was used for dichotomous items. The 2PL model for dichotomous items can be defined as:

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

where
i indexes the items,
j indexes students,
α represents item discrimination,
b represents item difficulty, and
D is a normalizing constant equal to 1.701.

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 ITEM RESPONSE THEORY RESULTS

The tables in Appendix F give the IRT item parameters for all the core items on the 2015–16 MSAA tests by grade and content area. The statistics for the core items are summarized in Tables 9-1 through 9-3. The mean item parameter estimates shown in the tables below are within generally acceptable and expected ranges. For easy reference, Table 9-1 displays the means and standard deviations averaged across all core items for each grade and content area.

Table 9-1. 2015–16 MSAA: IRT Summary Statistics Overall

Content Area	Grade	Number of Items	*a*	SD (*a*)	*b*	SD (*b*)
ELA	3	42	0.81	0.37	-0.29	1.19
	4	41	0.81	0.44	-0.45	0.60
	5	32	0.67	0.36	-0.03	1.87
	6	33	0.72	0.34	-0.44	0.8
	7	33	0.80	0.47	-0.49	0.61
	8	35	0.71	0.35	-0.57	0.65
	11	32	1.02	0.52	-0.58	1.08
Mathematics	3	35	0.68	0.31	-0.10	0.70
	4	33	0.57	0.25	0.15	0.90
	5	35	0.55	0.27	0.24	0.99
	6	35	0.62	0.27	-0.02	0.90
	7	34	0.59	0.18	-0.14	0.71
	8	35	0.60	0.18	0.01	0.66
	11	34	0.61	0.21	0.11	0.75

Although the IRT statistics appear slightly more variable than the classical statistics, they remain consistent with them; the difference between the content areas is somewhat expected and the design of the assessments calls for English language arts (ELA) items to be administered in sets while the mathematics design does not. Because the items were developed to correspond to different tiers, the item statistics have also been summarized by tier for ELA (Table 9-2) and mathematics (Table 9-3).

Table 9-2. 2015–16 MSAA: IRT Summary Statistics by Grade and Tier- ELA

Grade	Tier	Number of Items	*a*	SD (*a*)	*b*	SD (*b*)
3	1	11	0.81	0.28	-1.06	0.45
	2	12	0.88	0.38	-0.17	0.64
	3	18	0.77	0.43	0.10	1.57
	4	1	0.53		-0.33	
4	1	11	0.98	0.42	-1.15	0.47
	2	11	0.88	0.38	-0.16	0.29
	3	18	0.68	0.46	-0.24	0.47
	4	1	0.43		0.38	

continued

Grade	Tier	Number of Items	*a*	SD (*a*)	*b*	SD (*b*)
5	1	10	1.01	0.30	-1.06	0.34
	2	7	0.53	0.35	1.23	3.55
	3	14	0.52	0.23	0.05	0.87
	4	1	0.27		0.44	
6	1	12	0.91	0.38	-0.91	0.32
	2	7	0.72	0.27	-0.43	0.44
	3	13	0.55	0.24	-0.04	1.05
	4	1	0.44		0.03	
7	1	13	1.06	0.53	-0.96	0.42
	2	7	0.76	0.44	-0.34	0.57
	3	12	0.60	0.30	-0.09	0.50
	4	1	0.22		-0.36	
8	1	15	0.89	0.42	-1.03	0.33
	2	7	0.74	0.21	-0.56	0.29
	3	12	0.5	0.16	-0.09	0.69
	4	1	0.22		0.59	
11	1	19	1.17	0.52	-1.08	0.27
	2	6	1.15	0.23	-0.72	0.13
	3	6	0.46	0.38	1.05	1.66
	4	1	0.67		0.11	

Table 9-3. 2015–16 MSAA: IRT Summary Statistics by Grade and Tier- Mathematics

Grade	Tier	Number of Items	*a*	SD (*a*)	*b*	SD (*b*)
3	1	7	0.66	0.18	-0.96	0.26
	2	14	0.76	0.34	-0.08	0.41
	3	10	0.65	0.35	0.15	0.54
	4	4	0.45	0.25	0.73	0.95
4	1	8	0.49	0.20	-0.84	0.56
	2	11	0.64	0.25	0.30	0.52
	3	12	0.59	0.29	0.49	0.78
	4	2	0.39	0.05	1.32	1.34
5	1	8	0.71	0.25	-0.85	0.35
	2	11	0.48	0.20	0.19	0.41
	3	12	0.52	0.24	0.73	1.05
	4	4	0.5	0.47	1.10	1.02
6	1	7	0.62	0.24	-1.10	0.32
	2	12	0.74	0.33	-0.04	0.71
	3	12	0.57	0.23	0.26	0.72
	4	4	0.4	0.11	1.11	0.77
7	1	7	0.61	0.26	-1.22	0.35
	2	15	0.62	0.18	0.04	0.54
	3	10	0.54	0.12	0.34	0.27
	4	2	0.63	0.09	-0.03	0.13
8	1	6	0.58	0.17	-0.99	0.50
	2	12	0.69	0.16	0.00	0.41
	3	13	0.54	0.22	0.37	0.54
	4	4	0.59	0.08	0.32	0.32

continued

Grade	Tier	Number of Items	*a*	SD (*a*)	*b*	SD (*b*)
11	1	7	0.57	0.16	-0.93	0.66
	2	11	0.68	0.27	0.30	0.57
	3	14	0.61	0.15	0.32	0.30
	4	2	0.42	0.29	1.17	0.71

Item difficulty tends to have a positive relationship with tier; as the tier increases the items tend to be more difficult. Consistent with the classical statistics, the Tier 1 items appear to be less similar from the other tiers in terms of magnitude of difficulty, and the Tiers 2 and 3 items occasionally overlap. This reversal of difficulty (between Tiers 2 and 3) tends to happen more frequently in ELA than in mathematics. Further investigation may be warranted in grade 5 ELA where the Tier 2 items appear to be more difficult than anticipated.

The TCCs provide a more complete picture of the equivalence of the various forms. 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 items or item sets (for the Tier 1 writing prompt selected-response item sets) 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

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

where
X indexes total raw test score,
X_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 G shows graphs of the TCC and TIF for each grade/content area.

9.3 EQUATING

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year, as well as to equate one year's forms to those given in the previous year. Equating ensures that students are not advantaged or disadvantaged because the test form they took is easier or harder than those taken by other students.

All 2015–16 MSAA tests used item pre-equating methodology as described in Kolen and Brennan (2014). Item pre-equating allows the raw-to-scale score conversion to be produced before the form is administered, which in turn allows for faster reporting and turnaround times. In item pre-equating, new forms are built from a pool of preexisting IRT-calibrated items. In addition to these operational items, new nonoperational items can also be included on the forms. The operational items are then used as a set of common items for transforming the item parameters of the nonoperational items so that they are the same θ scale as the IRT-calibrated item pool. This allows for the item pool to be expanded continually.

However, with pre-equating there are a number of cautions that need to be taken into consideration. Kolen and Brennan (2014) state that to ensure items behave the same on each administration the items should appear in the same contexts and positions operationally as they did nonoperationally. Thus, care must to be taken to avoid significant shifts in position and context. Any drift must be carefully monitored and controlled to ensure comparability between forms of the test. Section 11.1 describes our scale validation, post-equated check procedures.

Item parameters for the 2015–16 operational administration were calibrated after the 2014–15 NCSC operational administration. Also, the 2015–16 operational forms were a straight reuse of one of the four 2014–15 operational forms with minor item replacements, and no new items were included on the 2015–16 forms. As such, no new calibrations were run for the operational items on these pre-equated tests. Raw score to scale score lookups are displayed in Appendix H.

9.4 ACHIEVEMENT STANDARDS

Cutpoints for MSAA in ELA and mathematics were set in August 2015. Details of the standard setting procedures can be found in the standard setting report (Measured Progress, 2015). The cuts on the theta scale that were established at those meetings are presented in Table 9-4. As alluded to in the discussion of equating above, the scale was established during that base year and the forms serve as the reference for subsequent equating. Also shown in the table are the cutpoints on the reporting score scale (described next).

Note that examinees classified in Levels 3 and 4 are considered "proficient." These cutpoints will remain fixed throughout the assessment program unless standards are reset for any reason.

Table 9-4. 2015–16 MSAA: Cut Scores on the Theta Metric and Reporting Scale

Content Area	Grade	Theta			Scale Score				
		Cut 1	*Cut 2*	*Cut 3*	*Minimum*	*Cut 1*	*Cut 2*	*Cut 3*	*Maximum*
ELA	3	-0.70	-0.18	0.72	1200	1234	1240	1251	1290
	4	-0.53	-0.01	1.43	1200	1234	1240	1258	1290
	5	-0.84	-0.13	1.16	1200	1232	1240	1256	1290
	6	-0.63	0.18	1.19	1200	1231	1240	1253	1290
	7	-0.59	-0.20	0.95	1200	1236	1240	1255	1290
	8	-0.75	0.04	0.78	1200	1230	1240	1250	1290
	11	-0.77	-0.37	0.90	1200	1236	1240	1255	1290
Mathematics	3	-0.65	-0.28	0.77	1200	1236	1240	1254	1290
	4	-0.55	0.01	0.82	1200	1233	1240	1251	1290
	5	-0.84	-0.11	0.99	1200	1231	1240	1255	1290
	6	-0.61	-0.10	0.53	1200	1234	1240	1249	1290
	7	-0.91	-0.25	0.77	1200	1232	1240	1254	1290
	8	-0.66	-0.18	0.44	1200	1234	1240	1249	1290
	11	-0.70	-0.19	0.44	1200	1234	1240	1249	1290

Table 9-5 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 percentage of Levels 3 and 4 within each grade and content are provided in the table.

Table 9-5. 2015–16 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	3,672	39.03	13.67	27.45	19.85	47.30	1237.26	19.77
	4	3,827	40.24	16.33	33.76	9.67	43.43	1236.88	18.85
	5	3,921	27.11	26.80	31.60	14.49	46.09	1238.01	18.65
	6	4,002	39.03	25.44	20.94	14.59	35.53	1235.04	18.75
	7	4,118	39.70	15.42	26.23	18.65	44.88	1238.47	19.01
	8	3,984	32.76	31.30	14.71	21.23	35.94	1235.91	18.49
	11	2,791	30.20	17.38	33.03	19.38	52.41	1242.26	20.31
Mathematics	3	3,691	36.17	14.74	31.89	17.20	49.09	1238.49	19.90
	4	3,846	36.97	20.85	24.60	17.58	42.18	1236.15	18.76
	5	3,943	24.50	26.81	36.80	11.89	48.69	1237.69	17.71
	6	4,019	36.55	23.49	19.06	20.90	39.96	1237.30	19.02
	7	4,119	20.27	33.82	28.14	17.77	45.91	1239.01	18.76
	8	3,993	30.80	21.31	24.37	23.52	47.89	1238.12	18.41
	11	1,964	28.26	27.85	24.64	19.25	43.89	1237.61	18.87

Additionally, graphs of the performance-level distributions are presented in Figures I-1 and I-2 in Appendix I.

9.5 Reported Scaled 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 combinations. The second cut is fixed at 1240 for each grade level. In other words, to be classified in Level 3 or above, a minimum scale score of 1240 was required at all grades.

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 2015–16 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 2015–16 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 comes from their being linear transformations of θ. Raw scores are not comparable from year to year 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 operational scaling, a scaling method with a proficient cut of 1240 and standard deviation of 15 with a lowest obtainable scale score (LOSS) of 1200 and a highest obtainable scale score (HOSS) of 1290 was adopted. A separate linear transformation is used for each grade and content-area combination. As previously stated, the transformation function was determined by fixing the Level 2/Level 3 cut score and the standard deviation of the scale—that is, the cut score set at 1240 and the scale score standard deviation of the

base year fixed at 15. Because only one point within the θ scale score space and the standard deviation of the scale is fixed, the scale score cutpoints between Level 1 and Level 2 and between Level 3 and Level 4 are free to vary across the grade and content-area combinations.

Table 9-6 shows the slope and intercept terms used to calculate the scale scores for each content area and grade. Note that the values in Table 9-6 will not change unless the standards are reset. Also, in a given year it may not be possible to attain a particular scale score, but the scale score cuts will remain the same.

Table 9-6. 2015–16 MSAA: Scale Score Slope and Intercept by Content Area and Grade

Content Area	Grade	Slope	Intercept
ELA	3	11.72	1242.05
	4	12.06	1240.09
	5	12.42	1241.61
	6	12.35	1237.81
	7	12.30	1242.43
	8	12.61	1239.46
	11	11.49	1244.22
Mathematics	3	13.06	1243.67
	4	13.10	1239.87
	5	13.08	1241.41
	6	12.82	1241.25
	7	12.91	1243.24
	8	13.02	1242.36
	11	12.99	1242.48

Appendix H contains raw score to scale score lookup tables for the 2015–16 MSAA tests. These are the actual tables used to determine student scale scores, error bands, and performance levels.

Appendix I presents the impact data for each grade by content area. Also, graphs of the scaled score cumulative frequency distributions for the last two years are presented in Appendix I. The cumulative graphs show the proportion of students at or below each scale score. Note that 2014–15 graphs included all students who took the 2014–15 NCSC assessment, and that the MSAA participating states differ slightly from those participating in the NCSC.

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 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. One possible approach is to give the same test to the same students at two different points in time. If students receive the same scores on each test, the extraneous factors affecting performance are small and the test is reliable. (This is referred to as "test-retest reliability.") A potential problem with this approach is that students may remember items from the first administration or may have gained (or lost) knowledge or skills in the interim between the two administrations. A solution to the remembering items problem is to give a different but parallel test at the second administration. If student scores on each test correlate highly, the test is considered reliable. (This is known as "alternate forms reliability" because an alternate form of the test is used in each administration.) This approach, however, does not address the problem that students may have gained (or lost) knowledge or skills in the interim between the two administrations. In addition, the practical challenges of developing and administering parallel forms generally preclude the use of parallel forms reliability indices. One way to address the latter two problems is to split the test in half and then correlate students' scores on the two half-tests; this in effect treats each half-test as a complete test. By doing this, the problems associated with an intervening time interval and with creating and administering two parallel form of the test are alleviated. This is known as a "split-half estimate of reliability." If the two half-test scores correlate highly, items on the two half-tests must be measuring very similar knowledge or skills. This is evidence that the items complement one another and function well as a group. This also suggests that measurement error will be minimal.

The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation, since each different possible split of the test into halves will result in a different correlation. Another problem with the split-half method of calculating

reliability is that it underestimates reliability, because test length is cut in half. All else being equal, a shorter test is less reliable than a longer test. Cronbach (1951) provided a statistic, α (alpha), that eliminates the problem of the split-half method by comparing individual item variances to total test variance. Cronbach's α was used to assess the reliability of the 2015–16 MSAA tests:

$$\alpha \equiv \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma^2_{(Y_i)}}{\sigma^2_x}\right],$$

where
i indexes the item,
n is the total number of items,
$\sigma^2_{(Y_i)}$ represents individual item variance, and
σ^2_x represents the total test variance.

10.1 Reliability and Standard Errors of Measurement

Tables 10-1 and 10-2 present descriptive statistics, Cronbach's α coefficient, and raw score standard errors of measurement (SEMs) for English language arts (ELA) and mathematics by grade. (Statistics are based on core items, which counted toward students' reported scores only.) The reliability of a test can also be exhibited in terms of the SEMs. SEMs can facilitate the interpretation of individual scores. With any given observed raw score point, the reasonable limits of the true score for the examinees can be calculated by using the SEMs. For more detailed description about the use of SEMs, the reader is referred to Gulliksen (1950) or Anastasi and Urbina (1997). SEM was also used to assess the reliability of the 2015–16 MSAA tests:

$$SEM \equiv \sigma^2_x \sqrt{1-\alpha},$$

where
σ^2_x represents the total test variance, and
α represents the reliability coefficient, Cronbach's alpha.

Table 10-1. 2015–16 MSAA: Reliability by Form—ELA

Grade	Number of Students	Raw Score			Alpha	SEM
		Maximum	*Mean*	*Standard Deviation*		
3	3,221	30	18.17	6.03	0.85	2.34
4	3,422	31	18.84	6.15	0.85	2.38
5	3,548	30	17.55	5.53	0.81	2.41
6	3,577	30	18.99	6.03	0.85	2.34
7	3,713	29	18.22	5.69	0.84	2.28
8	3,600	31	19.62	5.72	0.83	2.36
11	2,528	28	19.87	5.23	0.85	2.03

Table 10-2. 2015–16 MSAA: Reliability by Form—Mathematics

Grade	Number of Students	Raw Score			Alpha	SEM
		Maximum	*Mean*	*Standard Deviation*		
3	3,225	35	18.62	7.10	0.86	2.66
4	3,435	33	16.26	6.24	0.83	2.57
5	3,552	35	16.75	5.71	0.78	2.68
6	3,582	35	18.57	6.62	0.84	2.65
7	3,701	34	17.89	6.32	0.83	2.61
8	3,590	35	17.52	6.79	0.84	2.72
11	1,724	34	16.53	6.29	0.82	2.67

All of the reliability coefficients fell within acceptable ranges for large-scale summative assessments. An alpha coefficient that is 0.7 or higher is acceptable according to Kline (2000). Because different grades have different test designs, it is inappropriate to make inferences about the quality of one test by comparing its reliability to that of another test from a different grade.

10.2 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2015–16 MSAA tests. Appendix J presents reliabilities for various subgroups of interest. Subgroup Cronbach's α's and SEMs were calculated using the formula defined above based only on the members of the subgroup in question in the computations; values are calculated only for subgroups with 100 or more students.

For several reasons, the results of this section 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 J that subgroup sample sizes may vary considerably, which results in natural variation in reliability coefficients. Alternatively, α, which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper & Smith, 1998). 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.

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 were classified into those levels, empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications. For

MSAA, students are classified into one of four performance levels: Level 1, Level 2, Level 3, and 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. Instead, techniques have been developed to estimate both the accuracy and the consistency of classification decisions based on a single administration of a test. The Livingston and Lewis (1995) technique was used for the 2015–16 MSAA because it is easily adaptable to all types of testing formats, including mixed format tests.

The accuracy and consistency estimates reported in Appendix K make use of "true scores" in the classical test theory sense. A true score is the score that would be obtained if a test had no measurement error. Of course, true scores cannot be observed and so must be estimated. In the Livingston and Lewis (1995) method, estimated true scores are used to categorize students into their "true" classifications.

For the 2015–16 MSAA, after various technical adjustments (described in Livingston & Lewis, 1995), a four-by-four contingency table of accuracy was created for each content area by grade, where cell [i, j] represented the estimated proportion of students whose true score fell into classification i (where i = 1 to 4) and whose observed score fell into classification j (where j = 1 to 4). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

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

The decision accuracy and consistency (DAC) analyses described above are provided in Table K-1 of Appendix K. The table includes overall accuracy and consistency indices, including kappa. Accuracy and consistency values conditional on performance level are also given. For these calculations, the denominator is the proportion of students associated with a given performance level. For example, the conditional accuracy value is 0.87 for Level 1 for grade 3 ELA. This table indicates that among the students whose true scores placed them in this classification, 87% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.81 indicates that 81% of students with observed scores in Level 1 would be expected to score in this classification again if a second, parallel test form were used. The relatively lower accuracy and consistency values conditional on performance levels, in particular, the ones for Level 2 result from the relatively small number of students who fall into this category and the narrower score ranges of the performance level. However, it is also a clear indication of where test development should focus on to improve the quality of the tests, even though the overall indexes are satisfactory.

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 2015–16 MSAA, Table K-2 in Appendix K 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.

Note that, as with other methods of evaluating reliability, DAC statistics calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Appendix K should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare DAC statistics between grades and content areas.

CHAPTER 11 VALIDITY

11.1 SCALE VALIDATION AND POST-EQUATED CHECK OF PRE-EQUATED TESTS

As described in Section 9.3, all MSAA tests were equated using item pre-equating. However, with pre-equating there are a number of cautions that need to be taken into consideration. Kolen and Brennan (2014) state that to ensure that items behave the same on each administration the items should appear in the same contexts and positions operationally as they did nonoperationally. Thus, care must to be taken to avoid significant shifts in position and context. Any drift must be carefully monitored and controlled to ensure comparability between forms of the test. The MSAA items came from a large item bank whose scale is constantly monitored to ensure the stability of the items. This section describes the processes used to conduct this monitoring.

To provide scale validation evidence, Measured Progress performs a rigorous post-equated check of the test data. One primary aim of the check is to use item bank parameters selectively to exclude the adverse effect of parameter drift on the stability and health of the item bank. Another advantage of the check is the use of more recent calibration samples to get the most up-to-date and accurate parameter estimates.

The procedures for the post-equated check generally mirror the procedures for post-equating. 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 particular well-known violations of item response theory (IRT) equating assumptions and uses the research to estimate their effects on the reliability and validity of the equating. Additionally, the team analyzes the equating data 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.

Furthermore, special procedures are enacted during the calibration phase to check that the quality of the equating items is maintained consistently across years. Equating items that display lack of stability are flagged and removed from equating usage.

Item parameter estimates for the current year's tests are 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 future test form development.

11.2 Test Score Validation Evidence

Because interpretations of test scores, and not a test itself, are evaluated for validity, the purpose of the 2015–16 MSAA Technical Report is to describe several technical aspects of the MSAA tests in support of score interpretations (AERA et al., 2014). Each chapter contributes an important component in the investigation of score validation: test development and design; test administration; scoring, scaling, and equating; item analyses; reliability; and score reporting.

As stated in the overview chapter, Standards for Educational and Psychological Testing (AERA et al., 2014) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. The evidence around test content, response processes, internal structure, relationship to other variables, and consequences of testing speaks to different aspects of validity, but those aspects are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

Evidence on test content validity is meant to determine how well the assessment tasks represent the curriculum and standards for each content area. Content validation is informed by the item development process, including how the test blueprints and test items align to the curriculum and standards. Viewed through this lens provided by the standards, evidence based on test content was extensively described in Chapters 3 and 4. Item alignment with MSAA content standards; item bias, sensitivity, and content appropriateness review processes; adherence to the test blueprint; use of multiple item types; use of standardized administration procedures, with accommodated options for participation; and appropriate test administration training are all components of validity evidence based on test content. As discussed earlier, all MSAA questions undergo several rounds of review for content fidelity and appropriateness. Items are presented to students in multiple formats (constructed-response and selected-response). Finally, tests are

administered according to MSAA-mandated standardized procedures, with allowable accommodations, and all TAs are required to attend annual training sessions and pass a qualifying quiz prior to being allowed to administer tests.

Evidence based on internal structure is presented in great detail in the discussions of item analyses, reliability, and scaling and equating in Chapters 8 through 10. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation), differential item functioning (DIF) analyses, reliability, standard errors of measurement (SEM), and item response theory (IRT) parameters and procedures. Each test is equated to the same content test from the prior year to preserve the meaning of scores over time. In general, item difficulty and discrimination indices were in acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that most items were assessing consistent constructs, and students who performed well on individual items tended to perform well overall.

Evidence based on the consequences of testing is addressed in the scale scores information in Chapter 9 and the reporting information in Chapter 7, as well as in the test interpretation guide (*MSAA 2016 Guide for Score Report Interpretation*; see Appendix C), which is a separate document referenced in the discussion of reporting. Each of these chapters speaks to the efforts undertaken to promote accurate and clear information provided to the public regarding test scores. Scale scores offer the advantage of simplifying the reporting of results across content areas and subsequent years. Performance levels provide users with reference points for mastery at each content area, which is another useful and simple way to interpret scores. Several different standard reports are provided to stakeholders. Additional evidence of the consequences of testing could be supplemented with broader investigation of the impact of testing on student learning.

To further support the validation of the assessment program, additional studies might be considered to provide evidence regarding the relationship of MSAA results to other variables including the extent to which scores from MSAA converge with other measures of similar constructs, and the extent to which they diverge from measures of different constructs. Relationships among measures of the same or similar constructs can sharpen the meaning of scores and appropriate interpretations by refining the definition of the construct.

The evidence presented in this report supports inferences of student achievement on the content represented from the MSAA content standards for the MSAA assessments for the purposes of program and instructional improvement and as a component of school accountability.

REFERENCES

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

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.

Chicago Manual of Style (15th ed.). (2003). Chicago: University of Chicago Press.

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.

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.

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: Routledge.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). New York: 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.

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.

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

APPENDICES

APPENDIX A—TEST BLUEPRINTS

English Language Arts Test Blueprint NCSC

Operational English Language Arts Test Blueprint

The tables presented in this appendix constitute the NCSC operational English language arts (ELA) blueprint targets. The tables incorporate the overall content distributions targeted for the operational test. Each grade level/content area is represented by a table which first describes the content category (e.g., Reading: Literary), weights per CCC, standards (CCCs), item types, score point range, and reports the approximate overall scoring weights by content category by grade.

Please note that the content of the tables are targets and the eligible items in the bank affect how closely the test is constructed.

Table A1. NCSC Operational Test Blueprint – ELA Grade 3

Content Category	Weight	Core Content Connector	Item Type	Score Point Range	Passages
Reading: Literary	30-33%	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)	SR	2	2
		3.RL.i2 Answer literal questions and refer to text to support your answer	SR	4	
		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	SR	3-4	
Reading: Informational	30-33%	3.RI.h1** Identify the purpose of a variety of text features	SR	3-4	2
		3.RI.h4 Use illustrations (e.g., maps, photographs, diagrams, timelines) in informational texts to answer questions	SR	2	
		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	2	
		3.RI.k5** Determine the main idea of a text; recount the key details and explain how they support the main idea	SR	2–3	
Reading: Vocabulary	7-10%	3.RWL.i2 Use sentence context as a clue to the meaning of a new word, phrase, or multiple meaning word	SR	2-3	NA
Reading: Foundational	6-7%	3.RWL.h2** Identify grade level words with accuracy	SR	2	NA
Writing	19-20%	3.WI.l4 Sort evidence (e.g., graphic organizer) collected from print and/or digital sources into provided categories	SR	2	NA
		3.WI.p1 Include text features (e.g., numbers, labels, diagrams, charts, graphics) to enhance clarity and meaning	SR	2	
		3.WL.o1With 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	SR	2	
Total	100%			30-32	4

** CCCs require a multipart item to assess.

Table A2. NCSC Operational Test Blueprint – ELA Grade 4

Content Category	Weight	Core Content Connector	Item Type	Score Point Range	Passages
Reading: Literary	35-43%	4.RL.i1 Refer to details and examples in a text when explaining what the text says explicitly	SR	2-3	2
		4.RL.k2** Determine the theme of a story, drama, or poem; refer to text to support answer	SR	3–4	
		4.RL.l1** Describe character traits (e.g., actions, deeds, dialogue, description, motivation, interactions); use details from text to support description	SR	4–6	
Reading: Informational	23-30%	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	2–3	2
		4.RI.i3 Determine the main idea of an informational text	SR	2	
		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	3–4	
Reading: Vocabulary	10-12%	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	1–2	NA
		4.RWL.j1 Use general academic and domain specific words and phrases accurately	SR	1–2	
Reading: Foundational	6-8%	4.RWL.h2** Identify grade level words with accuracy and on successive attempts	SR	2	NA
Writing	19-20%	4.WI.q1 Provide a concluding statement or section to support the information presented	SR	2	NA
		4.WI.p1 Include formatting (e.g., headings, bulleted information), illustrations, and multimedia when useful to promote understanding	SR	2	
		4.WL.o1 Produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (e.g. to entertain), or audience	SR	2	
Total	100%			30-32	4

** CCCs require a multipart item to assess.

Table A3. NCSC Operational Test Blueprint – ELA Grade 5

Content Category	Weight	Core Content Connector	Item Type	Score Point Range	Passages
Reading: Literary	34-44%	5.RL.b1 Refer to details and examples in a text when explaining what the text says explicitly	SR	4	2
		5.RL.c2** Summarize a text from beginning to end in a few sentences	SR	4-6	
		5.RL.d1 Compare characters, settings, events within a story; provide or identify specific details in the text to support the comparison	SR	3–4	
Reading: Informational	28-34%	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	SR	3–4	2
		5.RI.c4** Determine the main idea, and identify key details to support the main idea	SR	3–4	
		5.RI.e2 Explain how an author uses reasons and evidence to support particular points in a text	SR	3-4	
Reading: Vocabulary	9-11%	5.RWL.a2 Use context to determine the meaning of unknown or multiple meaning words or phrases	SR	3	NA
Writing	19-21%	5.WI.b3 Organize ideas, concepts, and information (using definition, classification, comparison/contrast, and cause/effect)	SR	2	NA
		5.WI.d1 Support a topic with relevant facts, definitions, concrete details, quotations, or other information and examples	SR	2	
		5.WL.h1 Produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (e.g. to entertain), or audience	SR	2	
Total*	100%			29-32	4

*The intended operational score is to be derived from 29-32 raw score points.
** CCCs require a multipart item to assess.

Also, note that paired passage sets are used for one of the Informational passages in grades 5–8 and 11.

Table A4. NCSC Operational Test Blueprint – ELA Grade 6

Content Category	Weight	Core Content Connector	Item Type	Score Point Range	Passages
Reading: Literary	33-35%	6.RL.b2 Refer to details and examples in a text when explaining what the text says explicitly	SR	2	2
		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	2	
		6.RL.c3** Summarize a text from beginning to end in a few sentences without including personal opinions	SR	5-6	
Reading: Informational	31-33%	6.RI.b4 Summarize information gained from a variety of sources including media or texts	SR	1	2
		6.RI.c2** Provide a summary of the text distinct from personal opinions or judgments	SR	1-2	
		6.RI.g4 Determine how key individuals, events, or ideas are elaborated or expanded on in a text	SR	4	
		6.RI.g6 Evaluate the claim or argument; determine if it is supported by evidence	SR	3	
Reading: Vocabulary	13-14%	6.RWL.a1 Use context to determine the meaning of unknown or multiple meaning words or phrases	SR	2	NA
		6.RWL.c1 Use general academic and domain specific words and phrases accurately	SR	2	
Writing	20-21%	6.WL.c1 Organize ideas and event so that they unfold naturally	SR	2	NA
		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	2	
		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)	SR	2	
Total*	100%			29-32	4

*The intended operational score is to be derived form 29-32 raw score points.
** CCCs require a multipart item to assess.

Also, note that paired passage sets are used for one of the Informational passages in grades 5–8 and 11.

Table A5. NCSC Operational Test Blueprint – ELA Grade 7

Content Category	Weight	Core Content Connector	Item Type	Score Point Range	Passages
Reading: Literary	30-32%	7.RL.i2** Use two or more pieces of textual evidence to support inferences, conclusions, or summaries of text	SR	4-8	2
		7.RL.j1 Analyze the development of the theme or central idea over the course of the text	SR	2	
Reading: Informational	35-38%	7.RI.j1** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR	4-6	2
		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	3	
		7.RI.l1** Compare/contrast how two or more authors write about the same topic	SR	2	
		7.RI.k4 Evaluate the claim or argument to determine if they are supported by evidence	SR	2	
Reading: Vocabulary	13-14%	7.RWL.g1 Use context as a clue to determine the meaning of a grade appropriate word or phrase	SR	4	NA
Writing	19-20%	7.WL.o1 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)	SR	2	NA
		7.WL.l1 Use precise words and phrases, relevant descriptive details, and sensory language to capture the action and convey experiences and events	SR	2	
		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)	SR	2	
Total*	100%			30-32	4

*The intended operational score is to be derived from 30-32 raw score points.
** CCCs require a multipart item to assess.

Also, note that paired passage sets are used for one of the Informational passages in grades 5–8 and 11.

Table A6. NCSC Operational Test Blueprint – ELA Grade 8

Content Category	Weight	Core Content Connector	Item Type	Score Point Range	Passages
Reading: Literary	26-32%	8.RL.i2** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR	6-8	2
		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	2	
Reading: Informational	35-42%	8.RI.j1** Use two or more pieces of evidence to support inferences, conclusions, or summaries of text	SR	4-6	2
		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	1	
		8.RI.k2 Determine how the information in each section contribute to the whole or to the development of ideas	SR	3	
		8.RI.k4 Identify an argument or claim that the author makes	SR	3	
Reading: Vocabulary	12-13%	8.RWL.g1 Use context as a clue to the meaning of a grade-appropriate word or phrase	SR	2	NA
		8.RWL.i1 Use general academic and domain specific words and phrases accurately	SR	2	
Writing	19-20%	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)	SR	2	NA
		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	2	
		8.WP.k2 Create an organizational structure in which ideas are logically grouped to support the writer's claim	SR	2	
Total	100%			31-32	4

*The intended operational score is to be derived from 31-33 raw score points.

** CCCs require a multipart item to assess.
Also, note that paired passage sets are used for one of the Informational passages in grades 5–8 and 11.

Table A7. NCSC Operational Test Blueprint – ELA Grade 11

Content Category	Weight	Core Content Connector	Item Type	Score Point Rang	Passages
Reading: Literary	24-27%	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	3-4	2
		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	3-4	
Reading: Informational	40-44%	1112.RI.b1** Use two or more pieces of evidence to support inferences, conclusions, or summaries or text	SR	4-6	2
		1112.RI.b5** Determine how key details support the development of the central idea of a text	SR	4-8	
		1112.RI.d1 Determine the author's point of view or purpose in a text	SR	3–4	
		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	1	
Reading: Vocabulary	13-17%	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	2-3	NA
		1112.RWL.c3 Develop and explain ideas for why authors made specific word choices within text	SR	1-3	
Writing	19-21%	1112.WI.b2 Create an organizational structure for writing that groups information logically (e.g., cause/effect, compare/contrast, descriptions and examples) to support paragraph focus	SR	2	NA
		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	2	

		1112.WP.f1 Produce a clear, coherent, permanent product that is appropriate to the specific task, purpose (to persuade), and audience	SR	2	
Total	100%			29-32	4

*The intended operational score is to be derived from 31-33 raw score points.
** CCCs require a multipart item to assess.

Also, note that paired passage sets are used for one of the Informational passages in grades 5–8 and 11.

Mathematics Blueprint NCSC Operational

Mathematics Test Blueprint

The tables presented in this appendix constitute the NCSC operational mathematics blueprint targets. The tables incorporate the overall content distributions targeted for the operational test. Each grade level/content area is represented by a table, which first describes the content category (e.g., Number and Operations Base 10), weights per CCC, standards (CCCs), item types, score point range, and reports the approximate overall scoring weights by content category by grade.

Please note that the content of the tables are targets and the eligible items in the bank affect how closely the test is constructed.

Table A8. NCSC Operational Test Blueprint – Mathematics Grade 3

Content Category	Weight	Core Content Connector	Item Type	Score Point Range
Operations and Algebraic Thinking	28-32%	3.NO.2d3 Solve multiplication problems with neither number greater than 5	SR	3–4
		3.NO.2e1* Solve or solve and check one- or two-step word problems requiring addition, subtraction, or multiplication with answers up to 100	SR	3–4
		3.PRF.2d1 Identify multiplication patterns in a real word setting	SR	3–4
Number and Operations Base Ten	17-23%	3.NO.1j3 Use place value to round to the nearest 10 or 100	SR	3–4
		3.NO.2c1 Solve multi-step addition and subtraction problems up to 100	SR	3–4
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	3–4
		3.SE.1g1 Use =, <, or > to compare 2 fractions with the same numerator or denominator	SR	3–4
Measurement and Data	17-23%	3.DPS.1g1 Collect data; organize into picture or bar graph	SR/CR	3–4
		3.ME.1d2 Measure area of rectilinear figures by counting squares	SR	3–4
Geometry	9-11%	3.GM.1i1 Partition rectangles into equal parts with equal area	SR	3–4
Total**	100%			30-40

* This CCC requires a pair of math item versions.

**The intended operational score is to be derived from 35 raw score points.

Table A9. NCSC Operational Test Blueprint – Mathematics Grade 4

Content Category	Weight	Core Content Connector	Item Type	Score Point 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	3–4
		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])	SR	3–4
		4.NO.2e2* Solve or solve and check one or two step word problems requiring addition, subtraction, or multiplication with answers up to 100	SR	3–4
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	3–4
		4.NO.1n2 Compare up to 2 given fractions that have different denominators	SR	3–4
		4.SE.1g2 Use =, <, or > to compare 2 fractions (fractions with a denominator or 10 or less)	SR	3–4
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	3–4
		4.DPS.1g3 Collect data; organize in graph (e.g. picture graph, line plot, bar graph)	SR/CR	3–4
Geometry	9-11%	4.GM.1h2 Classify two-dimensional shapes based on attributes (# of angles)	SR/CR	3–4
Total**	100%			30-40

* This CCC requires a pair of math item versions.

**The intended operational score is to be derived from 35 raw score points.

Table A10. NCSC Operational Test Blueprint – Mathematics Grade 5

Content Category	Weight	Core Content Connector	Item Type	Score Point 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	3–4
		5.NO.1b4 Round decimals to the next whole number	SR	3–4
		5.NO.2c1 Solve one-step problems using decimals	SR	3–4
		5.NO.2a5 Solve word problems that require multiplication or division	SR	3–4
Number and Operations Fractions	17-23%	5.NO.2c2 Solve word problems involving the addition, subtraction, multiplication, or division of fractions	SR	3–4
		5.PRF.1a1 Determine whether the product will increase or decrease based on the multiplier	SR	3–4
Measurement and Data	17-23%	5.ME.1b2 Convert standard measurements of length	SR	3–4
		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	SR	3–4
Geometry	9-11%	5.GM.1c3 Use order pairs to graph given points	SR/CR	3–4
Total*	100%			30-40

*The intended operational score is to be derived from 35 raw score points.

Table A11. NCSC Operational Test Blueprint – Mathematics Grade 6

Content Category	Weight	Core Content Connector	Item Type	Score Point Range
Ratio and Proportions	28-32%	6.PRF.1c1 Describe the ratio relationship between two quantities for a given situation	SR	3–4
		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?)	SR	3–4
		6.NO.1f1 Find a percent of a quantity as rate per 100	SR	3–4
Expressions and Equations	17-23%	6.PRF.1d1 Solve real world single-step linear equations	SR	3–4
		6.NO.2a6 Solve problems or word problems using up to three-digit numbers and any of the four operations	SR	3–4
The Number System	28-32%	6.NO.2c3 Solve one-step, addition, subtraction, multiplication, or division problems with fractions or decimals	SR	3–4
		6.NO.1d4 Select the appropriate meaning of a negative number in a real world situation	SR	3–4
		6.NO.1d2* Locate positive and negative numbers on a number line	SR	3–4
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%			30-40

* This CCC requires a pair of math item versions.

**The intended operational score is to be derived from 35 raw score points.

Table A12. NCSC Operational Test Blueprint – Mathematics Grade 7

Content Category	Weight	Core Content Connector	Item Type	Score Point 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	3–4
		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	SR	3–4
		7.NO.2f6 Solve word problems involving ratios	SR	3–4
		7.PRF.1f1 Use proportional relationships to solve multistep percent problems in real world situations.	SR	3–4
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	3–4
		7.NO.2i2 Solve division problems with positive/negative numbers	SR	3–4
Statistics and Probability	9-11%	7.DPS.1k1 Analyze graphs to determine or select appropriate comparative inferences about two samples or populations	SR	3–4
Geometry	17-23%	7.ME.2d1 Apply formula to measure area and circumference of circles	SR	3–4
		7.GM.1h2 Find the surface area of three-dimensional figures using nets of rectangles or triangles	SR	3–4
Total*	100%			30-40

*The intended operational score is to be derived from 35 raw score points.

Table A13. NCSC Operational Test Blueprint – Mathematics Grade 8

Content Category	Weight	Core Content Connector	Item Type	Score Point Range
Functions	17-23%	8.PRF.2e2* Identify the rate of change (slope) and initial value (*y*-intercept) from graphs	SR	3–4
		8.PRF.1f2 Describe or select the relationship between the two quantities given a line graph of a situation	SR	3–4
Expressions and Equations	17-23%	8.PRF.1e2 Represent proportional relationships on a line graph	SR	3–4
		8.PRF.1g3 Solve linear equations with 1 variable	SR	3–4
The Number System	9-11%	8.NO.1k3 Use approximations of irrational numbers to locate them on a number line	SR	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	3–4
		8.DPS.1k2 Analyze displays of bivariate data to develop or select appropriate claims about those data	SR	3–4
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	3–4
		8.GM.1g1* Recognize congruent and similar figures	SR	3–4
		8.ME.2d2 Apply the formula to find the volume of 3-dimensional shapes (i.e., cubes, spheres, and cylinders)	SR	3–4
Total**	100%			30-40

* This CCC requires a pair of math item versions.

