New Mexico MSSA & ASR 2023 Technical Report

Grades 3–8 ELA and Mathematics

Grades 5, 8, and 11 Science

Prepared by Cognia and the New Mexico Public Education Department

NM-MSSA

NM-ASR

NEW MEXICO ASSESSMENT OF SCIENCE READINESS

cognia

Table of Contents

Chapter 1. Introduction to the Assessment Programs

1.1 Purposes and Uses of the NM-MSSA and NM-ASR Assessment Programs

The New Mexico Measures of Student Success and Achievement (NM-MSSA—see Appendix A for a list of acronyms) is New Mexico's statewide summative assessment for English Language Arts (ELA) and Mathematics, administered at the end of grades 3–8. The NM-MSSA also includes the statewide summative assessments administered in Spanish for qualifying students, i.e., Spanish Language Arts (SLA) and Mathematics. The NM-MSSA is designed to provide evidence to determine a student's grade-level proficiency and progress toward college and/or career readiness, as defined by the State, by showing he or she has mastered the New Mexico Common Core State Standards. Similarly, the New Mexico Assessment of Science Readiness (NM-ASR—see Appendix A for a list of acronyms) is New Mexico's statewide summative assessment for science in grades 5, 8, and 11. The NM-ASR is designed to provide evidence to determine a student's grade-level proficiency and progress toward college and/or career readiness, as defined by the State, by showing he or she has mastered the New Mexico STEM Ready! Science Standards. Like NM-MSSA, the NM-ASR assessments have both English and Spanish forms. Both the NM-MSSA and NM-ASR are key components of New Mexico's ESSA (Every Student Succeeds Act) plan to meet ESSA's general assessment requirements.

As the NM-MSSA and NM-ASR is a single measure at the end of a grade, interpretations and uses of test scores should be supplemented with additional measures, including information from classroom summative, interim, and formative assessments in ELA and mathematics and science. In keeping with the practices set forth in *Standards for Educational and Psychological Testing*, each student's test score should be used as part of a body of evidence regarding mastery and should not be used in isolation to make high-stakes decisions (AERA, APA, & NCME, 2014). Hence, aggregation of student scores on the NM-MSSA and NM-ASR at the school, district, or state levels is generally a more reliable indicator of program success, particularly when monitored over the course of several years.

The New Mexico MSSA and ASR Assessments were administered statewide in an operational setting for the first time in spring 2022. As a result, standards could be set after administering the first operational test. Due to COVID-19, the standard setting could not happen earlier. The Standard Setting workshops were held in July 2022. As a result, we now have set performance standards and cut scores for MSSA and ASR. These cut scores determine the level of performance on each test that corresponds to the knowledge, skills, and abilities (KSAs) that students must demonstrate to be classified into each of the performance levels: Advanced, Proficient, Nearing Proficiency, and Novice.

The NM-MSSA and NM-ASR Assessments are part of New Mexico's Balanced Assessment System, designed to provide point-in-time information about the academic achievement and progress of New Mexico students. Student results are reported according to academic achievement descriptors utilizing scale scores for each of four performance levels: Advanced, Proficient, Nearing Proficiency, and Novice. The results from these assessments provide educators and the public with information to guide the creation of future educational practices to meet the needs of students, while monitoring the continuous improvement efforts of schools, districts, and the state in achieving a world-class education system for all students.

The NM-MSSA English Language Arts (ELA) and Spanish Language Arts (SLA) assessments focused on reading skills related to the comprehension and analysis of texts, the analysis of pieces of writing and knowledge of standard language conventions, and the production of writing while using standard language conventions. Mathematics assessments focused on applying skills and concepts and understanding multi-step problems that require abstract reasoning and modeling real-word problems, precision, perseverance, and strategic use of tools. In both content areas, students were to demonstrate their acquired skills and knowledge by answering various types of questions such as selected-response items, multiple-select items, evidence-based selected-response items, and open-response items. Given that the number of students per grade who took an SLA assessment was at most 35, a mode study comparing ELA to SLA assessments was not feasible. Additionally, a mode study comparing CBT to PBT was not feasible given the small number of PBTs (i.e., fewer than 15 per content area in grades 5, 7, and 8; fewer than 50 per content area in grades 4 and 6; and fewer than 150 per content area in grade 3).

The Assessment of Science Readiness focused on the integration and application of disciplinary core ideas, science and engineering practices, and crosscutting concepts in order to engage in sense-making around scientific phenomena and engineering design problems. Students were to demonstrate their acquired skills and ability by answering various types of questions such as multiple-choice items, multiple-select items, technology-enhanced items, and open-ended items. Many of the items were grouped together in clusters with a common stimulus, to allow for better assessment of the depth of the constructs in the standards.

1.2 Statements of Intended Score Interpretations and Uses (SIUs)

The phrase “intended score interpretations for uses” appears several times in *Standards for Educational and Psychological Testing* and is the core of the field’s views on validity and validation. For the NM-MSSA, NM-ASR, and other assessment programs, the phrase refers broadly to test scores (e.g., total test scale scores, aggregations of test scores, the percentages of students at or above Standard), and other test performance information elements, such as the definition of “novice,” “nearing proficiency,” “proficient,” and “advanced” in the performance level descriptors (PLDs). For a complete list of all PLDs for both programs, please see Appendix B.

1.2.1 Primary Intended NM-MSSA and ASR Score Interpretations and Uses

- Educators, administrators, and other stakeholders at the state, district, and school levels can use the NM-MSSA and ASR and their results to (a) monitor trends in student performance, (b) design professional development for teachers, and (c) drive accountability results.
- Teachers can use the NM-MSSA and ASR and their results to better integrate assessment with their instructional planning.
- Parents can use the NM-MSSA and ASR and their results to get information about what their child knows and can do regarding the New Mexico Common Core State Standards and the New Mexico STEM Ready! Science Standards.

The intended score interpretation and uses stated here align with the original statements of intended score interpretations and uses in the National Center and State Collaborative 2015 Operational Assessment Technical Manual (see the “claim” and “uses” statements on page 8).

The NM-MSSA and ASR Assessments are designed, developed, and implemented to support three intended SIUs, according to the broad interpretation of the phrase above. These interpretations and uses

are applicable to assessments in general and to specific applications with individual students and groups of students, as described below.

SIU 1: Intended Score Interpretation

The NM-MSSA and ASR Assessments provide reliable and valid information about important knowledge and skills in grade-level reading, language usage, mathematics, and science attained by general education students.

- Claim 1.1: The content of the tests represents the content of the standards.
- Claim 1.2: The test items are construct-relevant.
- Claim 1.3: Test scores on the NM-MSSA and ASR Assessments provide reliable information about student performance and accurate classifications into performance levels.
- Claim 1.4: Item and test scoring are implemented accurately; approved scoring rules are implemented accurately.

SIU 2: Intended Score Use for Individual Students

Scale scores can be used to compare an individual student's performance to the performance of other students in the school, district, and state.

- Claim 2.1: Educators and school and district administrators can use results from the NM-MSSA and ASR Assessments to describe and monitor student achievement status with respect to mastery of the content standards.

SIU 3: Intended Score Use for Groups of Students

SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.

- Claim 3.1: Educators can use results from the NM-MSSA and ASR Assessments to support instructional planning for groups of students.
- Claim 3.2: Schools, districts, and state-level stakeholders can use results from the NM-MSSA and ASR Assessments to make comparisons between organizations (e.g., schools, districts).

Claims, subclaims, and evidence that support the intended interpretations and uses of NM-MSSA and ASR scores are provided in Chapter 11.

1.2.2 Unintended Score Interpretations and Uses

Where unintended interpretations and uses may be in use, it is the responsibility of that user to provide supporting evidence (as specified in *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-MSSA and ASR scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations, misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following.

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and proficiency-level classifications—are estimates accompanied by a standard error. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Interpreting and using NM MSSA and ASR scores correctly requires considering the width of the margin of error around a score. For example, students with a scale score 2 points below the cut score for the Proficient level could, hypothetically, have scored above the Proficient cut score on a different day because the NM MSSA and ASR scale score standard errors are expected to be 2–3 points. Interpretations of NM MSSA and ASR scores should account for the margin of error around each score estimate.

Drawing Conclusions and Making Decisions Based Solely on NM MSSA and ASR Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less-than-optimal decisions, such as students failing to receive additional instruction based solely on their NM MSSA and ASR score or teacher teams not being eligible for additional professional development based solely on their students' NM MSSA and ASR scores. Interpretations and uses of NM MSSA and ASR scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators and Item-Level Performance Information

Subdomain indicators (e.g., Literary Text, Operations and Algebraic Thinking, Practices and Crosscutting Concepts in Life Sciences) are based on fewer items than are NM MSSA and ASR total test scores. As a result, they are less-stable estimates of student achievement and learning needs in that subdomain. In addition, because the performance indicators for subdomains are highly correlated, differences in those performance indicators may be smaller than the proficiency level labels may suggest. Interpretations and uses of indicator scores should be supplemented with additional information.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. We assume that students who currently are performing at the Proficient and Advanced levels will continue at these levels only with sustained effort and support. It would be unwise—and unfair—to assume that students who currently are performing at the Novice and Nearing Proficiency levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.