**The intended operational score is to be derived from 35 raw score points.

Table A14. NCSC Operational Test Blueprint – Mathematics Grade 11

Content Category	Weight	Core Content Connector	Item Type	Score Point Range
Algebra And Functions	47-52%	H.PRF.2b1 Translate a real-world problem into a one-variable linear equation	SR	3–4
		H.PRF.2b2 Solve equations with one or two variables using equations or graphs	SR	3–4
		H.ME.1b2 Solve a linear equation to find a missing attribute given the area, surface area, or volume and the other attribute	SR	3–4
		H.PRF.1c1 Select the appropriate graphical representation of a linear model based on real world events	SR	3–4
		H.PRF.2c1 Make predictions based on a given model (for example, a weather model, data for athletes over years)	SR	3–4
Number and Quantity	17-23%	H.ME.1a2 Solve real world problems involving units of measurement	SR	3–4
		H.NO.1a1 Simplify expressions that include exponents	SR	3–4
Statistics and Probability	17-23%	H.DPS.1b1 Complete a graph given the data, using dot plots, histograms, or box plots	SR/CR	3–4
		H.DPS.1c1 Use descriptive stats, range, median, mode, mean, outliers/gaps, to describe data set	SR	3–4
Geometry	9-11%	H.GM.1b1 Use definitions to demonstrate congruency and similarity in figures	SR	3–4
Total*	100%			30-40

*The intended operational score is to be derived from 35 raw score points.

APPENDIX B—DECISION RULES

Multi-State Alternate Assessment (MSAA) Spring 2016 Analysis and Reporting Decision Rules (Approved 2016)

Multi-State Alternate Assessment (MSAA) Spring 2016 i
Analysis and Reporting Decision Rules (Approved 2016) i
Multi-State Alternate Assessment (MSAA) Spring 2016 1
I. Contract Overview 1
A. Test Administration(s) 1
B. Deliverables 1
II. Data Sources 1
A. Student Demographic Cleanup 1
B. Student Test Cleanup 2
C. Student Build Outs 2
D. Organization Cleanup 3
E. Scoring 3
III. Student Participation and Reporting Status 4
A. Overview 4
B. Participation Status Assignment Hierarchy (by subject: Math, ELA) 4
C. Participation Status Summary 5
IV. Calculations 6
A. Raw Score 6
B. Writing Trait Raw Scores 6
C. Scaling and Equating 6
D. Performance Levels and Cut Scores 6
E. Aggregate Calculations 6
V. Data and Reporting Deliverable Requirements 7
A. General (all deliverables) 7
B. Student Report Specifics 8
C. School Roster Report Specifics 10
E. Student Results Data File Specifics 12
F. State Duplicate / Void Data File 13
G. Early Release Data File 13

Multi-State Alternate Assessment (MSAA) Spring 2016

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.

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; in 15-16 scores for the Tier 2 writing prompt do not contribute to the overall ELA score and are not reported.

B. Deliverables

1. Student Report
   a. State Option - Print Copies (Parent): AR, TN
   b. State Option – Print Copies (Parent & School): MD, RI
2. Roster Report (School)
3. Summary Report (School, District, State)
4. Student Results Data File (School, District, State)
   a. Includes: State Level Proctor ID Lookup File
5. Duplicate/Void Test Data (State)
6. State Option - Early Release Data File (State): AR, ID

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 TAO (except those moved to the Do Not Process school: “0000-000”), regardless of test status or completion status.
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 final Math test and the final 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
      iv. Closed – Withdrew
      v. Closed – No Longer Eligible
   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

a. 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 Breakthrough's reporting platform to enable access to the reports. States may restrict access through control of the user-accounts associated with each organization.

E. Scoring

1. The Tier 2 (or above) 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 field-test in 15-16 and not included in the overall ELA score.
   c. Valid Scores for each trait:

Score / Code	Description	Score Value (Not Applicable in 15-16)	Reported Value
0, 1-3	Final Score	0, 1-3	0 = Unrelated Evidence 1 = Limited Evidence 2 = Partial Evidence 3 = Full Evidence
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

2. All other item scores are taken from the Breakthrough 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 below *(Section B)*.

B. 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, if the test is *Closed – Withdrew* then the Participation Status is **Withdrew**.
6. Otherwise if the test is *Closed – No Longer Eligible* then the Participation Status is **No Longer Eligible**.
7. Otherwise the Participation Status is **Did Not Test**.
8. Duplicate/Void tests, including off-grade tests, are not assigned participation statuses and are excluded from the *Analysis and Reporting dataset*.

C. Participation Status Summary

Participation Status	Description	Abbrev.	MP Code	State Data File (All Scores[1])	School, District Data Files:		In Agg. Calcs
					Scaled Score	Perf. Level	
Tested	Tested	TES	A	Yes	Yes	Yes	Yes
Early Stopping Rule	Closed – No Observable Communication Mode, no responses.	ESR	B	Yes	Yes	Yes	Yes
Early Stopping Rule Misadministration	Closed – No Observable Communication Mode with at least 1 response.	ESM	C	Yes	Yes	No	No
Incomplete	In Progress with at least 1 response.	INC	D	Yes	Yes	No	No
* Administration Irregularity	Administration Irregularity was reported but the does not necessitate an invalidation. Scores should be interpreted with caution.	IRR	E	Yes	Yes	Yes	Yes
* Invalidated	Student-based *or* administration-based irregularity resulting in invalidation.	INV	F	Yes	Yes	No	No
* Parental Refusal	Parental Refusal	PRF	G	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	No	No
* Exempt (Emergency, Medical, Other)	Student meets the requirements for exemption from the test.	EXE	I	No	No	No	No
Did Not Test	No test or an In Progress test with no responses.	DNT	J	No	No	No	No
** Withdrew	Student withdrew	WDR	K	No	No	No	No
** No Longer Eligible	Student is no longer eligible for testing.	NLE	L	No	No	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.					

** Only available through a state-supplied status code.*
*** State-Supplied or identified by MP using the Test Closure Reason.*
[1] *All Scores: State Student Results Files include Item Responses, Raw Scores, Scaled Scores, and Performance Levels, as applicable by status.*

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. For 15-16 the Tier 2 writing prompt is not eligible to be included as a "core" item.

B. Writing Trait Raw Scores

1. For 15-16, 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.

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

2. MSAA cut scores for each performance level were generated during NCSC 2015 standard setting.

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.

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 reports are marked "Confidential" on all pages.

B. Student Report Specifics

1. Each student report consists of a 1-page cover letter followed by 1-page with results for ELA and Math. Both pages are marked "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 – Student Report table below.
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. The scaled score and performance level earned are printed at the top.
   b. The sentence explaining the standard error of measurement associated with the student's scaled score is displayed.
   c. The performance level description associated with the earned performance level is printed below the graph.
   d. For students classified as Early Stopping Rule (ESR):
      i. An asterisk (*) is added to the earned performance level at the top: Level 1*
      ii. The asterisk corresponds to the alternate text to be displayed below the bar graph – see Participation Status Summary Table: Alternate Text.
      iii. 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.
7. 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.

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

Participation Status	Abbrev.	MP Code	Student Report Specifics		
			Scaled Score	Perf Level	Alternate Text
Tested	TES	A	Yes	Yes	
Early Stopping Rule	ESR	B	Yes (lowest)	Yes (Level 1) PLD 1 Text is NOT Shown.	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.
Early Stopping Rule Misadministration	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.
Tested – Incomplete	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	IRR	E	Yes	Yes	
* Invalidated	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	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)	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)	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	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	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	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
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	Yes	Yes	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			

D. 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), Invalid (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.

E. 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.
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 are 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 (See Decision Rules Page 5).
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. For 15-16, the field-test Tier 2 writing prompt scores are not available at the time of reporting and, if left in the layout, will be set to blank. They will be excluded from all school and district data files entirely.

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.

F. 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.

G. Early Release Data File

1. A preliminary results file will be generated for each state selecting option. The file will be produced after removing students moved the Do Not Process School "9999".
2. The following issues may be present in the preliminary 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. States are required to follow the standard demographic cleanup process (separate from this preliminary results file), and preliminary results are subject to change as a result of cleanup.
4. The preliminary 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.

APPENDIX C—MSAA 2016 GUIDE FOR SCORE REPORT INTERPRETATION

Multi-State Alternate Assessment

(MSAA)

2016 Guide
for
Score Report Interpretation

State Specific Information

Listed below is the contact information for each State MSAA Coordinator:

Arizona Audra Ahumada 602-542-4061 Audra.Ahumada@azed.gov	**Arkansas** Jared Hogue 501-682-4946 Jared.Hogue@arkansas.gov	**District of Columbia** LaNysha Adams 202-297-4193 Lanysha.Adams@dc.gov
Maine Sue Nay 207-624-6774 Sue.Nay@maine.gov	**Maryland** Ann Herrmann 410-767-0086 Ann.Herrmann@maryland.gov Nancy Schmitt 410-767-0743 Nancy.Schmitt@maryland.gov	**Montana** Yvonne Field 406-444-0748 yfield@mt.gov Judy Snow 406-444-3656 jsnow@mt.gov
Rhode Island Heather Heineke 401-222-8493 Heather.Heineke@ride.ri.gov	**South Dakota** Jan Martin 605-773-3246 Jan.Martin@state.sd.us Ben Morrison 605-773-6119 Ben.Morrison@state.sd.us	**Tennessee** Lori Nixon 615-741-5113 Lori.Nixon@tn.gov
United States Virgin Islands Alexandria Baltimore-Hookfin 340-773-1095 ext. 7084 alexandria.baltimore@stx.k12.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 Laura Brown (CNMI) 670-237-3022 Laura.Brown@cnmipss.org	

Table of Contents

Introduction to the MSAA

Purpose 1
Student Participation 2
Overview of the MSAA 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
Student Results File CSV 7
Testing Participation Requirements by Content Area 7
District Summary Report 8

Reports for the School

School Summary Report 9
Individual Student Report 11

Appendix A: Individual Student Report 12

Appendix B: Performance Level Descriptors 16

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 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 eleven participating states: Arizona, Arkansas, Maine, Maryland, Montana, the Pacific Assessment Consortium (PAC-6)[1], Rhode Island, South Dakota, Tennessee and Washington, DC.

This guide provides information regarding the administration and results of the spring 2016 MSAA to district and school personal.

[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 measureable 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

The MSAA assesses English language arts (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 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 designed to be administered one-on-one, online, or 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.

The writing prompts at each grade level were field tested this year. Student responses are hand-scored. Results from the writing prompts will not be included on score reports and are not part of the overall ELA score for the 2015-16 year.

MSAA Score Reports

Overview

This guide describes the types of score reports provided for the 2015-16 MSAA administration. The data in the sample reports are for illustrative purposes only and are not intended to reflect performance of any student(s).

Please remember that test data constitute a single source of information and should be used along with other relevant information on student performance, e.g., IEP progress reports, diagnostic assessments, class work, and report cards.

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. Performance level descriptors for each content area and grade level can be found in Appendix B 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. Table 1 on page 16 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.

Interpreting and Using the MSAA Scores

The MSAA tests student performance in English language arts (ELA) and mathematics, based on States 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

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

- The MSAA is a new or fairly new alternate assessment this school year for states, and we recognize that student achievement may differ between MSAA and the previous state assessments for ELA, Mathematics, and Writing.
- Previous state assessments measured the old state standards whereas MSAA measures progress toward post-secondary options using the new Core Content Connectors; which are aligned to the States' Content Standards.
- Do not compare results in ELA and mathematics from previous state alternate assessments, unless you were part of the NCSC assessment, with the results of MSAA because they are different tests that measure different standards.
- The MSAA assessments are based on higher learning standards than states have had before, and the MSAA assessment results are still a new baseline for all states.
- 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 State MSAA Coordinator.

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 TA
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 State MSAA Coordinator. 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 State MSAA Coordinator.

Testing Participation Requirements by Content Area

All students in grades 3 – 8 and 11 are required to be assessed in English language arts (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 State MSAA Coordinator. 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

SUMMARY REPORT
2 Demonstration State
Demonstration District B

3		4 Enrolled	Tested	Did Not Test	Average Scale Score	5 Performance Level							
						Level 1		Level 2		Level 3		Level 4	
						N	%	N	%	N	%	N	%
Grade 03	State	9	2	7	1241	1	50	0	0	1	50	0	0
	District	1	0	1									
Grade 04	State	10	5	5	1232	3	60	2	40	0	0	0	0
	District	3	1	2	1236	0	0	1	100	0	0	0	0
Grade 05	State	10	8	2	1241	3	38	1	13	3	38	1	13
	District	2	2	0	1226	1	50	1	50	0	0	0	0
Grade 06	State	10	5	5	1236	2	40	2	40	1	20	0	0
	District	3	1	2	1228	1	100	0	0	0	0	0	0
Grade 07	State	10	8	2	1249	3	38	0	0	1	13	4	50
	District	2	2	0	1255	1	50	0	0	0	0	1	50
Grade 08	State	10	9	1	1247	1	11	3	33	2	22	3	33
	District	3	2	1	1256	0	0	1	50	0	0	1	50
Grade 11	State	10	10	0	1261	1	10	1	10	3	30	5	50
	District	2	2	0	1252	0	0	0	0	1	50	1	50

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. Summary of results by Grade Level. The state and district data shown here are other third graders in the state and district.
4. Number of students Enrolled, Tested, Invalid and 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.
5. The number and percentage of students at each performance level by grade in the state, district.

Reports for the School

School Summary Report

Figure 2 – Sample School Summary Report

CONFIDENTIAL

msaa
Multi-State Alternate Assessment

1 English Language Arts

SUMMARY REPORT
2 Demonstration State
Demonstration District A
Demonstration School 1

3		4 Enrolled	Tested	Did Not Test	Average Scale Score	5 Performance Level: Level 1		Level 2		Level 3		Level 4	
						N	%	N	%	N	%	N	%
Grade 03	State	9	2	7	1241	1	50	0	0	1	50	0	0
	District	8	2	6	1241	1	50	0	0	1	50	0	0
	School	3	0	3									
Grade 04	State	10	5	5	1232	3	60	2	40	0	0	0	0
	District	7	4	3	1231	3	75	1	25	0	0	0	0
	School	4	2	2	1233	1	50	1	50	0	0	0	0
Grade 05	State	10	8	2	1241	3	38	1	13	3	38	1	13
	District	8	6	2	1245	2	33	0	0	3	50	1	17
	School	3	2	1	1246	1	50	0	0	0	0	1	50
Grade 06	State	10	5	5	1236	2	40	2	40	1	20	0	0
	District	7	4	3	1238	1	25	2	50	1	25	0	0
	School	3	1	2	1236	0	0	1	100	0	0	0	0
Grade 07	State	10	8	2	1249	3	38	0	0	1	13	4	50
	District	8	6	2	1247	2	33	0	0	1	17	3	50
	School	4	2	2	1244	1	50	0	0	0	0	1	50
Grade 08	State	10	9	1	1247	1	11	3	33	2	22	3	33
	District	7	7	0	1244	1	14	2	29	2	29	2	29
	School	3	3	0	1243	1	33	1	33	0	0	1	33
Grade 11	State	10	10	0	1261	1	10	1	10	3	30	5	50
	District	8	8	0	1263	1	13	1	13	2	25	4	50
	School	3	3	0	1283	0	0	0	0	0	0	3	100

© 2016 MSAA. All rights reserved.

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. Summary of results by Grade Level. The state and district data shown here are other third graders in the state, district and school.
4. Number of students Enrolled, Tested, Invalid and 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.
5. 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 2 below.

Figure 3 – Sample School Roster Report

CONFIDENTIAL

msaa
Multi-State Alternate Assessment

SCHOOL ROSTER REPORT
Demonstration State
Demonstration District A
Demonstration School 1
Grade 04

1

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	10	5	1232	60	40	0	0	7	1230	57	43	0	0
District	7	4	1231	75	25	0	0	6	1231	50	50	0	0
School	4	2	1233	50	50	0	0	4	1229	50	50	0	0

Spring 2016

5

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
LastName11, First11 17669	DNT					-	1220	Level 1
LastName13, First13 17665		-	1226	Level 1		+	1238	Level 2
LastName17, First17 17558	DNT					+	1238	Level 2
LastName19, First19 17560		+	1239	Level 2		-	1220	Level 1

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.

© 2016 MSAA. All rights reserved.

Page 1

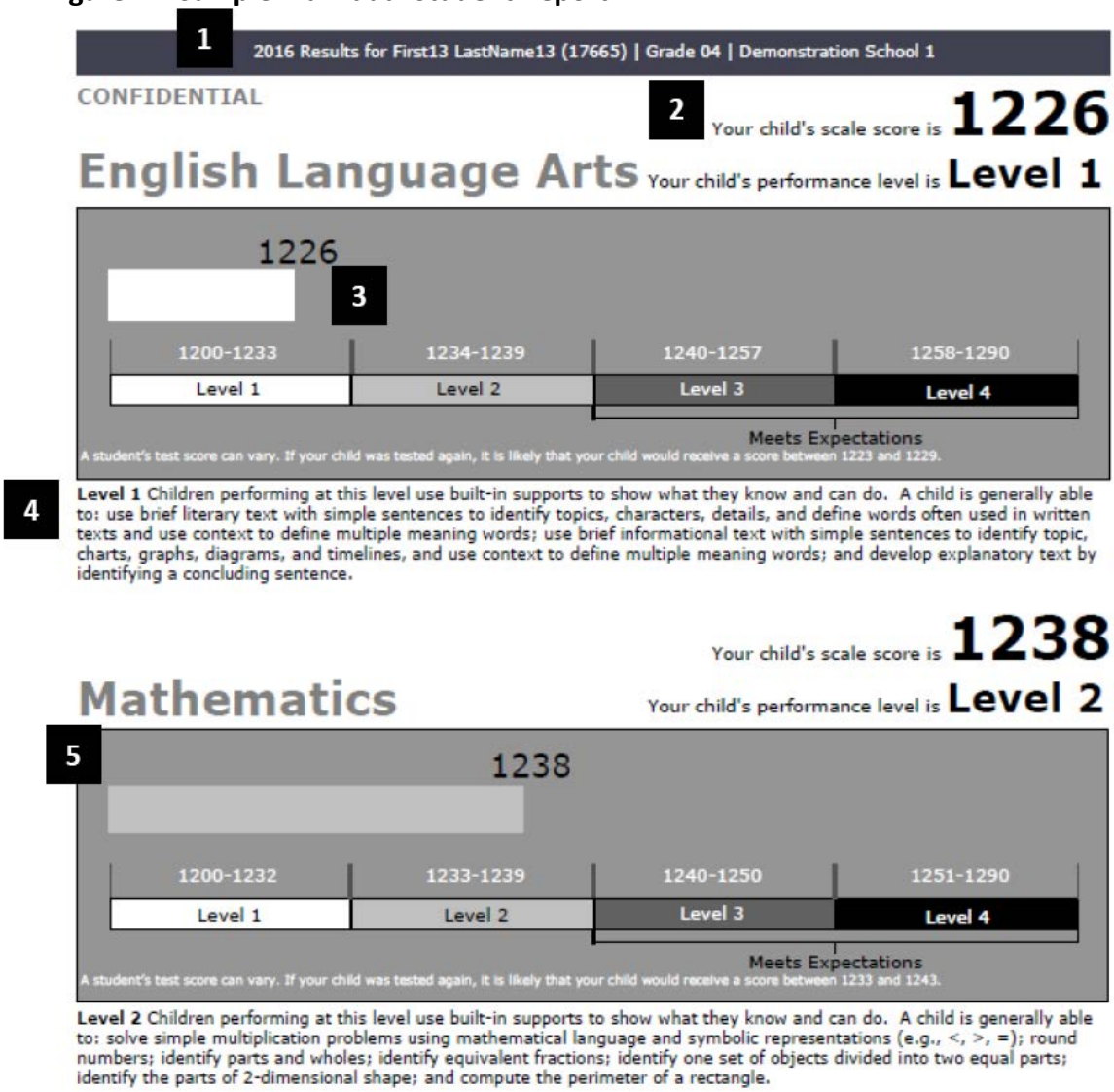
The School Roster Report contains the following features, highlighted above:

1. The state, district and school included in the report.
2. The results are displayed by Content Area.
3. A summary of enrolled and tested students and the average scale score for the state, district and reported school.
4. This section of the report includes all students tested at the school for the specified grade.
5. 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.
6. This key shows symbols used in the "State Compare" column.

Individual Student Report

The Individual Student Report provides scale score and performance level information for a specific student. Figure 3 shows page 2 of the Individual Student Report. A full sample is included in Appendix A.

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 student's scale score and performance level for each content area is shown.
3. This display shows the student's score compared to the performance level scale.
4. This text shows the performance level descriptor for the student's performance level.
5. The results for each content area are displayed separately on the report.

Appendix A

Individual Student Report

Spring 2016 English Language Arts and Mathematics Results for [STUDENT NAME]

Dear Parents and Guardians,

This report shows your child's scale score and performance level for the 2016 Multi-State Alternate Assessment (MSAA) in Mathematics and English Language Arts (ELA).

The MSAA is a group of states that have partnered to develop and administer your state's online alternate assessment for Mathematics and ELA for grades 3 – 8 and 11. The MSAA is designed to assess students 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 as independently as possible. 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 that students can use during the ELA and mathematics tests
- common geometric shapes such as circles, triangles, and squares
- smaller numbers on the mathematics tests
- the option to have the entire test read aloud

In order 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.

On the following pages, the scale score and performance levels for each content area summarizes your child's performance on the ELA and mathematics. The performance level descriptors describe the knowledge and skills that children who perform at this level generally demonstrate.

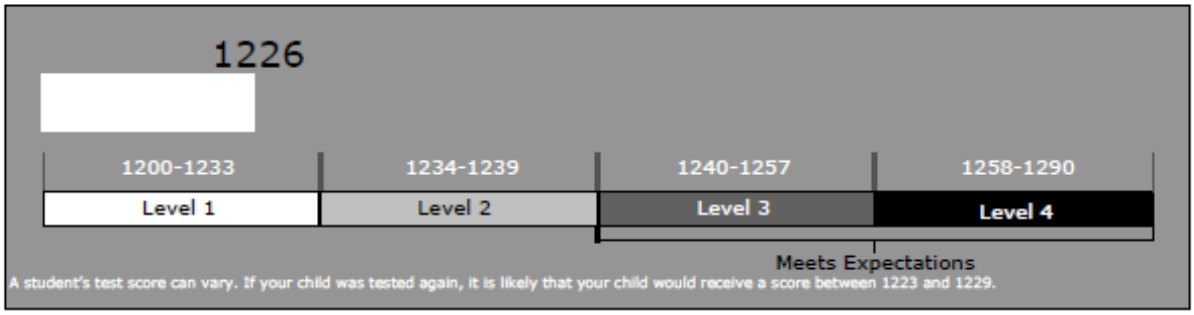
You can find more information and resources for helping your child by talking to your child's teacher or by going to your state's alternate assessment web page.

CONFIDENTIAL

Your child's scale score is **1226**

English Language Arts

Your child's performance level is **Level 1**

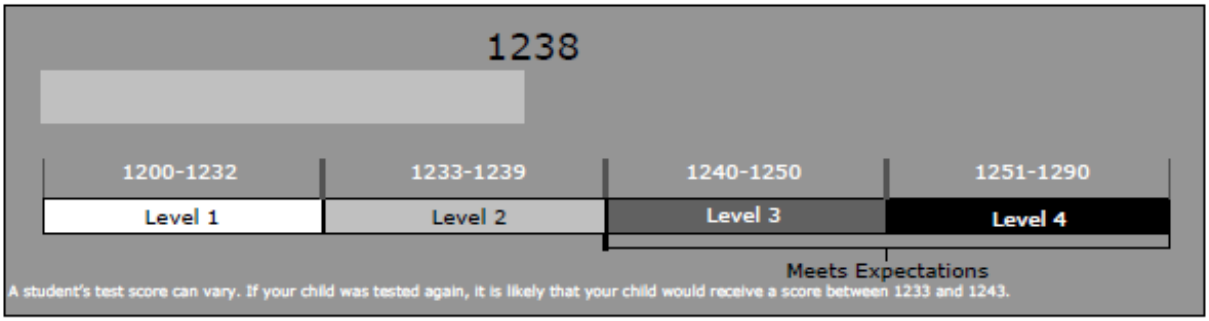


Level 1 Children performing at this level use built-in supports to show what they know and can do. A child is generally able to: use brief literary text with simple sentences to identify topics, characters, details, and define words often used in written texts and use context to define multiple meaning words; use brief informational text with simple sentences to identify topic, charts, graphs, diagrams, and timelines, and use context to define multiple meaning words; and develop explanatory text by identifying a concluding sentence.

Your child's scale score is **1238**

Mathematics

Your child's performance level is **Level 2**



Level 2 Children performing at this level use built-in supports to show what they know and can do. A child is generally able to: solve simple multiplication problems using mathematical language and symbolic representations (e.g., <, >, =); round numbers; identify parts and wholes; identify equivalent fractions; identify one set of objects divided into two equal parts; identify the parts of 2-dimensional shape; and compute the perimeter of a rectangle.

© 2016 MSAA. All rights reserved.

Table 1

2016 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	1251-1290	1258-1290	1256-1290	1253-1290	1255-1290	1250-1290	1255-1290
Level 3	1240-1250	1240-1257	1240-1255	1240-1252	1240-1254	1240-1249	1240-1254
Level 2	1234-1239	1234-1239	1232-1239	1231-1239	1236-1239	1230-1239	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	1255-1290	1249-1290	1254-1290	1249-1290	1249-1290
Level 3	1240-1253	1240-1250	1240-1254	1240-1248	1240-1253	1240-1248	1240-1248
Level 2	1236-1239	1233-1239	1231-1239	1234-1239	1232-1239	1234-1239	1234-1239
Level 1	1200-1235	1200-1232	1200-1230	1200-1233	1200-1231	1200-1233	1200-1233

Appendix B

Performance Level Descriptors

Performance Level Descriptors for ELA and Mathematics

MSAA developed Performance Level Descriptors (PLDs) for mathematics and English language arts (ELA) 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 B 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 help teachers, schools, parents and guardians build a path to student learning.

Grade 3 ELA Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="4">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)</td><td>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</td><td>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</td><td rowspan="4">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</td></tr>
<tr><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td>• use details from a literary text to answer specific questions
• describe the relationship between characters, and character and setting in literary text</td><td>• use details from a literary text to answer specific questions
• describe the relationship between characters, and character and setting in literary text</td></tr>
<tr><td>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)</td><td>AND with accuracy, he/she is able to:
• identify grade level words</td></tr>
<tr><td>AND in writing, he/she is able to:
• identify a statement related to an everyday topic</td><td>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</td><td>AND in writing, he/she is able to:
• identify a text feature (e.g., captions, graphs or diagrams) to present information in explanatory text</td></tr>
</table>

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="4">In reading, he/she 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</td><td>In reading, he/she 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</td><td>In reading, he/she 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</td><td rowspan="4">In reading, he/she 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</td></tr>
<tr><td></td><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td></td><td>• use details from a literary text to answer specific questions
• use context to identify the meaning of multiple meaning words</td><td>• 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</td></tr>
<tr><td></td><td>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)</td><td>AND with accuracy, he/she is able to:
• identify grade level words</td></tr>
<tr><td>AND in writing, he/she is able to:
• identify the concluding sentence in a short explanatory text</td><td>AND in writing, he/she 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</td><td>AND in writing, he/she is able to:
• identify a text feature (e.g., headings, charts, or diagrams) to present information in explanatory text</td><td></td></tr>
</table>

Grade 5 ELA Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="4">In reading, he/she 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</td><td>In reading, he/she 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</td><td>In reading, he/she 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</td><td rowspan="4">In reading, he/she 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</td></tr>
<tr><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td>• summarize a literary text from beginning to end
• use details from a literary text to answer specific questions</td><td>• summarize a literary text from beginning to end
• use details from a literary text to answer specific questions</td></tr>
<tr><td></td><td></td></tr>
<tr><td>AND in writing, he/she is able to:
• identify the category related to a set of common nouns</td><td>AND in writing, he/she 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</td><td>AND in writing, he/she is able to:
• support an explanatory text topic with relevant information</td></tr>
</table>

Grade 6 ELA Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="4">In reading, he/she 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</td><td>In reading, he/she 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</td><td>In reading, he/she 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</td><td rowspan="4">In reading, he/she 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</td></tr>
<tr><td></td><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td></td><td>• use details from a literary text to answer specific questions
• use context to identify the meaning of multiple meaning words</td><td>• use details from a literary text to answer specific questions
• use context to identify the meaning of multiple meaning words</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>AND in writing, he/she is able to:
• identify an everyday order of events</td><td>AND in writing, he/she is able to:
• identify elements of an explanatory text to include introduction, body, and conclusion
• identify the next event in a brief narrative</td><td>AND in writing, he/she is able to:
• identify transition words and phrases to convey a sequence of events in narrative text</td><td></td></tr>
</table>

Grade 7 ELA Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="4">In reading, he/she 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</td><td>In reading, he/she 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 informational text</td><td>In reading, he/she 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</td><td rowspan="4">In reading, he/she 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</td></tr>
<tr><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td>• use details to support themes from literary text
• use details to support inferences from literary text</td><td>• use details to support themes from literary text
• use details to support inferences from literary text</td></tr>
<tr><td>AND in writing, he/she is able to:
• identify a graphic that includes an event as described in a text</td><td>AND in writing, he/she is able to:
• identify elements of an explanatory text to include introduction, body, and conclusion
• identify the next event in a brief narrative</td><td>AND in writing, he/she is able to:
• identify a sentence that provides a conclusion in narrative text</td></tr>
</table>

Grade 8 ELA Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="4">In reading, he/she 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</td><td>In reading, he/she 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</td><td>In reading, he/she 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</td><td rowspan="4">In reading, he/she 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</td></tr>
<tr><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td>• 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</td><td>• 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</td></tr>
<tr><td>AND in writing, he/she is able to:
• identify a writer’s opinion</td><td>AND in writing, he/she is able to:
• identify elements of an explanatory text to include introduction, body, and conclusion
• identify an idea relevant to a claim</td><td>AND in writing, he/she is able to:
• identify relevant information to support a claim</td></tr>
</table>

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Low text complexity -
Brief text with straightforward ideas and relationships; short, simple sentences.</td><td>Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td rowspan="3">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</td><td>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</td><td>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</td><td rowspan="4">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</td></tr>
<tr><td>AND with Moderate text complexity -
Text with clear, complex ideas and relationships and simple; compound sentences.</td><td>AND with High text complexity -
Text with detailed and implied complex ideas and relationships; a variety of sentence types including phrases and transition words.</td></tr>
<tr><td>• 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</td><td>• 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</td></tr>
<tr><td>AND in writing, he/she is able to:
• identify information which is unrelated to a given topic</td><td>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</td><td>AND in writing, he/she is able to:
• identify relevant information to address a given topic and support the purpose of a text</td></tr>
</table>

Grade 3 Mathematics Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>High task complexity -
Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</td></tr>
<tr><td rowspan="4">He/she 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</td><td>He/she 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</td><td>He/she 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</td><td rowspan="4">He/she 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</td></tr>
<tr><td>AND with Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>AND with High task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td></tr>
<tr><td>• identify geometric figures which are divided into equal parts</td><td>• round numbers to nearest 10
• identify geometric figures which are divided into equal parts
• count unit squares to compute the area of a rectangle</td></tr>
</table>

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*
He/she 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	**He/she 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	**He/she 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	**He/she 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	

Grade 5 Mathematics Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>High task complexity -
Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</td></tr>
<tr><td rowspan="3">He/she 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</td><td>He/she 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</td><td>He/she 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</td><td rowspan="3">He/she 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</td></tr>
<tr><td>AND with Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>AND with High task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td></tr>
<tr><td>• compare the values of two products based upon multipliers
• round decimals to nearest whole number</td><td>• compare the values of two products based upon multipliers
• round decimals to nearest whole number</td></tr>
</table>

Grade 6 Mathematics Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>High task complexity -
Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</td></tr>
<tr><td rowspan="3">He/she 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)</td><td>He/she 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</td><td>He/she 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</td><td rowspan="3">He/she 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</td></tr>
<tr><td>AND with Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>AND with High task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td></tr>
<tr><td>• perform one-step operations with two decimal numbers
• solve word problems using a percent</td><td>• perform one-step operations with two decimal numbers
• solve word problems using a percent
• solve word problems using ratios and rates</td></tr>
</table>

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*
He/she 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	**He/she 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	**He/she 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	**He/she 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	

Grade 8 Mathematics Performance Level Descriptors

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Level 4</th></tr>
<tr><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Low task complexity -
Simple problems using common mathematical terms and symbols</td><td>Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>High task complexity -
Multiple mathematical ideas presented in problems using various mathematical terms and symbolic representations of numbers, variables, and other item elements</td></tr>
<tr><td rowspan="4">He/she 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</td><td>He/she 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</td><td>He/she 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</td><td rowspan="4">He/she 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</td></tr>
<tr><td>AND with Moderate task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td><td>AND with High task complexity -
Common problems presented in mathematical context using various mathematical terms and symbols</td></tr>
<tr><td>• identify congruent figures
• use properties of similarity to identify similar figures
• interpret data tables to identify the relationship between variables</td><td>• 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</td></tr>
</table>

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*
He/she 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	**He/she 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	**He/she 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	**He/she 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	

APPENDIX D—ITEM-LEVEL CLASSICAL STATISTICS

Table D-1. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 3

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
114008A	MC	0.67	0.50	1
114010A	MC	0.72	0.33	1
114011A	MC	0.43	0.11	1
114957A	MC	0.80	0.44	2
114958A	MC	0.83	0.39	1
114960A	MC	0.63	0.28	1
115985A	MC	0.64	0.44	1
115986A	MC	0.58	0.48	1
115987A	MC	0.63	0.24	1
115988A	MC	0.73	0.30	1
116202A	MC	0.52	0.34	0
116203A	MC	0.55	0.47	1
116204A	MC	0.62	0.53	0
116205A	MC	0.74	0.35	1
120879A	MC	0.38	0.35	1
120880A	MC	0.50	0.14	1
120922B	MC	0.48	0.34	1
120923A	MC	0.34	0.02	1
120967A	MC	0.86	0.35	1
121545A	MC	0.50	0.43	1
121726A	MC	0.64	0.31	1
121731B	MC	0.34	0.25	1
124168A	MC	0.82	0.36	1
124170A	MC	0.53	0.25	1
124175A	MC	0.76	0.40	2
124181A	MC	0.57	0.35	1
125942A	MC	0.71	0.39	1
125943A	MC	0.65	0.17	1
125945A	MC	0.61	0.45	1
125947B	MC	0.73	0.43	1
125948A	MC	0.59	0.41	1
125949B	MC	0.70	0.38	0

Table D-2. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 4

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
113280A	MC	0.65	0.40	0
113281A	MC	0.57	0.38	1
113282A	MC	0.51	0.04	1
113283A	MC	0.65	0.45	0
114053A	MC	0.60	0.51	1
114054B	MC	0.52	0.39	1
114055A	MC	0.47	0.23	1
114056A	MC	0.68	0.44	1
116574A	MC	0.82	0.40	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
116575A	MC	0.72	0.53	1
116576A	MC	0.83	0.38	1
116577A	MC	0.83	0.30	1
116605A	MC	0.46	0.25	1
116606A	MC	0.51	0.20	1
116607A	MC	0.44	0.23	1
116608A	MC	0.59	0.18	1
121279A	MC	0.80	0.46	1
121426A	MC	0.75	0.48	1
121550A	MC	0.58	0.37	1
121551A	MC	0.57	0.19	1
121575A	MC	0.44	0.20	1
121982A	MC	0.59	0.21	1
122582A	MC	0.54	0.15	1
124194A	MC	0.72	0.47	1
124196A	MC	0.58	0.26	1
124199A	MC	0.45	0.29	1
124205A	MC	0.59	0.34	1
126141A	MC	0.70	0.36	0
126142A	MC	0.79	0.25	0
126143A	MC	0.73	0.20	0
126144B	MC	0.64	0.45	0

Table D-3. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 5

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
115053A	MC	0.73	0.44	1
115054A	MC	0.85	0.36	1
115055A	MC	0.82	0.34	1
115056A	MC	0.73	0.41	1
117109A	MC	0.50	0.23	0
117110B	MC	0.32	0.12	1
117111A	MC	0.66	0.37	0
117112A	MC	0.48	0.31	3
119268A	MC	0.52	0.30	1
119269A	MC	0.47	0.18	1
119270A	MC	0.46	0.35	1
119271A	MC	0.63	0.39	1
119970B	MC	0.38	0.20	1
119971A	MC	0.43	0.18	1
119972B	MC	0.34	-0.05	1
119973A	MC	0.59	0.48	1
121222A	MC	0.79	0.42	1
121478B	MC	0.55	0.39	1
121479A	MC	0.62	0.31	6
121564A	MC	0.61	0.26	1
121568A	MC	0.53	0.40	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
121571B	MC	0.42	0.12	1
121672A	MC	0.72	0.45	0
122062A	MC	0.49	0.19	1
124213A	MC	0.43	0.35	0
124219A	MC	0.83	0.41	1
124228A	MC	0.47	0.21	1
124234A	MC	0.70	0.45	0
126984B	MC	0.73	0.34	0
126985B	MC	0.63	0.35	1
126986A	MC	0.80	0.31	0
126987B	MC	0.68	0.41	0

Table D-4. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 6

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
113533A	MC	0.74	0.48	0
113536A	MC	0.58	0.41	1
113537A	MC	0.65	0.40	1
113611A	MC	0.65	0.49	0
113612A	MC	0.56	0.35	1
113614A	MC	0.53	0.41	1
115181A	MC	0.61	0.24	0
115183A	MC	0.61	0.44	0
115947A	MC	0.55	0.25	0
120042A	MC	0.90	0.33	1
120043A	MC	0.84	0.38	1
120044A	MC	0.77	0.45	0
121225A	MC	0.60	0.38	1
121226A	MC	0.57	0.32	1
121373A	MC	0.71	0.40	0
121374A	MC	0.75	0.49	0
121375A	MC	0.75	0.40	1
121521A	MC	0.56	0.28	1
121522B	MC	0.44	0.31	0
121529B	MC	0.49	0.25	1
121764A	MC	0.64	0.43	1
121768A	MC	0.67	0.45	1
121775A	MC	0.63	0.34	1
123356A	MC	0.33	0.09	1
124240A	MC	0.68	0.31	1
124242A	MC	0.64	0.44	1
124257A	MC	0.52	0.25	1
124263A	MC	0.46	0.17	1
127272B	MC	0.63	0.38	3
127273A	MC	0.69	0.41	1
127274A	MC	0.77	0.27	0
127276A	MC	0.60	0.16	2
127277B	MC	0.65	0.43	0

Table D-5. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 7

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
114593A	MC	0.69	0.44	1
114594A	MC	0.64	0.46	1
114596A	MC	0.65	0.48	1
115372A	MC	0.50	0.13	1
115373A	MC	0.53	0.31	0
115431A	MC	0.65	0.54	0
115432A	MC	0.55	0.35	0
115433A	MC	0.62	0.43	0
115434A	MC	0.52	0.42	1
120098A	MC	0.86	0.39	1
121313A	MC	0.60	0.43	1
121315A	MC	0.53	0.39	0
121343A	MC	0.48	0.20	1
121347A	MC	0.54	0.31	1
121497A	MC	0.77	0.35	1
121501A	MC	0.78	0.39	1
121505A	MC	0.86	0.41	1
121509A	MC	0.78	0.42	1
121512A	MC	0.75	0.43	1
121997A	MC	0.58	0.44	1
121998A	MC	0.53	0.21	1
122175B	MC	0.40	0.18	1
122235A	MC	0.45	0.22	1
124269A	MC	0.83	0.37	1
124271B	MC	0.35	0.24	1
124284A	MC	0.58	0.28	1
124286A	MC	0.51	0.18	1
127690A	MC	0.49	0.15	3
127691A	MC	0.74	0.36	1
127692B	MC	0.70	0.32	1
127693A	MC	0.64	0.33	1
127694A	MC	0.72	0.28	1
127695B	MC	0.73	0.37	0

Table D-6. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 8

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
114228A	MC	0.51	0.19	0
114229A	MC	0.49	0.26	0
114230B	MC	0.54	0.32	1
114231A	MC	0.51	0.32	0
114687A	MC	0.60	0.44	0
114688A	MC	0.56	0.29	1
114689A	MC	0.67	0.46	0
114690B	MC	0.71	0.41	0
114876A	MC	0.65	0.40	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
114877A	MC	0.39	0.19	1
114878A	MC	0.56	0.32	0
114879A	MC	0.66	0.41	0
121030A	MC	0.84	0.44	1
121031A	MC	0.63	0.42	1
121032A	MC	0.79	0.32	0
121033B	MC	0.58	0.16	1
121105A	MC	0.75	0.44	1
121106B	MC	0.61	0.28	1
121107A	MC	0.77	0.45	0
121108A	MC	0.84	0.38	1
121805A	MC	0.53	0.26	1
121890A	MC	0.58	0.31	1
121891A	MC	0.72	0.42	1
122082A	MC	0.52	0.29	1
122086A	MC	0.35	0.16	1
124300A	MC	0.65	0.30	0
124302A	MC	0.86	0.35	1
124309A	MC	0.57	0.33	0
124311A	MC	0.42	0.17	0
127781A	MC	0.65	0.28	3
127782A	MC	0.75	0.33	0
127783A	MC	0.82	0.39	1
127784A	MC	0.56	0.13	0
127785A	MC	0.70	0.32	1
127786B	MC	0.83	0.42	1

Table D-7. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—ELA Grade 11

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
116348A	MC	0.84	0.44	0
116349A	MC	0.83	0.50	0
116350A	MC	0.84	0.47	0
116351A	MC	0.86	0.42	1
117167A	MC	0.61	0.41	0
117168A	MC	0.68	0.47	1
117169B	MC	0.39	0.10	0
119078A	MC	0.67	0.50	6
119079A	MC	0.72	0.51	1
119080A	MC	0.73	0.47	1
119081A	MC	0.74	0.48	1
120148B	MC	0.87	0.39	0
120149A	MC	0.74	0.41	1
120150A	MC	0.71	0.32	1
120151A	MC	0.82	0.36	2
121065A	MC	0.37	0.06	1
121130A	MC	0.51	0.26	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
121227A	MC	0.85	0.43	1
121228A	MC	0.89	0.35	1
121229A	MC	0.71	0.43	1
121230A	MC	0.82	0.45	1
121953B	MC	0.30	0.10	1
122000A	MC	0.63	0.46	1
124319A	MC	0.70	0.15	1
124328A	MC	0.48	0.33	0
124334A	MC	0.75	0.42	0
126773A	MC	0.84	0.43	1
126774B	MC	0.84	0.43	0
126775A	MC	0.69	0.30	0
126776B	MC	0.58	0.29	4
126777B	MC	0.81	0.39	3
126778B	MC	0.79	0.42	0

Table D-8. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 3

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
110855A	MC	0.46	0.42	1
110862A	MC	0.60	0.24	1
110865A	MC	0.39	0.08	1
110876A	MC	0.37	0.36	1
110923A	MC	0.71	0.43	1
110959A	MC	0.68	0.23	1
110964A	MC	0.71	0.33	1
110974A	MC	0.46	0.46	1
110975A	MC	0.51	0.32	1
111377A	MC	0.60	0.31	1
111382A	MC	0.60	0.38	1
111386A	MC	0.52	0.45	1
111391A	MC	0.54	0.33	1
111397A	MC	0.54	0.31	1
111411A	MC	0.46	0.20	1
111416A	MC	0.63	0.34	1
111420A	MC	0.41	0.36	1
111422A	MC	0.42	0.22	1
111424A	MC	0.75	0.30	1
111434A	MC	0.75	0.30	2
111883A	MC	0.35	0.24	1
112552A	MC	0.58	0.35	0
112555A	MC	0.72	0.35	2
112564A	MC	0.67	0.31	1
112565A	MC	0.38	0.37	1
112571A	MC	0.47	0.49	1
112575A	MC	0.60	0.39	1
112595A	MC	0.56	0.49	1
112601A	MC	0.50	0.48	1
112615B	MC	0.45	0.25	1
112616B	MC	0.31	0.20	2
112621B	MC	0.45	0.24	1
120682A	MC	0.59	0.37	1
122091A	MC	0.45	0.58	1
122105A	MC	0.52	0.55	1

Table D-9. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 4

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
111135A	MC	0.60	0.21	1
111136A	MC	0.58	0.45	1
111160A	MC	0.57	0.25	2
111166A	MC	0.50	0.43	1
111178B	MC	0.59	0.17	1
111663A	MC	0.76	0.33	1
111667A	MC	0.79	0.26	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
111676A	MC	0.45	0.42	1
111677A	MC	0.49	0.48	1
111682A	MC	0.47	0.45	1
111686A	MC	0.44	0.35	1
111688A	MC	0.51	0.32	1
111695B	MC	0.74	0.21	1
111705B	MC	0.49	0.20	1
111712B	MC	0.42	0.25	1
111716A	MC	0.49	0.38	1
111717A	MC	0.58	0.37	1
111721A	MC	0.41	0.28	1
111727A	MC	0.43	0.33	1
111731B	MC	0.47	0.35	1
112782A	MC	0.49	0.45	1
112797B	MC	0.42	0.10	1
112801A	MC	0.71	0.15	1
112803B	MC	0.56	0.38	1
112813A	MC	0.29	0.14	1
112832B	MC	0.42	0.16	1
120697A	MC	0.68	0.36	1
121665A	MC	0.37	0.25	1
121691B	MC	0.39	0.28	1
122265A	MC	0.22	0.20	1
122267A	MC	0.51	0.45	1
122394A	MC	0.28	0.43	1
122426A	MC	0.23	0.38	1

Table D-10. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 5

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
111234A	MC	0.56	0.12	1
111244B	MC	0.43	0.39	1
111257A	MC	0.76	0.21	1
111260A	MC	0.50	0.08	1
111276A	MC	0.47	0.30	1
111291A	MC	0.74	0.33	1
111299A	MC	0.33	0.17	1
111303A	MC	0.53	0.23	1
112340A	MC	0.66	0.30	1
112342A	MC	0.48	0.24	1
112346A	MC	0.68	0.34	1
112358A	MC	0.34	0.27	1
112363A	MC	0.35	0.25	1
112372A	MC	0.74	0.37	1
112373B	MC	0.52	0.21	1
112374B	MC	0.52	0.19	1
112378A	MC	0.55	0.39	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
112385A	MC	0.45	0.27	1
112409A	MC	0.37	0.36	1
112416A	MC	0.53	0.13	1
113838A	MC	0.44	0.24	1
113843B	MC	0.44	0.27	1
113844A	MC	0.79	0.32	1
113852A	MC	0.52	0.38	1
113856A	MC	0.31	0.14	1
113862A	MC	0.32	0.18	1
113867A	MC	0.33	0.13	1
113877A	MC	0.28	0.06	1
113883A	MC	0.49	0.10	1
113884B	MC	0.60	0.25	1
113889A	MC	0.61	0.21	1
113892A	MC	0.46	0.31	1
113901A	MC	0.28	0.15	1
120737A	MC	0.24	0.42	1
121514A	MC	0.25	0.45	1

Table D-11. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 6

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
110891A	MC	0.76	0.28	2
110909A	MC	0.75	0.27	1
110910A	MC	0.49	0.27	1
110938A	MC	0.51	0.32	1
110943A	MC	0.30	0.28	1
110978A	MC	0.46	0.34	1
110986A	MC	0.59	0.38	1
110996A	MC	0.48	0.42	1
111027A	MC	0.34	0.14	1
111444A	MC	0.74	0.34	0
111450A	MC	0.78	0.33	1
111457A	MC	0.38	0.30	1
111461A	MC	0.48	0.28	1
111478A	MC	0.61	0.44	1
111482A	MC	0.66	0.44	1
111487A	MC	0.59	0.22	1
111492A	MC	0.53	0.27	1
111496A	MC	0.71	0.25	3
111507A	MC	0.50	0.38	1
111514A	MC	0.57	0.41	1
111517A	MC	0.61	0.47	0
111518A	MC	0.53	0.31	1
111632A	MC	0.38	0.07	1
112645A	MC	0.48	0.41	1
112654A	MC	0.58	0.30	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
112657A	MC	0.63	0.35	1
112658A	MC	0.54	0.27	1
112663A	MC	0.45	0.26	1
112667A	MC	0.32	0.14	1
112676A	MC	0.66	0.23	1
112678A	MC	0.52	0.41	0
112699A	MC	0.43	0.22	1
120855A	MC	0.56	0.38	0
121487A	MC	0.51	0.44	1
121520A	MC	0.32	0.08	1

Table D-12. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 7

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
111046A	MC	0.87	0.24	1
111055A	MC	0.50	0.33	1
111059B	MC	0.47	0.35	1
111076A	MC	0.42	0.26	0
111080A	MC	0.43	0.34	1
111093A	MC	0.49	0.33	1
111098A	MC	0.71	0.29	1
111101A	MC	0.52	0.33	1
111113A	MC	0.40	0.36	1
111127A	MC	0.44	0.30	1
111131A	MC	0.38	0.23	1
111734A	MC	0.57	0.42	1
111735A	MC	0.42	0.31	1
111744A	MC	0.65	0.32	1
111748A	MC	0.74	0.16	3
111749A	MC	0.71	0.34	1
111754B	MC	0.50	0.38	1
111758A	MC	0.70	0.23	1
111765A	MC	0.40	0.34	1
111766A	MC	0.45	0.33	0
111769A	MC	0.62	0.36	1
111778A	MC	0.83	0.28	1
111779A	MC	0.54	0.39	1
111804A	MC	0.50	0.30	1
112853B	MC	0.50	0.40	1
112870A	MC	0.45	0.24	0
112871A	MC	0.48	0.16	1
112882A	MC	0.46	0.28	1
112886A	MC	0.41	0.30	1
112889A	MC	0.62	0.22	1
112890A	MC	0.38	0.33	1
112899A	MC	0.64	0.19	1
112901A	MC	0.41	0.12	1
112911A	MC	0.42	0.35	1

Table D-13. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 8

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
111247A	MC	0.72	0.39	1
111286A	MC	0.44	0.37	1
111336A	MC	0.42	0.27	1
111339A	MC	0.41	0.29	1
111581A	MC	0.60	0.24	1
111583A	MC	0.37	0.21	1
111594A	MC	0.47	0.35	1
111622A	MC	0.45	0.38	1
112464A	MC	0.69	0.34	1
112466A	MC	0.52	0.33	1
112470A	MC	0.61	0.24	1
112476A	MC	0.48	0.06	0
112480A	MC	0.43	0.29	0
112481A	MC	0.39	0.26	1
112484A	MC	0.71	0.24	1
112495A	MC	0.51	0.43	1
112506A	MC	0.53	0.36	1
112515B	MC	0.41	0.30	1
113908A	MC	0.55	0.42	1
113918A	MC	0.68	0.35	0
113922A	MC	0.51	0.42	1
113926A	MC	0.61	0.34	1
113931A	MC	0.38	0.25	1
113934A	MC	0.79	0.17	1
113938A	MC	0.40	0.33	1
113942A	MC	0.55	0.43	1
113952A	MC	0.41	0.32	1
113968A	MC	0.52	0.46	1
113972A	MC	0.33	0.14	1
113978A	MC	0.52	0.36	2
117072A	MC	0.46	0.43	1
120569B	MC	0.40	0.18	1
122050A	MC	0.47	0.38	1
122051A	MC	0.47	0.39	1
122099A	MC	0.47	0.37	1

Table D-14. 2015–16 MSAA: Item-Level Classical Test Theory Statistics—Mathematics Grade 11

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
110847A	MC	0.47	0.28	2
110851A	MC	0.43	0.29	1
110914A	MC	0.43	0.42	1
110921A	MC	0.47	0.34	1
110936A	MC	0.59	0.32	1
110969B	MC	0.52	0.21	1
111016A	MC	0.43	0.34	1

continued

Item ID	Item Type	*p*-values	Item-Total Correlation	Omit Rates
111032A	MC	0.37	0.37	1
111109B	MC	0.44	0.16	2
111533A	MC	0.76	0.31	2
111537A	MC	0.37	0.30	1
111538A	MC	0.40	0.43	1
111544A	MC	0.37	0.25	1
111554B	MC	0.48	0.39	1
111557A	MC	0.42	0.37	1
111810A	MC	0.42	0.49	1
111817A	MC	0.51	0.17	1
111824A	MC	0.49	0.36	1
111830A	MC	0.67	0.21	2
111832B	MC	0.80	0.12	1
112703B	MC	0.39	0.42	1
112708A	MC	0.47	0.40	1
112717A	MC	0.47	0.43	1
112722A	MC	0.60	0.42	1
112726A	MC	0.76	0.24	4
112921B	MC	0.34	0.14	2
112931A	MC	0.42	0.20	1
112936A	MC	0.41	0.32	0
112941A	MC	0.49	0.34	1
112944A	MC	0.79	0.23	1
112945A	MC	0.54	0.25	1
112946A	MC	0.50	0.29	2
122054A	MC	0.46	0.08	0
122055A	MC	0.45	0.14	0

APPENDIX E—DIFFERENTIAL ITEM FUNCTIONING RESULTS

Table E-1. 2015–16 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*
3	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	4	1	3	1	0	1
	White	American Indian or Alaska Native	MC	35	9	5	4	0	0	0
		Black or African American	MC	35	3	3	0	0	0	0
		Hispanic or Latino	MC	35	0	0	0	0	0	0
4	Male	Female	MC	33	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	33	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	7	2	5	0	0	0
	White	American Indian or Alaska Native	MC	33	10	6	4	2	1	1
		Black or African American	MC	33	7	6	1	0	0	0
		Hispanic or Latino	MC	33	1	1	0	0	0	0
5	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	9	4	5	3	2	1
	White	American Indian or Alaska Native	MC	35	7	5	2	2	0	2
		Black or African American	MC	35	2	1	1	0	0	0
		Hispanic or Latino	MC	35	1	1	0	0	0	0
6	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	7	1	6	3	2	1
	White	American Indian or Alaska Native	MC	35	8	4	4	2	1	1

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	*Reference*			*Total*	*Favoring*		*Total*	*Favoring*	
						Reference	*Focal*		*Reference*	*Focal*
6	White	Black or African American	MC	35	5	4	1	0	0	0
		Hispanic or Latino	MC	35	1	1	0	0	0	0
7	Male	Female	MC	34	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	34	0	0	0	0	0	0
	Non-LEP	LEP	MC	34	10	3	7	0	0	0
	White	American Indian or Alaska Native	MC	34	5	3	2	1	0	1
		Black or African American	MC	34	4	2	2	0	0	0
		Hispanic or Latino	MC	34	7	5	2	0	0	0
8	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	12	2	10	2	1	1
	White	American Indian or Alaska Native	MC	35	10	7	3	3	1	2
		Black or African American	MC	35	1	0	1	0	0	0
		Hispanic or Latino	MC	35	1	0	1	0	0	0
11	Male	Female	MC	34	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	34	8	1	7	0	0	0
	White	Black or African American	MC	34	6	4	2	0	0	0
		Hispanic or Latino	MC	34	4	3	1	0	0	0

Table E-2. 2015–16 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*
3	Male	Female	MC	32	0	0	0	0	0	0
			OR	10	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	32	0	0	0	0	0	0
			OR	10	1	1	0	0	0	0
	Non-LEP	LEP	MC	32	7	5	2	0	0	0
			OR	10	5	0	5	0	0	0
	White	American Indian or Alaska Native	MC	32	12	5	7	1	1	0
			OR	10	1	0	1	0	0	0
		Black or African American	MC	32	1	1	0	0	0	0
			OR	10	0	0	0	0	0	0
		Hispanic or Latino	MC	32	0	0	0	0	0	0
			OR	10	1	0	1	0	0	0
4	Male	Female	MC	31	1	0	1	0	0	0
			OR	10	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	31	1	0	1	0	0	0
			OR	10	0	0	0	0	0	0
	Non-LEP	LEP	MC	31	10	2	8	3	2	1
			OR	10	0	0	0	0	0	0
	White	American Indian or Alaska Native	MC	31	6	4	2	1	1	0
			OR	10	0	0	0	0	0	0
		Black or African American	MC	31	0	0	0	0	0	0
			OR	10	0	0	0	0	0	0
		Hispanic or Latino	MC	31	2	1	1	0	0	0
			OR	10	0	0	0	0	0	0
5	Male	Female	MC	32	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	32	0	0	0	0	0	0
	Non-LEP	LEP	MC	32	4	2	2	0	0	0
	White	American Indian or Alaska Native	MC	32	4	1	3	3	3	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	*Focal*			*Total*	*Favoring*		*Total*	*Favoring*	
						Reference	*Focal*		*Reference*	*Focal*
5	White	Black or African American	MC	32	2	2	0	0	0	0
		Hispanic or Latino	MC	32	1	1	0	0	0	0
6	Male	Female	MC	33	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	33	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	5	4	1	2	1	1
	White	American Indian or Alaska Native	MC	33	9	3	6	1	1	0
		Black or African American	MC	33	1	0	1	0	0	0
		Hispanic or Latino	MC	33	3	2	1	0	0	0
7	Male	Female	MC	33	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	33	0	0	0	0	0	0
	Non-LEP	LEP	MC	33	9	2	7	1	1	0
	White	American Indian or Alaska Native	MC	33	3	1	2	0	0	0
		Black or African American	MC	33	2	2	0	0	0	0
		Hispanic or Latino	MC	33	2	1	1	0	0	0
8	Male	Female	MC	35	0	0	0	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	6	2	4	2	1	1
	White	American Indian or Alaska Native	MC	35	6	3	3	1	1	0
		Black or African American	MC	35	1	0	1	0	0	0
		Hispanic or Latino	MC	35	1	1	0	0	0	0
11	Male	Female	MC	32	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	32	2	0	2	0	0	0
	White	Black or African American	MC	32	2	2	0	0	0	0
		Hispanic or Latino	MC	32	2	2	0	0	0	0

APPENDIX F—ITEM RESPONSE THEORY PARAMETERS

Table F-1. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 3

IREF	*a*	SE (*a*)	*b*	SE (*b*)
114008A	1.02880	0.03787	-0.81752	0.02740
114010A	0.63698	0.02888	-1.10469	0.04919
114011A	0.19478	0.01881	0.56227	0.11007
114957A	1.28518	0.10023	-1.11420	0.05586
114958A	1.10583	0.09645	-1.40089	0.08008
114960A	0.43026	0.04474	-0.78655	0.11378
115985A	0.66327	0.05407	-0.68183	0.07314
115986A	0.76233	0.05617	-0.35806	0.05713
115987A	0.40442	0.04339	-0.69090	0.11471
115988A	0.56553	0.05288	-1.11532	0.10821
116202A	0.52703	0.02397	-0.32882	0.03983
116203A	0.75440	0.02798	-0.17970	0.02850
116204A	1.13772	0.03802	-0.45046	0.02173
116205A	0.71472	0.03186	-1.21888	0.04825
120879A	0.53051	0.04735	0.87748	0.09868
120880A	0.17818	0.03192	-0.12028	0.20878
120922B	0.54963	0.02400	-0.03269	0.03687
120923A	0.06654	0.01312	6.09144	1.22854
120967A	0.82828	0.08349	-1.79217	0.13465
121545A	0.77928	0.02849	-0.18231	0.02777
121587A	1.29463	0.07949	0.00724	0.03726
121588A	1.28590	0.07950	0.35496	0.03978
121589A	1.37265	0.08376	0.00191	0.03583
121590B	1.05662	0.03436	0.16072	0.02239
121591A	1.30562	0.04104	-0.09552	0.01893
121593A	1.23822	0.03904	0.02400	0.01973
121605A	1.37211	0.08369	0.01386	0.03587
121608B	1.37119	0.08377	-0.02189	0.03579
121610B	1.28893	0.04051	-0.05552	0.01911
121613A	1.13965	0.03647	0.19669	0.02136
121726A	0.40995	0.02255	-0.70020	0.05807
121731B	0.40457	0.02222	1.00592	0.06859
124168A	0.85375	0.03964	-1.49734	0.05203
124170A	0.35290	0.02122	-0.34326	0.05784
124175A	1.05569	0.04083	-1.01048	0.03029
124181A	0.53299	0.02408	-0.33284	0.03947
125942A	0.81820	0.03201	-0.82223	0.03331
125943A	0.28200	0.02146	-1.54561	0.12884
125945A	0.87972	0.03182	-0.51539	0.02716
125947B	0.95356	0.03676	-0.93654	0.03135
125948A	0.72144	0.02775	-0.37120	0.03064
125949B	0.71726	0.02974	-0.87899	0.03857

Table F-2. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 4

IREF	*a*	SE (*a*)	*b*	SE (*b*)
113280A	0.75172	0.02914	-0.71741	0.03319
113281A	0.64594	0.02563	-0.32689	0.03256
113282A	0.08237	0.01431	-0.52931	0.24050
113283A	0.80801	0.03003	-0.62975	0.02990
114053A	1.14949	0.03701	-0.37377	0.02072
114054B	0.54526	0.02346	0.03553	0.03633
114055A	0.39095	0.02113	0.23362	0.05018
114056A	0.88968	0.03274	-0.75065	0.02922

continued

IREF	*a*	SE (*a*)	*b*	SE (*b*)
116574A	0.93793	0.04087	-1.39253	0.04392
116575A	1.63857	0.05476	-0.68140	0.01749
116576A	0.87019	0.03935	-1.48253	0.04988
116577A	0.68325	0.03435	-1.71979	0.07059
116605A	0.35963	0.04071	0.34205	0.11336
116606A	0.23898	0.03569	0.17266	0.16028
116607A	0.40706	0.04247	0.48668	0.10635
116608A	0.27522	0.03804	-0.78507	0.17061
121279A	1.38608	0.05134	-0.98791	0.02354
121426A	1.56911	0.05451	-0.80607	0.01910
121550A	0.67702	0.02610	-0.27339	0.03090
121551A	0.29487	0.02041	-0.84976	0.08359
121575A	0.37857	0.04143	0.43589	0.11147
121624A	1.50564	0.09096	-0.16130	0.03343
121626A	1.11844	0.06921	0.18130	0.04205
121627A	1.40268	0.08425	-0.00022	0.03528
121628B	1.16186	0.03614	-0.00966	0.02007
121630A	0.97392	0.03167	0.06092	0.02287

IREF	*a*	SE (*a*)	*b*	SE (*b*)
121631A	1.11588	0.03496	0.10742	0.02086
121643A	1.44797	0.08747	-0.14110	0.03429
121647B	1.28027	0.03968	-0.21909	0.01880
121650A	1.33371	0.04066	-0.06892	0.01825
121982A	0.26979	0.02006	-0.92566	0.09452
122582A	0.21062	0.01877	-0.27206	0.09095
124194A	1.05515	0.03725	-0.75190	0.02537
124196A	0.37218	0.02147	-0.67624	0.06152
124199A	0.42742	0.02167	0.38440	0.04823
124205A	0.59716	0.02503	-0.46696	0.03640
126141A	0.72232	0.02911	-0.84891	0.03692
126142A	0.48979	0.02779	-1.82924	0.09531
126143A	0.37277	0.02346	-1.63627	0.10481
126144B	1.03228	0.03494	-0.53100	0.02348
200626A	1.15373	0.07093	0.11998	0.04074

Table F-3. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 5

IREF	*a*	SE (*a*)	*b*	SE (*b*)
115053A	1.44516	0.05081	-0.76197	0.01917
115054A	1.20298	0.05262	-1.28394	0.03437
115055A	0.69835	0.03445	-1.52388	0.06147
115056A	1.29788	0.04604	-0.76830	0.02099
117109A	0.33139	0.03731	0.10520	0.11119
117110B	0.22684	0.03468	2.26600	0.37093
117111A	0.63156	0.05184	-0.85088	0.07751
117112A	0.54105	0.04381	0.13698	0.07214
119268A	0.61338	0.02467	-0.13396	0.03255
119269A	0.34599	0.02028	0.26542	0.05628
119270A	0.72458	0.02625	0.16307	0.02909

IREF	*a*	SE (*a*)	*b*	SE (*b*)
119271A	0.70879	0.02828	-0.64668	0.03304
119970B	0.35202	0.02061	1.05350	0.07757
119971A	0.24676	0.01904	0.71727	0.09146
119972B	0.04116	0.00891	9.10447	2.01886
119973A	1.04712	0.03467	-0.34639	0.02154
121222A	0.99373	0.04159	-1.16417	0.03524
121478B	0.59694	0.02474	-0.30524	0.03411
121479A	0.54260	0.02507	-0.77696	0.04508
121564A	0.33047	0.02126	-1.00374	0.08088
121568A	0.80687	0.02842	-0.15254	0.02597
121571B	0.15166	0.01747	1.20791	0.18133

continued

IREF	*a*	SE (*a*)	*b*	SE (*b*)
121672A	0.90496	0.03564	-0.93065	0.03159
122062A	0.30641	0.03652	0.03546	0.11882
124213A	0.63385	0.02455	0.29163	0.03338
124219A	1.29733	0.05520	-1.21494	0.03036
124228A	0.26546	0.01925	0.44346	0.07615

IREF	*a*	SE (*a*)	*b*	SE (*b*)
124234A	0.85474	0.03344	-0.85412	0.03150
126984B	0.62330	0.02852	-1.08687	0.04839
126985B	0.83096	0.03076	-0.56934	0.02774
126986A	0.63435	0.03203	-1.53259	0.06607
126987B	1.10956	0.03955	-0.72145	0.02329

Table F-4. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 6

IREF	*a*	SE (*a*)	*b*	SE (*b*)
113533A	1.09303	0.03897	-0.79416	0.02462
113536A	0.69462	0.02586	-0.21223	0.02948
113537A	0.64633	0.02621	-0.62850	0.03566
113611A	0.99001	0.06760	-0.61275	0.04869
113612A	0.55634	0.04584	-0.20920	0.07111
113614A	0.61608	0.04787	-0.18430	0.06489
115181A	0.41301	0.02236	-0.92229	0.06281
115183A	0.81267	0.02913	-0.46899	0.02738
115947A	0.31390	0.03768	-0.24135	0.12078
120042A	1.41351	0.13984	-1.56889	0.08129
120043A	1.36128	0.11257	-1.23256	0.05758
120044A	1.30949	0.09139	-0.82433	0.04289
121225A	0.60954	0.04851	-0.40168	0.06849
121226A	0.45902	0.04307	-0.42307	0.08900
121373A	0.84460	0.06267	-0.79925	0.06100
121374A	1.42557	0.09814	-0.79093	0.03935
121375A	0.91395	0.06939	-0.96233	0.06308

IREF	*a*	SE (*a*)	*b*	SE (*b*)
121521A	0.47406	0.02231	-0.34806	0.04247
121522B	0.46582	0.02179	0.62973	0.04822
121529B	0.36706	0.02046	-0.03348	0.05106
121764A	0.86715	0.03050	-0.49607	0.02615
121768A	0.90187	0.03221	-0.63755	0.02674
121775A	0.60815	0.02568	-0.70052	0.03903
123356A	0.15728	0.00000	3.22859	0.38671
124240A	0.53225	0.02488	-0.93124	0.05012
124242A	0.88884	0.03160	-0.59551	0.02658
124257A	0.44123	0.02142	0.03167	0.04318
124263A	0.25519	0.01900	0.46564	0.07874
127272B	0.70014	0.02665	-0.44671	0.03089
127273A	0.76439	0.02927	-0.70921	0.03196
127274A	0.62182	0.02919	-1.31365	0.05639
127276A	0.28137	0.01987	-0.81607	0.08493
127277B	0.80063	0.02935	-0.56852	0.02876

Table F-5. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 7

IREF	*a*	SE (*a*)	*b*	SE (*b*)
114593A	1.46403	0.05125	-0.77077	0.01928
114594A	0.90022	0.03285	-0.63961	0.02688
114596A	1.07575	0.03722	-0.60564	0.02278
115372A	0.26797	0.01972	-0.61515	0.08086
115373A	0.47055	0.02231	-0.11949	0.04091
115431A	1.17260	0.07689	-0.55101	0.04137
115432A	0.50862	0.04537	-0.07750	0.07582
115433A	0.97336	0.06645	-0.52800	0.04769
115434A	0.86999	0.05841	-0.06452	0.04852
120098A	1.16721	0.10383	-1.39032	0.07656
121313A	0.71572	0.05486	-0.52790	0.06194
121315A	0.74465	0.05407	-0.20711	0.05520
121343A	0.27699	0.01946	0.22651	0.06827
121347A	0.47054	0.02234	-0.14742	0.04103
121497A	0.62652	0.05936	-1.36880	0.11629
121501A	1.91570	0.14519	-0.96024	0.03487
121505A	1.36346	0.11731	-1.28688	0.06197
121509A	1.88027	0.14033	-0.93539	0.03470
121512A	1.61864	0.11591	-0.87805	0.03746
121997A	0.82477	0.02968	-0.39577	0.02641
121998A	0.32622	0.02029	-0.22406	0.05839
122175B	0.24387	0.01911	0.88584	0.09972
122235A	0.31262	0.01991	0.47329	0.06562
124269A	1.37871	0.05708	-1.17146	0.02797
124271B	0.35907	0.02087	1.11609	0.07836
124284A	0.46963	0.02314	-0.62212	0.04774
124286A	0.21950	0.01882	-0.35626	0.08850
127690A	0.34607	0.02034	0.26791	0.05594
127691A	0.74526	0.03137	-1.01073	0.03973
127692B	0.57713	0.02713	-1.08176	0.05183
127693A	0.69750	0.02786	-0.62304	0.03332
127694A	0.69628	0.02971	-0.96924	0.04091
127695B	0.72733	0.03092	-1.02016	0.04082

Table F-6. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 8

IREF	*a*	SE (*a*)	*b*	SE (*b*)
114228A	0.39878	0.02061	0.17625	0.04731
114229A	0.47432	0.02175	-0.00168	0.03972
114230B	0.53892	0.02332	-0.41292	0.03776
114231A	0.61739	0.02403	-0.02746	0.03159
114687A	0.90580	0.03071	-0.42540	0.02420
114688A	0.51718	0.02282	-0.34265	0.03837
114689A	0.98433	0.03428	-0.69301	0.02503
114690B	0.94016	0.03485	-0.89656	0.02928
114876A	0.66293	0.05276	-0.92630	0.07871
114877A	0.26818	0.03571	1.31822	0.21082
114878A	0.46082	0.04164	-0.24083	0.08106
114879A	0.87207	0.05856	-0.51351	0.05040
121030A	1.63768	0.13201	-1.26680	0.05051
121031A	0.91927	0.06114	-0.55952	0.04911
121032A	0.77726	0.06329	-1.27128	0.08720
121033B	0.32661	0.03797	-0.76744	0.13567

continued

IREF	*a*	SE (*a*)	*b*	SE (*b*)
121105A	1.19068	0.08333	-0.97103	0.04987
121106B	0.56413	0.04643	-0.61032	0.07655
121107A	1.47956	0.10569	-1.03112	0.04401
121108A	1.14517	0.09012	-1.29132	0.06653
121805A	0.37472	0.02045	-0.14819	0.04959
121890A	0.50290	0.02261	-0.35913	0.03954
121891A	0.76454	0.03062	-0.97953	0.03691
122082A	0.40358	0.03994	-0.34414	0.09391
122086A	0.39102	0.03988	0.99932	0.12924
124300A	0.51313	0.02448	-1.00695	0.05321

IREF	*a*	SE (*a*)	*b*	SE (*b*)
124302A	1.01366	0.04780	-1.59709	0.04992
124309A	0.57954	0.02371	-0.24888	0.03393
124311A	0.22201	0.01837	0.59219	0.09349
127781A	0.42133	0.02263	-1.05450	0.06533
127782A	0.64347	0.02900	-1.24392	0.05129
127783A	1.26834	0.05014	-1.17000	0.02843
127784A	0.26911	0.01917	-0.39824	0.07211
127785A	0.61521	0.02673	-0.97412	0.04441
127786B	1.11973	0.04586	-1.25474	0.03374

Table F-7. 2015–16 MSAA: IRT Parameters for Dichotomous Items—ELA Grade 11

IREF	*a*	SE (*a*)	*b*	SE (*b*)
116348A	1.76135	0.15689	-1.19121	0.04573
116349A	2.37845	0.18398	-0.90399	0.02705
116350A	2.20431	0.17778	-0.98694	0.03053
116351A	1.36446	0.12568	-1.32057	0.06409
117167A	0.81439	0.03054	-0.46245	0.02796
117168A	1.05217	0.03696	-0.53162	0.02310
117169B	0.15140	0.01794	1.52924	0.21536
119078A	1.16348	0.04084	-0.60865	0.02190
119079A	1.55274	0.05443	-0.68751	0.01798
119080A	1.18504	0.04351	-0.76428	0.02330
119081A	1.13483	0.04437	-0.93043	0.02705
120148B	1.31108	0.05985	-1.28146	0.03381
120149A	0.83716	0.03492	-0.98431	0.03627
120150A	0.75028	0.03142	-0.87452	0.03694
120151A	0.76387	0.03653	-1.40269	0.05402
121065A	0.17976	0.00000	2.58168	0.28776
121130A	0.40779	0.02189	-0.11153	0.04791

IREF	*a*	SE (*a*)	*b*	SE (*b*)
121227A	1.22237	0.05543	-1.28428	0.03549
121228A	1.19706	0.06107	-1.54617	0.04764
121229A	0.92079	0.03536	-0.76067	0.02882
121230A	1.33457	0.05553	-1.08622	0.02721
121953B	0.17387	0.00000	3.29016	0.37745
122000A	0.91744	0.03366	-0.56223	0.02620
124319A	0.39674	0.02415	-1.40048	0.08902
124328A	0.67261	0.02606	0.10750	0.03122
124334A	0.96101	0.03683	-0.78976	0.02827
126773A	1.26275	0.05246	-1.08765	0.02842
126774B	1.38524	0.05801	-1.09486	0.02667
126775A	0.56647	0.02698	-0.96473	0.05001
126776B	0.55060	0.02472	-0.39337	0.03859
126777B	1.05479	0.04382	-1.07986	0.03244
126778B	1.05820	0.04221	-0.96804	0.02951

Table F-8. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 3

IREF	*a*	SE (*a*)	*b*	SE (*b*)
110855A	0.77145	0.05629	0.18147	0.05586
110862A	0.48051	0.02450	-0.55273	0.04678
110865A	0.14503	0.02889	1.95790	0.46164
110876A	0.50493	0.04596	0.76195	0.09936
110923A	1.16630	0.04386	-0.75820	0.02425
110959A	0.38199	0.02434	-1.37179	0.09222
110964A	0.73265	0.06203	-0.86513	0.07401
110974A	0.86953	0.03121	0.13648	0.02566
110975A	0.57699	0.04919	-0.03715	0.06954
111377A	0.64090	0.02766	-0.45480	0.03454
111382A	0.64148	0.05448	-0.56364	0.06993
111386A	0.80352	0.03011	-0.06147	0.02657
111391A	0.44117	0.02314	-0.19302	0.04510
111397A	0.43435	0.02306	-0.21925	0.04596
111411A	0.33362	0.02115	0.17375	0.05875
111416A	0.53928	0.02600	-0.64590	0.04406
111420A	0.60666	0.02561	0.44087	0.03755
111422A	0.39827	0.02207	0.54026	0.05625
111424A	0.78260	0.03489	-1.09165	0.04266
111434A	0.79824	0.03557	-1.11301	0.04269
111883A	0.44624	0.04405	0.90805	0.11872
112552A	0.57381	0.05038	-0.38238	0.07207
112555A	0.86430	0.06992	-0.88349	0.06524
112564A	0.59296	0.05416	-0.82708	0.08668
112565A	0.59326	0.02542	0.62487	0.04138
112571A	0.97016	0.06704	0.19846	0.04835
112575A	0.63439	0.05270	-0.33508	0.06501
112595A	1.05800	0.07353	-0.28452	0.04249
112601A	1.01309	0.03494	-0.04588	0.02221
112615B	0.29887	0.03847	0.23880	0.12911
112616B	0.40801	0.02274	1.25282	0.07867
112621B	0.35449	0.04006	0.26517	0.11251
120682A	0.75995	0.03001	-0.35661	0.02880
122091A	1.45434	0.04654	0.02949	0.01739
122105A	1.59082	0.09980	-0.15787	0.03118

Table F-9. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 4

IREF	*a*	SE (*a*)	*b*	SE (*b*)
111135A	0.38126	0.04231	-0.55859	0.11251
111136A	0.67436	0.05570	-0.22280	0.06204
111160A	0.33895	0.02151	-0.53318	0.06320
111166A	0.62587	0.02609	0.00874	0.03193
111178B	0.35738	0.04153	-0.65864	0.12491
111663A	0.84323	0.03503	-1.07198	0.03904
111667A	0.75174	0.03467	-1.40203	0.05428
111676A	0.60992	0.02568	0.23576	0.03387
111677A	0.93791	0.06428	-0.00288	0.04561
111682A	0.62112	0.02592	0.13554	0.03262
111686A	0.40392	0.02214	0.47607	0.05287
111688A	0.35378	0.04052	0.37162	0.11346
111695B	0.52067	0.02740	-1.37958	0.07062
111705B	0.35730	0.04092	0.21224	0.10814
111712B	0.37429	0.02169	0.54743	0.05843
111716A	0.61381	0.05029	-0.00609	0.06456
111717A	0.52585	0.02477	-0.44075	0.04048
111721A	0.44328	0.04430	0.80890	0.11227
111727A	0.71731	0.05493	0.36297	0.06164
111731B	0.68560	0.05316	0.01151	0.05882

continued

IREF	*a*	SE (*a*)	*b*	SE (*b*)
112782A	0.90734	0.06276	0.08608	0.04724
112797B	0.17767	0.01868	1.18854	0.15958
112801A	0.39847	0.04569	-1.33040	0.16525
112803B	0.61618	0.05073	-0.19702	0.06513
112813A	0.22734	0.02063	2.35965	0.22057
112832B	0.33086	0.02105	0.61285	0.06742
120697A	0.90013	0.06567	-0.62311	0.05480

IREF	*a*	SE (*a*)	*b*	SE (*b*)
121665A	0.24893	0.03649	1.17199	0.21981
121691B	0.40225	0.04275	0.82661	0.12329
122265A	0.41773	0.04913	2.26323	0.25281
122267A	1.22141	0.03992	-0.11220	0.01871
122394A	1.00656	0.07024	0.79916	0.05757
122426A	0.81752	0.03165	1.06891	0.03970

Table F-10. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 5

IREF	*a*	SE (*a*)	*b*	SE (*b*)
111234A	0.23946	0.01986	-0.80545	0.09927
111244B	0.82749	0.02993	0.27387	0.02643
111257A	0.56343	0.05561	-1.46844	0.13803
111260A	0.18116	0.01854	-0.18387	0.10214
111276A	0.68309	0.05225	0.11839	0.05817
111291A	0.78636	0.03223	-0.92281	0.03691
111299A	0.43599	0.04553	1.21020	0.14185
111303A	0.42299	0.02256	-0.29704	0.04656
112340A	0.58516	0.02691	-0.86233	0.04551
112342A	0.39842	0.04215	0.24085	0.09538
112346A	0.83233	0.06254	-0.70097	0.05846
112358A	0.63557	0.05193	0.79572	0.07995
112363A	0.44769	0.02307	0.95389	0.06082
112372A	1.17641	0.08206	-0.78777	0.04676
112373B	0.34364	0.02118	-0.17377	0.05500
112374B	0.44919	0.04522	-0.07096	0.08616
112378A	0.86926	0.06092	-0.17278	0.04674
112385A	0.39688	0.02187	0.39530	0.05160

IREF	*a*	SE (*a*)	*b*	SE (*b*)
112409A	0.72492	0.02787	0.47581	0.03171
112416A	0.21989	0.01926	-0.21820	0.08512
113838A	0.50351	0.02362	0.39892	0.04178
113843B	0.52019	0.04575	0.28230	0.07660
113844A	0.83570	0.03537	-1.17618	0.04215
113852A	0.82684	0.05852	0.00710	0.04884
113856A	0.29965	0.04025	1.78340	0.25258
113862A	0.36028	0.02213	1.42771	0.09536
113867A	0.35396	0.04065	1.37542	0.18012
113877A	0.16013	0.01920	3.30415	0.40396
113883A	0.19236	0.01872	0.14679	0.09586
113884B	0.42293	0.04372	-0.44019	0.09696
113889A	0.46079	0.02342	-0.45628	0.04548
113892A	0.65872	0.05121	0.17427	0.06074
113901A	0.30210	0.02190	2.04107	0.15094
120737A	0.85653	0.03223	1.01830	0.03698
121514A	1.19766	0.07952	0.77438	0.05055