Misinterpretations about students' current proficiency levels and future performance is not really a misinterpretation of NM MSSA and ASR scores. It is a logical error in concluding that current performance determines future performance.

Overinterpreting NM MSSA and ASR Scores as Indicators of College and Career Readiness

The New Mexico content standards are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM MSSA and ASR indicates readiness for college and career is supported only by the evidence contained in the content standards. NM MSSA and ASR scores can also be interpreted as predictors of future performance in college and

career training. However, until empirical prediction studies are completed, this interpretation should be made with caution and with attention to the strong, but limited, evidence in the content standards.

Claims, subclaims, and evidence that support the intended interpretations and uses of test scores are provided in Chapter 11. For additional information regarding the score interpretations and uses, refer to the published SIU statements in Appendix B.

1.3 Introduction to Validity Arguments for the Program: Rationales for the Approach

This report documents test development procedures and psychometric outcomes for the 2023 NM-MSSA and NM-ASR. These technical aspects of the 2023 NM-MSSA and NM-ASR programs contribute to the accumulation of validity evidence to support the NM-MSSA and NM-ASR score interpretations and uses. Because the intended interpretations and uses of test scores, not the test itself, are evaluated for validity, this report presents documentation to substantiate intended interpretations and uses (AERA et al., 2014). Subsequent chapters of this report discuss test development, test alignment, test administration, scoring, equating, item analyses, reliability, scale scores, performance levels, and reporting. Each of these topics contributes important information toward establishing the validity of intended score interpretations and uses of the reported scores from these assessment programs. *Standards for Educational and Psychological Testing* (AERA et al., 2014) also gives a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. These sources address different aspects of supporting evidence for validity arguments; they are not distinct types of validity. Instead, each contributes to a body of evidence about the overall validity and supportability of intended score interpretations and uses. Moreover, these sources represent only a partial list of sources of evidence from the design, development, test administration, analysis, and reporting processes that are relevant to the overall validity arguments for intended interpretations and uses of NM-MSSA and NM-ASR scores and other information. This report does not include certain aspects of an even more comprehensive validity argument that could be important to consider when drawing conclusions about validity of interpretations and uses of NM-MSAA and NM-ASR scores. For example, additional sources of validity evidence might speak to the extent to which NM-MSSA and NM-ASR scores converge with other measures of the same or similar constructs and diverge from measures of different constructs and consequences that arise from scores at the student, school, district, and state levels.

Chapter 2. Overview of the Program

2.1 History of the Programs

This chapter provides a general overview of both NM-MSSA and NM-ASR assessment programs in the state of New Mexico that took place in school year 2022-2023.

2.1.1 NM-MSSA

The creation of the New Mexico Measures of Student Success and Achievement Balanced Assessment System began with the New Mexico Task Force for Student Success. In March 2019, The New Mexico Public Education Department (NM PED) convened 13 statewide community engagements followed by a taskforce made up of key stakeholders to gather public input to reimagine the state assessment system. This task force held a series of in-person and virtual meetings between April 2019 and June 2019 to deliberate over technical, policy, and practical issues associated with implementing an improved assessment system. The resulting recommendations and an overview of the proposed assessment system were published in a report that was shared with the public in October of 2019. That report can be found on the NM PED website: https://webnew.ped.state.nm.us/wp-content/uploads/2019/11/Student_Success_Task_Force_Report_Balanced_Assessment_System_-October_2019.pdf. Working together, the NM PED and Cognia™ have used these recommendations to create the current assessment system.

The NM-MSSA was scheduled to have its first administration in the spring of 2020. Due to the impact of COVID-19, that administration was canceled. This made the Spring 2021 administration the first one for NM-MSSA. With COVID-19 still impacting students' ability to be in school in-person, the NM PED implemented a flexible testing model in the state, allowing districts and schools the opportunity to opt into the spring summative testing administration. As such, student participation rates were much lower in Spring 2021 than in a typical spring administration (see section 2.3 below). The state was able to set standards for NM-MSSA ELA and mathematics, grades 3-8, in July 2022 for the first time after their first operational assessment in spring 2021.

2.1.2 NM-ASR

With the beginning of the four-year contract with the state in September 2018, Cognia (Measured Progress then, before the merger) developed a new summative science test starting with a Stand-Alone Field Test (SAFT) in spring 2019 when the Science Standards Based Assessment (SBA) had its last operational administration as the state was transitioning into the NGSS and NM STEM Ready! Science Standards. The NM-ASR was originally scheduled to have its first operational administration in the spring of 2020. However, the state was able to obtain a waiver from the Department of Education to extend the opportunity to learn and have a Census Field Test (CFT) in spring 2020 instead. The NM-ASR CFT was able to be administered until March 14, 2020, when the PED made the decision to stop all assessment activities due to COVID-19. With COVID-19 still impacting students' ability to be in school in-person for the 2020-2021 school year, the NM PED implemented a flexible testing model in the state, allowing districts and schools the opportunity to opt into the spring summative testing administration. As such, student participation rates were much lower in Spring 2021 than in a typical spring administration. Therefore, the spring 2021 NM-ASR testing administration remained a field test and the standard setting that was scheduled to happen in 2021, after the first extension, got moved to 2022 when the first operational NM-ASR test was administered. Although COVID-19 continued impacting schools across the state, NM PED was able to maintain the policy of requiring all schools to participate in the spring 2022 NM-ASR in grades 5, 8, and 11 and successfully completed its administration. There was a decision to use the same spring 2021 test for spring 2022; in other words, the test items and test forms were the same since very few took the test in 2021. Following the

spring 2022 administration, a standard setting was conducted that established the NM-ASR score scales for grades 5, 8, and 11 with three cut scores for each test that are used for classifying students into the four performance levels. The results from spring 2023 NM-ASR were the first set of students' results to apply the newly set cut scores.

The administration window for both NM-MSSA and NM-ASR spring 2023 testing administrations was 3/27/2023–5/5/2023.

2.2 Stakeholder Involvement

Cognia and the NM PED work together on all aspects of the implementation of the NM-MSSA and NM-ASR programs. The NM PED also works with several stakeholder groups for input into the implementation of the NM-MSSA and NM-ASR programs.

2.2.1 AAAC

The AAAC is a group of district test coordinators from across the state who meet monthly to advise the Assessment and the Research, Evaluation and Accountability (REA) Bureaus on issues of policy and program matters related to assessment and accountability. The members of the 2022-23 AAAC are listed in Table 2-1.

Table 2-1. AAAC Members 2022–23

Member Name	Member Job Title	Organization
Melissa Adkins	School Counselor	Cloudcroft Municipal Schools
Sandy Beery	Executive Director	New Mexico Connections Academy
Kenneth Bewley	Director of Data Support, Assessment and Research	Roswell Independent School District
LaShawn Byrd	Deputy Director of Data Analysis and Assessment	Hobbs Municipal Schools
Samuel Constant	Coordinator for District Testing	Gadsden Independent School District
Linda Kerr	District Assessment Coordinator	Farmington Municipal Schools
Boyd Lewis	Director of Curriculum and Instruction	Zuni Public School District
Lea Leyba	District Coordinator and Liaison	Chama Valley Independent School District
Dr. Happy Miller	Executive Director, RADA	Rio Rancho Public Schools
Carrie Nigreville	Executive Director of Strategic Planning and School Support	Clovis Municipal School District
James Olivas	Director of Operations and Data	Bloomfield Schools
Danny Parker	Assistant Superintendent	Artesia Public Schools
Edward Pena	District Coordinator and High School Counselor	Cobre Consolidated Schools
Nina Smith	Continuous School Improvement Director	Santa Fe Indian School
Frank Telge	Senior Director of Assessment	Albuquerque Public Schools
Teri Trejo	Director of Assessment, Research and Student Success	Deming Public Schools
Leandro Venturina	Data & Assessment Coordinator	Central Consolidated School District
Sharon West	TriStar Coordinator and SRCL/CLSD Literacy Coordinator	Santa Rosa Consolidated Schools

2.2.2 Educator Committees

In Chapter 4 we will detail the different educator committees that were convened for the purpose of content development. The committees include those listed below, with the details of each committee found in chapter 4.

2.2.2.1 NM-MSSA

- Item-Writing Committees: A group of New Mexico educators convened for a virtual workshop to create unique writing prompts for the NM-MSSA ELA Assessment.
- National Item Review Committee: Cognia convened a national item review committee to review the content of the items that are created. New Mexico educators comprise two seats per grade/content span for those committees.

- International Bias Review Committee: Cognia convened an international bias committee to look for bias and sensitivity concerns in the content that is created. New Mexico educators comprise two seats on that committee.