Table F-11. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 6

IREF	*a*	SE (*a*)	*b*	SE (*b*)
110891A	0.69789	0.06569	-1.37672	0.11329
110909A	0.54414	0.02834	-1.36486	0.06728
110910A	0.42152	0.02190	0.11911	0.04528
110938A	0.49547	0.02323	-0.00247	0.03867
110943A	0.42995	0.02243	1.28196	0.07393
110978A	0.54740	0.02397	0.16787	0.03629
110986A	0.89062	0.06349	-0.26475	0.04767
110996A	0.76772	0.02826	0.12194	0.02717
111027A	0.30906	0.04011	1.54374	0.22114
111444A	0.85717	0.06981	-0.94886	0.07035
111450A	1.02217	0.08205	-1.02435	0.06521
111457A	0.49029	0.02290	0.69876	0.04877
111461A	0.38561	0.02133	0.12769	0.04916
111478A	0.85259	0.03215	-0.48417	0.02624
111482A	1.36299	0.09369	-0.58544	0.03795
111487A	0.38634	0.02203	-0.55068	0.05525
111492A	0.37212	0.04166	-0.18800	0.10189
111496A	0.44560	0.02571	-1.45493	0.08322
111507A	0.69655	0.02711	-0.04762	0.02875
111514A	0.75389	0.02920	-0.34994	0.02784
111517A	1.14380	0.07711	-0.35807	0.03980
111518A	0.53582	0.02407	-0.09284	0.03605
111632A	0.15522	0.01784	1.89064	0.24223
112645A	0.75935	0.02813	0.09922	0.02728
112654A	0.70948	0.05592	-0.36569	0.05938
112657A	0.91626	0.06668	-0.47225	0.04964
112658A	0.42521	0.02224	-0.18820	0.04491
112663A	0.49236	0.04463	0.36975	0.08300
112667A	0.25008	0.02014	2.08240	0.17521
112676A	0.40792	0.02338	-1.01419	0.06792
112678A	0.85576	0.06082	-0.10178	0.04842
112699A	0.44698	0.04312	0.33926	0.08930
120855A	0.66741	0.02720	-0.32404	0.03074
121487A	0.92687	0.06376	-0.05543	0.04547
121520A	0.23168	0.03672	2.11409	0.35713

Table F-12. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 7

IREF	*a*	SE (*a*)	*b*	SE (*b*)
111046A	0.85514	0.04485	-1.72353	0.06795
111055A	0.53971	0.04757	0.11637	0.07280
111059B	0.61979	0.05064	0.07108	0.06497
111076A	0.51259	0.02349	0.40028	0.04156
111080A	0.57027	0.02466	0.21891	0.03572
111093A	0.62625	0.05097	0.09625	0.06388
111098A	0.50884	0.05352	-1.32957	0.13701
111101A	0.70002	0.02785	-0.12099	0.02862
111113A	0.60880	0.02521	0.43387	0.03651
111127A	0.64027	0.02601	0.20406	0.03230
111131A	0.34122	0.02080	0.97748	0.07799
111734A	0.96718	0.06968	-0.42072	0.04583
111735A	0.50885	0.02356	0.20504	0.03921
111744A	0.67281	0.05754	-0.71709	0.07253
111748A	0.43420	0.02585	-1.44108	0.08540
111749A	0.83575	0.03473	-0.83956	0.03304
111754B	0.58418	0.02513	0.05338	0.03377
111758A	0.42382	0.04838	-1.33976	0.15854

continued

IREF	*a*	SE (*a*)	*b*	SE (*b*)
111765A	0.63889	0.02580	0.39442	0.03456
111766A	0.56934	0.04881	0.06820	0.06895
111769A	0.64803	0.02846	-0.63708	0.03604
111778A	1.09040	0.04768	-1.20865	0.03699
111779A	0.91429	0.06545	-0.25916	0.04617
111804A	0.49766	0.02360	-0.00506	0.03862
112853B	0.58344	0.04916	0.15464	0.06864
112870A	0.40414	0.04153	0.63295	0.11223

IREF	*a*	SE (*a*)	*b*	SE (*b*)
112871A	0.38416	0.02158	0.05947	0.04909
112882A	0.72351	0.05459	0.22071	0.05853
112886A	0.44139	0.02226	0.55706	0.05048
112889A	0.42497	0.02354	-0.74641	0.05616
112890A	0.54639	0.04750	0.68431	0.08894
112899A	0.52824	0.05065	-0.77435	0.09273
112901A	0.30496	0.03884	1.01264	0.17267
112911A	0.53918	0.04722	0.36459	0.07820

Table F-13. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 8

IREF	*a*	SE (*a*)	*b*	SE (*b*)
111247A	0.89734	0.06542	-0.84746	0.06020
111286A	0.66017	0.05030	0.28598	0.06065
111336A	0.54113	0.02380	0.47409	0.03983
111339A	0.64601	0.04994	0.33297	0.06293
111581A	0.40410	0.02213	-0.61855	0.05394
111583A	0.44982	0.02239	0.87483	0.05692
111594A	0.57692	0.02455	0.08912	0.03350
111622A	0.64221	0.02580	0.12054	0.03073
112464A	0.60663	0.05270	-0.85739	0.08689
112466A	0.50960	0.02352	-0.16832	0.03733
112470A	0.43390	0.02283	-0.69654	0.05262
112476A	0.18838	0.01811	0.02500	0.09477
112480A	0.65462	0.02597	0.26854	0.03141
112481A	0.50821	0.04498	0.63729	0.08800
112484A	0.47606	0.02532	-1.32948	0.07119
112495A	0.86378	0.05918	-0.03222	0.04599
112506A	0.52685	0.04798	-0.08416	0.07399
112515B	0.49138	0.02291	0.45609	0.04294

IREF	*a*	SE (*a*)	*b*	SE (*b*)
113908A	0.88142	0.03164	-0.27566	0.02370
113918A	0.87765	0.06326	-0.72251	0.05670
113922A	0.85369	0.05876	-0.04143	0.04641
113926A	0.57338	0.02525	-0.46976	0.03663
113931A	0.47632	0.04401	0.77473	0.09978
113934A	0.51722	0.05257	-1.81069	0.16864
113938A	0.53573	0.02369	0.43073	0.03945
113942A	0.73339	0.05360	-0.04378	0.05260
113952A	0.56850	0.04701	0.44674	0.07326
113968A	0.95589	0.06337	-0.05423	0.04240
113972A	0.26253	0.01979	1.56261	0.13127
113978A	0.64770	0.02628	-0.19390	0.03030
117072A	0.77650	0.05547	0.12455	0.05125
120569B	0.29007	0.03716	1.13143	0.18412
122050A	0.66074	0.02619	0.10128	0.02990
122051A	0.75202	0.05422	0.12279	0.05236
122099A	0.68859	0.02670	0.16941	0.02929

Table F-14. 2015–16 MSAA: IRT Parameters for Dichotomous Items—Mathematics Grade 11

IREF	*a*	SE (*a*)	*b*	SE (*b*)
110847A	0.52335	0.04955	0.24490	0.08237
110851A	0.50175	0.02538	0.48296	0.04666
110914A	0.85590	0.06558	0.22509	0.05511
110921A	0.63123	0.02843	0.12770	0.03393
110936A	0.76043	0.06417	-0.45160	0.05968
110969B	0.57199	0.05429	-0.07634	0.07329
111016A	0.59160	0.05266	0.22855	0.07392
111032A	0.72595	0.05860	0.45005	0.06870
111109B	0.34785	0.02250	0.60611	0.06801
111533A	0.66821	0.03444	-1.23317	0.05792
111537A	0.62223	0.05367	0.67081	0.08601
111538A	0.89549	0.03437	0.41712	0.02865
111544A	0.58151	0.05189	0.76643	0.09515
111554B	0.88687	0.06738	0.17084	0.05259
111557A	0.73230	0.03059	0.24271	0.03106
111810A	0.98683	0.07288	0.13556	0.04794
111817A	0.29975	0.02168	0.33581	0.06977
111824A	0.53831	0.02646	0.07184	0.03851
111830A	0.45131	0.02685	-1.07160	0.07015
111832B	0.48840	0.05581	-1.64894	0.17541
112703B	0.63842	0.02824	0.47670	0.03816
112708A	0.79502	0.03242	0.07294	0.02770
112717A	0.86976	0.03414	0.13464	0.02620
112722A	1.09849	0.08216	-0.28227	0.04176
112726A	0.59978	0.03176	-1.16694	0.05975
112921B	0.24063	0.02114	1.80295	0.17168
112931A	0.45136	0.02438	0.64101	0.05502
112936A	0.56124	0.02661	0.43261	0.04155
112941A	0.61304	0.05425	-0.00583	0.06794
112944A	0.71153	0.06868	-1.29627	0.10800
112945A	0.44951	0.02507	-0.29705	0.04662
112946A	0.57911	0.02757	-0.06244	0.03574
122054A	0.21120	0.02039	1.67276	0.18094
122055A	0.36278	0.04255	0.83696	0.14525

APPENDIX G—TEST CHARACTERISTIC CURVES & TEST INFORMATION FUNCTIONS

Figure G-1. Test Characteristic Curve for Grade 3 Mathematics

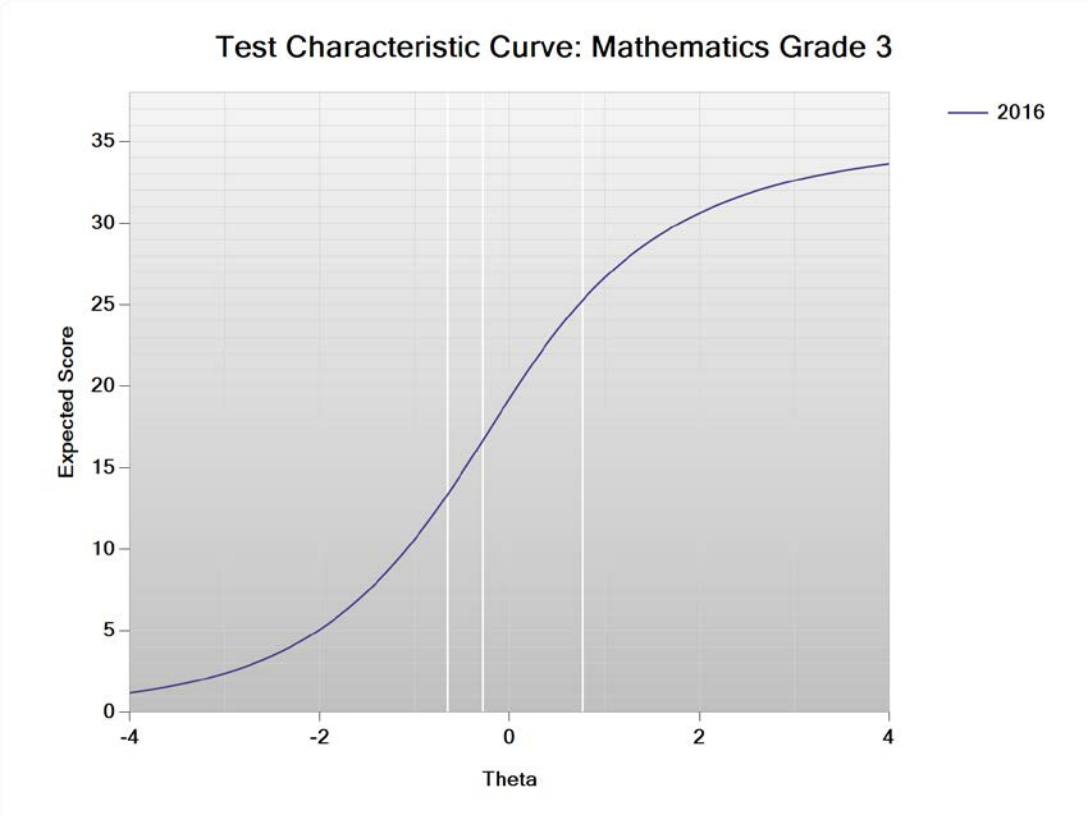


Figure G-2. Test Information Function and Standard Error for Grade 3 Mathematics

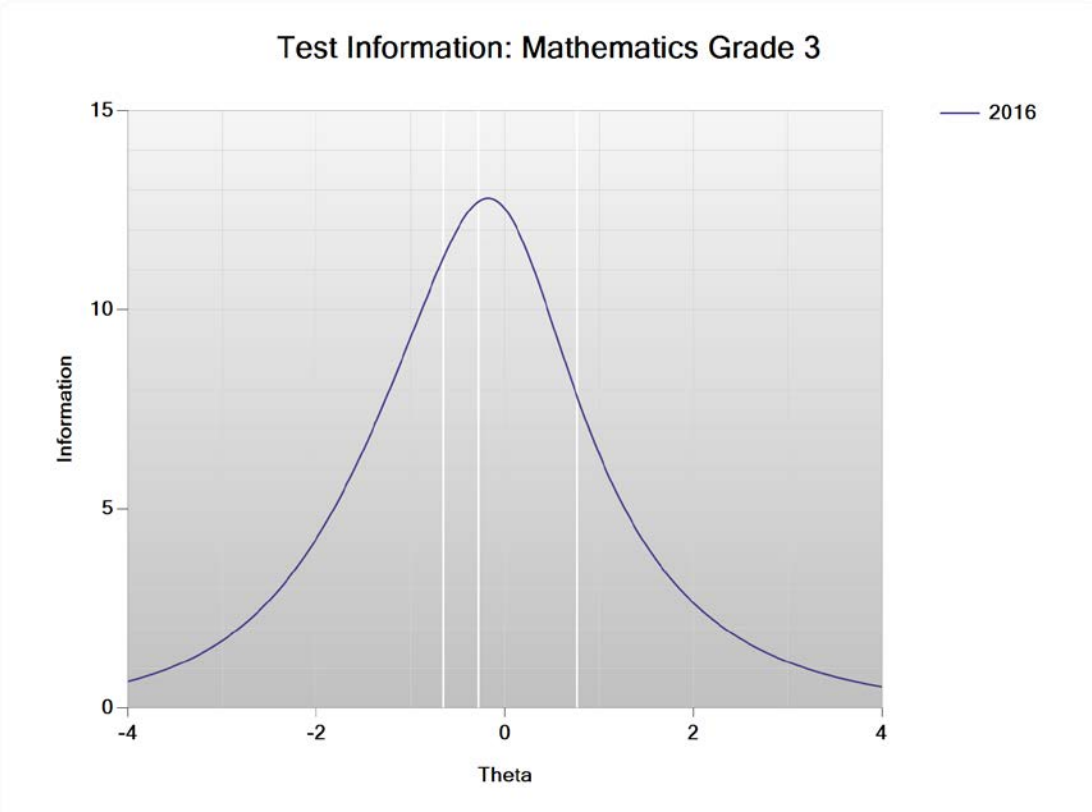


Figure G-3. Test Characteristic Curve for Grade 4 Mathematics

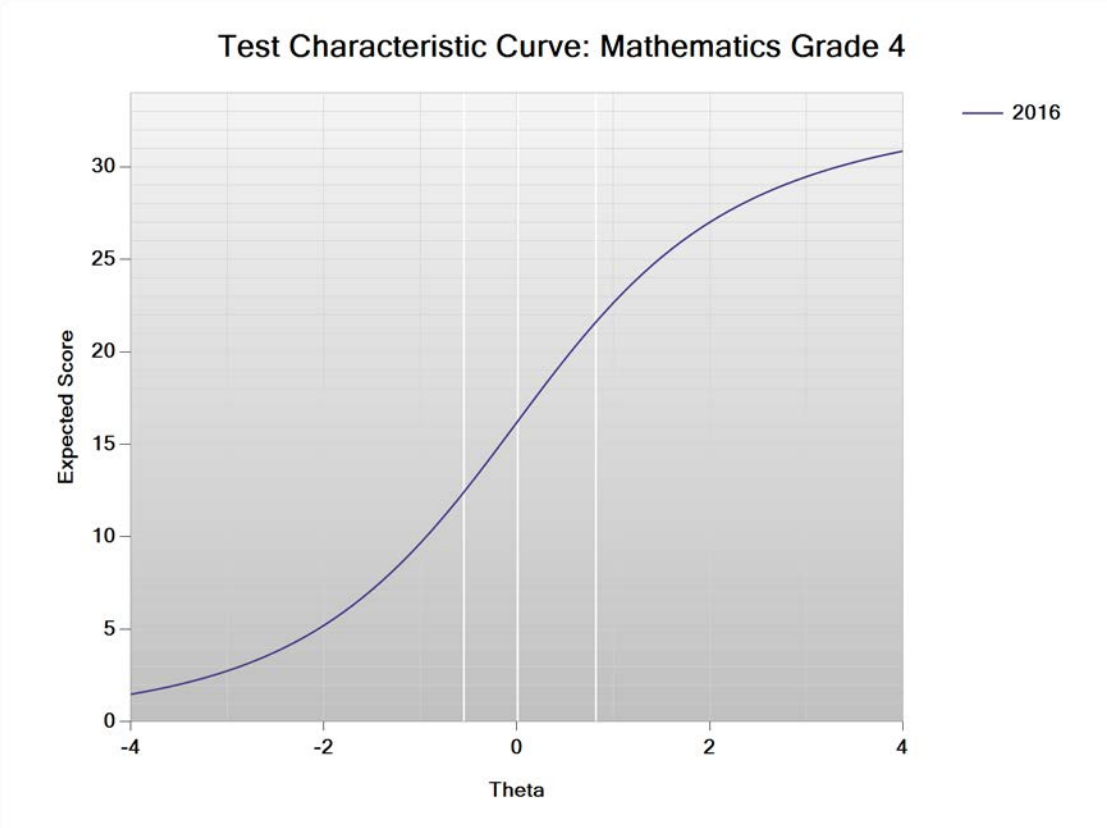


Figure G-4. Test Information Function and Standard Error for Grade 4 Mathematics

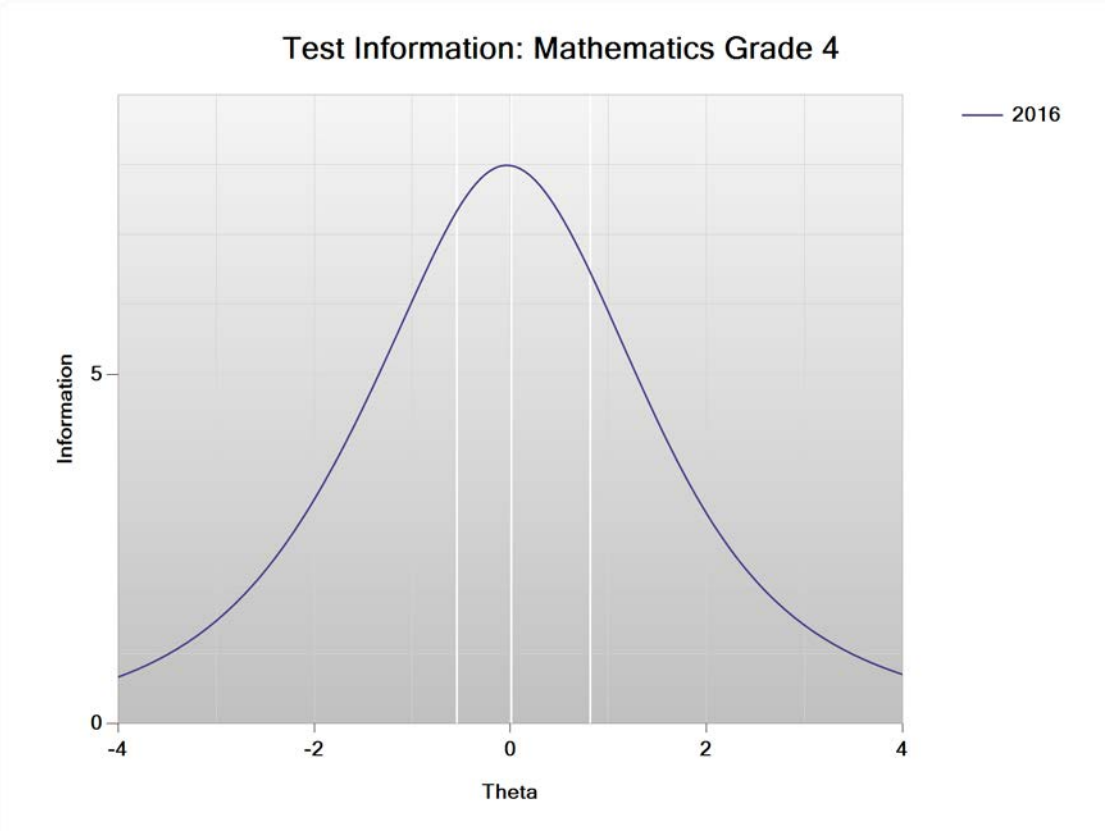


Figure G-5. Test Characteristic Curve for Grade 5 Mathematics

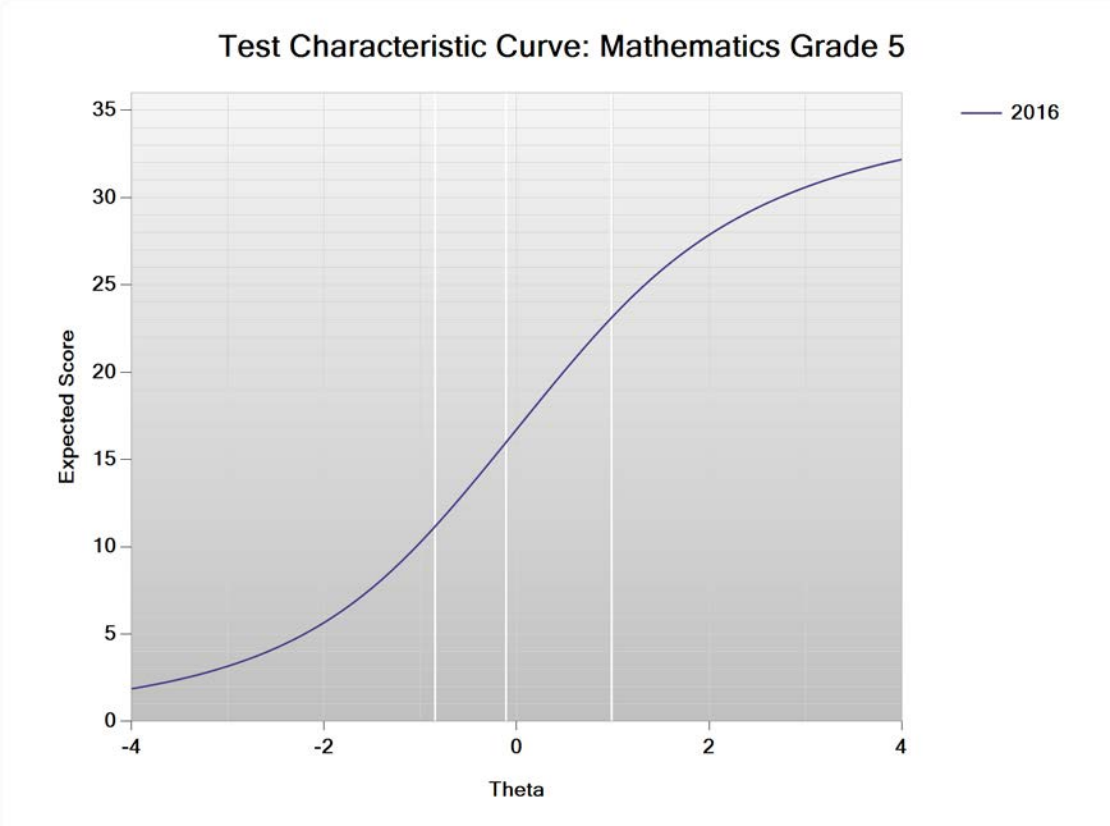


Figure G-6. Test Information Function and Standard Error for Grade 5 Mathematics

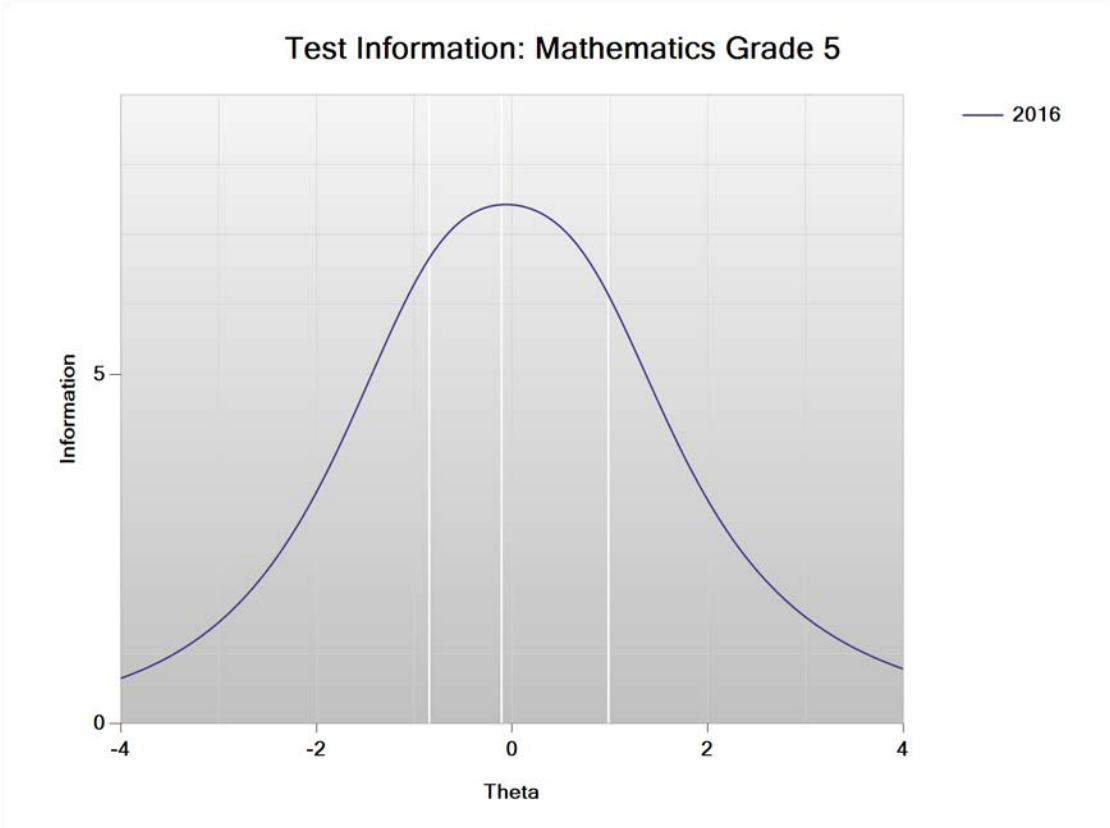


Figure G-7. Test Characteristic Curve for Grade 6 Mathematics

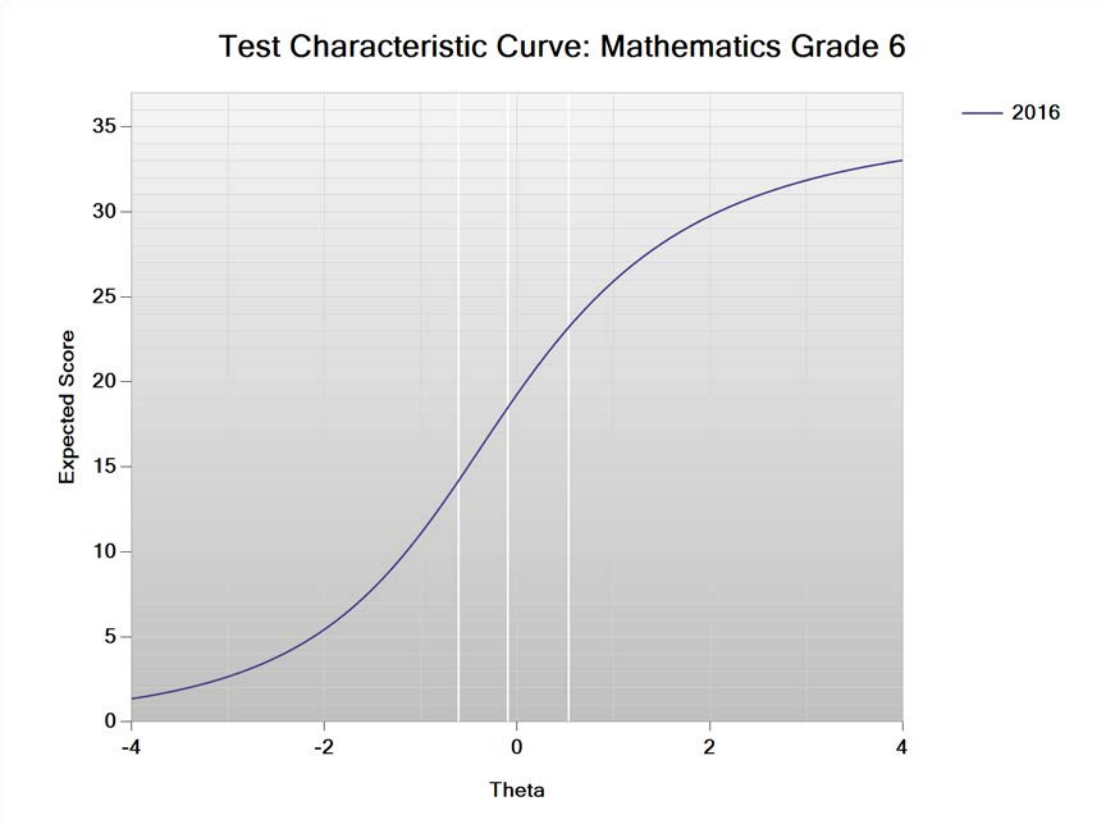


Figure G-8. Test Information Function and Standard Error for Grade 6 Mathematics

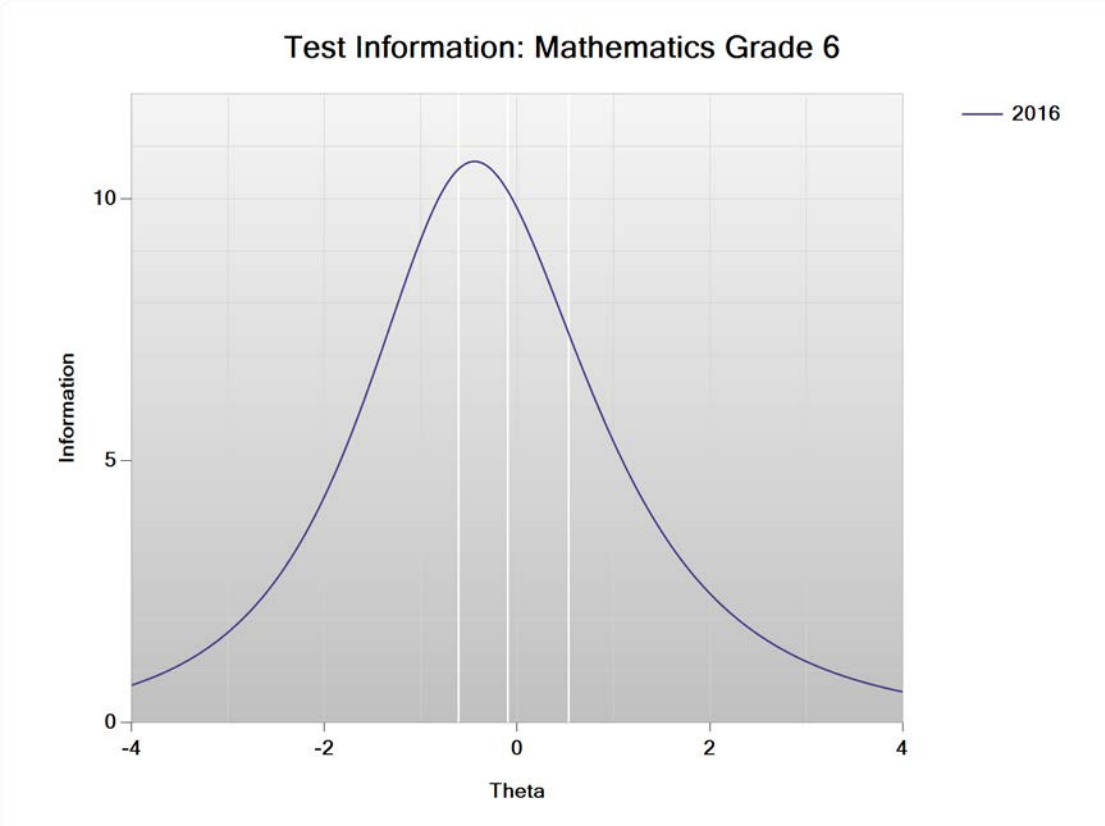


Figure G-9. Test Characteristic Curve for Grade 7 Mathematics

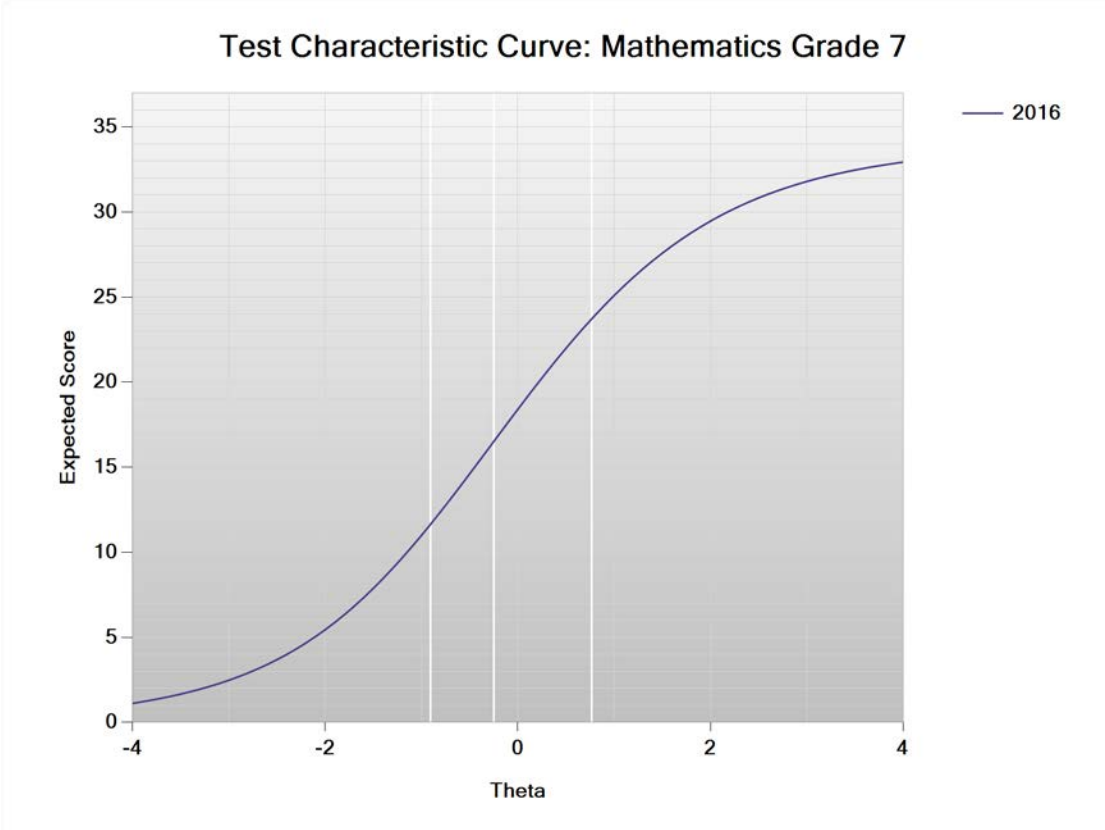


Figure G-10. Test Information Function and Standard Error for Grade 7 Mathematics

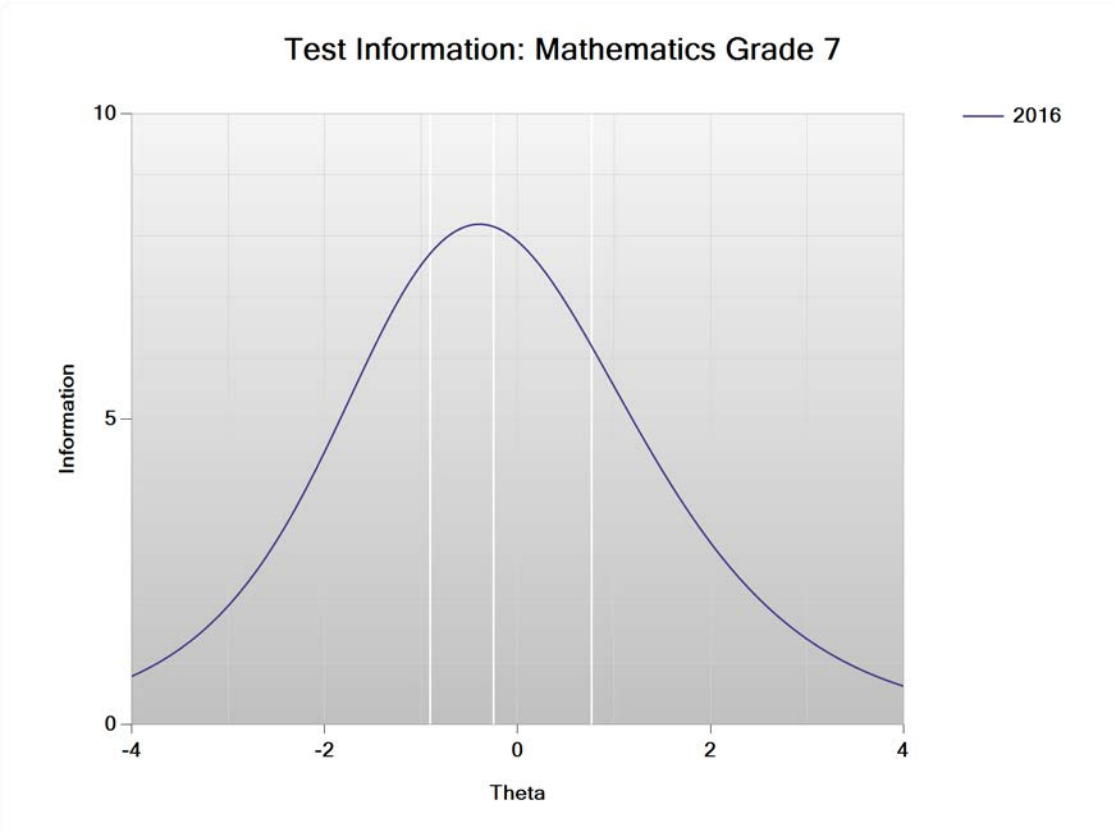


Figure G-11. Test Characteristic Curve for Grade 8 Mathematics

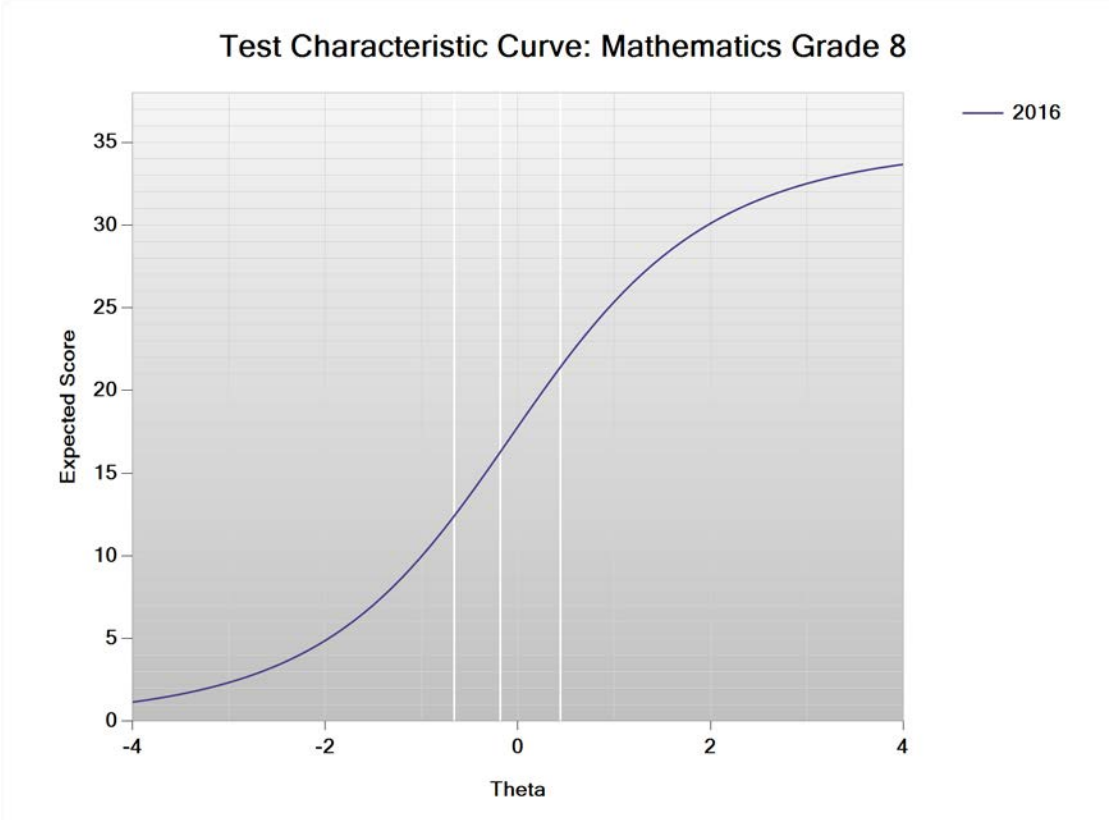


Figure G-12. Test Information Function and Standard Error for Grade 8 Mathematics

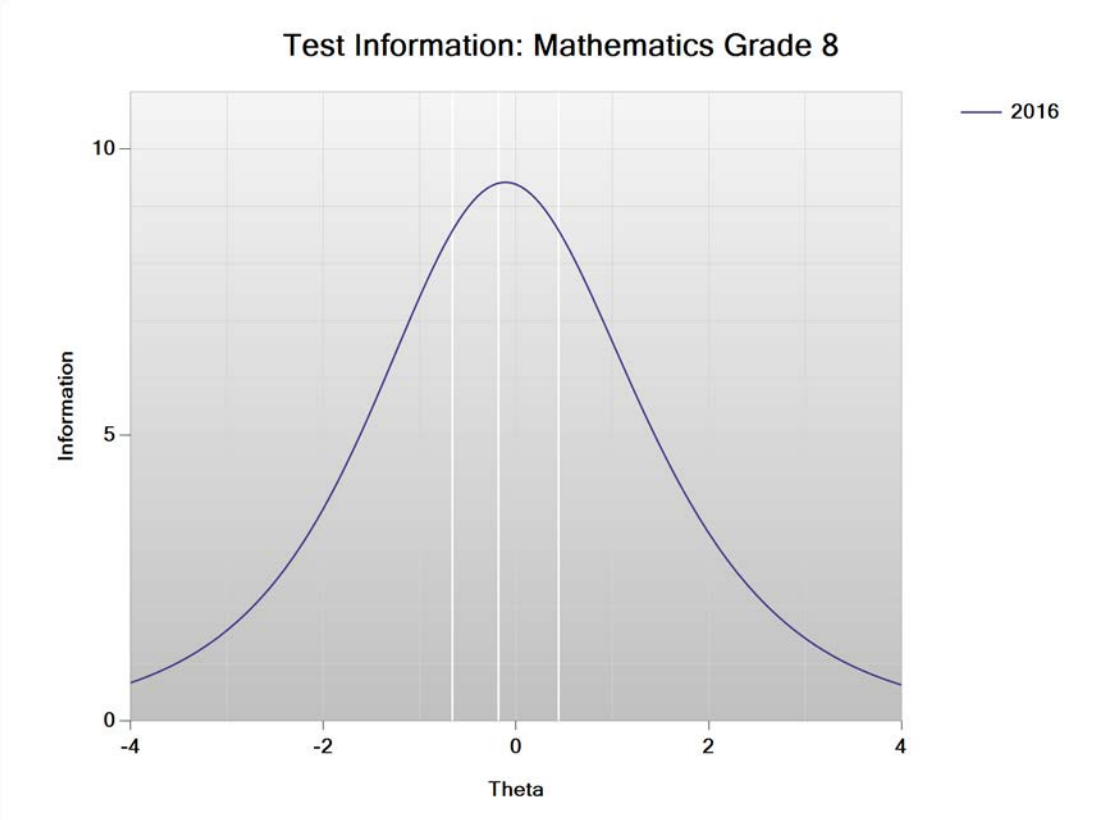


Figure G-13. Test Characteristic Curve for Grade 11 Mathematics

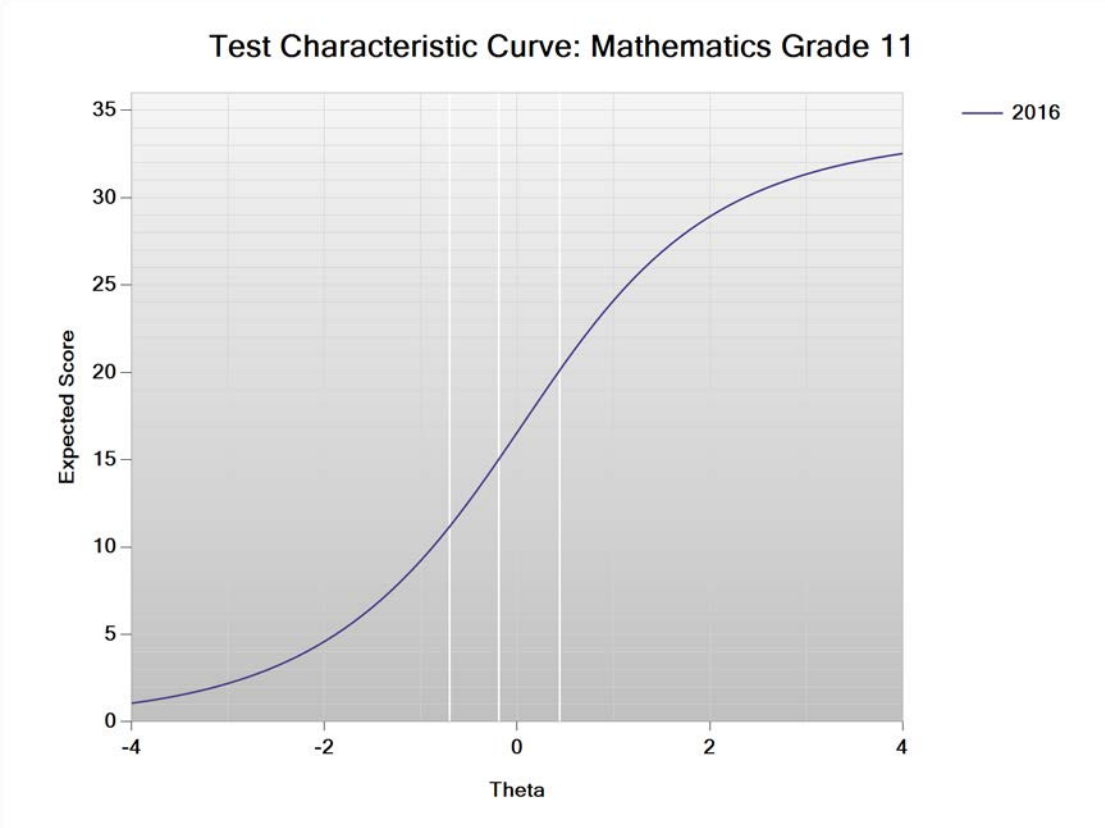


Figure G-14. Test Information Function and Standard Error for Grade 11 Mathematics