2.2.2.2 NM-ASR

- Item Review Committee: A group of New Mexico educators convened in 2020 to review newly created items field tested in spring 2021 and spring 2022. (Note that for science, there was no development during the 21-22 development year and therefore no educator committees met for science item review.)
- Bias and Sensitivity Committee: A group of New Mexico educators convened in 2020 to review newly created items field tested in spring 2021 and spring 2022.
- Data Review Committee: A group of New Mexico educators reviewed field test item statistics from spring 2022 to determine what items would be eligible to use operationally in spring 2023 or beyond.

2.2.3 Technical Advisory Committee

The NM PED consults with their technical advisory committee (TAC) to provide feedback and recommendations on program implementation. The NM TAC includes the following members:

Edynn Sato, PhD (Chair)

Edynn Sato has more than 25 years of experience in education research and development, evaluation, training, technical assistance, and management. Her focus has been on making learning inclusive, accessible, and equitable, and her research, development, and consultation have affected practice and policy in the U.S. and abroad. Currently, she works as an independent consultant for her own company, Sato Education Consulting LLC. Additionally, she is a research faculty in the School of Education and Information Sciences at UCLA. Recent and current work include peer review of evidence for state assessment systems; management and development of English language proficiency standards for English learners with significant cognitive disabilities; facilitation and development of a Theory of Action, Logic Model, and technical manual related to English language development; and evaluation of accommodations for English learners.

Tony Albano, PhD

Tony Albano is an Associate Professor in the School of Education at UC Davis, where he teaches courses in testing, assessment, and data science. His research aims to improve teaching and learning via effective educational and psychological measurement, including improvements in psychometric methods for scaling and modeling assessment results and enhanced accessibility in test design and implementation. In addition, Dr. Albano collaborates on projects examining best practices in instruction and assessment with emergent bilingual students, computerized adaptive testing technologies in early education, and culturally responsive pedagogy in higher education.

Scott Marion, PhD

Scott Marion partners with Associate Director Chris Domaleski to manage the operations of the Center for Assessment, and he works closely with the Center Board of Directors to establish the long- and short-term strategic direction of the organization. He is also actively engaged with Center clients; his projects include designing and supporting states in implementing assessment and accountability reforms, developing and implementing educator evaluation systems, and designing and implementing high-quality, locally designed performance-based assessments. He is a national leader in designing innovative and comprehensive assessment systems to support instructional and accountability uses, including helping states and districts design systems of assessments for evaluating student learning of identified competencies. Scott coordinates and/or serves on five district or state Technical Advisory Committees (TACs).

Richard Brown, PhD

Richard S. Brown is the Founder and CEO of West Coast Analytics, a research and consulting firm, and the Chief Research Scientist with the National Math + Science Initiative. Formerly, he held faculty posts in the USC Rossier School of Education and the Department of Education, University of California, Irvine. At both USC and UCI, he taught courses in educational measurement, advanced statistics, and research methodology. Previously, he worked as Senior Researcher at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA. His work at CRESST involved providing technical expertise on two large-scale public school assessment initiatives, specifically in the areas of test development, measurement, and performance standard setting.

Sheryl Lazarus, PhD

Sheryl Lazarus is Director of the National Center on Educational Outcomes (NCEO) at the University of Minnesota. She provides technical assistance to states and conducts research on issues related to the inclusion of all students, including students with disabilities, English learners (ELs), and ELs with disabilities in assessments. Her areas of focus include student participation, accessibility and accommodations, alternate assessments, technology-enhanced assessments, graduation policies, and diploma options. Her work covers the span of assessments in a comprehensive assessment system (e.g., large-scale assessments, interim/benchmark assessments, formative assessments). She has a PhD in Educational Policy and Administration from the University of Minnesota, with a minor in Agricultural and Applied Economics. Dr. Lazarus also holds a K–12 Minnesota principal's license. She has published numerous journal articles, book chapters, reports, and training materials. Dr. Lazarus serves on the assessment Technical Advisory Committees (TACs) of several states.

2.3 Student Participation

2.3.1 NM-MSSA & ASR

NM PED policy defines student participation on a NM-MSSA or ASR Assessment as attempting five or more items on the given assessment. Appendix C provides participation rates as a function of assessment content area (ELA, Mathematics, and Science), test form language (English and Spanish), accommodation/accessibility feature, and background/demographic variable.

The NM-MSSA and ASR Assessments were administered in either computer-based or paper-based formats. Most students utilized the computer-based administration as paper-based is reserved as an accommodation. Tables containing the number of students utilizing accommodation(s)/accessibility feature(s), as a function of content area and grade are available in Appendix D. Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in the aforementioned tables.

Of the students that participated in the Spring 2023 administration, Table 2-2 indicates numbers of students who were assessed in each language mode. The total number of students participating in NM-MSSA and ASR per content area and grade in 2023 ranged from approximately 20,000 to 24,000, which is similar to the numbers of students participating in NM-MSSA and ASR in 2022.

Table 2-2. Number of Participating Students, as a Function of Content Area, Grade, Administration Format, and Test Form Language, NM-MSSA & NM-ASR

	English-Language Forms		Spanish-Language Forms	
Grade	Computer-Based	Paper-Based	Computer-Based	Paper-Based
ELA				
3	20,864	187	664	0
4	21,350	196	611	1
5	21,877	166	274	0
6	21,978	233	269	0
7	22,445	188	259	0
8	23,717	183	280	0
Mathematics				
3	20,868	185	682	0
4	21,358	197	616	1
5	21,884	163	286	0
6	21,997	232	272	0
7	22,413	189	270	0
8	23,704	182	290	0
Science				
5	21,861	166	277	0
8	23,712	181	283	0
11	21,395	76	301	0

Chapter 3. Test Content

3.1 Content Standards

Test content, including items and passages, for the New Mexico MSSA Assessment was developed according to the college- and career-readiness standards for each content area and grade. Content area standards are the basis for the test designs developed for each content area and are used to inform the development of items. Each item is designed to measure a specific standard; however, many Mathematics items assess a mathematical practice standard in addition to a conceptual or procedural standard.

Test content, including items and stimuli, for the New Mexico ASR Assessment was developed according to the New Mexico STEM Ready! Science Standards, which are comprised of the Next Generation Science Standards and a small number of New Mexico-specific standards. These standards are the basis for the test designs developed for each grade and are used to inform the development of items. Each item is designed to measure a specific standard, or performance expectation, and align to multiple dimensions of the standard (Disciplinary Core Idea, Science and Engineering Practice, Crosscutting Concept).

The specific content standards were subsequently grouped into categories for the purpose of communicating with students, families, and educators. The content standards that are eligible to be included in the ELA and Mathematics portions of the NM-MSSA Assessment, as well as the Spanish Language Arts and Spanish Mathematics assessment, and the NM-ASR assessment, are described in the following sections.

3.1.1 Eligible Standards

3.1.1.1 NM-MSSA

Mathematics

The NM-MSSA Mathematics assessment and Spanish Mathematics assessment may assess any of the Common Core State Standards for Mathematics at each grade level, 3–8. While all grade-level standards are eligible to be used on the assessment in their respective grade, not all standards are included in every administration of the assessment. Cognia content specialists strive to include a breadth of standards within and across test administrations while still meeting the reporting category constraints outlined in the test blueprints. This approach allows for the test to meet the requirements of various stakeholders while also maintaining a reasonable test length, and thus testing time.

English Language Arts

The NM-MSSA assesses the Common Core State Standards for English Language Arts. On the Reading portion of the ELA test, at all grade levels, the Reading standards for Literature that may be assessed include RL.1–7 and RL.9, and the Reading standards for Informational Text that may be assessed include RI.1–RI.9. On the Writing and Language portion of the test, the Writing standards that may be assessed at Grades 3–5 are W.2 and W.3 (including all associated sub-standards) and the Writing standards that may be assessed at Grades 6–8 are W.1 and W.2 (including all associated sub-standards). The Language standards that may be assessed at all grade levels are L.1–6 (including all associated sub-standards).

Spanish Language Arts

The Spanish Language Arts (SLA) Assessment may assess the Common Core State Standards for English Language Arts and/or the Common Core State Standards en Español. The items on the SLA are transadapted from the English Language Arts assessment, so the eligible standards for both tests are the same.

3.1.1.2 NM-ASR

Science

The NM-ASR Science assessment and Spanish Science assess students in grades 5, 8, and 11 on the New Mexico STEM Ready! Science Standards. Almost all standards are eligible for assessment as noted below:

- Grade 5: All standards in grades 3, 4, and 5, *except* 5-SS-1 NM.
- Grade 8: All standards in the middle school grade band (6-8), **including** MS-ESS3-3 NM.
- Grade 11: All standards in the high school grade band (9-12), *except* HS-LS2-7 NM and HS-SS-1 NM (but **including** HS-SS-2 NM).