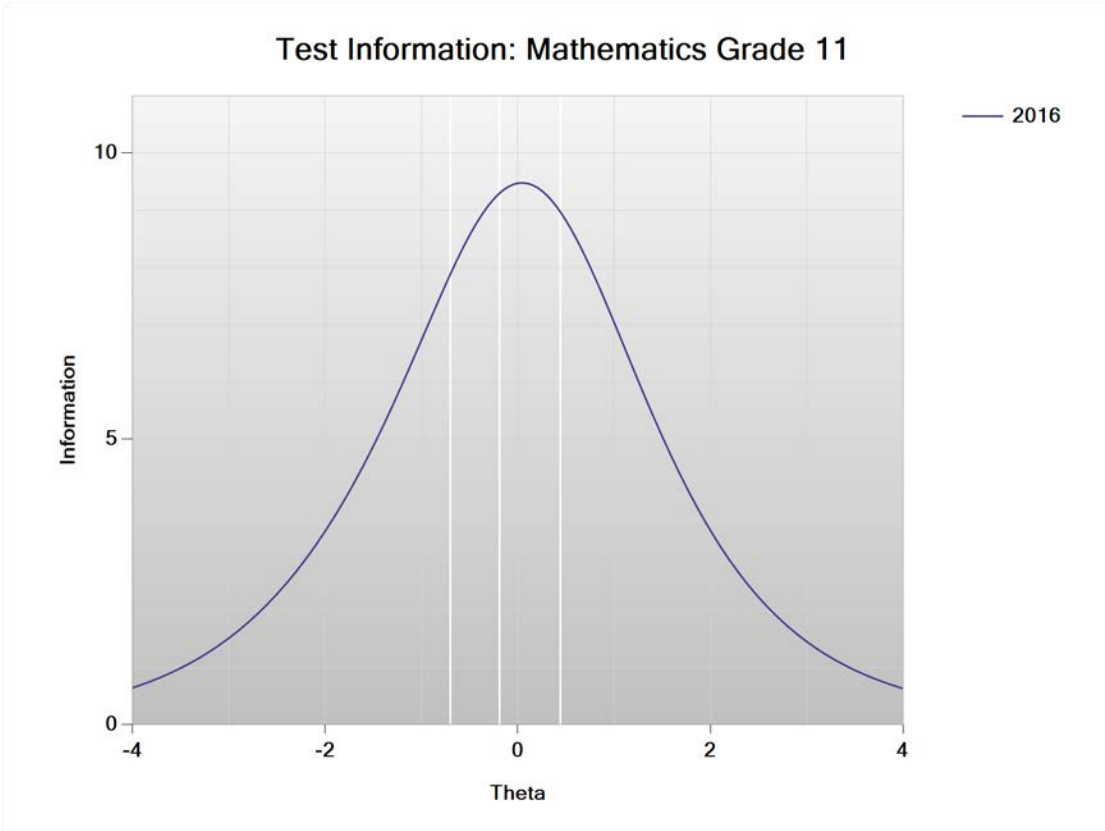


Figure G-15. Test Characteristic Curve for Grade 3 ELA

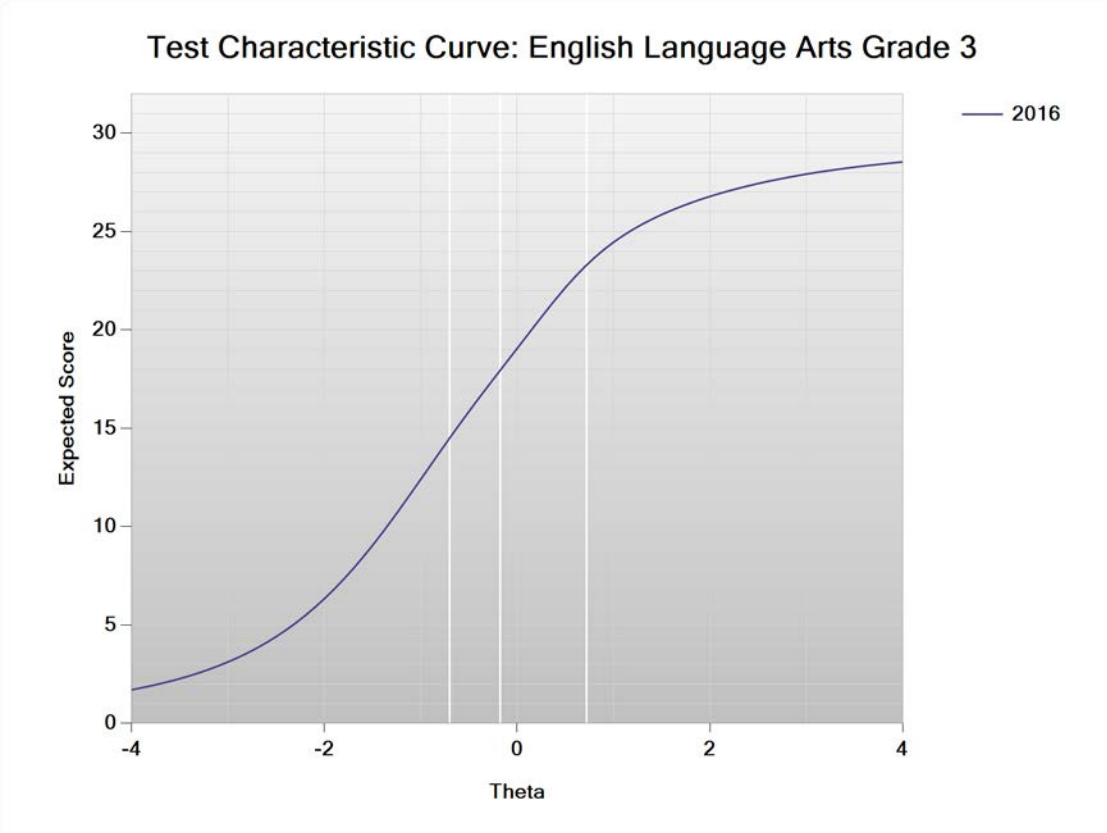


Figure G-16. Test Information Function and Standard Error for Grade 3 ELA

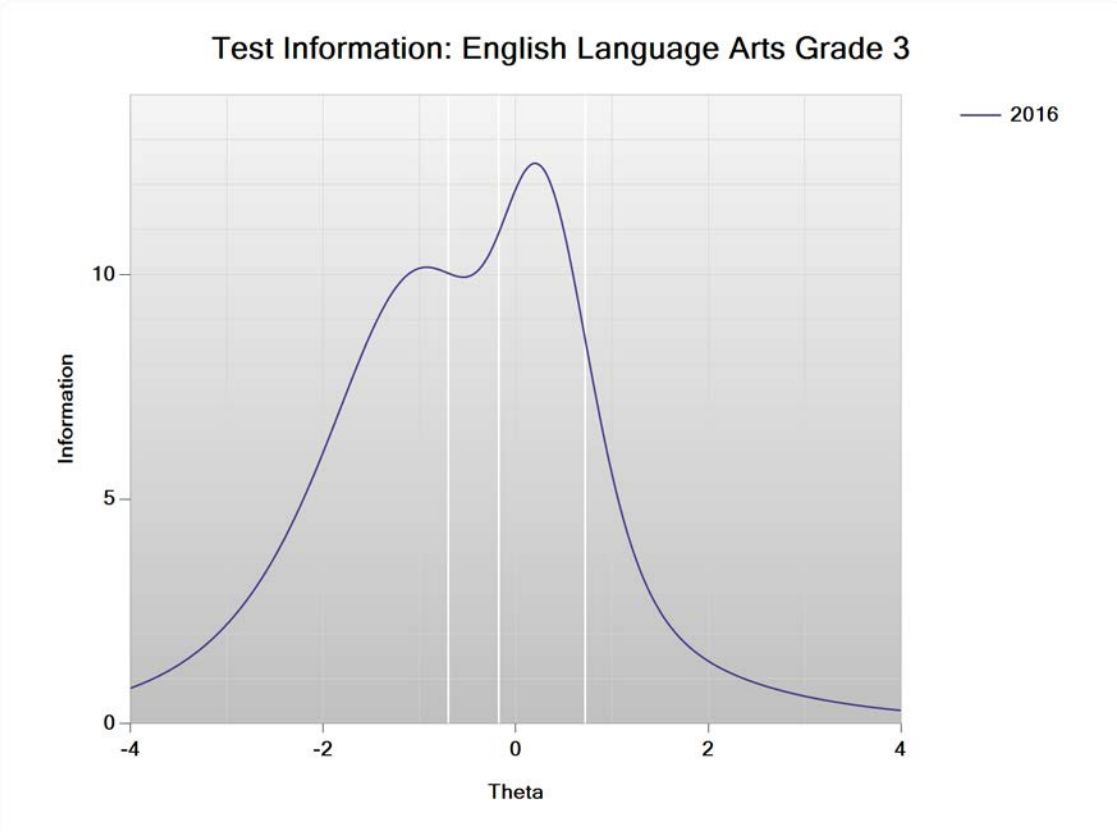


Figure G-17. Test Characteristic Curve for Grade 4 ELA

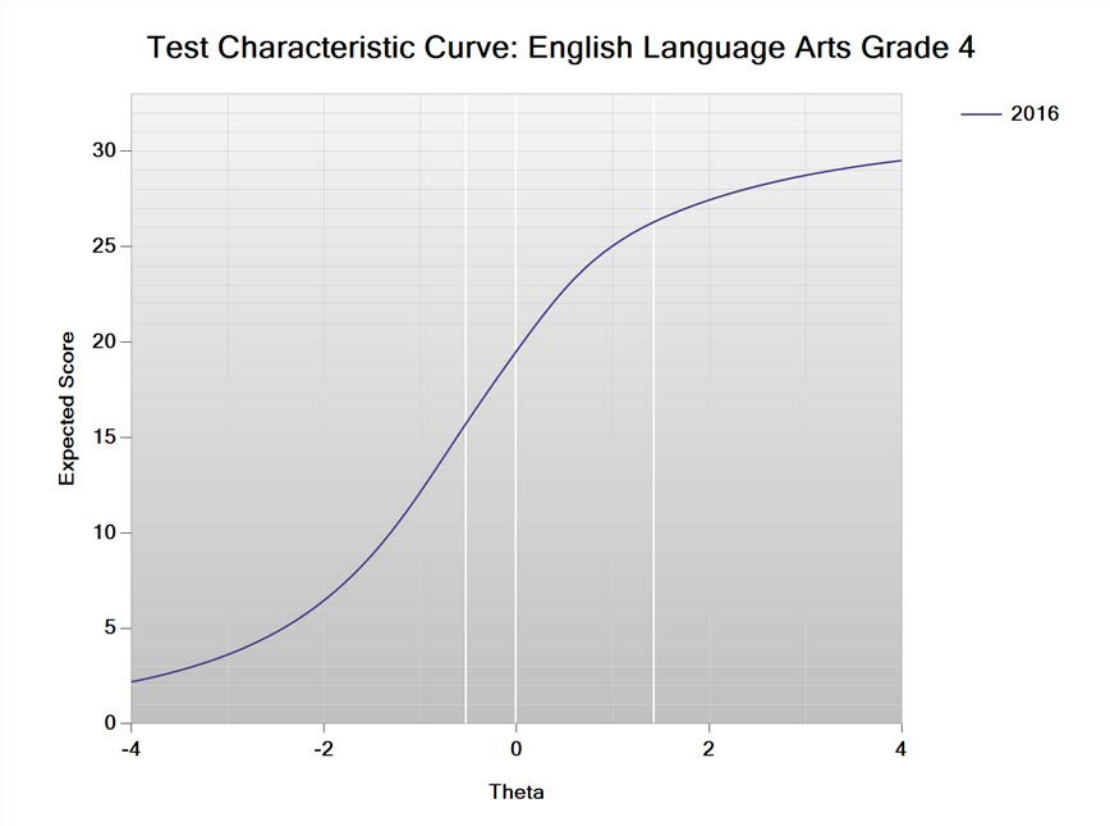


Figure G-18. Test Information Function and Standard Error for Grade 4 ELA

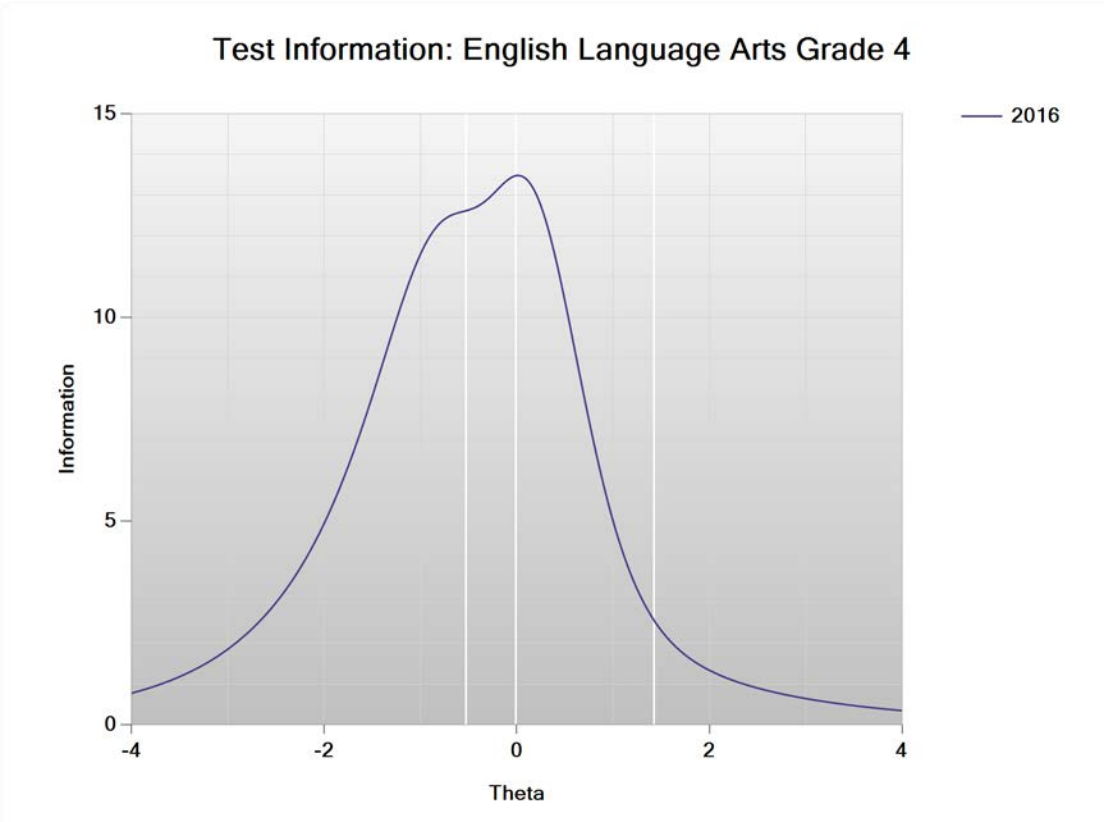


Figure G-19. Test Characteristic Curve for Grade 5 ELA

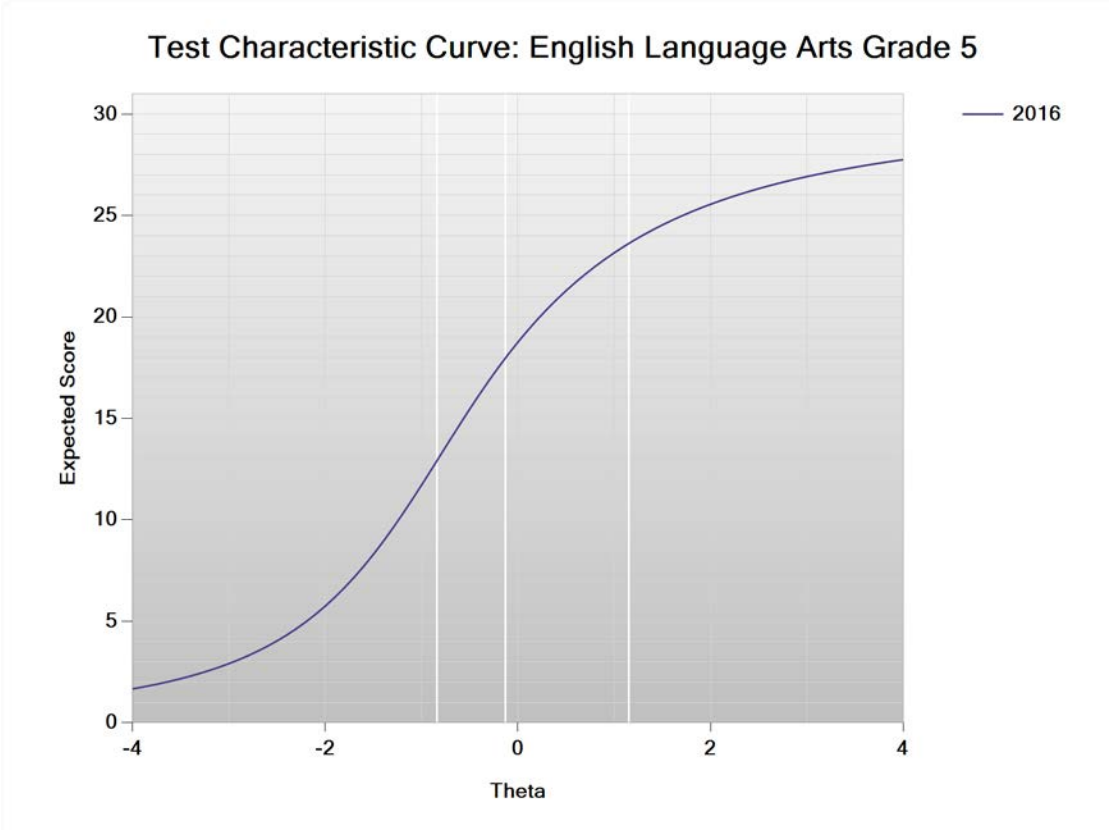


Figure G-20. Test Information Function and Standard Error for Grade 5 ELA

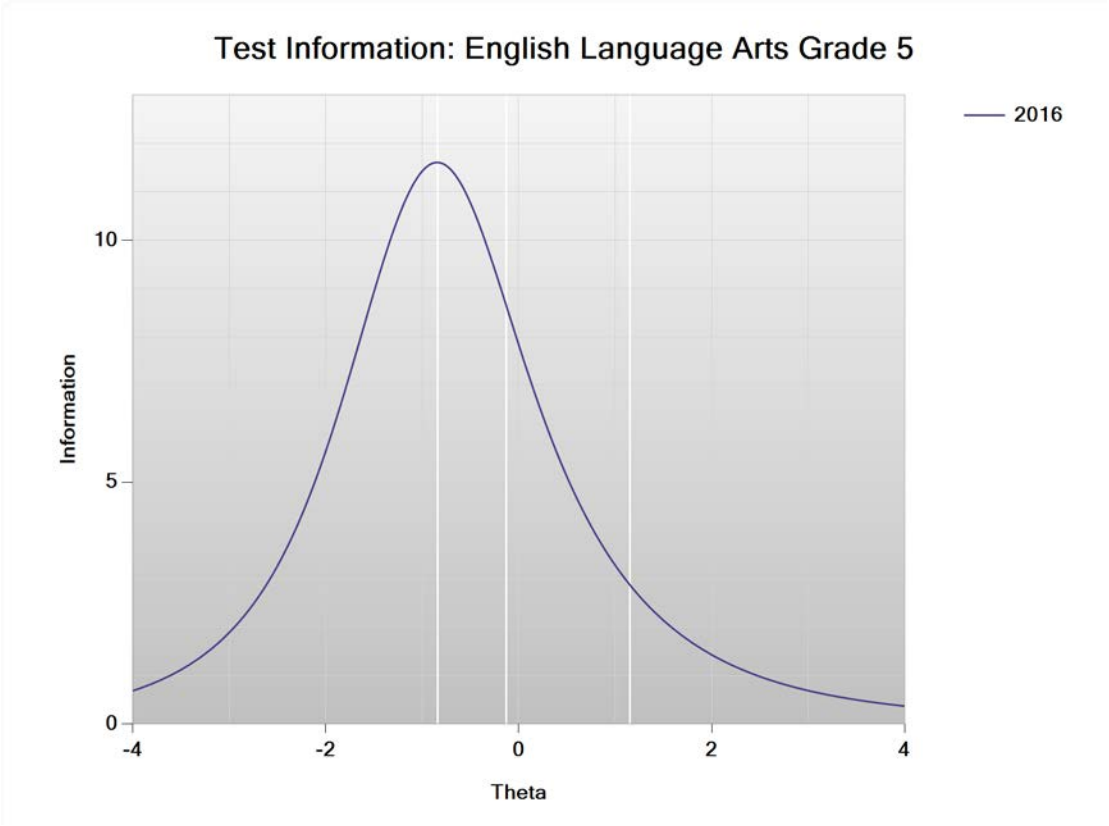


Figure G-21. Test Characteristic Curve for Grade 6 ELA

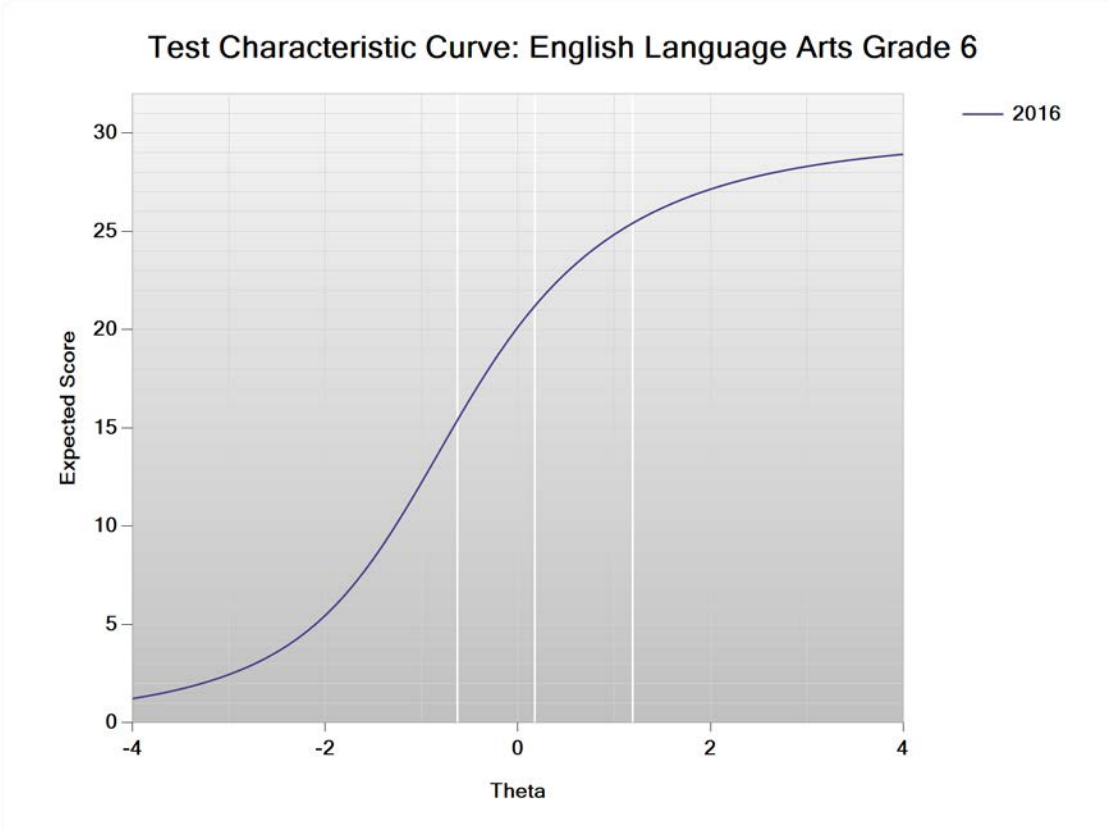


Figure G-22. Test Information Function and Standard Error for Grade 6 ELA

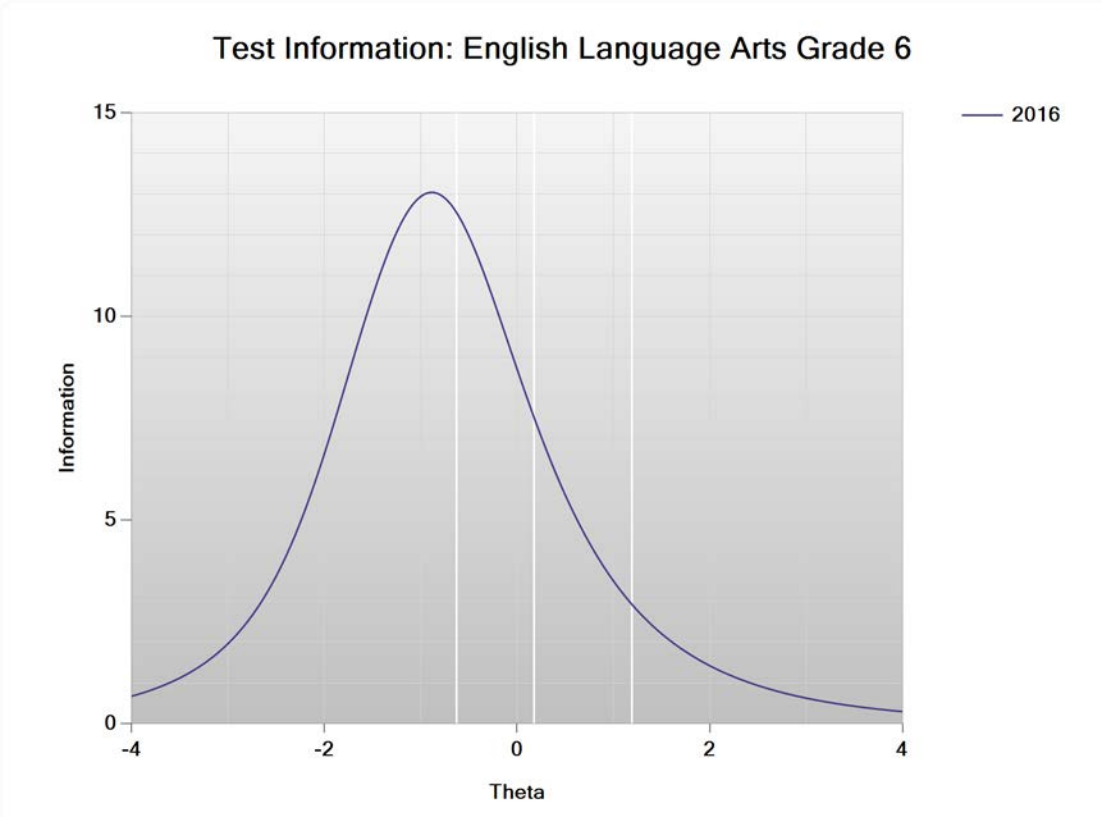


Figure G-23. Test Characteristic Curve for Grade 7 ELA

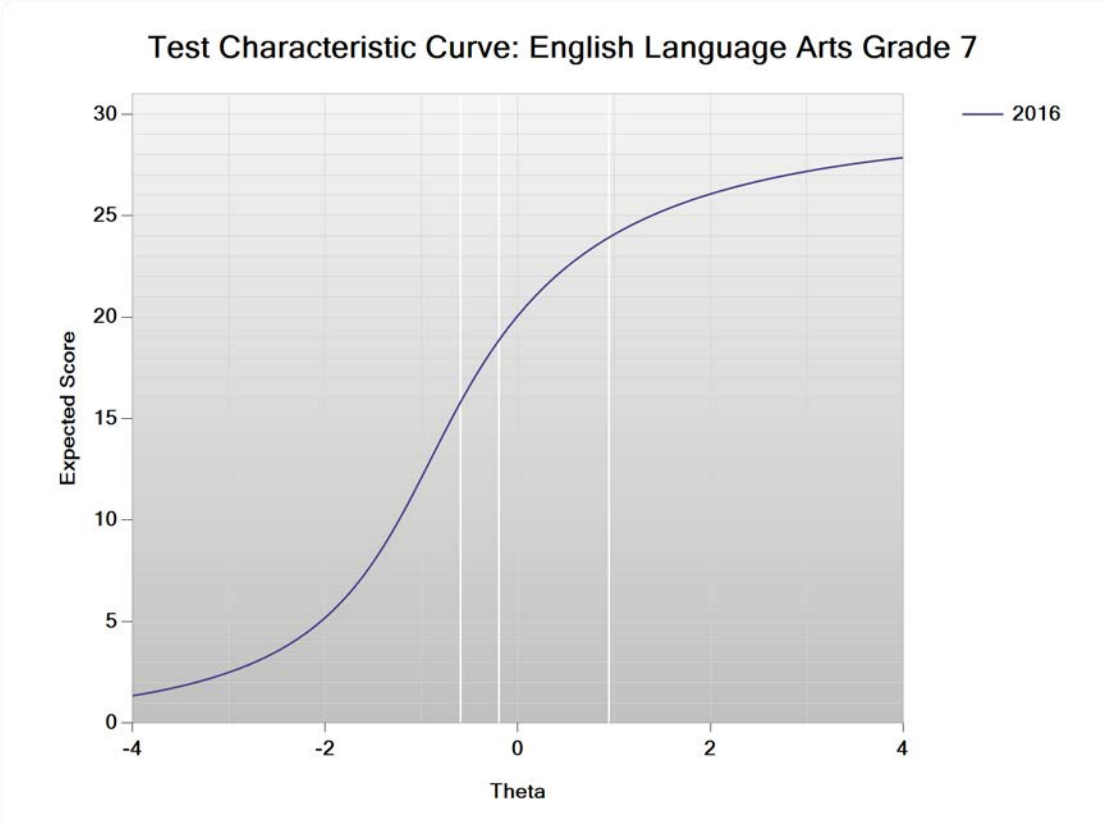


Figure G-24. Test Information Function and Standard Error for Grade 7 ELA

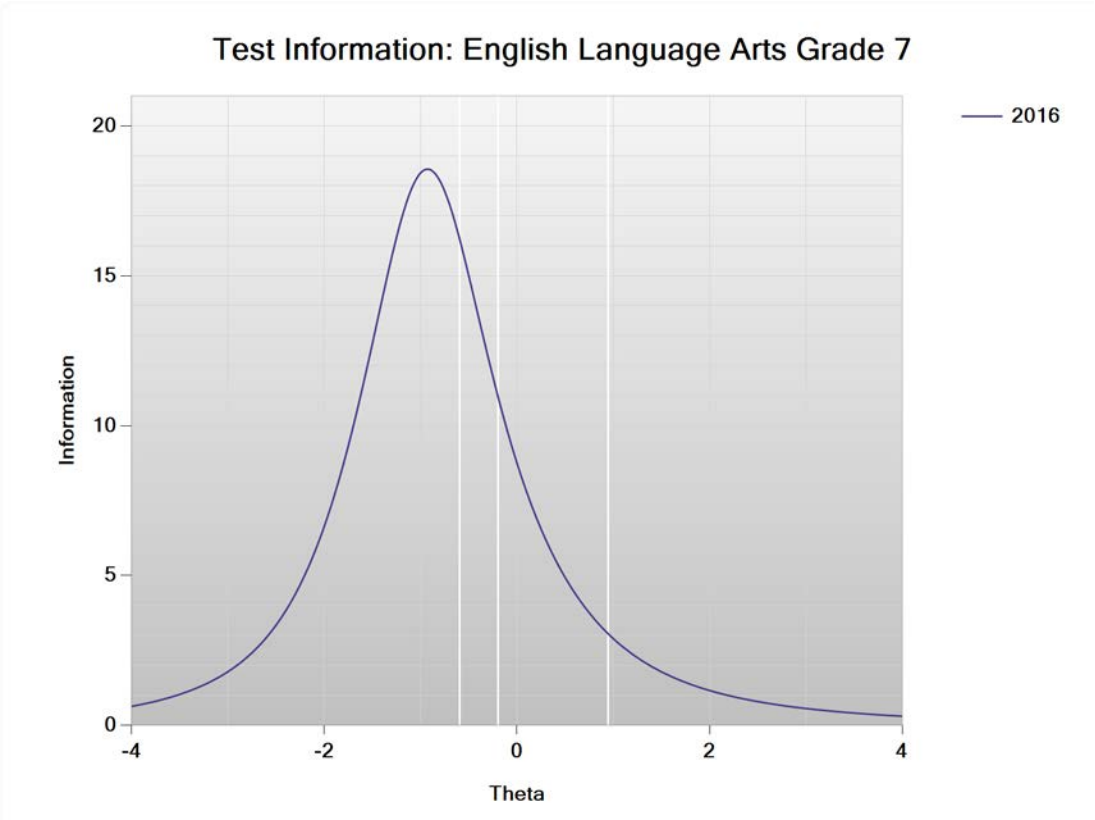


Figure G-25. Test Characteristic Curve for Grade 8 ELA

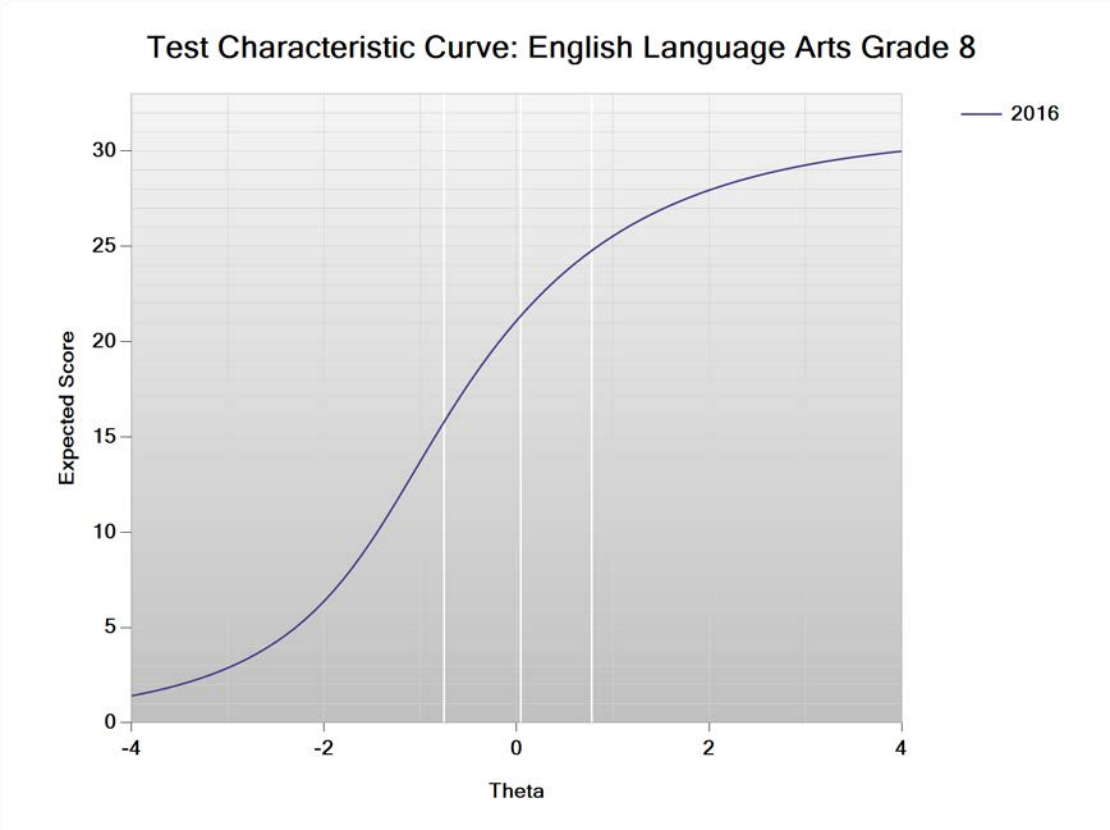


Figure G-26. Test Information Function and Standard Error for Grade 8 ELA

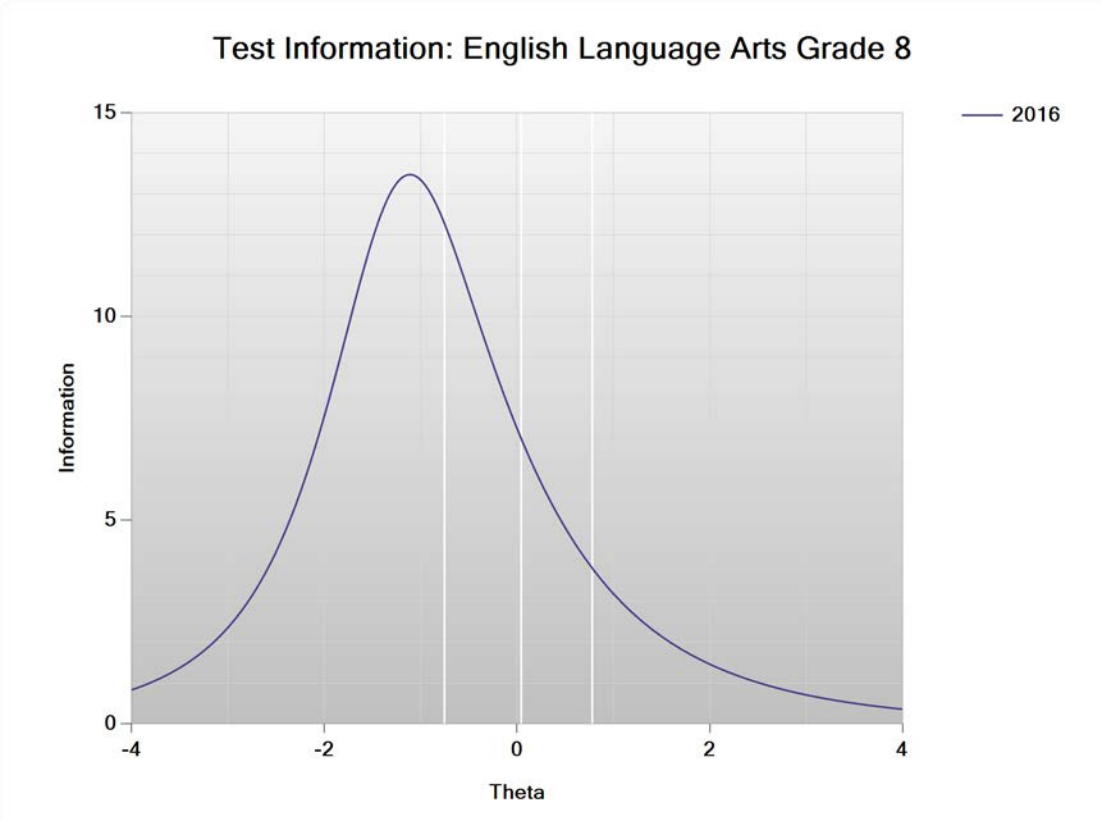


Figure G-27. Test Characteristic Curve for Grade 11 ELA

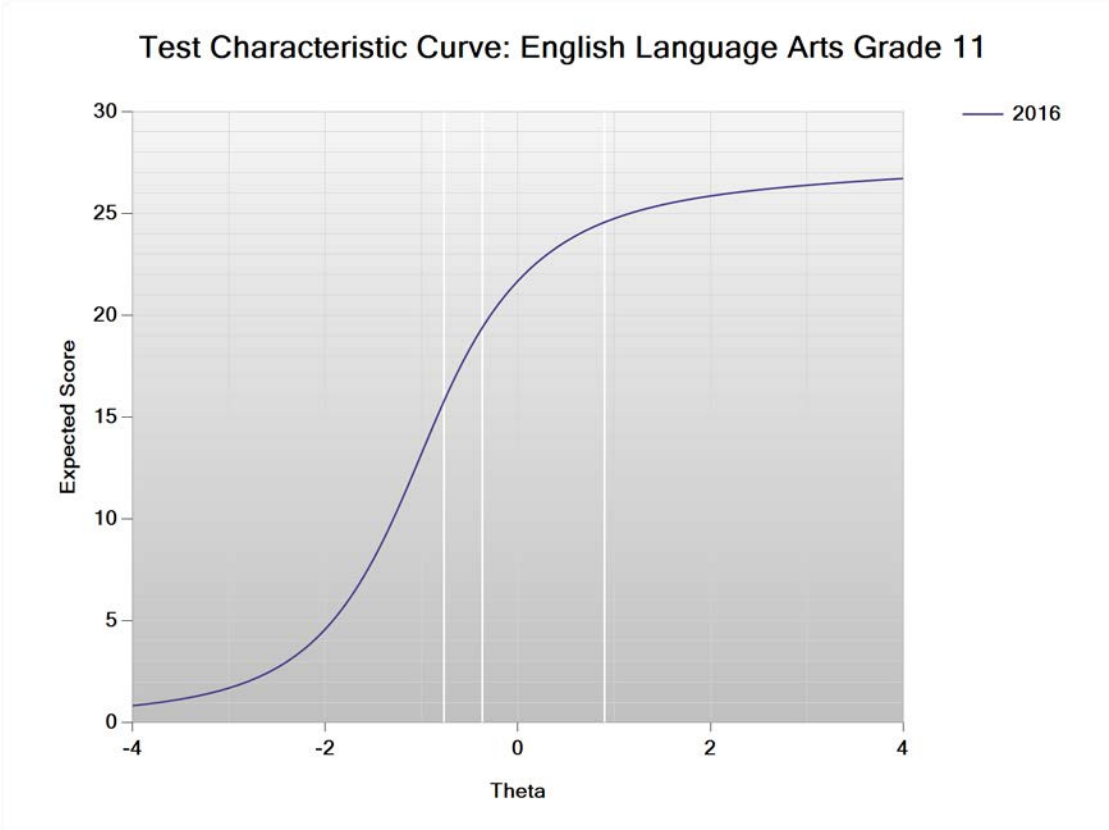
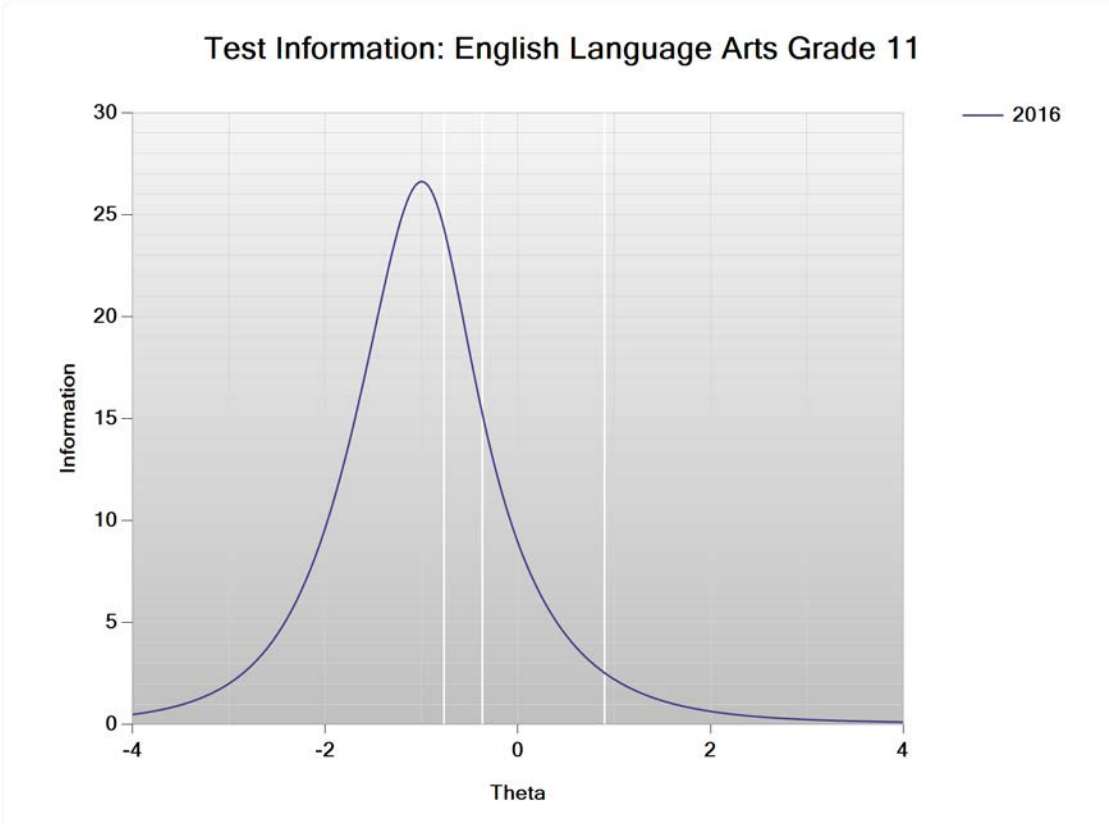


Figure G-28. Test Information Function and Standard Error for Grade 11 ELA



APPENDIX H—RAW TO SCALED SCORE LOOK-UP TABLES

Table H-1. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 3

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	22.2	1
1	1200	16.2	1
2	1200	11.0	1
3	1206	7.7	1
4	1211	6.1	1
5	1215	5.2	1
6	1218	4.6	1
7	1220	4.2	1
8	1222	3.8	1
9	1224	3.6	1
10	1226	3.4	1
11	1228	3.2	1
12	1230	3.1	1
13	1231	3.0	1
14	1233	2.9	1
15	1235	2.8	2
16	1236	2.7	2
17	1238	2.7	2
18	1240	2.6	3
19	1242	2.6	3
20	1244	2.7	3
21	1246	2.8	3
22	1248	3.0	3
23	1250	3.3	3
24	1252	3.7	4
25	1256	4.5	4
26	1260	5.9	4
27	1267	8.5	4
28	1278	13.8	4
29	1289	20.1	4
30	1290	20.1	4

Table H-2. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 4

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	23.5	1
1	1200	18.4	1
2	1200	14.1	1
3	1200	10.0	1
4	1206	7.7	1
5	1211	6.3	1

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
6	1215	5.3	1
7	1218	4.6	1
8	1220	4.1	1
9	1222	3.7	1
10	1224	3.4	1
11	1226	3.2	1
12	1228	3.0	1
13	1229	2.9	1
14	1231	2.8	1
15	1233	2.7	1
16	1234	2.6	2
17	1236	2.6	2
18	1237	2.6	2
19	1239	2.7	2
20	1241	2.8	3
21	1243	2.9	3
22	1245	3.1	3
23	1247	3.4	3
24	1249	3.8	3
25	1252	4.5	3
26	1256	5.5	3
27	1261	7.3	4
28	1269	10.3	4
29	1280	15.2	4
30	1288	19.4	4
31	1290	19.4	4

Table H-3. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 5

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	26.5	1
1	1200	19.1	1
2	1200	12.6	1
3	1205	8.7	1
4	1210	6.8	1
5	1214	5.7	1
6	1217	5.0	1
7	1220	4.5	1
8	1222	4.1	1
9	1224	3.9	1
10	1226	3.8	1
11	1228	3.6	1
12	1230	3.6	1
13	1232	3.6	2

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
14	1233	3.6	2
15	1235	3.7	2
16	1236	3.8	2
17	1238	3.9	2
18	1240	4.1	3
19	1242	4.4	3
20	1244	4.7	3
21	1247	5.1	3
22	1250	5.7	3
23	1253	6.4	3
24	1258	7.4	4
25	1263	8.9	4
26	1270	11.1	4
27	1280	14.9	4
28	1290	20.0	4
29	1290	20.0	4
30	1290	20.0	4

Table H-4. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 6

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	22.8	1
1	1200	15.9	1
2	1200	10.0	1
3	1204	7.3	1
4	1209	5.8	1
5	1212	5.0	1
6	1214	4.4	1
7	1217	4.1	1
8	1219	3.8	1
9	1220	3.6	1
10	1222	3.5	1
11	1224	3.4	1
12	1225	3.4	1
13	1227	3.4	1
14	1228	3.4	1
15	1229	3.4	1
16	1231	3.5	2
17	1232	3.5	2
18	1234	3.7	2
19	1236	3.8	2
20	1238	4.0	2
21	1239	4.3	2
22	1242	4.6	3

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
23	1244	5.0	3
24	1247	5.5	3
25	1251	6.3	3
26	1255	7.4	4
27	1261	9.2	4
28	1271	12.7	4
29	1287	21.2	4
30	1290	21.2	4

Table H-5. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 7

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	24.5	1
1	1200	17.2	1
2	1201	10.8	1
3	1209	7.6	1
4	1214	5.9	1
5	1217	4.8	1
6	1220	4.2	1
7	1222	3.7	1
8	1224	3.4	1
9	1226	3.2	1
10	1227	3.0	1
11	1229	2.9	1
12	1230	2.8	1
13	1231	2.8	1
14	1233	2.8	1
15	1234	2.9	1
16	1236	3.0	2
17	1237	3.2	2
18	1238	3.4	2
19	1240	3.6	3
20	1242	3.9	3
21	1245	4.3	3
22	1247	4.8	3
23	1250	5.5	3
24	1255	6.5	4
25	1259	7.9	4
26	1266	10.1	4
27	1277	13.9	4
28	1290	20.4	4
29	1290	20.4	4

Table H-6. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 8

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	20.6	1
1	1200	15.3	1
2	1200	10.3	1
3	1202	7.6	1
4	1207	6.1	1
5	1211	5.2	1
6	1213	4.6	1
7	1216	4.1	1
8	1218	3.8	1
9	1220	3.6	1
10	1221	3.5	1
11	1223	3.4	1
12	1224	3.3	1
13	1226	3.3	1
14	1227	3.3	1
15	1229	3.4	1
16	1230	3.4	2
17	1232	3.5	2
18	1234	3.7	2
19	1235	3.9	2
20	1237	4.1	2
21	1239	4.4	2
22	1241	4.7	3
23	1244	5.1	3
24	1247	5.6	3
25	1250	6.2	4
26	1254	7.1	4
27	1259	8.2	4
28	1265	10.0	4
29	1274	12.9	4
30	1290	19.7	4
31	1290	19.7	4

Table H-7. 2015–16 MSAA: Raw to Scaled Score Look-up Table—ELA Grade 11

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	27.7	1
1	1201	13.4	1
2	1212	6.8	1
3	1217	4.9	1
4	1220	3.9	1
5	1222	3.4	1

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
6	1224	3.0	1
7	1226	2.8	1
8	1227	2.6	1
9	1228	2.4	1
10	1229	2.3	1
11	1230	2.2	1
12	1231	2.2	1
13	1233	2.2	1
14	1234	2.2	1
15	1235	2.2	1
16	1236	2.3	2
17	1237	2.3	2
18	1238	2.5	2
19	1239	2.7	2
20	1241	2.9	3
21	1243	3.2	3
22	1245	3.7	3
23	1248	4.4	3
24	1252	5.5	3
25	1258	7.7	4
26	1270	14.3	4
27	1290	29.7	4
28	1290	29.7	4

Table H-8. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 3

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	24.3	1
1	1200	17.0	1
2	1201	11.3	1
3	1209	8.7	1
4	1213	7.3	1
5	1217	6.4	1
6	1220	5.7	1
7	1223	5.3	1
8	1225	4.9	1
9	1228	4.6	1
10	1229	4.4	1
11	1231	4.2	1
12	1233	4.1	1
13	1235	3.9	1
14	1236	3.8	2
15	1238	3.7	2
16	1239	3.7	2
17	1240	3.7	3
18	1242	3.7	3
19	1243	3.7	3
20	1245	3.7	3
21	1246	3.8	3
22	1248	3.9	3
23	1249	4.1	3
24	1251	4.3	3
25	1253	4.6	3
26	1255	4.9	4
27	1258	5.3	4
28	1260	5.8	4
29	1263	6.5	4
30	1267	7.3	4
31	1272	8.5	4
32	1278	10.4	4
33	1287	13.7	4
34	1290	17.8	4
35	1290	17.8	4

Table H-9. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 4

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	24.7	1
1	1200	18.6	1

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
2	1200	13.3	1
3	1202	10.2	1
4	1208	8.6	1
5	1213	7.5	1
6	1217	6.8	1
7	1220	6.2	1
8	1223	5.8	1
9	1225	5.5	1
10	1228	5.2	1
11	1230	5.0	1
12	1232	4.9	1
13	1234	4.8	2
14	1236	4.7	2
15	1238	4.6	2
16	1239	4.6	2
17	1241	4.7	3
18	1243	4.7	3
19	1245	4.8	3
20	1247	4.9	3
21	1249	5.0	3
22	1251	5.2	4
23	1254	5.5	4
24	1256	5.8	4
25	1259	6.2	4
26	1262	6.8	4
27	1266	7.5	4
28	1270	8.6	4
29	1276	10.1	4
30	1283	12.4	4
31	1290	15.6	4
32	1290	15.6	4
33	1290	15.6	4

Table H-10. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 5

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	26.9	1
1	1200	20.8	1
2	1200	15.4	1
3	1201	11.3	1
4	1208	9.1	1
5	1213	7.8	1
6	1217	6.9	1
7	1220	6.3	1

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
8	1223	5.8	1
9	1225	5.5	1
10	1228	5.3	1
11	1230	5.1	1
12	1232	5.0	2
13	1234	4.9	2
14	1236	4.8	2
15	1238	4.8	2
16	1240	4.8	3
17	1242	4.8	3
18	1244	4.8	3
19	1246	4.8	3
20	1248	4.9	3
21	1250	5.0	3
22	1252	5.1	3
23	1254	5.3	3
24	1256	5.5	4
25	1259	5.8	4
26	1262	6.2	4
27	1265	6.7	4
28	1268	7.4	4
29	1272	8.3	4
30	1277	9.6	4
31	1283	11.4	4
32	1290	14.1	4
33	1290	14.7	4
34	1290	14.7	4
35	1290	14.7	4

Table H-11. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 6

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	23.1	1
1	1200	17.0	1
2	1200	11.8	1
3	1205	9.0	1
4	1210	7.5	1
5	1214	6.5	1
6	1217	5.8	1
7	1220	5.3	1
8	1222	4.9	1
9	1225	4.6	1
10	1226	4.4	1
11	1228	4.2	1

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
12	1230	4.1	1
13	1232	4.0	1
14	1233	4.0	1
15	1235	3.9	2
16	1236	3.9	2
17	1238	3.9	2
18	1239	4.0	2
19	1241	4.1	3
20	1242	4.2	3
21	1244	4.3	3
22	1246	4.5	3
23	1248	4.7	3
24	1250	4.9	4
25	1252	5.2	4
26	1254	5.6	4
27	1257	6.0	4
28	1260	6.6	4
29	1264	7.4	4
30	1268	8.4	4
31	1274	9.9	4
32	1281	12.3	4
33	1290	16.5	4
34	1290	16.7	4
35	1290	16.7	4

Table H-12. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 7

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	20.8	1
1	1200	15.1	1
2	1201	10.4	1
3	1208	8.3	1
4	1212	7.1	1
5	1216	6.4	1
6	1219	5.8	1
7	1222	5.5	1
8	1224	5.2	1
9	1226	5.0	1
10	1228	4.8	1
11	1230	4.7	1
12	1232	4.6	2
13	1234	4.6	2
14	1236	4.5	2
15	1237	4.5	2

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
16	1239	4.5	2
17	1241	4.5	3
18	1243	4.6	3
19	1244	4.6	3
20	1246	4.7	3
21	1248	4.8	3
22	1250	4.9	3
23	1252	5.1	3
24	1254	5.2	4
25	1256	5.5	4
26	1258	5.7	4
27	1261	6.1	4
28	1264	6.5	4
29	1267	7.1	4
30	1271	8.0	4
31	1277	9.3	4
32	1284	11.5	4
33	1290	16.2	4
34	1290	16.2	4

Table H-13. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 8

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	23.9	1
1	1200	16.8	1
2	1200	11.4	1
3	1208	8.9	1
4	1213	7.6	1
5	1217	6.7	1
6	1220	6.0	1
7	1223	5.6	1
8	1225	5.2	1
9	1227	5.0	1
10	1229	4.8	1
11	1231	4.6	1
12	1233	4.5	1
13	1235	4.4	2
14	1236	4.3	2
15	1238	4.3	2
16	1239	4.3	2
17	1241	4.2	3
18	1243	4.3	3
19	1244	4.3	3
20	1246	4.3	3

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
21	1247	4.4	3
22	1249	4.5	4
23	1251	4.6	4
24	1253	4.8	4
25	1255	5.0	4
26	1257	5.2	4
27	1259	5.5	4
28	1262	5.9	4
29	1265	6.4	4
30	1268	7.1	4
31	1272	8.1	4
32	1278	9.6	4
33	1286	12.5	4
34	1290	16.3	4
35	1290	16.3	4

Table H-14. 2015–16 MSAA: Raw to Scaled Score Look-up Table—Mathematics Grade 11

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
0	1200	24.0	1
1	1200	16.5	1
2	1202	11.1	1
3	1209	8.9	1
4	1214	7.6	1
5	1218	6.8	1
6	1221	6.2	1
7	1224	5.7	1
8	1227	5.3	1
9	1229	5.1	1
10	1231	4.8	1
11	1233	4.7	1
12	1235	4.5	2
13	1237	4.4	2
14	1238	4.3	2
15	1239	4.3	2
16	1242	4.2	3
17	1243	4.2	3
18	1245	4.2	3
19	1246	4.3	3
20	1248	4.3	3
21	1250	4.4	4
22	1251	4.5	4
23	1253	4.7	4
24	1255	4.9	4

continued

Raw Score	2016		
	Scaled Score	*Standard Error*	*Performance Level*
25	1257	5.1	4
26	1260	5.4	4
27	1262	5.8	4
28	1265	6.4	4
29	1269	7.1	4
30	1273	8.2	4
31	1279	9.9	4
32	1288	13.1	4
33	1290	16.3	4
34	1290	16.3	4

APPENDIX I—SCORE DISTRIBUTIONS

Figure I-1. 2015–16 MSAA: Performance Level Distributions Graph—ELA

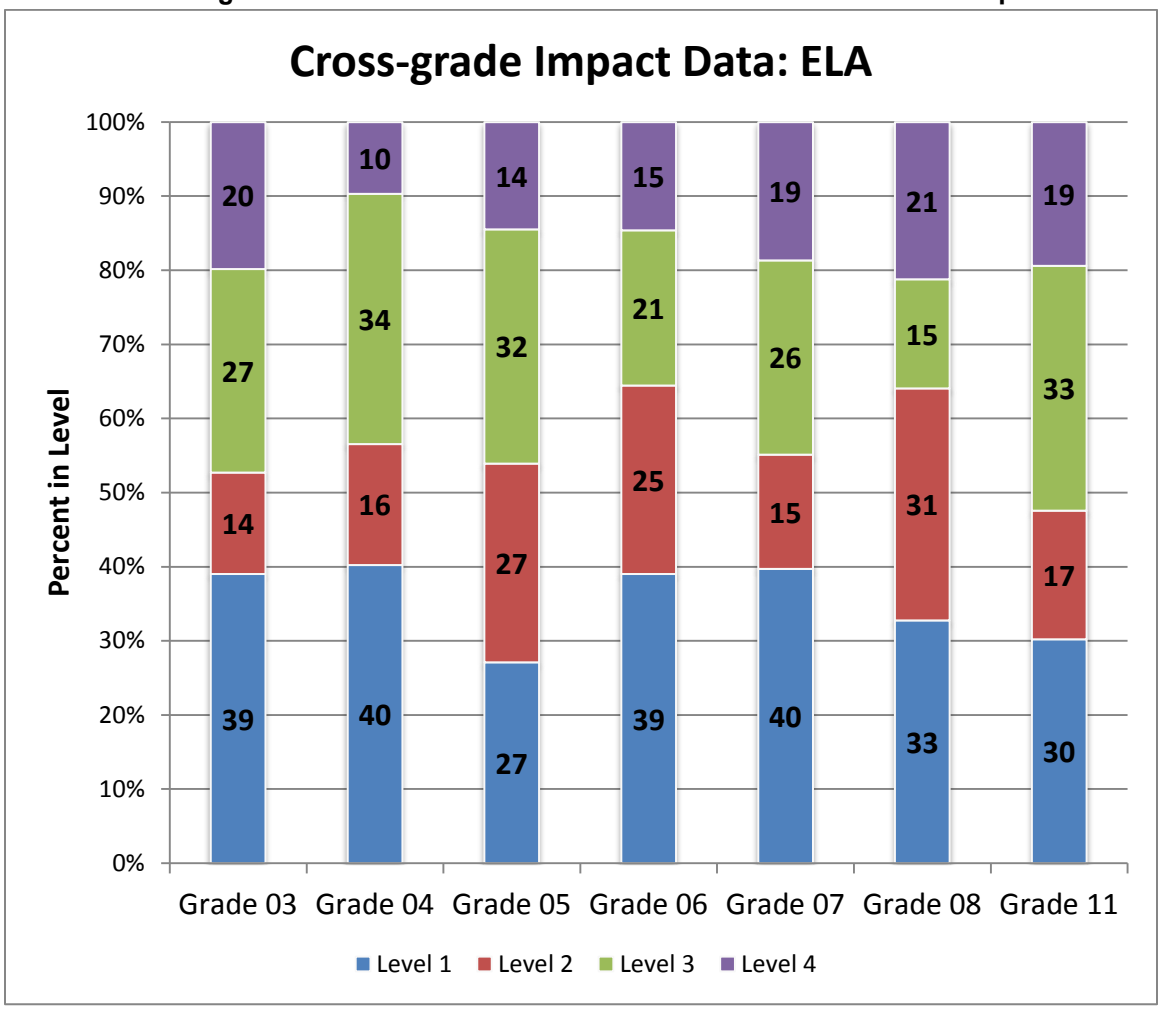


Figure I-2. 2015–16 MSAA: Performance Level Distributions Graph—Mathematics

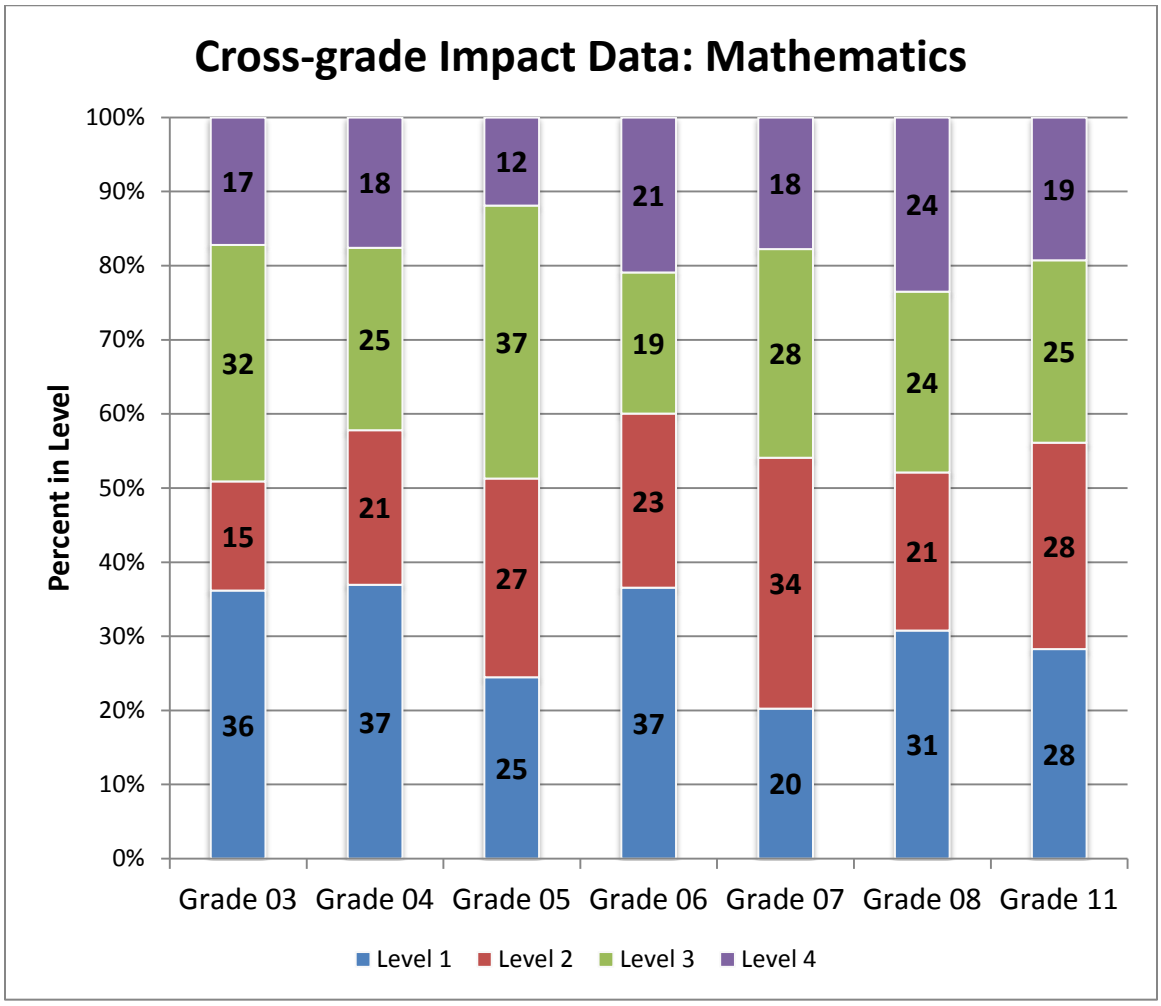
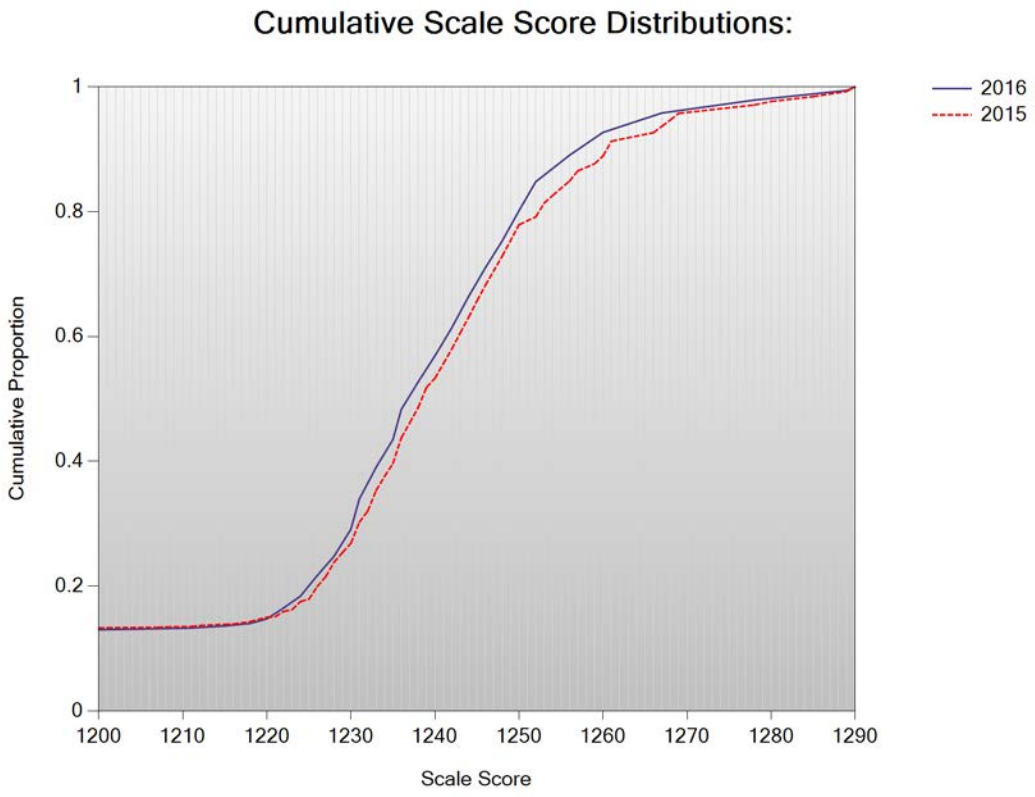


Figure I-3. 2015–16 MSAA: Cumulative Score Distribution—
Top: ELA Grade 3 Bottom: ELA Grade 4



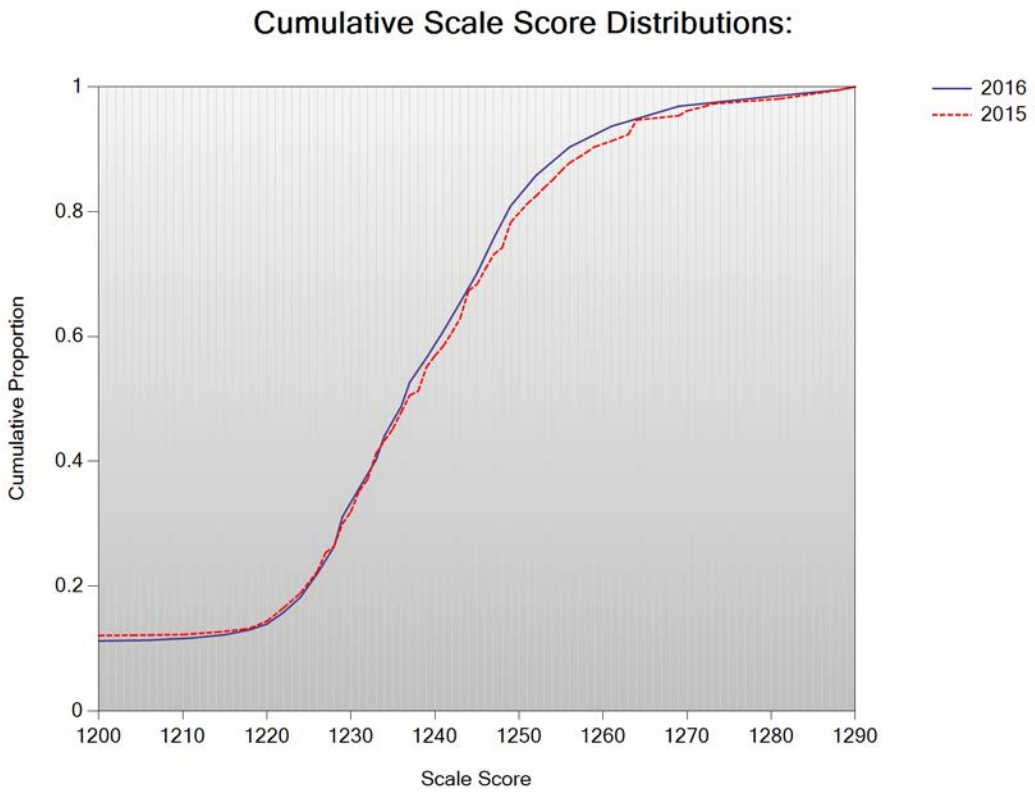
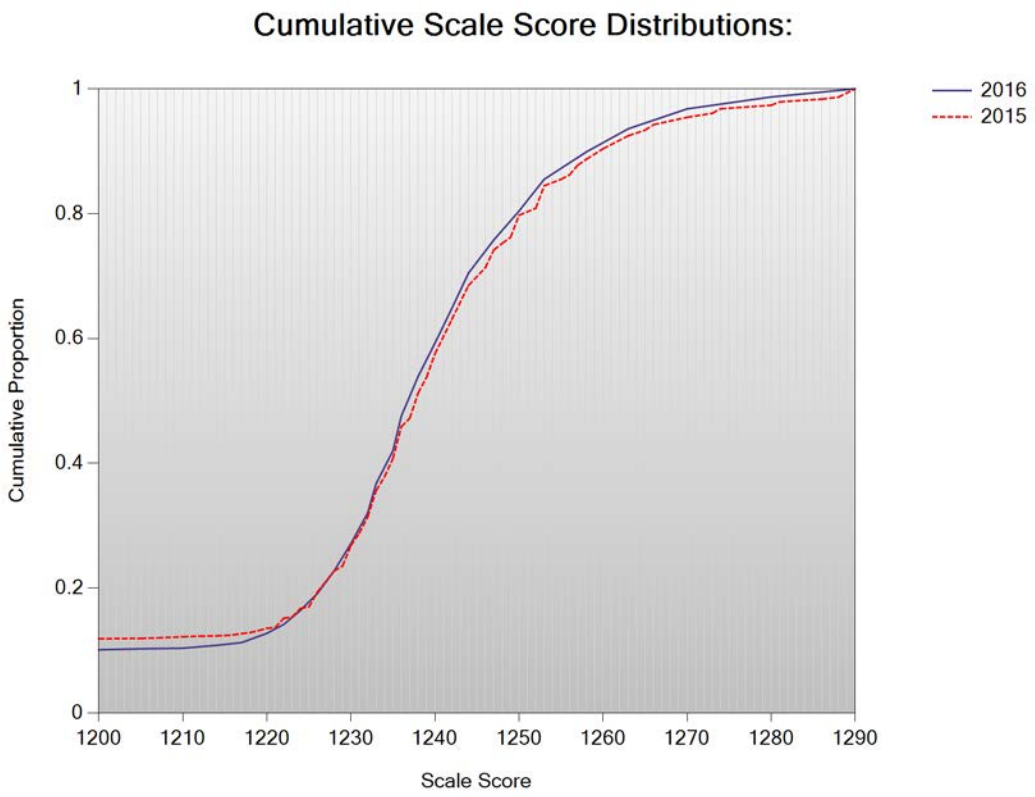


Figure I-4. 2015–16 MSAA: Cumulative Score Distribution
Top: ELA Grade 5 Bottom: ELA Grade 6



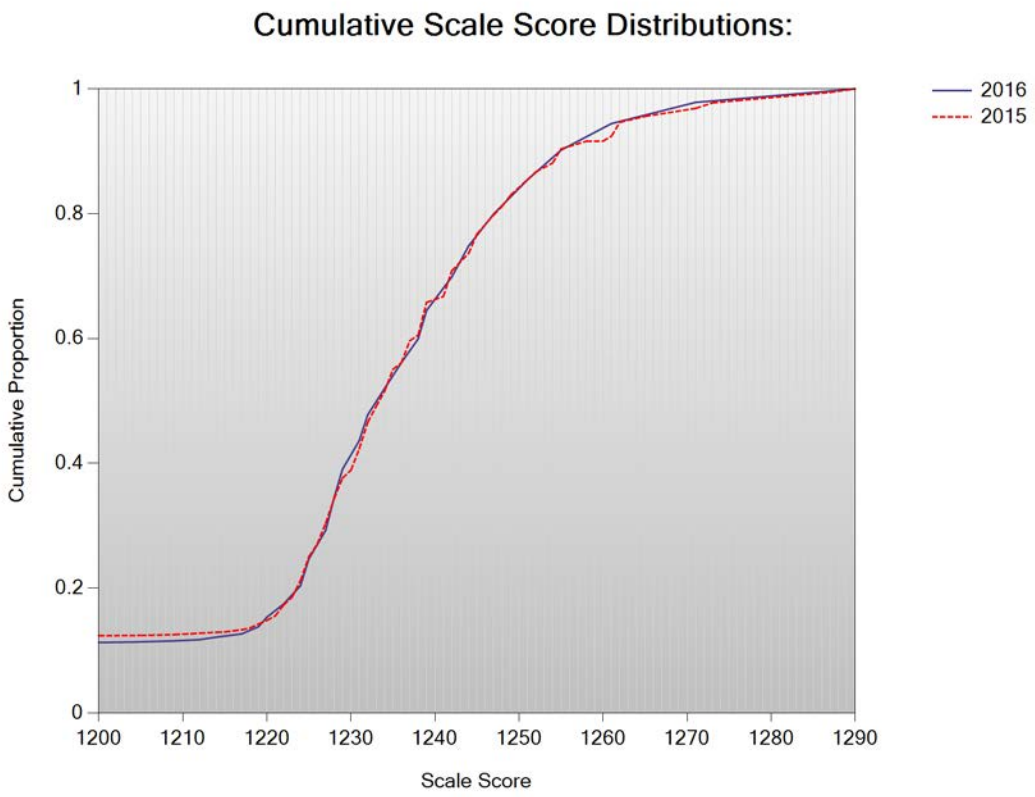
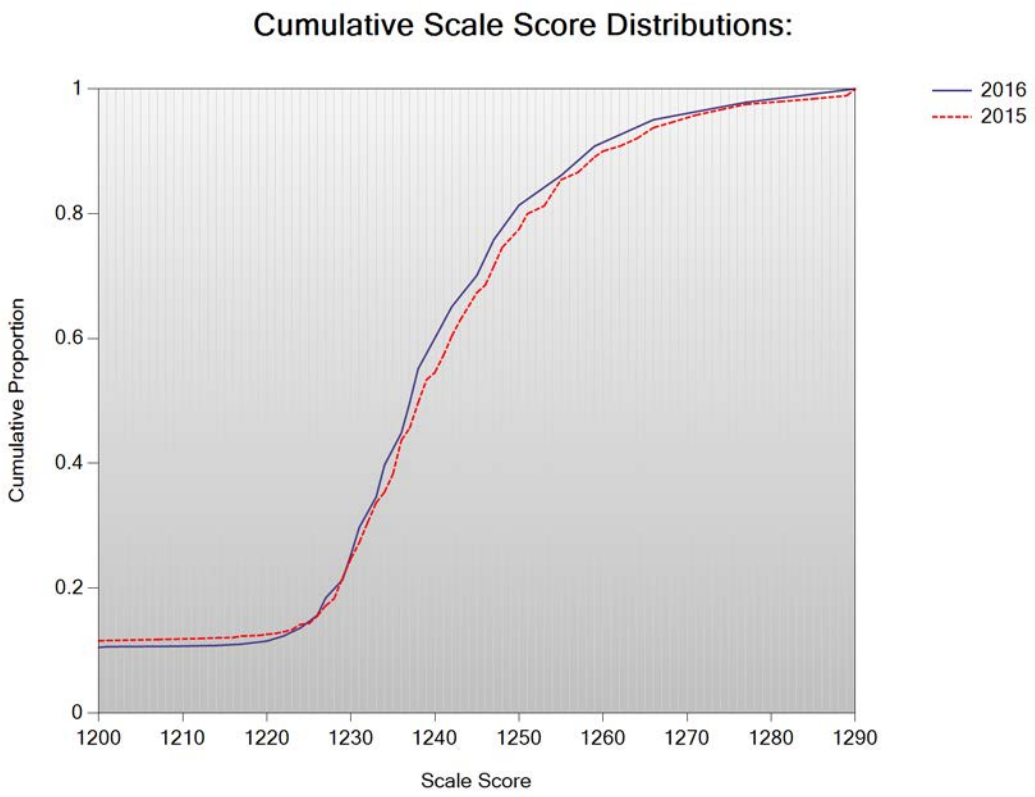


Figure I-5. 2015–16 MSAA: Cumulative Score Distribution
Top: ELA Grade 7 Bottom: ELA Grade 8



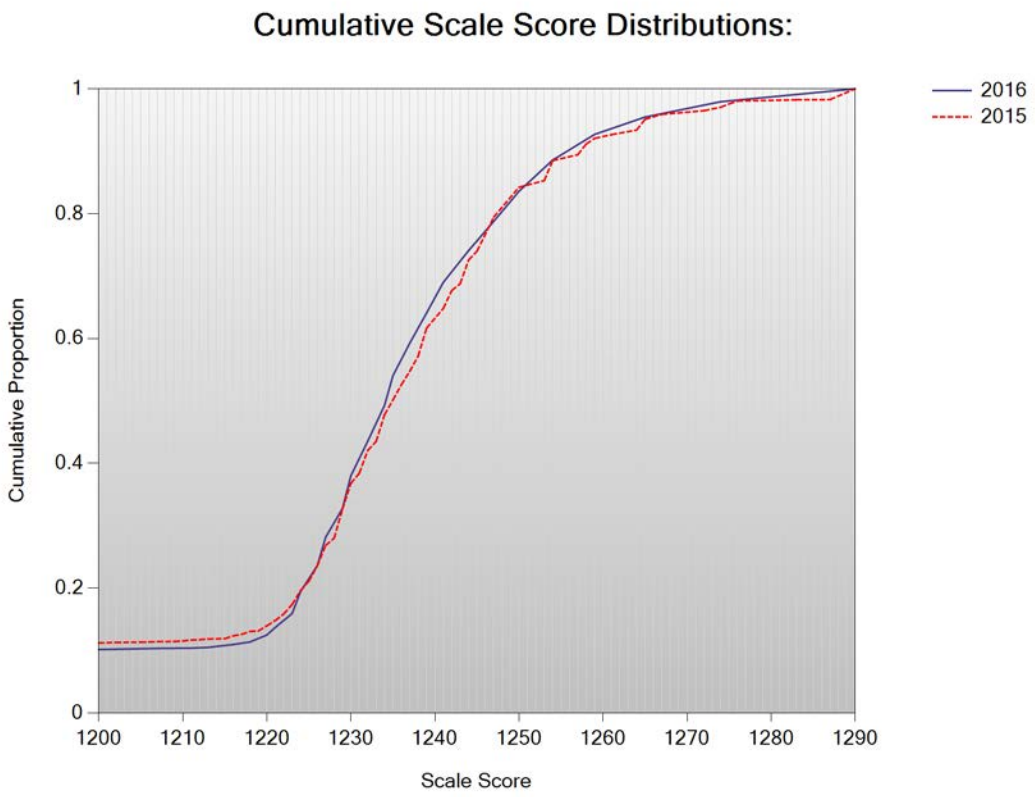


Figure I-6. 2015–16 MSAA: Cumulative Score Distribution
ELA Grade 11

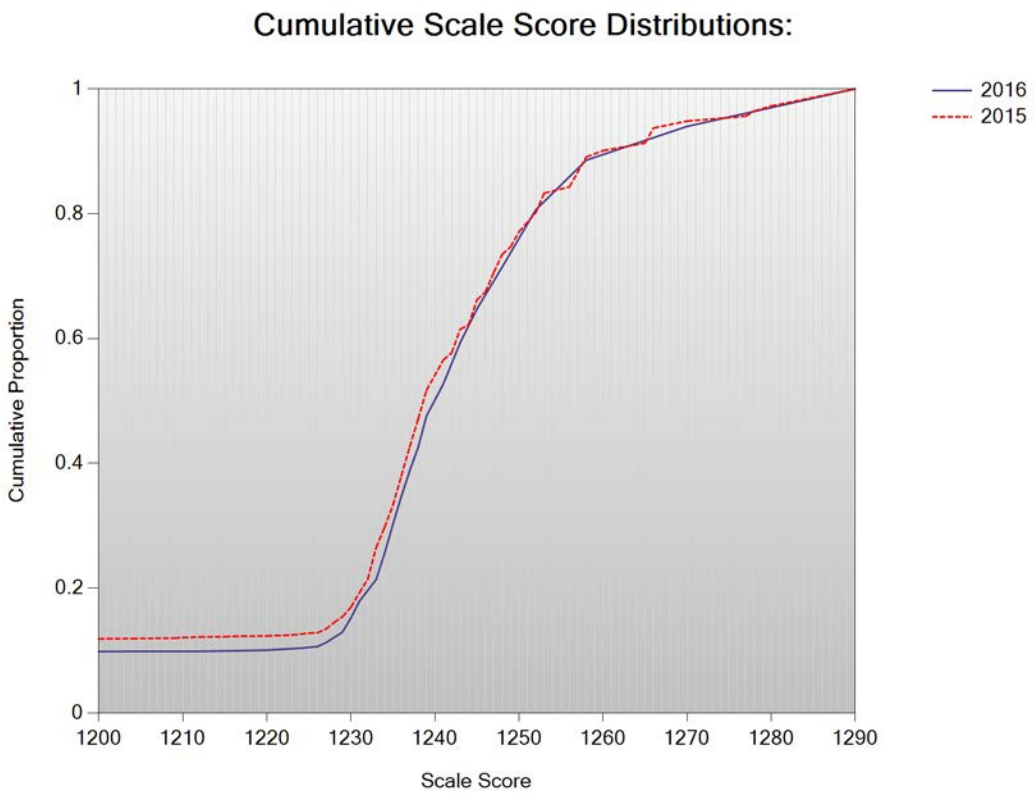
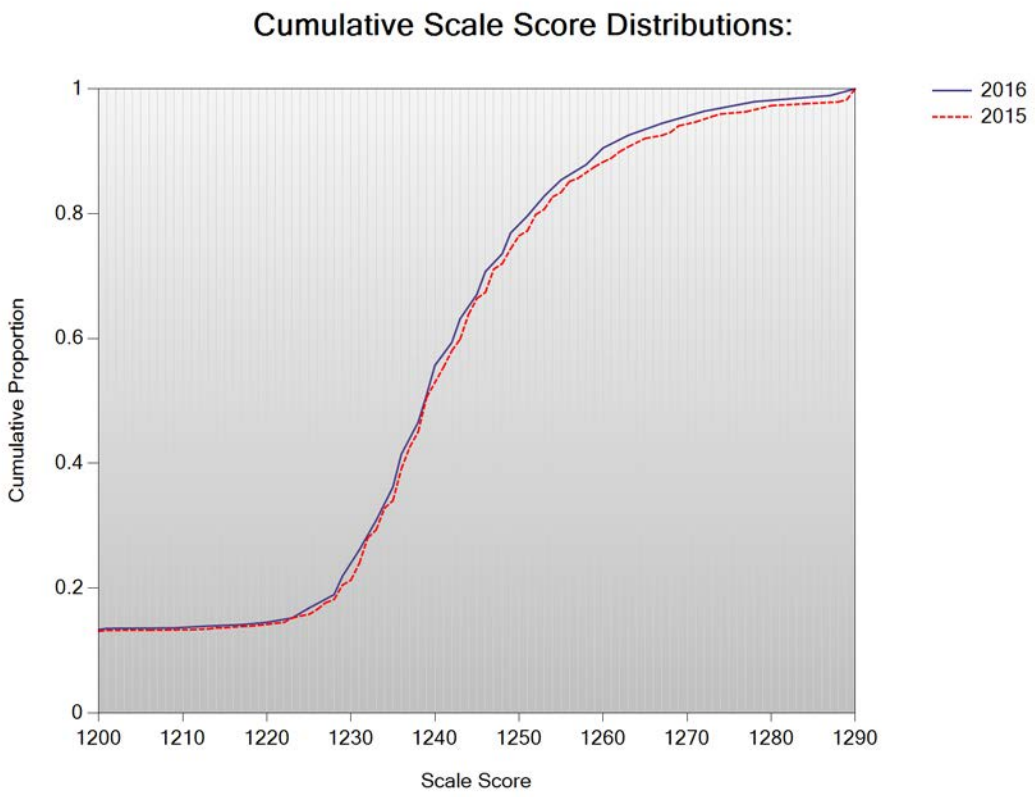


Figure I-7. 2015–16 MSAA: Cumulative Score Distribution
Top: Mathematics Grade 3 Bottom: Mathematics Grade 4



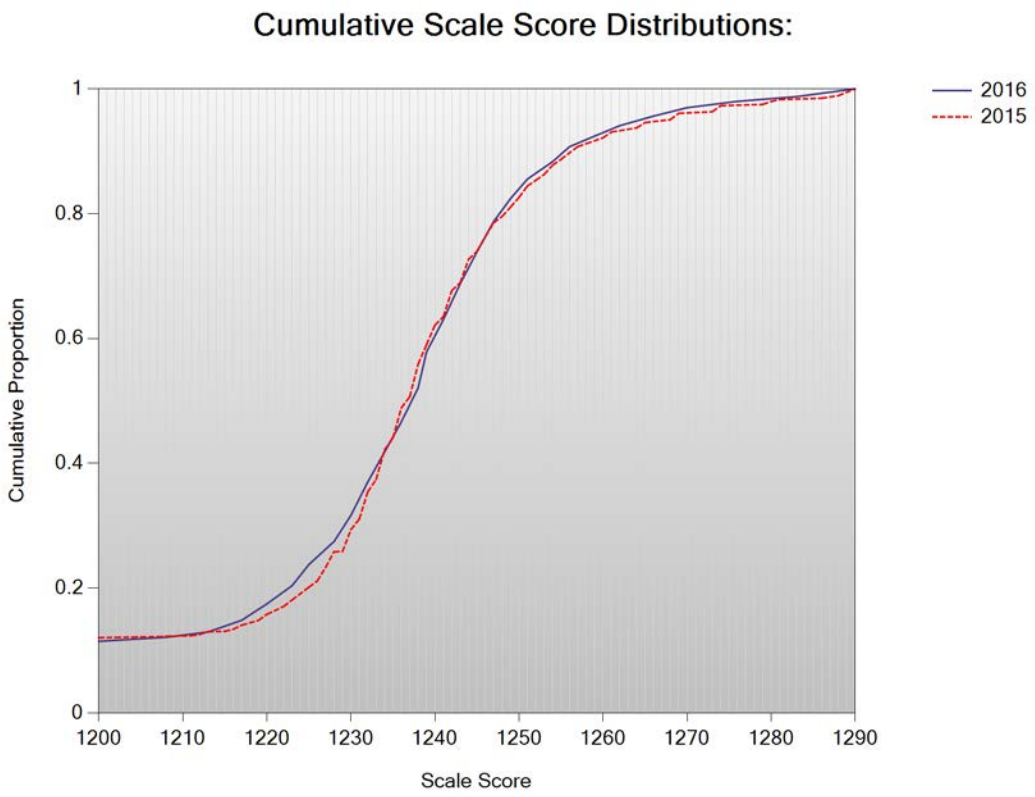
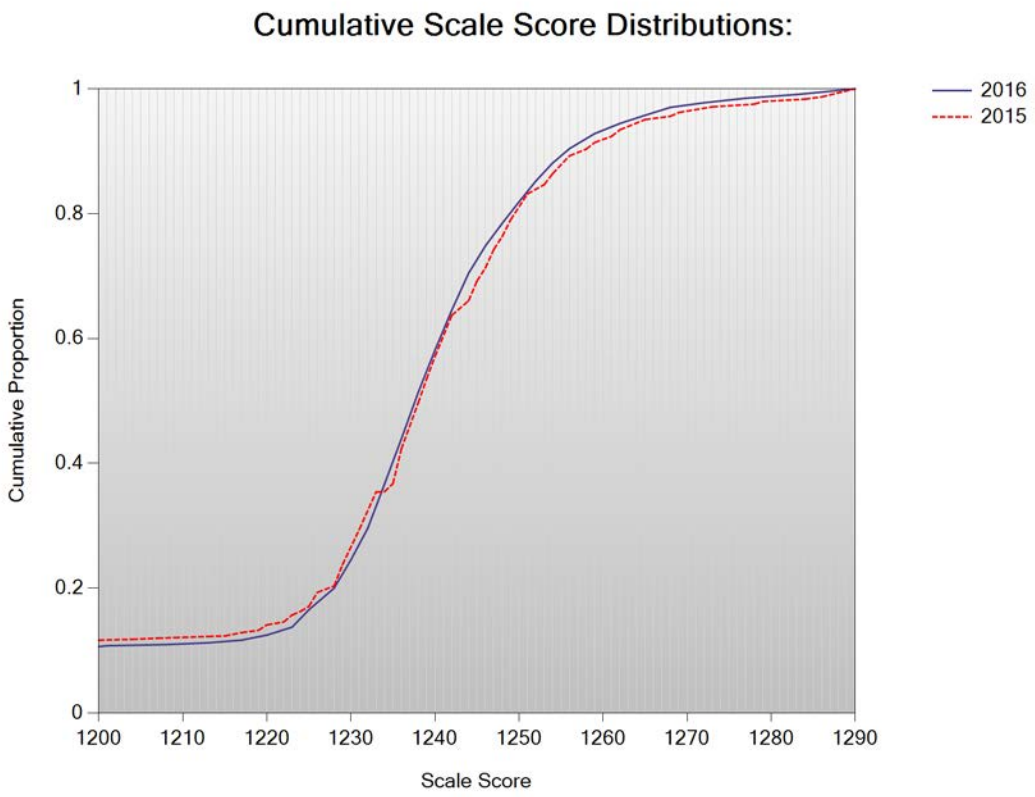


Figure I-8. 2015–16 MSAA: Cumulative Score Distribution
Top: Mathematics Grade 5 Bottom: Mathematics Grade 6



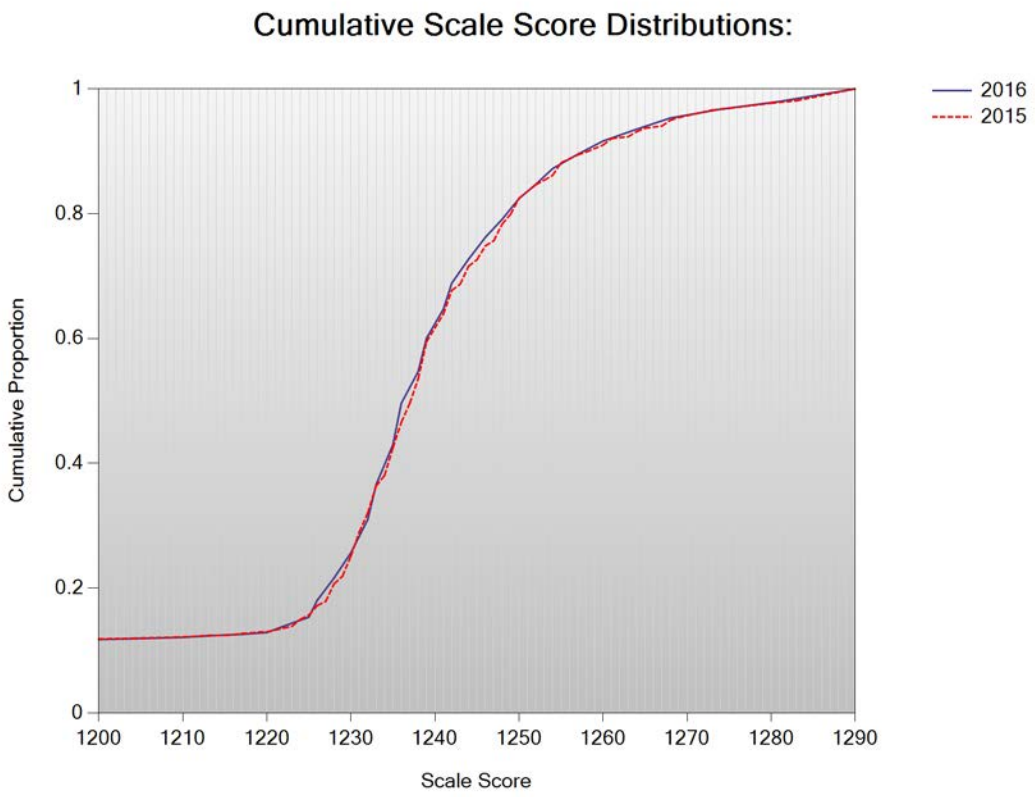
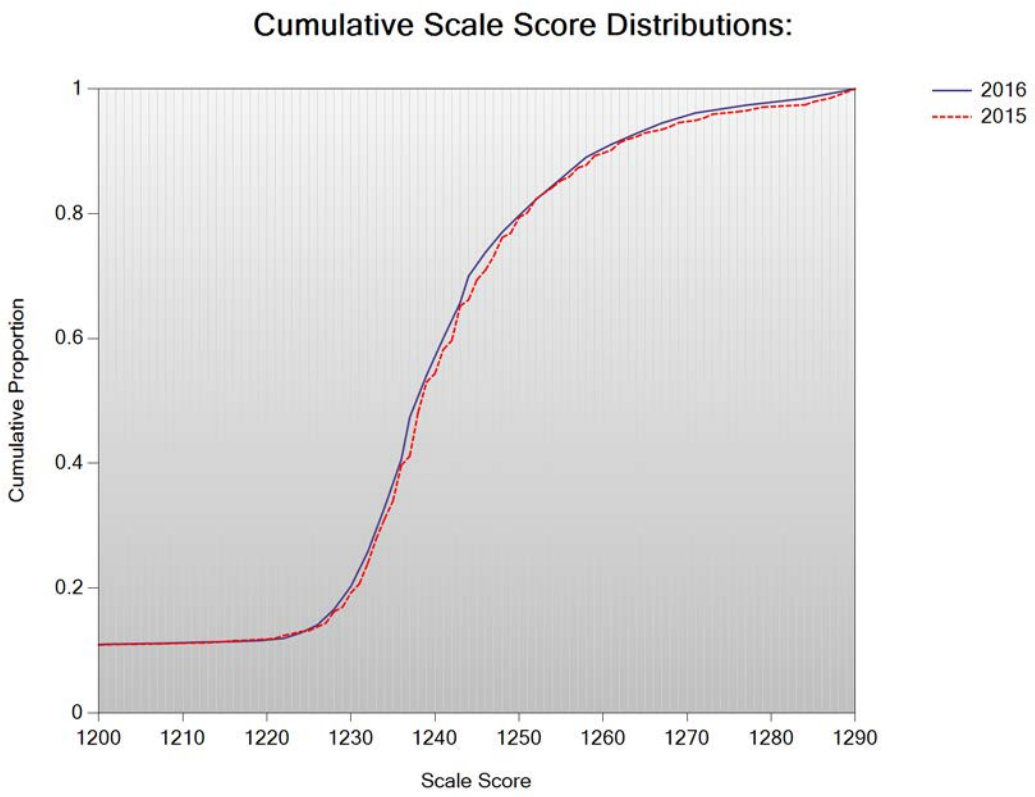


Figure I-9. 2015–16 MSAA: Cumulative Score Distribution
Top: Mathematics Grade 7 Bottom: Mathematics Grade 8



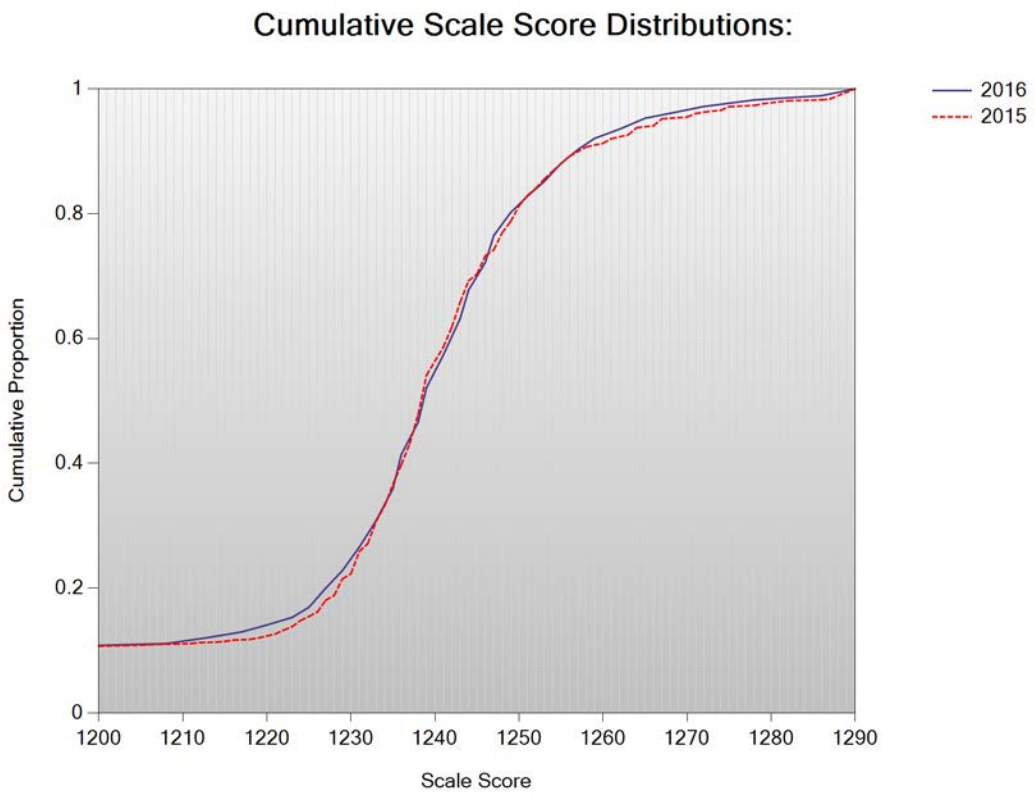
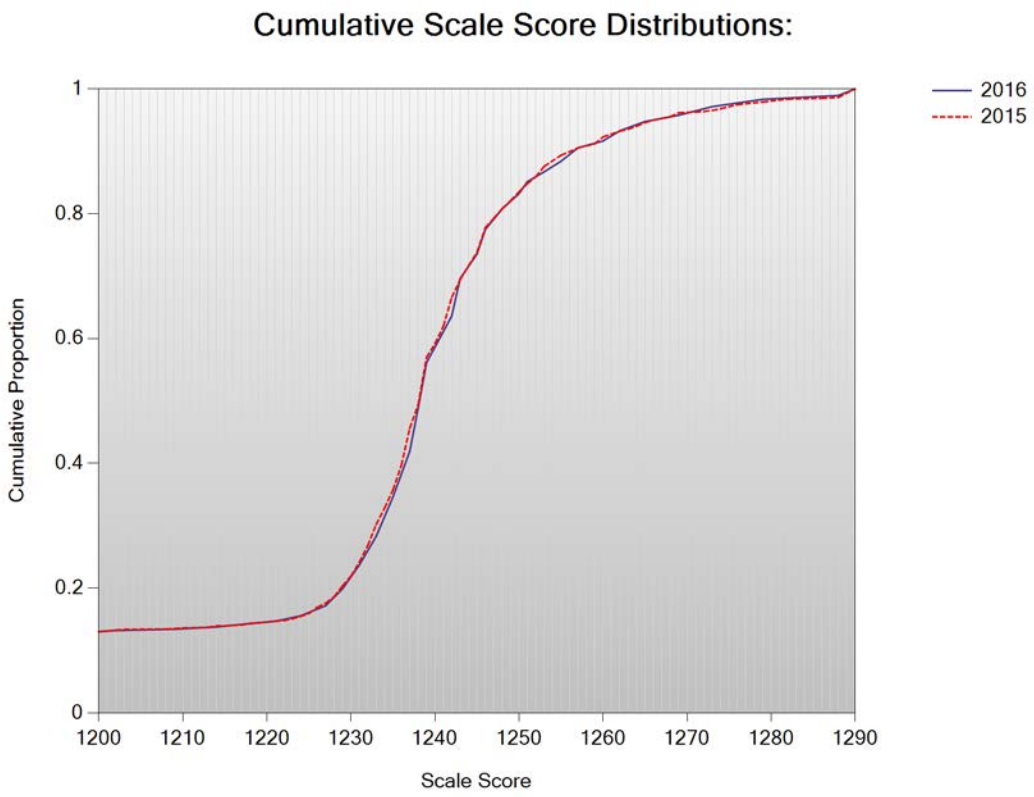


Figure I-10. 2015–16 MSAA: Cumulative Score Distribution—Mathematics Grade 11



APPENDIX J—CLASSICAL RELIABILITY

Table J-1. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 3

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,221	30	18.17	6.03	0.85	2.34
Female	1,016	30	18.00	6.01	0.85	2.33
Male	2,050	30	18.34	6.01	0.85	2.33
Gender Undefined	155	30	17.10	6.36	0.87	2.29
Hispanic or Latino	694	30	17.66	6.00	0.85	2.32
American Indian or Alaska Native	109	30	17.17	6.18	0.86	2.31
Asian						
Black or African American	668	30	17.65	5.82	0.84	2.33
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,331	30	18.98	6.07	0.86	2.27
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	254	30	17.61	6.23	0.86	2.33
Currently receiving LEP services	158	30	18.49	5.63	0.83	2.32
Not receiving LEP services	2,187	30	18.35	6.04	0.86	2.26
LEP: All Other Students	876	30	17.66	6.04	0.85	2.34
Economically Disadvantaged Students	1,020	30	19.07	6.04	0.86	2.26
Non-economically Disadvantaged Students	1,306	30	17.84	5.96	0.85	2.31
SES: All Other Students	895	30	17.64	6.01	0.85	2.33
Migrant						
Non-migrant	915	30	17.78	5.90	0.84	2.36
Undefined Migrant Status	2,306	30	18.33	6.07	0.86	2.27
Augmentative Communication	445	30	14.74	5.01	0.76	2.45
No Augmentative Communication	2,753	30	18.73	6.00	0.86	2.24
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,138	30	18.25	6.02	0.85	2.33
Undefined Hearing Loss						
Visual Impairment						
Within Normal Limits	3,104	30	18.27	6.02	0.85	2.33
Undefined Visual Impairment						
Sensory Stimuli Response	209	30	12.32	5.72	0.82	2.43

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Follow Directions	3,003	30	18.58	5.84	0.85	2.26
Undefined Receptive Language						
Special School	236	30	15.09	5.91	0.83	2.44
Regular School Self-contained	2,151	30	17.72	6.03	0.85	2.34
Regular School Resource Room	460	30	19.79	5.37	0.83	2.21
Regular School Primarily Self-contained	241	30	21.39	5.39	0.84	2.16
Regular School General Education	124	30	19.61	5.54	0.83	2.28
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	143	30	10.64	5.17	0.79	2.37
Uses Intentional Communication	713	30	15.13	5.36	0.79	2.46
Uses Symbolic Language	2,356	30	19.55	5.58	0.84	2.23
Undefined Expressive Communication						

Table J-2. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 4

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,422	31	18.84	6.15	0.85	2.38
Female	1,077	31	19.20	6.20	0.85	2.40
Male	2,161	31	18.67	6.15	0.85	2.38
Gender Undefined	184	31	18.75	5.84	0.83	2.41
Hispanic or Latino	688	31	18.39	5.82	0.82	2.47
American Indian or Alaska Native	118	31	19.84	6.03	0.85	2.34
Asian						
Black or African American	690	31	17.84	6.17	0.84	2.47
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,438	31	19.50	6.27	0.86	2.35
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	304	31	18.97	5.88	0.83	2.42
Currently receiving LEP services	132	31	20.44	5.17	0.79	2.37
Not receiving LEP services	2,412	31	18.84	6.23	0.85	2.41

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
LEP: All Other Students	878	31	18.59	6.01	0.84	2.40
Economically Disadvantaged Students	1,099	31	19.75	6.23	0.86	2.33
Non-economically Disadvantaged Students	1,436	31	18.31	6.10	0.84	2.44
SES: All Other Students	887	31	18.57	6.01	0.84	2.40
Migrant						
Non-migrant	1,044	31	18.14	6.22	0.85	2.41
Undefined Migrant Status	2,376	31	19.14	6.09	0.85	2.36
Augmentative Communication	482	31	14.62	4.96	0.73	2.58
No Augmentative Communication	2,923	31	19.54	6.04	0.85	2.34
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,326	31	18.93	6.16	0.85	2.39
Undefined Hearing Loss						
Visual Impairment	108	31	15.66	6.76	0.87	2.44
Within Normal Limits	3,293	31	18.95	6.10	0.84	2.44
Undefined Visual Impairment						
Sensory Stimuli Response	192	31	13.07	5.52	0.79	2.53
Follow Directions	3,221	31	19.20	6.01	0.84	2.40
Undefined Receptive Language						
Special School	246	31	15.01	5.46	0.78	2.56
Regular School Self-contained	2,223	31	18.34	6.10	0.84	2.44
Regular School Resource Room	535	31	20.50	5.79	0.84	2.32
Regular School Primarily Self-contained	294	31	21.92	5.25	0.82	2.23
Regular School General Education	115	31	21.32	6.03	0.86	2.26
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	155	31	12.01	5.12	0.76	2.51
Uses Intentional Communication	693	31	15.07	5.30	0.77	2.54
Uses Symbolic Language	2,563	31	20.28	5.69	0.83	2.35
Undefined Expressive Communication						

Table J-3. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 5

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,548	30	17.55	5.53	0.81	2.41
Female	1,170	30	17.49	5.60	0.82	2.38
Male	2,193	30	17.59	5.49	0.81	2.39
Gender Undefined	185	30	17.49	5.62	0.82	2.38
Hispanic or Latino	711	30	17.32	5.34	0.80	2.39
American Indian or Alaska Native	124	30	17.44	5.89	0.84	2.36
Asian						
Black or African American	709	30	16.92	5.52	0.81	2.41
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,518	30	18.13	5.59	0.82	2.37
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	289	30	17.53	5.53	0.81	2.41
Currently receiving LEP services	118	30	17.80	5.62	0.82	2.38
Not receiving LEP services	2,457	30	17.69	5.49	0.81	2.39
LEP: All Other Students	973	30	17.17	5.61	0.82	2.38
Economically Disadvantaged Students	1,091	30	18.54	5.48	0.82	2.32
Non-economically Disadvantaged Students	1,468	30	17.10	5.43	0.80	2.43
SES: All Other Students	989	30	17.14	5.59	0.82	2.37
Migrant						
Non-migrant	1,107	30	17.15	5.47	0.81	2.38
Undefined Migrant Status	2,440	30	17.73	5.54	0.82	2.35
Augmentative Communication	504	30	14.11	4.87	0.73	2.53
No Augmentative Communication	3,020	30	18.13	5.43	0.81	2.37
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,435	30	17.63	5.49	0.81	2.39
Undefined Hearing Loss						
Visual Impairment	144	30	13.99	6.12	0.84	2.45
Within Normal Limits	3,382	30	17.71	5.46	0.81	2.38
Undefined Visual Impairment						
Sensory Stimuli Response	205	30	12.22	5.37	0.79	2.46
Follow Directions	3,336	30	17.88	5.37	0.80	2.40

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Undefined Receptive Language						
Special School	311	30	14.24	5.49	0.80	2.46
Regular School Self-contained	2,252	30	17.18	5.43	0.80	2.43
Regular School Resource Room	559	30	19.47	5.08	0.80	2.27
Regular School Primarily Self-contained	298	30	19.65	5.00	0.79	2.29
Regular School General Education	121	30	19.03	5.24	0.80	2.34
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	156	30	11.33	5.20	0.78	2.44
Uses Intentional Communication	634	30	14.22	5.09	0.76	2.49
Uses Symbolic Language	2,751	30	18.68	5.09	0.79	2.33
Undefined Expressive Communication						

Table J-4. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 6

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,577	30	18.99	6.03	0.85	2.34
Female	1,215	30	19.01	5.81	0.84	2.32
Male	2,176	30	19.04	6.17	0.86	2.31
Gender Undefined	186	30	18.13	5.72	0.82	2.43
Hispanic or Latino	652	30	18.36	6.16	0.85	2.39
American Indian or Alaska Native	132	30	19.20	5.33	0.80	2.38
Asian						
Black or African American	756	30	18.54	5.88	0.84	2.35
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,564	30	19.64	6.09	0.86	2.28
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	271	30	18.72	5.71	0.83	2.35
Currently receiving LEP services	133	30	19.30	6.23	0.87	2.25
Not receiving LEP services	2,442	30	19.21	6.03	0.85	2.34
LEP: All Other Students	1,002	30	18.38	5.98	0.84	2.39

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Economically Disadvantaged Students	1,068	30	20.34	5.71	0.84	2.28
Non-economically Disadvantaged Students	1,485	30	18.48	6.14	0.85	2.38
SES: All Other Students	1,024	30	18.28	5.96	0.84	2.38
Migrant						
Non-migrant	1,099	30	18.62	5.97	0.84	2.39
Undefined Migrant Status	2,477	30	19.14	6.05	0.85	2.34
Augmentative Communication	436	30	14.84	5.55	0.80	2.48
No Augmentative Communication	3,106	30	19.56	5.87	0.85	2.27
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,449	30	19.07	6.02	0.85	2.33
Undefined Hearing Loss						
Visual Impairment	137	30	16.58	6.35	0.86	2.38
Within Normal Limits	3,406	30	19.08	6.01	0.85	2.33
Undefined Visual Impairment						
Sensory Stimuli Response	166	30	12.90	5.96	0.83	2.46
Follow Directions	3,391	30	19.28	5.88	0.84	2.35
Undefined Receptive Language						
Special School	287	30	15.93	5.80	0.82	2.46
Regular School Self-contained	2,428	30	18.50	6.01	0.85	2.33
Regular School Resource Room	514	30	21.25	5.65	0.85	2.19
Regular School Primarily Self-contained	248	30	21.89	4.77	0.79	2.19
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	119	30	12.90	6.23	0.84	2.49
Uses Intentional Communication	641	30	15.26	5.39	0.78	2.53
Uses Symbolic Language	2,797	30	20.09	5.66	0.84	2.26
Undefined Expressive Communication						

Table J-5. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 7

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,713	29	18.22	5.69	0.84	2.28
Female	1,223	29	18.43	5.79	0.85	2.24
Male	2,307	29	18.08	5.66	0.84	2.26
Gender Undefined	183	29	18.63	5.39	0.82	2.29
Hispanic or Latino	713	29	17.98	5.67	0.84	2.27
American Indian or Alaska Native	119	29	18.08	5.77	0.84	2.31
Asian						
Black or African American	798	29	17.60	5.61	0.83	2.31
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,616	29	18.80	5.77	0.85	2.23
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	262	29	18.25	5.35	0.82	2.27
Currently receiving LEP services	120	29	18.87	5.65	0.84	2.26
Not receiving LEP services	2,580	29	18.35	5.62	0.84	2.25
LEP: All Other Students	1,013	29	17.81	5.85	0.85	2.27
Economically Disadvantaged Students	1,156	29	19.34	5.61	0.84	2.24
Non-economically Disadvantaged Students	1,523	29	17.69	5.53	0.83	2.28
SES: All Other Students	1,034	29	17.74	5.84	0.85	2.26
Migrant						
Non-migrant	1,196	29	17.92	5.66	0.84	2.26
Undefined Migrant Status	2,515	29	18.36	5.70	0.84	2.28
Augmentative Communication	433	29	14.04	5.06	0.76	2.48
No Augmentative Communication	3,246	29	18.77	5.54	0.84	2.22
Undefined Augmentative Communications						
Hearing Loss	104	29	15.01	6.15	0.85	2.38
Within Normal Limits	3,585	29	18.29	5.65	0.84	2.26
Undefined Hearing Loss						
Visual Impairment	138	29	14.33	6.61	0.87	2.38
Within Normal Limits	3,553	29	18.37	5.61	0.84	2.24
Undefined Visual Impairment						
Sensory Stimuli Response	169	29	12.15	5.78	0.83	2.38
Follow Directions	3,532	29	18.51	5.53	0.83	2.28