Because of the number of standards per grade, not all standards can be tested every year. The design of the NM-ASR allows for all assessable standards to be included on the NM-ASR at least once within a three-year time period.

3.2 Assessment Design

3.2.1 NM-MSSA Assessment Summary

Tables 3-1 and 3-2 provide a summary of the number of items and points by item type, usage (i.e., operational items or field-test items), and estimated testing time for each grade level and content area of the NM-MSSA Assessment.

Each NM-MSSA content-area test is administered in two sessions. Test forms contain core operational items, matrix operational items, and matrix field-test items. Matrix operational items are items administered to subsets of students as a means of randomly spiraling operational content. Matrix field-test items are items administered to subsets of students to "try out" performance (with different students receiving different field-test items), and therefore do not count toward student score.

English Language Arts

The types of items on the ELA portion of the NM-MSSA Assessment are 1–point machine-scored items (MS-1), 2–point machine-scored items (MS-2), and 6–point writing prompts (WP). Additional item-type descriptions can be found in section 3.2.4.

Table 3-1. Student Testing Experience—ELA (Full Form)

Grades 3–8	Discrete Items				Total	TotalPoints	
	Passage Sets	MS-1	MS-2	WP	Items	Min	Max
Core Operational Items	5	27	5	0	32	37	37
Matrix Operational Items	1	0	0	1	1	6	6
Matrix Field-Test Items	2	10	2	0	12	14	14
Total Student Experience	**8**	**37**	**7**	**1**	**45**	**57**	**57**
				Estimated Test Time (min)			**150 (60/90)**

Mathematics

The types of items on the mathematics portion of the NM-MSSA Assessment are 1-point machine-scored items (MS–1), 3–point constructed-response items (CR-3), and 6–point constructed-response items (CR-6). Additional item-type descriptions can be found in section 3.2.4.

Table 3-2. Student Testing Experience—Mathematics (Full Form)

Grades 3–5	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	33	2	2	37	51	51
Matrix Operational Items	0	0	0	0	0	0
Matrix Field-Test Items	5	1		6	8	11
Total Student Experience	**38**	**5**		**43**	**59**	**62**
			Estimated Test Time (min)			**120**
Grades 6, 7	**Discrete Items**			**Total Items**	**Total Points**	
	MS-1	**CR-3**	**CR-6**		**Min**	**Max**
Core Operational Items	36	2	2	40	54	54
Matrix Operational Items	0	0	0	0	0	0
Matrix Field-Test Items	5	1		6	8	11
Total Student Experience	**41**	**5**		**46**	**62**	**65**
			Estimated Test Time (min)			**120**
Grade 8	**Discrete Items**			**Total Items**	**Total Points**	
	MS-1	**CR-3**	**CR-6**		**Min**	**Max**
Core Operational Items	37	2	2	41	55	55
Matrix Operational Items	0	0	0	0	0	0
Matrix Field-Test Items	5	1		6	8	11
Total Student Experience	**42**	**5**		**47**	**63**	**66**
			Estimated Test Time (min)			**120**

3.2.2 NM-ASR Assessment Summary

Table 3.3 provides a summary of the number of items and points by item type, usage (i.e., operational items or field-test items), and estimated testing time for each grade level and content area of the NM-ASR Assessment. The NM-ASR test is administered in three sessions. Test forms contain core operational items, matrix operational items, and matrix field-test items. All operational items count toward student score, with the core operational items being common across all forms and the matrix operational items being administered across different operational forms. Matrix field-test items are items administered to subsets of students to “try out” performance (with different students receiving different field-test items), and therefore do not count toward student score.

The types of items on the NM-ASR Assessment are 1-point machine-scored items (MS–1), 2-point machine scored items (MS-2), and 4–point constructed-response items (OE-4). Additional item-type descriptions can be found in section 3.2.4.

Table 3-3. Student Testing Experience—Science (Full Form)

Grades 5, 8	Cluster/Passage Items			Standalone Items		Total Items	Total Points
	Stimulus/Passage	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	0	3	27	48
Matrix Operational Items	0	0	0	8	0	8	16
Matrix Field Test Items	2	4	4	4	1	13	24
Total Student Experience	**8**	**16**	**16**	**12**	**4**	**48**	**88**
				Estimated Testing Time (min)			**150 (50/50/50)**
Grade 11	**Cluster/Passage Items**			**Standalone Items**		**Total Items**	**Total Points**
	Stimulus/Passage	**MS-1**	**MS-2**	**MS-2**	**OE**		
Core Operational Items	6	12	12	0	3	27	48
Matrix Operational Items	0	0	0	10	0	10	20
Matrix Field Test Items	2	4	4	5	1	14	26
Total Student Experience	**8**	**16**	**16**	**15**	**4**	**51**	**94**
				Estimated Testing Time (min)			**165 (55/55/55)**

3.2.3 NM-MSSA Assessment Specifications

The reporting categories for the NM-MSSA Assessment are based on the clusters of standards found in the Common Core State Standards. Target percentages for the distribution of operational (core) test points for each of the reporting categories reflect the distribution in the standards, so as not to overrepresent or underrepresent content. These percentages are shown in the tables in the next two sections.

English Language Arts

Specifications for the full test blueprints for the construction of the core forms reflect the reporting category specifications, as well as percentage requirements for each cluster. These constructs represent key aspects of the standards to which items are aligned; as such, the percentage of operational (core) test points for each should be maintained from year to year. Note that percentages in Reading for (a) text type and (b) reading strategy are calculated independently. An individual item may contribute to multiple parts of the blueprint.

For the English Language Arts assessment, there are a total of 17 forms. The operational items are common across all forms, and then sets of field test items are embedded to create 17 matrix forms. The operational items in Form 1 are modified for students who require a PBT form, Large-Print form, or Braille. Form 1 is also the form that is transadapted into Spanish for the SLA assessment and administered under the various allowed accommodations including ASL (see Appendix E for more information about accommodations). Additionally, Form 1 is the form used for Text-To-Speech for the computer-based test for students with that specific accommodation.

Table 3.4. ELA Operational Test Blueprint

English Language Arts			Grades 3–5 Ideal # of Core Pts	Grades 3–5 Ideal % of Core Pts	Grades 6–8 Ideal # of Core Pts	Grades 6–8 Ideal % of Core Pts
Reading	Text Type	Literary Text	15	65%	8	35%
		Informational Text	8	35%	15	65%
	Reading Strategy	Comprehension	12–14	52–60%	9–12	39–52%
		Analysis and Interpretation	9–11	39–47%	13–16	56–70%
	Cluster	Key Ideas and Details	9–11	7–11%	7–11	9–13%
		Craft and Structure	7–9	30–39%	6–10	26–43%
		Integration of Knowledge & Ideas	4–6	17–26%	4–6	17–26%
		Total	**23***	**100%***	**23***	**100%**
Writing & Language	Language & Writing Passage Sets	Writing Analysis	9-11	42-52%	9-11	42-52%
		English Language Conventions	3-5	14-23%	3-5	14-23%
	Writing Prompt	Production of Writing	3	15%	3	15%
		Use of Conventions	3	15%	3	15%
		Total	**20**	**100%**	**20**	**100%**

**All items align to a text type (Literary, Informational), reading strategy (Comprehension and Analysis and Interpretation) and a cluster (Key Ideas and Details, Craft and Structure, and Integration of Knowledge and Ideas).*

Mathematics

Specifications for the full test blueprints for the construction of the core forms reflect the reporting category specifications, as well as percentage requirements for each cluster. These constructs represent key aspects of the standards to which items are aligned; as such, the percentage of operational (core) test points for each should be maintained from year to year. Note that percentages for (a) content clusters and (b) mathematical practices are calculated independently. An individual item may contribute to multiple parts of the blueprint.

Most multiple-choice (MC) Mathematics items are dually coded to both a Concepts and Procedures (CP) standard as well as a Mathematical Practice (MP). While the MC items are coded to both CP and MP, each MC item is scored as 1 point toward the student's overall score in CP. Each constructed-response (CR) item is scored on a rubric in which points are assigned to both CP and MP. Across all CR items, there are a total of 12 points for CP and a total of 6 points for MP.

For the Mathematics assessment, there are a total of 12 forms. The operational items are common across all forms, and then sets of field test items are embedded to create 12 matrix forms. The operational items in Form 1 are modified for students who require a PBT form, Large-Print form, or Braille. Form 1 is also the form that is translated into Spanish and administered under the various allowed accommodations including ASL (see Appendix E for more information about accommodations). Additionally, Form 1 is the form used for Text-To-Speech for the computer-based test.

Table 3-5. Mathematics Grades 3–5 Operational Test Blueprint

	Mathematics Grades 3–5	Grade 3 Ideal # of Core Pts	Grade 3 Ideal % of Core Pts	Grade 4 Ideal # of Core Pts	Grade 4 Ideal % of Core Pts	Grade 5 Ideal # of Core Pts	Grade 5 Ideal % of Core Pts
Concepts & Procedures	Operations & Algebraic Thinking	12–18	27–40%	10–16	22–36%	7–11	16–24%
	Number & Operations in Base Ten	5–7	11–16%	8–10	17–22%	7–13	16–29%
	Number & Operations – Fractions	8–10	18–22%	10–16	22–36%	11–15	24–33%
	Measurement & Data	11–15	24–33%	6–10	13–22%	10–14	22–31%
	Geometry	3–5	7–11%	3–5	7–11%	4–8	9–18%
	Subtotal	**45**	**100%**	**45**	**100%**	**45**	**100%**
Mathematical Practices	Problem Solving* Reasoning* & Argument	≥ 8	≥ 17%	≥ 8	≥ 17%	≥ 8	≥ 17%
	Modeling Structure & Repeated Reasoning*	≥ 8	≥ 17%	≥ 8	≥ 17%	≥ 8	≥ 17%
	Total	**51****		**51****		**51****	

**All or most items are dually coded to Concepts and Procedures and Mathematical Practice Standards.*
***Constructed-response items are scored for both Concepts and Procedures and Mathematical Practices. A total of 6 points from the Mathematical Practices rubric contributes to a student's overall score.*

Table 3-6. Mathematics Grades 6–8 Operational Test Blueprint

	Mathematics Grades 6 & 7	Grade 6 Ideal # of Core Pts	Grade 6 Ideal % of Core Pts	Grade 7 Ideal # of Core Pts	Grade 7 Ideal % of Core Pts	Mathematics Grade 8	Grade 8 Ideal # of Core Pts	Grade 8 Ideal % of Core Pts
Concepts & Procedures	Ratios & Proportional Relationships	8–12	17–25%	8–12	17–25%	Functions	10–16	20–33%
	The Number System	8–12	17–25%	6	13%	The Number System	4	8%
	Expressions & Equations	8–12	17–25%	8–16	17–33%	Expressions	11–17	22–35%
	Geometry	6–10	13–21%	6–10	13–21%	Geometry	10–16	20–33%
	Statistics & Probability	6–10	13–21%	10–12	21–25%	Statistics & Probability	10–12	20–24%
	Subtotal	**48**	**100%**	**48**	**100%**	**Subtotal**	**49**	**100%**
Mathematical Practices	Problem Solving* Reasoning* & Argument	≥ 8	≥ 16%	≥ 8	≥ 16%	Problem Solving* Reasoning* & Argument	≥ 8	≥ 16%
	Modeling* Structure & Repeated Reasoning*	≥ 8	≥ 16%	≥ 8	≥ 16%	Modeling* Structure & Repeated Reasoning*	≥ 8	≥ 16%
	Total	**54****		**54****		**Total**	**55****	

**All or most items are dually coded to Concepts and Procedures and Mathematical Practice Standards.*
***Constructed-response items are scored for both Concepts and Procedures and Mathematical Practices. A total of 6 points from the Mathematical Practices rubric contributes to a student's overall score.*

3.2.4 NM-ASR Assessment Specifications

The reporting categories for the NM-ASR Assessment are based on the science domains in the New Mexico STEM Ready! Science Standards. Target percentages for the distribution of operational test points for each of the reporting categories reflect the distribution in the standards, so as not to overrepresent or underrepresent content. These percentages are shown in the tables in this section.

Specifications for the full test blueprints for the construction of the operational forms reflect the reporting category specifications. These constructs represent key aspects of the standards to which items are aligned; as such, the percentage of operational test points for each should be maintained from year to year. Note that some of the points for each reporting category come from clusters (a grouping of four

items—2 MS-1 and 2 MS-2—all associated with a common stimulus), and some points come from standalone/discrete items.

Table 3-7. Grades 5, 8, 11 – NM-ASR Operational Test Blueprint

Grade 5 NM-ASR					
Reporting Category	**Ideal # of Clusters**	**Ideal # of Standalone MS-2**	**Ideal # of Standalone OE**	**Ideal # of Core Points**	**Ideal % of Core Points (+/- 4%)**
Practices and Crosscutting Concepts in Physical Sciences	2	4-6	1	24-28	40%
Practices and Crosscutting Concepts in Life Sciences	2	1-3	1	18-22	30%
Practices and Crosscutting Concepts in Earth and Space Sciences	2	1-3	1	18-22	30%
Grade 8 NM-ASR					
Reporting Category	**Ideal # of Clusters**	**Ideal # of Standalone MS-2**	**Ideal # of Standalone OE**	**Ideal # of Core Points**	**Ideal % of Core Points (+/- 4%)**
Practices and Crosscutting Concepts in Physical Sciences	2	2-4	1	20-24	35%
Practices and Crosscutting Concepts in Life Sciences	2	2-4	1	20-24	35%
Practices and Crosscutting Concepts in Earth and Space Sciences	2	1-3	1	18-22	30%
Grade 11 NM-ASR					
Reporting Category	**Ideal # of Clusters**	**Ideal # of Standalone MS-2**	**Ideal # of Standalone OE**	**Ideal # of Core Points**	**Ideal % of Core Points (+/- 4%)**
Practices and Crosscutting Concepts in Physical Sciences	2	3-5	1	22-26	35%
Practices and Crosscutting Concepts in Life Sciences	2	3-5	1	22-26	35%
Practices and Crosscutting Concepts in Earth and Space Sciences	2	1-3	1	18-22	30%

Note that items aligned to standards in Engineering, Technology, and Applications of Science as well as the NM-specific content domain of Science and Society are reported under the reporting category domain that matches the context of the phenomenon or design problem presented.

For the Science assessment, there are a total of 7 forms. There are two sets of operational items, set A and set B, differing in the standalone MS-2 items that are in the set (but still following the same content blueprint), in order to support sufficient assessment of all content standards over time. Three sets of field-test items are embedded with one of the operational sets, and then another three sets of field-test items are embedded with the other operational set, for a total of 6 matrix forms. A seventh matrix form, Form AX, is also created by modifying the set A operational items for students who require a PBT form. This Form AX is administered not only as PBT but also in CBT, to allow for calibration of the modified operational items.

Additionally, for NM-ASR, Form 1 is the form used for Text-To-Speech for both English and Spanish versions of the computer-based test. Form AX is the form that is specifically used for the paper version of the test as it modifies the TEI items that are seen on the computer-based test version. Form AX is the form used to produce both the English and Spanish PBT, Large Print, and Braille. As noted in the previous paragraph, Form AX is also included in the computer-based testing to see comparability of the same form between online and paper test mode. (See Appendix E for more information about NM state

assessment accommodations.) Below is a crosswalk table that explains the different naming conventions used for sets and forms.

Table 3-8. Naming Conventions Crosswalk

Subject	Grade	Form	OP Form
Science	5	A1	1
	5	A2	1
	5	A3	1
	5	B4	2
	5	B5	2
	5	B6	2
	5	AX	3
	8	A1	1
	8	A2	1
	8	A3	1
	8	B4	2
	8	B5	2
	8	B6	2
	8	AX	3
	11	A1	1
	11	A2	1
	11	A3	1
	11	B4	2
	11	B5	2
	11	B6	2
	11	AX	3

3.2.5 Content Coverage Blueprint

NM-MSSA

The distribution of emphasis for NM-MSSA content strands in English Language Arts is shown in Table 3-9; Mathematics for the Spring 2023 assessment is shown in Table 3-10.

Table 3-9. Distribution of Emphasis Across Content Strands in Terms of Percentage of Total Test Points by Grade—ELA Grades 3–8—Spring 2023

		Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
		Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points
Reading Clusters	Key Ideas and Details	11	48%	11	48%	12	52.2%	10	43.5%	9	39%	10	43%
	Craft and Structure	8	35%	7	30%	7	30.4%	7	30.4%	11	48%	11	48%
	Integration of Knowledge and Ideas	4	17%	5	22%	4	17.4%	6	26.1%	3	13%	2	9%
	Total	**23**	**100%**	**23**	**100%**	**23**	**100%**	**23**	**100%**	**23**	**100%**	**23**	**100%**
Writing & Language Strands	Writing	16	80%	16	80%	16	80%	16	80%	16	80%	16	80%
	Language	4	20%	4	20%	4	20%	4	20%	4	20%	4	20%
	Total	**20**	**100%**	**20**	**100%**	**20**	**100%**	**20**	**100%**	**20**	**100%**	**20**	**100%**

Table 3-10. Distribution of Emphasis Across Content Strands in Terms of Percentage of Total Test Points by Grade—Mathematics Grades 3–8—Spring 2023