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Undefined Receptive Language						
Special School	314	29	14.51	5.72	0.82	2.43
Regular School Self-contained	2,486	29	17.99	5.63	0.83	2.32
Regular School Resource Room	594	29	20.19	5.30	0.83	2.19
Regular School Primarily Self-contained	235	29	20.03	4.47	0.76	2.19
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	138	29	11.95	6.26	0.86	2.34
Uses Intentional Communication	625	29	14.74	5.30	0.79	2.43
Uses Symbolic Language	2,937	29	19.25	5.26	0.82	2.23
Undefined Expressive Communication						

Table J-6. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 8

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,600	31	19.62	5.72	0.83	2.36
Female	1,137	31	19.66	5.63	0.82	2.39
Male	2,278	31	19.63	5.79	0.83	2.39
Gender Undefined	185	31	19.18	5.41	0.80	2.42
Hispanic or Latino	640	31	18.81	5.57	0.81	2.43
American Indian or Alaska Native	129	31	19.46	5.40	0.81	2.35
Asian						
Black or African American	782	31	19.11	5.55	0.81	2.42
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,609	31	20.33	5.80	0.84	2.32
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	265	31	19.61	5.61	0.82	2.38
Currently receiving LEP services	101	31	20.37	4.65	0.73	2.42
Not receiving LEP services	2,492	31	19.75	5.89	0.84	2.36
LEP: All Other Students	1,007	31	19.22	5.36	0.79	2.46

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Economically Disadvantaged Students	1,062	31	21.01	5.66	0.84	2.26
Non-economically Disadvantaged Students	1,514	31	18.91	5.81	0.83	2.40
SES: All Other Students	1,024	31	19.22	5.38	0.80	2.41
Migrant						
Non-migrant	1,118	31	19.30	5.85	0.84	2.34
Undefined Migrant Status	2,480	31	19.76	5.65	0.82	2.40
Augmentative Communication	423	31	15.14	5.18	0.75	2.59
No Augmentative Communication	3,154	31	20.23	5.52	0.82	2.34
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,494	31	19.66	5.72	0.83	2.36
Undefined Hearing Loss						
Visual Impairment	108	31	16.18	6.26	0.84	2.50
Within Normal Limits	3,472	31	19.72	5.66	0.82	2.40
Undefined Visual Impairment						
Sensory Stimuli Response	169	31	14.01	5.29	0.77	2.54
Follow Directions	3,420	31	19.90	5.59	0.82	2.37
Undefined Receptive Language						
Special School	328	31	15.64	5.37	0.77	2.58
Regular School Self-contained	2,408	31	19.47	5.57	0.81	2.43
Regular School Resource Room	536	31	21.36	5.56	0.84	2.22
Regular School Primarily Self-contained	242	31	22.12	5.21	0.82	2.21
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	134	31	14.07	5.17	0.75	2.58
Uses Intentional Communication	596	31	16.25	5.42	0.78	2.54
Uses Symbolic Language	2,859	31	20.58	5.39	0.81	2.35
Undefined Expressive Communication						

Table J-7. 2015–16 MSAA: Reliability: Subgroup-ELA Grade 11

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	2,528	28	19.87	5.23	0.85	2.03
Female	859	28	20.07	4.93	0.83	2.03
Male	1,574	28	19.83	5.34	0.85	2.07
Gender Undefined						
Hispanic or Latino	427	28	18.96	5.15	0.83	2.12
American Indian or Alaska Native						
Asian						
Black or African American	498	28	19.18	5.27	0.84	2.11
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,259	28	20.65	5.10	0.85	1.98
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	177	28	19.14	5.55	0.86	2.08
Currently receiving LEP services						
Not receiving LEP services	1,697	28	20.18	5.12	0.85	1.98
LEP: All Other Students	779	28	19.14	5.43	0.85	2.10
Economically Disadvantaged Students	665	28	21.40	5.02	0.86	1.88
Non-economically Disadvantaged Students	1,078	28	19.46	5.03	0.83	2.07
SES: All Other Students	785	28	19.14	5.42	0.85	2.10
Migrant						
Non-migrant	518	28	20.51	5.15	0.85	1.99
Undefined Migrant Status	2,009	28	19.71	5.24	0.85	2.03
Augmentative Communication	204	28	15.44	5.22	0.79	2.39
No Augmentative Communication	2,313	28	20.26	5.06	0.84	2.02
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	2,443	28	19.94	5.19	0.85	2.01
Undefined Hearing Loss						
Visual Impairment						
Within Normal Limits	2,419	28	19.96	5.17	0.85	2.00
Undefined Visual Impairment						
Sensory Stimuli Response						
Follow Directions	2,449	28	20.07	5.09	0.84	2.04

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Undefined Receptive Language						
Special School	211	28	16.83	5.33	0.82	2.26
Regular School Self-contained	1,658	28	19.44	5.23	0.84	2.09
Regular School Resource Room	463	28	21.72	4.52	0.83	1.86
Regular School Primarily Self-contained	162	28	22.68	3.99	0.81	1.74
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries						
Uses Intentional Communication	284	28	15.73	5.16	0.79	2.36
Uses Symbolic Language	2,181	28	20.62	4.81	0.83	1.98
Undefined Expressive Communication						

Table J-8. 2015–16 MSAA: Reliability: Subgroup-Mathematics Grade 3

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,225	35	18.62	7.10	0.86	2.66
Female	1,017	35	18.53	6.88	0.85	2.66
Male	2,048	35	18.73	7.20	0.87	2.60
Gender Undefined	160	35	17.79	7.33	0.87	2.64
Hispanic or Latino	691	35	18.48	7.13	0.86	2.67
American Indian or Alaska Native	109	35	18.55	7.39	0.87	2.66
Asian						
Black or African American	668	35	17.79	6.84	0.85	2.65
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,332	35	19.24	7.14	0.86	2.67
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	260	35	18.36	7.28	0.87	2.62
Currently receiving LEP services	157	35	19.76	7.34	0.87	2.65
Not receiving LEP services	2,185	35	18.86	7.10	0.86	2.66
LEP: All Other Students	883	35	17.80	7.02	0.86	2.63

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Economically Disadvantaged Students	1,019	35	19.49	7.15	0.87	2.58
Non-economically Disadvantaged Students	1,304	35	18.54	7.08	0.86	2.65
SES: All Other Students	902	35	17.76	6.99	0.85	2.71
Migrant						
Non-migrant	916	35	17.98	6.98	0.85	2.70
Undefined Migrant Status	2,309	35	18.87	7.14	0.86	2.67
Augmentative Communication	446	35	14.55	5.47	0.75	2.73
No Augmentative Communication	2,756	35	19.28	7.11	0.86	2.66
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,141	35	18.68	7.11	0.86	2.66
Undefined Hearing Loss						
Visual Impairment						
Within Normal Limits	3,107	35	18.75	7.10	0.86	2.66
Undefined Visual Impairment						
Sensory Stimuli Response	208	35	12.27	5.86	0.80	2.62
Follow Directions	3,008	35	19.06	6.98	0.86	2.61
Undefined Receptive Language						
Special School	235	35	15.45	6.64	0.84	2.66
Regular School Self-contained	2,155	35	18.05	7.03	0.86	2.63
Regular School Resource Room	462	35	20.28	6.69	0.85	2.59
Regular School Primarily Self-contained	241	35	22.26	6.80	0.86	2.54
Regular School General Education	123	35	21.14	6.92	0.86	2.59
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	144	35	10.55	5.02	0.74	2.56
Uses Intentional Communication	712	35	15.15	6.00	0.80	2.68
Uses Symbolic Language	2,360	35	20.16	6.82	0.85	2.64
Undefined Expressive Communication						

Table J-9. 2015–16 MSAA: Reliability: Subgroup- Mathematics Grade 4

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,435	33	16.26	6.24	0.83	2.57
Female	1,082	33	16.40	6.11	0.82	2.59
Male	2,162	33	16.24	6.31	0.83	2.60
Gender Undefined	191	33	15.73	6.20	0.83	2.56
Hispanic or Latino	688	33	16.13	5.90	0.81	2.57
American Indian or Alaska Native	118	33	16.67	5.92	0.80	2.65
Asian						
Black or African American	690	33	15.43	6.32	0.83	2.61
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,438	33	16.79	6.33	0.84	2.53
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	314	33	16.22	6.29	0.83	2.59
Currently receiving LEP services	132	33	18.16	5.68	0.79	2.60
Not receiving LEP services	2,410	33	16.36	6.22	0.83	2.56
LEP: All Other Students	893	33	15.70	6.31	0.83	2.60
Economically Disadvantaged Students	1,099	33	17.25	6.35	0.84	2.54
Non-economically Disadvantaged Students	1,433	33	15.88	6.04	0.82	2.56
SES: All Other Students	903	33	15.66	6.28	0.83	2.59
Migrant						
Non-migrant	1,046	33	15.80	6.08	0.82	2.58
Undefined Migrant Status	2,387	33	16.46	6.30	0.83	2.60
Augmentative Communication	491	33	12.79	5.14	0.74	2.62
No Augmentative Communication	2,927	33	16.84	6.22	0.83	2.56
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,338	33	16.31	6.25	0.83	2.58
Undefined Hearing Loss						
Visual Impairment	106	33	14.25	6.64	0.85	2.57
Within Normal Limits	3,308	33	16.33	6.20	0.83	2.56
Undefined Visual Impairment						
Sensory Stimuli Response	190	33	11.13	5.39	0.79	2.47
Follow Directions	3,236	33	16.56	6.16	0.82	2.61

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Undefined Receptive Language						
Special School	246	33	12.98	5.96	0.82	2.53
Regular School Self-contained	2,230	33	15.72	6.13	0.82	2.60
Regular School Resource Room	539	33	17.92	5.98	0.82	2.54
Regular School Primarily Self-contained	294	33	18.86	5.91	0.81	2.58
Regular School General Education	117	33	19.35	5.81	0.80	2.60
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	154	33	11.11	4.99	0.74	2.54
Uses Intentional Communication	703	33	13.29	5.86	0.81	2.55
Uses Symbolic Language	2,567	33	17.39	5.99	0.82	2.54
Undefined Expressive Communication						

Table J-10. 2015–16 MSAA: Reliability: Subgroup- Mathematics Grade 5

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,552	35	16.75	5.71	0.78	2.68
Female	1,171	35	16.53	5.62	0.77	2.70
Male	2,192	35	16.84	5.73	0.78	2.69
Gender Undefined	189	35	17.09	5.94	0.80	2.66
Hispanic or Latino	712	35	16.61	5.36	0.74	2.73
American Indian or Alaska Native	123	35	17.60	5.97	0.80	2.67
Asian						
Black or African American	708	35	16.28	5.58	0.76	2.73
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,516	35	17.10	5.94	0.80	2.66
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	294	35	16.80	5.43	0.75	2.71
Currently receiving LEP services	117	35	17.34	5.59	0.76	2.74
Not receiving LEP services	2,455	35	16.88	5.66	0.77	2.71
LEP: All Other Students	980	35	16.36	5.82	0.79	2.67

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Economically Disadvantaged Students	1,089	35	17.50	5.92	0.80	2.65
Non-economically Disadvantaged Students	1,467	35	16.48	5.44	0.75	2.72
SES: All Other Students	996	35	16.33	5.79	0.78	2.72
Migrant						
Non-migrant	1,106	35	16.20	5.54	0.76	2.71
Undefined Migrant Status	2,445	35	17.00	5.77	0.78	2.71
Augmentative Communication	506	35	13.64	4.80	0.68	2.72
No Augmentative Communication	3,022	35	17.26	5.69	0.78	2.67
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,439	35	16.78	5.69	0.78	2.67
Undefined Hearing Loss						
Visual Impairment	145	35	12.32	5.73	0.79	2.63
Within Normal Limits	3,384	35	16.94	5.64	0.77	2.70
Undefined Visual Impairment						
Sensory Stimuli Response	202	35	12.00	5.26	0.75	2.63
Follow Directions	3,343	35	17.03	5.61	0.77	2.69
Undefined Receptive Language						
Special School	316	35	13.97	5.49	0.76	2.69
Regular School Self-contained	2,256	35	16.23	5.47	0.76	2.68
Regular School Resource Room	558	35	18.73	5.79	0.79	2.65
Regular School Primarily Self-contained	294	35	19.20	5.58	0.77	2.68
Regular School General Education	121	35	18.40	5.46	0.76	2.67
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	155	35	11.47	5.47	0.78	2.57
Uses Intentional Communication	633	35	13.78	4.82	0.68	2.73
Uses Symbolic Language	2,757	35	17.72	5.51	0.76	2.70
Undefined Expressive Communication						

Table J-11. 2015–16 MSAA: Reliability: Subgroup-Mathematics Grade 6

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,582	35	18.57	6.62	0.84	2.65
Female	1,219	35	18.40	6.27	0.81	2.73
Male	2,174	35	18.72	6.84	0.85	2.65
Gender Undefined	189	35	17.87	6.23	0.81	2.72
Hispanic or Latino	650	35	18.22	6.61	0.84	2.64
American Indian or Alaska Native	131	35	19.02	6.49	0.83	2.68
Asian						
Black or African American	755	35	17.86	6.40	0.82	2.72
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,569	35	19.18	6.79	0.85	2.63
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	275	35	18.09	6.28	0.81	2.74
Currently receiving LEP services	135	35	19.58	6.93	0.86	2.59
Not receiving LEP services	2,438	35	18.80	6.60	0.84	2.64
LEP: All Other Students	1,009	35	17.86	6.56	0.83	2.70
Economically Disadvantaged Students	1,068	35	19.64	6.59	0.84	2.64
Non-economically Disadvantaged Students	1,483	35	18.31	6.62	0.84	2.65
SES: All Other Students	1,031	35	17.83	6.52	0.83	2.69
Migrant						
Non-migrant	1,098	35	18.10	6.41	0.82	2.72
Undefined Migrant Status	2,483	35	18.77	6.70	0.84	2.68
Augmentative Communication	436	35	15.21	6.16	0.80	2.75
No Augmentative Communication	3,111	35	19.02	6.55	0.84	2.62
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,456	35	18.61	6.62	0.84	2.65
Undefined Hearing Loss						
Visual Impairment	135	35	16.07	5.86	0.78	2.75
Within Normal Limits	3,413	35	18.66	6.64	0.84	2.66
Undefined Visual Impairment						
Sensory Stimuli Response	165	35	13.14	6.00	0.80	2.68
Follow Directions	3,397	35	18.82	6.54	0.83	2.70

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Undefined Receptive Language						
Special School	288	35	15.72	6.14	0.80	2.75
Regular School Self-contained	2,431	35	18.05	6.38	0.82	2.71
Regular School Resource Room	515	35	20.55	7.06	0.87	2.55
Regular School Primarily Self-contained	248	35	21.49	6.32	0.84	2.53
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	117	35	12.67	6.73	0.85	2.61
Uses Intentional Communication	642	35	15.51	5.67	0.76	2.78
Uses Symbolic Language	2,803	35	19.51	6.49	0.83	2.68
Undefined Expressive Communication						

Table J-12. 2015–16 MSAA: Reliability: Subgroup- Mathematics Grade 7

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,701	34	17.89	6.32	0.83	2.61
Female	1,219	34	17.85	6.29	0.83	2.59
Male	2,301	34	17.93	6.36	0.83	2.62
Gender Undefined	181	34	17.62	6.01	0.81	2.62
Hispanic or Latino	711	34	17.71	6.41	0.83	2.64
American Indian or Alaska Native	119	34	18.13	5.97	0.80	2.67
Asian						
Black or African American	793	34	17.22	6.12	0.81	2.67
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,612	34	18.45	6.55	0.84	2.62
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	262	34	17.30	5.64	0.77	2.70
Currently receiving LEP services	120	34	19.52	6.16	0.82	2.61
Not receiving LEP services	2,576	34	18.09	6.31	0.83	2.60
LEP: All Other Students	1,005	34	17.17	6.31	0.82	2.68
Economically Disadvantaged Students	1,154	34	18.95	6.53	0.84	2.61

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Non-economically Disadvantaged Students	1,521	34	17.59	6.11	0.81	2.66
SES: All Other Students	1,026	34	17.12	6.26	0.82	2.66
Migrant						
Non-migrant	1,192	34	17.66	6.10	0.81	2.66
Undefined Migrant Status	2,507	34	17.99	6.43	0.83	2.65
Augmentative Communication	433	34	14.35	4.99	0.71	2.69
No Augmentative Communication	3,234	34	18.34	6.34	0.83	2.61
Undefined Augmentative Communications						
Hearing Loss	102	34	15.63	6.59	0.84	2.64
Within Normal Limits	3,575	34	17.94	6.30	0.83	2.60
Undefined Hearing Loss						
Visual Impairment	133	34	14.66	6.17	0.82	2.62
Within Normal Limits	3,546	34	18.01	6.29	0.82	2.67
Undefined Visual Impairment						
Sensory Stimuli Response	165	34	12.71	5.52	0.77	2.65
Follow Directions	3,524	34	18.13	6.25	0.82	2.65
Undefined Receptive Language						
Special School	307	34	14.45	5.51	0.76	2.70
Regular School Self-contained	2,480	34	17.62	6.12	0.81	2.67
Regular School Resource Room	595	34	19.68	6.78	0.86	2.54
Regular School Primarily Self-contained	236	34	19.92	5.88	0.80	2.63
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	133	34	12.41	5.91	0.80	2.64
Uses Intentional Communication	627	34	15.09	5.36	0.75	2.68
Uses Symbolic Language	2,928	34	18.73	6.24	0.82	2.65
Undefined Expressive Communication						

Table J-13. 2015–16 MSAA: Reliability: Subgroup- Mathematics Grade 8

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	3,590	35	17.52	6.79	0.84	2.72
Female	1,136	35	17.26	6.52	0.83	2.69
Male	2,271	35	17.70	6.96	0.85	2.70
Gender Undefined	183	35	16.99	6.32	0.82	2.68
Hispanic or Latino	638	35	17.37	6.55	0.83	2.70
American Indian or Alaska Native	129	35	16.53	6.31	0.81	2.75
Asian						
Black or African American	777	35	16.94	6.45	0.82	2.74
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	1,609	35	18.02	7.08	0.86	2.65
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	263	35	17.48	6.48	0.83	2.67
Currently receiving LEP services	101	35	19.93	6.03	0.80	2.70
Not receiving LEP services	2,481	35	17.59	6.88	0.85	2.66
LEP: All Other Students	1,008	35	17.09	6.58	0.83	2.71
Economically Disadvantaged Students	1,057	35	19.02	7.04	0.86	2.63
Non-economically Disadvantaged Students	1,508	35	16.78	6.59	0.83	2.72
SES: All Other Students	1,025	35	17.06	6.57	0.83	2.71
Migrant						
Non-migrant	1,111	35	17.01	6.83	0.85	2.65
Undefined Migrant Status	2,477	35	17.75	6.76	0.84	2.70
Augmentative Communication	419	35	13.37	6.12	0.81	2.67
No Augmentative Communication	3,148	35	18.07	6.68	0.84	2.67
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	3,485	35	17.53	6.80	0.84	2.72
Undefined Hearing Loss						
Visual Impairment	106	35	12.95	7.09	0.87	2.56
Within Normal Limits	3,464	35	17.65	6.73	0.84	2.69
Undefined Visual Impairment						
Sensory Stimuli Response	165	35	12.20	5.65	0.78	2.65
Follow Directions	3,414	35	17.78	6.74	0.84	2.70

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Undefined Receptive Language						
Special School	326	35	13.58	5.83	0.78	2.73
Regular School Self-contained	2,402	35	17.29	6.62	0.83	2.73
Regular School Resource Room	534	35	19.32	6.87	0.85	2.66
Regular School Primarily Self-contained	242	35	20.13	6.92	0.86	2.59
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries	131	35	11.94	5.53	0.77	2.65
Uses Intentional Communication	594	35	14.77	6.33	0.82	2.69
Uses Symbolic Language	2,854	35	18.34	6.67	0.84	2.67
Undefined Expressive Communication						

Table J-14. 2015–16 MSAA: Reliability: Subgroup-Mathematics Grade 11

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
All Students	1,724	34	16.53	6.29	0.82	2.67
Female	606	34	16.78	6.43	0.83	2.65
Male	1,062	34	16.39	6.23	0.82	2.64
Gender Undefined						
Hispanic or Latino	380	34	15.71	5.97	0.80	2.67
American Indian or Alaska Native						
Asian						
Black or African American	255	34	16.23	5.87	0.79	2.69
Native Hawaiian or Pacific Islander						
White (non-Hispanic)	823	34	17.05	6.52	0.84	2.61
Two or More Races (non-Hispanic)						
No Primary race/Ethnicity Undefined	111	34	16.00	5.59	0.77	2.68
Currently receiving LEP services						
Not receiving LEP services	1,325	34	16.70	6.35	0.82	2.69
LEP: All Other Students	370	34	15.60	5.99	0.80	2.68

continued

Description	Number of Students	Raw Score			Alpha	Standard Error
		Maximum	*Mean*	*Standard Deviation*		
Economically Disadvantaged Students	372	34	18.47	6.68	0.85	2.59
Non-economically Disadvantaged Students	976	34	16.12	6.09	0.81	2.65
SES: All Other Students	376	34	15.68	6.04	0.81	2.63
Migrant						
Non-migrant	516	34	17.61	6.54	0.84	2.62
Undefined Migrant Status	1,207	34	16.06	6.13	0.81	2.67
Augmentative Communication	164	34	13.61	5.68	0.78	2.66
No Augmentative Communication	1,553	34	16.83	6.27	0.82	2.66
Undefined Augmentative Communications						
Hearing Loss						
Within Normal Limits	1,662	34	16.51	6.23	0.82	2.64
Undefined Hearing Loss						
Visual Impairment						
Within Normal Limits	1,654	34	16.67	6.24	0.82	2.65
Undefined Visual Impairment						
Sensory Stimuli Response						
Follow Directions	1,677	34	16.69	6.22	0.82	2.64
Undefined Receptive Language						
Special School	195	34	14.68	5.50	0.76	2.69
Regular School Self-contained	1,064	34	16.07	6.14	0.81	2.68
Regular School Resource Room	325	34	18.24	6.58	0.84	2.63
Regular School Primarily Self-contained	110	34	19.21	6.38	0.83	2.63
Regular School General Education						
Undefined Classroom Setting						
Student Communicates Primarily Through Cries						
Uses Intentional Communication	189	34	13.24	5.04	0.72	2.67
Uses Symbolic Language	1,485	34	17.13	6.22	0.82	2.64
Undefined Expressive Communication						

APPENDIX K—DECISION ACCURACY AND CONSISTENCY RESULTS

Table K-1. 2015–16 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	0.72 (0.65)	0.49	0.87 (0.81)	0.27 (0.20)	0.45 (0.35)	0.84 (0.74)
	4	0.73 (0.65)	0.48	0.87 (0.82)	0.29 (0.21)	0.64 (0.54)	0.81 (0.67)
	5	0.68 (0.59)	0.43	0.85 (0.77)	0.35 (0.27)	0.58 (0.47)	0.80 (0.67)
	6	0.71 (0.64)	0.48	0.87 (0.81)	0.42 (0.32)	0.54 (0.43)	0.82 (0.71)
	7	0.72 (0.65)	0.48	0.87 (0.81)	0.19 (0.14)	0.57 (0.46)	0.82 (0.70)
	8	0.67 (0.59)	0.43	0.85 (0.77)	0.42 (0.33)	0.41 (0.32)	0.81 (0.70)
	11	0.73 (0.66)	0.51	0.87 (0.80)	0.19 (0.14)	0.57 (0.46)	0.85 (0.76)
Mathematics	3	0.74 (0.67)	0.51	0.88 (0.83)	0.19 (0.14)	0.55 (0.44)	0.84 (0.74)
	4	0.69 (0.62)	0.45	0.86 (0.80)	0.31 (0.24)	0.45 (0.35)	0.82 (0.71)
	5	0.64 (0.55)	0.39	0.82 (0.73)	0.37 (0.29)	0.54 (0.44)	0.78 (0.62)
	6	0.72 (0.65)	0.49	0.87 (0.81)	0.29 (0.21)	0.40 (0.31)	0.84 (0.75)
	7	0.67 (0.59)	0.43	0.84 (0.76)	0.35 (0.27)	0.53 (0.42)	0.81 (0.69)
	8	0.71 (0.64)	0.48	0.86 (0.80)	0.29 (0.22)	0.41 (0.32)	0.84 (0.75)
	11	0.70 (0.64)	0.47	0.86 (0.80)	0.28 (0.21)	0.39 (0.30)	0.83 (0.74)

Table K-2. 2015–16 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	0.88 (0.83)	0.06	0.06	0.88 (0.83)	0.06	0.06	0.90 (0.86)	0.06	0.04
	4	0.88 (0.83)	0.06	0.06	0.88 (0.83)	0.06	0.06	0.93 (0.90)	0.05	0.03
	5	0.87 (0.82)	0.06	0.07	0.87 (0.81)	0.07	0.07	0.91 (0.87)	0.06	0.03
	6	0.88 (0.84)	0.06	0.06	0.89 (0.84)	0.06	0.05	0.92 (0.89)	0.05	0.03
	7	0.87 (0.83)	0.06	0.07	0.87 (0.82)	0.06	0.06	0.91 (0.87)	0.06	0.04
	8	0.87 (0.82)	0.06	0.07	0.87 (0.82)	0.07	0.06	0.90 (0.85)	0.06	0.04
	11	0.89 (0.84)	0.05	0.06	0.88 (0.84)	0.06	0.06	0.90 (0.86)	0.06	0.04
Mathematics	3	0.88 (0.84)	0.06	0.06	0.88 (0.84)	0.06	0.06	0.91 (0.87)	0.05	0.04
	4	0.87 (0.82)	0.06	0.07	0.87 (0.82)	0.07	0.06	0.90 (0.86)	0.06	0.04
	5	0.86 (0.80)	0.06	0.08	0.85 (0.79)	0.08	0.07	0.90 (0.86)	0.07	0.03
	6	0.88 (0.83)	0.06	0.06	0.88 (0.83)	0.06	0.06	0.90 (0.86)	0.06	0.04
	7	0.87 (0.82)	0.06	0.07	0.87 (0.81)	0.07	0.07	0.90 (0.85)	0.06	0.04
	8	0.88 (0.83)	0.06	0.06	0.88 (0.83)	0.06	0.06	0.89 (0.85)	0.06	0.05
	11	0.87 (0.82)	0.06	0.07	0.87 (0.82)	0.06	0.06	0.89 (0.85)	0.06	0.05

APPENDIX L—ACCOMMODATION FREQUENCIES

Table L-1. 2015–16 MSAA: Accommodation Frequencies—Mathematics

Accommodations	Grades						
	3	*4*	*5*	*6*	*7*	*8*	*11*
LCI_Vision[1]	221	212	268	256	263	224	116
SAR_Assistive_Presentation_After[2]	220	263	243	244	273	237	155
SAR_Assistive_Response_After[2]	284	302	316	271	290	292	137
SAR_No_Accomm_Needed_After[3]	727	859	938	1,009	1,146	1,114	642
SAR_Paper_Version_After[4]	475	467	413	351	301	316	141
SAR_Scribe_After[5]	1,293	1,303	1,286	1,209	1,126	1,080	335
SAR_Sign_Interpretation_After[6]	53	51	60	34	35	26	28

[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_Presentation_After and 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 NCSC 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 NCSC Assessment System.
[6]: SAR_Sign_Interpretation_After - TA may communicate passages, items and response options using sign language to student.

Table L-2. 2015–16 MSAA: Accommodation Frequencies—ELA

Accommodations	Grades						
	3	*4*	*5*	*6*	*7*	*8*	*11*
LCI_Vision[1]	215	214	270	252	262	224	153
SAR_Assistive_Presentation_After[2]	219	259	242	241	270	235	181
SAR_Assistive_Response_After[2]	282	299	315	270	290	289	160
SAR_No_Accomm_Needed_After[3]	719	858	935	1,006	1,145	1,116	956
SAR_Paper_Version_After[4]	479	464	416	350	302	314	191
SAR_Scribe_After[5]	1,299	1,300	1,284	1,207	1,127	1,074	513
SAR_Sign_Interpretation_After[6]	52	50	62	34	35	24	34

[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_Presentation_After and 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 NCSC 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 NCSC Assessment System.
[6]: SAR_Sign_Interpretation_After - TA may communicate passages, items and response options using sign language to student.

Table L-3. 2015–16 MSAA: Accommodation Summary

Content Area	Grade	Number of Students Tested	
		With	*Without*
ELA	03	2,314	1,358
	04	2,518	1,309
	05	2,589	1,332
	06	2,605	1,397
	07	2,677	1,441
	08	2,562	1,422
	11	1,775	1,016
Mathematics	03	2,324	1,367
	04	2,524	1,322
	05	2,593	1,350
	06	2,613	1,406
	07	2,681	1,438
	08	2,570	1,423
	11	1,227	737

APPENDIX M—TECHNICAL ADVISORY COMMITTEE MEMBERS

Table M-1. 2015-16 MSAA: Technical Advisory Committee Members

Name	Organization	Expertise
Mike Russell	Boston College	Technical / Accessibility
Rachel Quenemoen	National Center on Educational Outcomes	Student With Significant Cognitive Disabilities/NCSC Awareness
Martha Thurlow	University of Minnesota/NCEO	Martha: Special Education/Accessibility Brian:
Joseph Martineau	The National Center for the Improvement of Educational Assessment	Psychometrics /Computer Adaptive Testing
Derek Briggs	University of Colorado	Assessment / Growth / Psychometrics

APPENDIX N—ITEM REVIEW AND BIAS AND SENSITIVITY REVIEW COMMITTEE MEMBERS

MSAA 2016 Item Content and Bias Review Meeting
Final Panelist List

Reading Content Grades 3-6

Name	City
Beringer, Amanda	Falls Church, VA
Charns, Daniel	Phoenix, AZ
Connor, Chantal	St. Croix, VI
Cruz, Helene	Hagatna, GU
Goodman, Genevive	Germantown, MD
Griffin, Doretta	Hamburg, AR
Hinners, Heather	Sioux Falls, SD
Lemme, Michelle	Cranston, RI
Lowney, Lisa	Helena, MT
Wishart, Erica	Rising Sun, MD

Reading Content Grades 7, 8 & 11

Name	City
Benamati, Joseph	Salisbury, MD
Charleswell, Renee	St. Thomas, VI
Del Rosario, Tracy Lynn	Saipan, MP
Foreman, Heidi	Helena, MT
Franklin, Rebecca	Kingman, AZ
Gordon, Benita	Bowie, MD
Maitland, Melody	Washington, DC
Nelson, Matthew	Nashville, TN
Verrill, Meredith	Scarborough, ME
Volpe, Kendra	N. Providence, RI

Reading Bias All Grades

Name	City
Cropper, Bess	Snow Hill, MD
Drake, Donald	Poland, ME
Gross, Rhonda Doris	Saipan, MP
Lane, Melissa	Belgrade, MT
Manley, Elmie	Hagatna, GU
Mugavin, Marybeth	Memphis, TN
Murtagh, Hilary	Nashvillet, TN
Salcido, Cecilia	Glendale, AZ
Stabnow, Erin	Sioux Falls, SD
Turbanada-Umlas, Tara Ann	St. Thomas, US

Writing Content All Grades

Name	City
Angles-Figueroa, Lourdes	Christiansted, VI
Bowerman, Margaret	Phoenix, AZ
Brown, Jessie	Fruitland, MD
Borden, Ryan	Warwick, RI
Gutierrez, Christine	Phoenix, AZ
Merchen, Lisa	Spearfish, SD
Rybka, Tiffany	Riverside, RI
Satterfield, Susan	Knoxville, TN
Stripling, Robin	Little Rock, AR
Toothman, Kristie	Portland, ME

Writing Bias All Grades

Name	City
Barrett, Kari	Spring Hill, TN
Bell, Angela	Manchester, MD
Fox, Maureen	Peaks Island, ME
Hanson, Melissa	Bozeman, MT
Karls, Marcia	Phoenix, AZ
Kelk, Pam	Phoenix, AZ
Moncure, Lora	Yarmouth, ME
Sjostrom, Kelly	Malta, MT
Smith, Charmelle	Arlington, VA
Underdown, Rachel	Springdale, AR

Math Content Grades 3-6

Name	City
Anderson, Ann	Brunswick, ME
Bahti, Jill	Phoenix, AZ
Davis, Sam	Arlington, TN
Falk, Elsbeth	Rapid City, SD
Hofschneider, Lizabeth	Saipan, MP
Howard, Carla	St. Thomas, VI
Kempler, Sara	Baltimore, MD
Ritchie, Audra	Baltimore, MD
Veazey, Joy	Brinkley, AR
Windsor, Lauren	Cumberland, RI

Math Content Grades 7, 8 & 11

Name	City
Bidwell, Renea	Broadus, MT
Durham, Murphy	Pleasant View, TN
open slot	
Hollinger, Carissa	Rehoboth, DE
McIntosh, Mary	Frederiksted, VI
Orcullo, Daniel	Washington, DC
Quiogue, Maria Elena	Yigo, GU
Skwirz, Kristen	Coventry, RI
Storey, Connie	Little Rock, AR
Whitlock, Becky	Harrisburg, SD

Math Bias All Grades

Name	City
Ayers, Jennifer	Elkton, MD
Brooks, Suzanne	Washington, DC
Browning, Adam	Washington, DC
Burrell, Andy	Kalispell, MT
Eilertson, Jessica	Phoenix, AZ
Guerrero, Valerie	Dededo, GU
Linder, Jennifer	Little Rock, AR
Phillips, Windy	Berlin, MD
Saraiva, Darlene	Chepachet, RI
St. Prix, Mackisha	Kingshill, VI

MSAA State Leads

Name	State
Adams, LaNysha	Washington DC
Ahumada, Audra	Arizona
Baltimore-Hookfin, Alexandria	US Virgin Islands
Field, Yvonne	Montana
Hogue, Jared	Arkansas
Morrison, Ben	South Dakota
Nay, Sue	Maine
Nixon, Lori	Tennessee
Schmitt, Nancy	Maryland

APPENDIX O—PARTICIPATION RATES

Table O-1. 2015–16 MSAA: Summary of Participation by Demographic Category—Mathematics

Description	Tested		Total Tested	Total Percent
	# Complete	*# No Observable Mode of Communication*[1]		
All Students	22,809	2,766	25,575	100.00
Female	7,450	1,005	8,455	33.06
Male	14,210	1,625	15,835	61.92
Gender Undefined	1,149	136	1,285	5.02
Hispanic or Latino	4,470	554	5,024	19.64
American Indian or Alaska Native	805	84	889	3.48
Asian	512	83	595	2.33
Black or African American	4,646	572	5,218	20.40
Native Hawaiian or Pacific Islander	177	13	190	0.74
White (non-Hispanic)	9,899	1,206	11,105	43.42
Two or More Races (non-Hispanic)	521	58	579	2.26
No Primary race/Ethnicity Undefined	1,779	196	1,975	7.72
Currently receiving LEP services	791	58	849	3.32
Not receiving LEP services	15,870	1,852	17,722	69.29
LEP: All Other Students	6,148	856	7,004	27.39
Economically Disadvantaged Students	6,858	696	7,554	29.54
Non-economically Disadvantaged Students	9,692	1,207	10,899	42.62
SES: All Other Students	6,259	863	7,122	27.85
Migrant	9	0	9	0.04
Non- migrant	6,985	806	7,791	30.46
Undefined Migrant Status	15,815	1,960	17,775	69.5
Augmentative Communication	2,895	907	3,802	14.87
No Augmentative Communication	19,751	1,821	21,572	84.35
Undefined Augmentative Communications	163	38	201	0.79
Hearing Loss	576	353	929	3.63
Within Normal Limits	22,096	2,382	24,478	95.71
Undefined Hearing Loss	137	31	168	0.66
Visual Impairment	778	782	1,560	6.1
Within Normal Limits	21,876	1,942	23,818	93.13
Undefined Visual Impairment	155	42	197	0.77

continued

Description	Tested		Total Tested	Total Percent
	# Complete	*# No Observable Mode of Communication*[1]		
Sensory Stimuli Response	1,140	1,821	2,961	11.58
Follow Directions	21,599	915	22,514	88.03
Undefined Receptive Language	70	30	100	0.39
Special School	1,913	696	2,609	10.20
Regular School Self-contained	15,018	1,865	16,883	66.01
Regular School Resource Room	3,528	141	3,669	14.35
Regular School Primarily Self-contained	1,665	17	1,682	6.58
Regular School General Education	615	13	628	2.46
Undefined Classroom Setting	70	34	104	0.41
Student Communicates Primarily Through Cries	882	1,730	2,612	10.21
Uses Intentional Communication	4,100	747	4,847	18.95
Uses Symbolic Language	17,754	258	18,012	70.43
Undefined Expressive Communication	73	31	104	0.41

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

Table O-2. 2015–16 MSAA: Summary of Participation by Demographic Category—ELA

Description	Tested		Total Tested	Total Percent
	# Complete	*# No Observable Mode of Communication*[1]		
All Students	23,609	2,706	26,315	100.00
Female	7,697	996	8,693	33.03
Male	14,739	1,579	16,318	62.01
Gender Undefined	1,173	131	1,304	4.96
Hispanic or Latino	4,525	537	5,062	19.24
American Indian or Alaska Native	809	75	884	3.36
Asian	517	76	593	2.25
Black or African American	4,901	564	5,465	20.77
Native Hawaiian or Pacific Islander	176	13	189	0.72
White (non-Hispanic)	10,335	1,194	11,529	43.81
Two or More Races (non-Hispanic)	524	54	578	2.20
No Primary race/Ethnicity Undefined	1,822	193	2,015	7.66
Currently receiving LEP services	814	56	870	3.31
Not receiving LEP services	16,267	1,819	18,086	68.73
LEP: All Other Students	6,528	831	7,359	27.97
Economically Disadvantaged Students	7,161	682	7,843	29.80
Non-economically Disadvantaged Students	9,810	1,185	10,995	41.78
SES: All Other Students	6,638	839	7,477	28.41
Migrant	9	0	9	0.03
Non- migrant	6,997	782	7,779	29.56
Undefined Migrant Status	16,603	1,924	18,527	70.40
Augmentative Communication	2,927	875	3,802	14.45
No Augmentative Communication	20,515	1,793	22,308	84.77
Undefined Augmentative Communications	167	38	205	0.78
Hearing Loss	598	351	949	3.61
Within Normal Limits	22,870	2,324	25,194	95.74
Undefined Hearing Loss	141	31	172	0.65
Visual Impairment	814	776	1,590	6.04
Within Normal Limits	22,629	1,887	24,516	93.16
Undefined Visual Impairment	166	43	209	0.79

continued

Description	Tested		Total Tested	Total Percent
	# Complete	*# No Observable Mode of Communication*[1]		
Sensory Stimuli Response	1,186	1,794	2,980	11.32
Follow Directions	22,352	882	23,234	88.29
Undefined Receptive Language	71	30	101	0.38
Special School	1,933	669	2,602	9.89
Regular School Self-contained	15,606	1,838	17,444	66.29
Regular School Resource Room	3,661	137	3,798	14.43
Regular School Primarily Self-contained	1,720	16	1,736	6.60
Regular School General Education	618	12	630	2.39
Undefined Classroom Setting	71	34	105	0.40
Student Communicates Primarily Through Cries	905	1,703	2,608	9.91
Uses Intentional Communication	4,186	726	4,912	18.67
Uses Symbolic Language	18,444	246	18,690	71.02
Undefined Expressive Communication	74	31	105	0.40

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

Table O-3. 2015–16 MSAA: Participation Rates by Subgroup

Description	Total Tested	Invalidated	Did Not Test
ELA	26,315	231	1,138
Mathematics	25,575	149	1,878