Content Strand	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points
Operations and Algebraic Thinking	18	35%	12	24%	7	14%	--	--	--	--	--	--
Number & Operations in Base 10	5	10%	8	16%	9	18%	--	--	--	--	--	--
Number & Operations – Fractions	8	16%	16	31%	13	25%	--	--	--	--	--	--
Measurement & Data	11	22%	6	12%	12	24%	--	--	--	--	--	--
Geometry 3–5	3	6%	3	6%	4	8%	--	--	--	--	--	--
Ratios & Proportional Relationships	--	--	--	--	--	--	10	19%	12	22%	--	--
The Number System	--	--	--	--	--	--	12	22%	6	11%	4	7%
Expressions & Equations	--	--	--	--	--	--	12	22%	14	26%	13	24%
Geometry 6–8	--	--	--	--	--	--	6	11%	6	11%	10	18%
Statistics & Probability	--	--	--	--	--	--	8	15%	10	19%	10	18%
Functions	--	--	--	--	--	--	--	--	--	--	12	22%
Problem Solving	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
Reasoning & Argument	4	8%	4	8%	5	10%	4	7%	3	6%	1	2%
Modeling	0	0%	0	0%	0	0%	2	4%	3	6%	3	5%
Patterns & Structure	2	4%	2	4%	1	2%	0	0%	0	0%	2	4%
Total	**51**	**100%**	**51**	**100%**	**51**	**100%**	**54**	**100%**	**54**	**100%**	**55**	**100%**

NM-ASR

The distribution of emphasis for NM-ASR content standards in Science for the Spring 2023 assessment is shown in Table 3-11. Assessable standards cover physical science, life science, earth and space science, and engineering, technology, and applications of science (ETS), as well as science and society in the NM-specific portion of the standards in grades 8 and 11.

Table 3-11. Distribution of Emphasis Across Content Standards in Terms of Percentage of Total Test Points by Grade—Science Grades 5, 8, 11 Spring 2023

Standards Category	Grade 5		Grade 8		Grade 11	
	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points
Physical Science	24	37.50%	22	34.38%	22	32.35%
Life Science	20	31.25%	22	34.38%	24	35.29%
Earth and Space Science	20	31.25%	20	31.25%	22	32.35%
Grand Total	**64**	**100.00%**	**64**	**100.00%**	**68**	**100.00%**

3.2.6 Operational Section

NM-MSSA

Table 3-12 shows the reporting categories for English Language Arts in the NM-MSSA test design, and the maximum possible number of raw-score points students could earn in each reporting category. Note: Because only operational items are counted toward students' scale scores, only operational items are reflected in this table. The number of items and item types that are used to achieve these distributions are provided in the tables at the beginning of section 3.2.

Table 3-12. Distribution of Raw-Score Points Across Reporting Categories by Grade—English Language Arts Grades 3–8 Spring 2023

Reporting Category	Grade 3 Total Points	Grade 3 % of Total Core Points	Grade 4 Total Points	Grade 4 % of Total Core Points	Grade 5 Total Points	Grade 5 % of Total Core Points	Grade 6 Total Points	Grade 6 % of Total Core Points	Grade 7 Total Points	Grade 7 % of Total Core Points	Grade 8 Total Points	Grade 8 % of Total Core Points
English Language Arts	43	100%	43	100%	43	100%	43	100%	43	100%	43	100%
Reading	23	100%	23	100%	23	100%	23	100%	23	100%	23	100%
Literary Text	15	65%	15	65%	15	65%	8	35%	8	35%	8	35%
Informational Text	8	35%	8	35%	8	35%	15	65%	15	65%	15	65%
Comprehension	13	57%	13	57%	12	52%	9	39%	9	39%	9	39%
Analysis and Interpretation	10	43%	10	43%	11	48%	14	61%	14	61%	14	61%
Writing & Language	20	100%	20	100%	20	100%	20	100%	20	100%	20	100%
Writing Analysis & Language Conventions	14	70%	14	70%	14	70%	14	70%	14	70%	14	70%
Production of Writing	3	15%	3	15%	3	15%	3	15%	3	15%	3	15%
Use of Conventions	3	15%	3	15%	3	15%	3	15%	3	15%	3	15%

Table 3-13 shows the reporting categories for Mathematics in the NM-MSSA test design, and the maximum possible number of raw-score points students could earn in each reporting category on the Spring 2023 assessment. Note: Because only operational items are counted toward students' scale scores, only operational items are reflected in this table. The number of items and item types that are used to achieve these distributions are provided in the tables at the beginning of section 3.2.

Table 3-13. Distribution of Raw-Score Points Across Reporting Categories by Grade—Mathematics Grades 3–8 Spring 2023

Reporting Category	Grade 3 Total Points	Grade 3 % of Total Core Points	Grade 4 Total Points	Grade 4 % of Total Core Points	Grade 5 Total Points	Grade 5 % of Total Core Points	Grade 6 Total Points	Grade 6 % of Total Core Points	Grade 7 Total Points	Grade 7 % of Total Core Points	Grade 8 Total Points	Grade 8 % of Total Core Points
Operations and Algebraic Thinking	18	22%	12	14%	7	8%	--	--	--	--	--	--
Number & Operations in Base 10 / Number & Operations – Fractions	13	16%	24	29%	22	27%	--	--	--	--	--	--
Measurement & Data / Geometry	14	17%	9	11%	16	19%	--	--	--	--	--	--
Ratios & Proportional Relationships	--	--	--	--	--	--	10	11%	12	13%	--	--
The Number System / Expressions & Equations	--	--	--	--	--	--	24	27%	20	22%	17	19%
Geometry / Statistics & Probability	--	--	--	--	--	--	14	16%	16	18%	20	22%
Functions	--	--	--	--	--	--	--	--	--	--	12	13%
Problem Solving / Reasoning & Argument	22	27%	22	26%	23	28%	23	26%	26	29%	17	19%
Modeling / Patterns & Structure	15	18%	17	20%	15	18%	18	20%	16	18%	24	27%
Total	**82**	**100%**	**84**	**100%**	**83**	**100%**	**89**	**100%**	**90**	**100%**	**90**	**100%**

NM-ASR

Table 3-14 shows the reporting categories for Science in the NM-ASR test design, and the maximum possible number of raw-score points students could earn in each reporting category on the Spring 2023 assessment. Note: Because only operational items are counted toward students' scale scores, only operational items are reflected in this table. The number of items and item types that are used to achieve these distributions are provided in the tables at the beginning of section 3.2. Any items aligned to standards in Engineering, Technology, and Applications of Science as well as the NM-specific content domain of Science and Society are reported under the reporting category domain that matches the context of the phenomenon or design problem presented. The distribution of raw-score points in the table is applicable for both operational forms, as each of the two operational forms for the NM-ASR were built to the same specification for reporting categories.

Table 3-14. Distribution of Raw-Score Points Across Reporting Categories by Grade—Science Grades 5, 8, 11—Spring 2023

Reporting Category	Grade 5		Grade 8		Grade 11	
	Total Points	% of Total Core Points	Total Points	% of Total Core Points	Total Points	% of Total Core Points
Practices and Crosscutting Concepts in Physical Sciences	24	37.50%	22	34.38%	22	32.35%
Practices and Crosscutting Concepts in Life Sciences	20	31.25%	22	34.38%	24	35.29%
Practices and Crosscutting Concepts in Earth and Space Sciences	20	31.25%	20	31.25%	22	32.35%
Grand Total	**64**	**100.00%**	**64**	**100.00%**	**68**	**100.00%**

3.2.7 Field-Test Sections

All NM-MSSA and NM-ASR items are appropriately field tested prior to operational use. The NM-MSSA and NM-ASR Assessments employ a matrix design that embeds field-test items within each form. Matrix field-test items are items administered to subsets of students to "try out" performance (with different students receiving different field-test items), and therefore do not count toward student scores.

The ELA portion of the NM-MSSA Assessment contains a total of two field-test passage sets and 12 field-test items per grade-level form: 10 MS-1, and two MS-2. The mathematics portion contains a total of six field-test items per grade-level form: five MS-1 items and one CR-3 or 1 CR-6.

For Science, the NM-ASR Assessment contains a total of 13 field-test items for grades 5 and 8 per form: two clusters (with four items each), four MS-2 standalones, and one OE-4. The grade 11 test contains 14 field-test items per form: two clusters (with four items each), five MS-2 standalones, and one OE-4.

3.2.8 Item Types

Item types are chosen to best balance the desires for making efficient use of limited testing time and providing coverage of a broad range of knowledge and skills. The item types used on the NM-MSSA and NM-ASR Assessments and the functions of each are listed below.

English Language Arts

The Reading portion of the NM-MSSA ELA Assessment includes SR, MS, and evidence-based selected-response (EBSR) items.

SR and MS items each require students to demonstrate a wide range of knowledge and skills. MS items consist of a single prompt, much like standard SR items, but include up to a maximum of six answer choices. Of these answer choices, two or three choices make up the key. Students in grades 3–8 are directed to select a given number of answer choices. The MS items are scored correct only; partial credit is not awarded for partially correct responses.

EBSR items are selected-response items with two parts. The second part of an EBSR item asks students to select evidence that supports the response in the first part. Each part of an EBSR item is worth one point; however, students will only receive partial credit (one point) if they answer Part A correctly. Students will not receive a point for answering only Part B correctly.

Each type of item is worth a specific number of points in the student's total reading score, as shown in Table 3-15.

Table 3-15. Reading Item Types

Item Type	Maximum Number of Points Available
SR/MS	1
EBSR	2

The Writing and Language portion of the NM-MSSA ELA Assessment includes SR, MS, and EBSR items. Grades 3–8 Writing and Language passages feature an embedded-error format, in which deliberate errors are identified or introduced into passage text. Items associated with the passages are developed to address the specific errors identified or introduced into the passage text.

The Writing and Language portion of the NM-MSSA also includes a writing prompt (WP). Writing prompts require students to write an extended response to a single prompt. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The WP items are evaluated using a "Production of Writing" rubric on a scale from 1–3 and a "Use of Conventions" rubric on a scale from 1–3.

Each type of item is worth a specific number of points in the student's total Writing and Language score, as shown in Table 3-16.

Table 3-16. Writing and Language Item Types

Item Type	Maximum Number of Points Available
SR/MS	1
EBSR	2
WP	6

Mathematics

The NM-MSSA Mathematics tests include selected-response (SR), multi-select selected-response (MS), and constructed-response (CR) items.

SR and MS items each require students to demonstrate a wide range of knowledge and skills. MS items consist of a single prompt, much like standard SR items, but include at least five answer choices. Of these five+ answer choices, at least two choices make up the key. Students in grades 3–5 are directed to select a given number of answer choices for their response. Students in grades 6–8 are sometimes directed to select a given number of answer choices, but also may be asked to "select all that apply"

instead as their response. The MS items are scored correct only; partial credit is not awarded for partially correct responses.

There are two varieties of CR items: 3-point and 6-point items. These CR items require students to write an extended response to a prompt. The prompt may be a single prompt, or more typically, the items are written with multiple, scaffolded parts for students to respond to. The items are hand-scored, with scorers using a multi-trait rubric, scoring notes, and anchor exemplars to evaluate responses.

The 3–point items (CR-3) require students to perform a computation, write an expression, equation, or inequality, and/or solve a simple problem, and may include having the student provide written evidence of the understanding of the standard(s) being assessed. The CR-3 items are evaluated using a concepts and procedures rubric on a scale from 0–2 and a mathematical practices rubric on a scale from 0–1. The 6–point items (CR-6) are more complex and require students to provide written evidence of the understanding of the standard(s) being assessed. The CR-6 items are evaluated using a concepts and procedures rubric on a scale from 0–4 and a mathematical practices rubric on a scale from 0–2.

Each type of item on the assessment is worth a specific number of points in the student's total Mathematics score, as shown in Table 3-17.

Table 3-17. Mathematics Item Types

Item Type	Maximum Number of Points Available
SR/MS	1
CR	3 or 6

Science

The NM-ASR tests include machine-scored 1-point items (MS-1), machine-scored 2-point items (MS-2), and open-ended items (OE4). Some of the MS-1 and MS-2 items are grouped together in clusters.

MS-1 items may be multiple-choice, multiple select, or technology-enhanced (e.g., drag-and-drop, hot spot, drop-down selections). MS-1 items are only found in clusters. They are all machine-scored as correct only; partial credit is not awarded.

MS-2 items have two parts (Part a and Part b) for students to answer. These items may combine multiple choice, multiple select, and/or technology-enhanced interactions across the two parts. MS-2 items are included in clusters and as standalone items. They are all machine scored, and students may earn 2, 1, or 0 points across Part a and Part b.

An item cluster is a set of items all associated with a common stimulus. Clusters contain four items, with two of the items being worth 1 point (MS-1) and two of the items being worth 2 points (MS-1). The clusters typically align to two PEs, and all clusters measure all three dimensions of the PEs being assessed.

Open-ended items are worth 4 points. These items require students to write an extended response to a prompt. The prompt may be a single prompt, or more typically, the items are written with multiple, scaffolded parts for students to respond to. These items are hand-scored, with scorers using a rubric and scoring notes to evaluate responses on a scale from 0–4.

Each type of item on the assessment is worth a specific number of points in the student's total Science score, as shown in Table 3-18.

Table 3-18. Science Item Types

Item Type	Maximum Number of Points Available
MS-1	1
MS-2	2
OE-4	4

3.2.9 Passage Types

All NM-MSSA ELA items, for both Reading and Writing and Language, are based on passages. The configuration of texts on the ELA assessment seeks to balance national high-quality assessment guidance (e.g., NAEP, CCSSO) with considerations around test length.

Some NM-ASR items are connected to an extended, rich stimulus that presents a phenomenon or design problem to frame the set of items. The content of the Science stimuli reflects best practice as recommended by the CCSSO SACI, NRC, and Achieve.

Reading

The reading comprehension portion of the ELA test design incorporates as much of a 50/50 split of literary and informational texts as possible in the elementary grades while still maintaining a limited summative test footprint. Beginning at grade 6, there is a shift in emphasis to informational texts at the upper grade band. For grades 3–5, item sets are based on single literary passages, paired literary passages, and paired informational passages. For grades 6–8, item sets are based on paired literary passages, single informational passages, and paired informational passages.

The reading passages on the NM-MSSA assessment are selected from the following categories:

- Literary passages, representing a variety of forms, including drama, poetry, excerpts from novels, short stories, and traditional narratives such as fables and folktales.
- Informational passages, often about science- and social studies-related topics. These passages are often from news sources, magazines, and book excerpts. The passages are authentic texts selected from grade-level-appropriate reading sources that students would be likely to encounter in the classroom and when reading independently.

All passages are collected from published works.

Writing and Language

The Writing and Language embedded-error passages on the NM-MSSA Assessment are developed to conform to the following categories:

- Narrative passages, representing a variety of forms including short stories and traditional narratives such as fables and folktales. Narrative passages succinctly and lucidly describe a fictional event and feature many or all the hallmarks of the narrative form—plot/conflict, climax/epiphany, conclusion, dialogue, characters' thoughts, action, and description.
- Informative/explanatory passages, representing one of three content areas: social studies/history; science/social science/technical subjects; and, to a lesser extent, the humanities. Although written with the general reader in mind, passages strive to present compelling information that responds to relevant issues in each field—a new interpretation of an event or phenomenon; an

examination of an overlooked (or misunderstood) movement, moment, or figure; an introduction to foundational knowledge in any of the three disciplines; etc.

- Argument passages, representing cogent argumentation. Argument passages tend to be informed by issues in the social sciences or current events. Argument passages establish a position; provide claims, supported by evidence, which develop that position; introduce and rebut a counterclaim (in grades 7 and 8); and, throughout, use rhetorical techniques (persuasive transitions, rhetorical questions, appeals to reason or personal experience, etc.) to advance the position.

All embedded-error passages are commissioned texts, which are passages developed specifically for the purpose of the assessment.

Writing Prompts

The passages paired with the NM-MSSA writing prompts were developed by educators from across the state of New Mexico to support student writing for each of the three purposes for writing: narrative, informative/explanatory, and opinion (grades 3–5) or argument (grades 6–8). The teachers selected passage topics that would be engaging and culturally relevant for New Mexico students.

All writing prompts are partnered with one to three brief text stimuli. These may be complete texts or excerpts from a more extended text. Some possible text types include stories, memoirs, biographies, articles, websites, letters, and brochures.

The number of text stimuli varies depending on the purpose for writing. Narrative prompts are associated with one or two text stimuli, while informative/explanatory and opinion/argument prompts are associated with two or three text stimuli. The passages may be either previously published texts or commissioned texts composed by New Mexico educators specifically for the associated writing prompts and grade levels.

Science

On the NM-ASR, all clusters are written with an extended, rich stimulus. The stimulus must present a single, rich science phenomenon or engineering design problem aligned to the standards/performance expectations being assessed. The phenomenon or problem must launch and support a single storyline, or sequence of sense-making, which is carried out in the items.

The stimulus may present any variety of elements to provide the necessary information related to the phenomenon or problem and the storyline: text paragraphs, passages, graphs, data tables, models, drawings, etc. All information in the stimulus should be necessary, but not conceptually sufficient, for students to respond (i.e., students must also use their own knowledge of the constructs in the standards to answer the items, rather than simply identify given information), and the stimulus must provide enough information to allow students to engage in the SEPs, DCIs, and CCCs of the targeted standards as they respond to items.

Chapter 4. Test Development

4.1 Overview of General Approach

This chapter provides an overview of the development of the NM-MSSA and NM-ASR Assessments, including test and item specifications, item reviews, and test construction.

According to *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014), "important validity evidence can be obtained from an analysis of the relationship between a test's content and the construct it is intended to measure" (p. 14). Accordingly, the descriptions of the test development procedures included in this chapter provide evidence that supports both the content and construct validity of the assessments.

4.2 Item Specifications

English Language Arts

ELA items on the NM-MSSA Assessment are primarily developed by Cognia content specialists using item development best practices. First and foremost, among these is ensuring that all newly developed items align to the intended content standard. Cognia specialists use detailed internal specifications documents to develop items with accurate content alignments. The items also undergo review by nationally representative panels of content and assessment experts, including educators from across the state of New Mexico, with the explicit intent of verifying that the items align with the assigned content standard(s).

The English Language Arts portion of the NM-MSSA Assessment comprises two categories: Reading and Writing and Language.

Each Reading item is designed to measure either students' comprehension of what they have read or their ability to analyze and/or interpret what they have read. All Reading items align to a text type (Literary, Informational) and Reading Strategy (Comprehension, Analysis and Interpretation). The items are organized into three main clusters as further described by the New Mexico Common Core State Standards:

- **Key Ideas and Details** (comprehension or analysis/interpretation): In grades 3–8, students refer to texts solely to demonstrate understanding. At increasing levels of complexity as they advance through the grades, students also draw inferences from texts; show their ability to comprehend or analyze the central events, central ideas, and/or themes of texts; and analyze and interpret the relationships between aspects of a text (e.g., causes and effects in informational texts, or character traits and the plot of literary texts).
- **Craft and Structure** (comprehension or analysis/interpretation): At increasing levels of complexity through the grades, students demonstrate the ability to comprehend and analyze the meanings of words and phrases in texts (including figurative language in grades 5–8), as well as analyze the impact of an author's words (in grades 6–8); identify and analyze the structure of texts, including how certain portions of text affect meaning; and identify and analyze how point of view and purpose shape the content and style of a text.
- **Integration of Knowledge and Ideas** (analysis/interpretation): At increasing levels of complexity through the grades, students integrate knowledge and ideas in texts. Specifically, students integrate:

- visual information (e.g., pictures) and textual information;
- evidence provided in informational texts to support ideas and/or claims; and
- important aspects (e.g., main ideas, characters, settings, themes, structures) of paired texts.

Each Writing and Language item is designed to measure students' ability to evaluate the content and context of text in order to correctly apply the targeted writing skill or language convention. The items are organized into two main categories. Each category contains a unique set of clusters:

Writing

- **Text Types and Purposes**: In grades 3–8, students interact with a variety of texts to demonstrate increasing sophistication with demanding content and sources. At increasing levels of complexity across the grades, students analyze and revise informative/explanatory texts to examine a topic and convey ideas and information clearly or analyze and revise argumentative pieces on topics or texts to help support a point of view with reasons and information or analyze and revise narrative texts to develop real or imagined experiences or events using effective technique, descriptive details, and clear event sequences.

Language

- **Conventions of Standard English**: In grades 3–8, students demonstrate command of the conventions of standard English grammar and usage. At increasing levels of complexity across the grades, students move from simple identification of conventions (e.g., identifying uppercase and lowercase letters or applying the rules of capitalization) to more complex applications of conventions (e.g., recognizing and correcting inappropriate shifts in pronoun number or recognizing and correcting misplaced and dangling modifiers).
- **Knowledge of Language**: In grades 3–8, students apply knowledge of language and conventions to convey ideas or to create a specific effect. At increasing levels of complexity across the grades, students move from conveying ideas or creating a desired effect to focusing on developing and maintaining style and tone by choosing language that expresses ideas precisely and concisely.
- **Vocabulary Acquisition and Use**: In grades 3–8, students apply knowledge of vocabulary structure (e.g., affixes and roots) to understanding the meaning of grade-level vocabulary. At increasing levels of complexity across the grades, students use the context of passage text to determine the concrete and inferred meaning of vocabulary. Additionally, students move from using basic reference materials (e.g., glossary and dictionary) to using more complex references (e.g., thesaurus).

Mathematics

The test designs for Mathematics address the standards within the mathematics domains, or concepts and procedures, as well as the mathematical practices standards.

The mathematics items at grades 3–5 are organized into three concepts and procedures reporting categories:

- **Operations and Algebraic Thinking**: Students represent and solve problems, understand and apply the properties of operations, and generate and analyze patterns and relationships.

- **Number and Operations in Base Ten and Fractions**: Students understand and demonstrate a sense of what whole numbers, fractions, and decimal numbers mean and how they are used. Students understand and demonstrate computation skills.
- **Measurement and Data and Geometry**: Students understand and demonstrate measurement skills, including geometric measurement, by accurately measuring and estimating, solving problems, and converting between units within a measurement system. Students represent and interpret data using picture graphs, bar graphs, and line plots. Students reason with shapes and their attributes, classify shapes based on their properties, and graph points on the coordinate plane to solve problems.

The mathematics items at grades 6 and 7 are organized into three concepts and procedures reporting categories:

- **Ratios and Proportional Relationships**: Students understand ratio concepts and proportional relationships and use them to solve real-world problems.
- **The Number System and Expressions and Equations**: Students extend their previous number sense and computation of whole numbers, fractions, and decimal numbers to the entire system of rational numbers. Students write and evaluate expressions, apply the properties of operations to generate equivalent expressions, and solve problems using algebraic expressions, equations, and inequalities.
- **Geometry and Statistics and Probability**: Students solve problems involving area, surface area, volume, and angle measures. Students draw, construct, and describe geometric figures and describe the relationships between figures. Students understand statistical variability, summarize and describe distributions, and use random sampling to draw inferences about a population or comparative inferences between populations. Students develop an understanding of probability and use and evaluate probability models.

The mathematics items at grade 8 are organized into three concepts and procedures reporting categories:

- **Functions**: Students define, evaluate, and compare functions and use functions to model relationships between quantities.
- **The Number System and Expressions and Equations**: Students extend their previous number sense to include the system of irrational numbers. Students work with radicals and integer exponents. Students understand the connections between proportional relationships, lines, and linear equations, and analyze and solve linear equations and pairs of simultaneous linear equations.
- **Geometry and Statistics and Probability**: Students understand congruence and similarity, understand and apply the Pythagorean Theorem, and solve problems involving volume of three-dimensional figures. Students investigate the patterns of association in bivariate data.

Additionally, the Mathematics items at each of the grades 3–8 have embedded in them the processes and proficiencies associated with the following mathematical practice standards:

- **Problem Solving/Reasoning and Argument**: Students apply grade-level appropriate mathematical concepts and procedures and quantitative and logical reasoning to solve standard and nonstandard real-world and mathematical problems. Students critique the mathematical reasoning of others and construct viable arguments.
- **Modeling/Structure and Repeated Reasoning**: Students use grade-appropriate quantitative reasoning to interpret mathematical representations, represent real-world mathematical situations using mathematical models, and use mathematical models to solve real-world and mathematical

problems. Students look for and make use of mathematical structure. Students look for and make use of repeated reasoning in mathematics.

Mathematics Content Supports and Considerations

Calculators

While the team of assessment content specialists who designed the mathematics test acknowledge the importance of mastering arithmetic algorithms, they understand that the use of calculators is a necessary and important skill. Calculators can save time and prevent errors in the measurement of some higher-order thinking skills, allowing students to work on more sophisticated and intricate problems. For these reasons, it was decided that, at grades 3–8, calculators should be prohibited in the first of the two sessions of the NM-MSSA Assessment mathematics tests and permitted in the second session. Students in grades 3–6 who are taking the paper-pencil test can use their own four-function calculator with a square root key during Session 2. Students in grades 7 and 8 who are taking the paper-pencil test can use their own scientific calculator during Session 2. Students taking the online test use the calculator tools provided in the online platform.

Reference Sheets

Reference sheets are not provided to students at grades 3–8. To properly assess the applicable standards, some items are written so that students will need to know the formulas to answer the question, whereas other items are written so that knowledge of the formula is not being assessed, and thus the formulas may be provided within the item. Guidance from grade-level mathematics educators is used to help guide the inclusion or exclusion of formulas.

A ruler or protractor will be embedded within a graphic for items that require students to measure lengths of objects or angles.

Science

The NM-ASR test design is based on the three content domains of Physical Sciences, Life Sciences, and Earth and Space Sciences. Items are expected to align to the multiple dimensions of the standards (Disciplinary Core Ideas, Science and Engineering Practices, Crosscutting Concepts) in each domain, such that every item is at least two-dimensional, if not three-dimensional. To emphasize this multi-dimensional nature of the items, the names of the reporting categories incorporate the three dimensions (Practices and Crosscutting Concepts in Physical Sciences, Practices and Crosscutting Concepts in Life Sciences, Practices and Crosscutting Concepts in Earth and Space Sciences). Students are expected to demonstrate sense-making by using core ideas, practices, and crosscutting concepts together to respond to items.

Items assessing Engineering, Technology, and Applications of Science as well as the New Mexico-specific content domain of Science and Society are reported within the Physical, Life, or Earth and Space Sciences category, depending on the content match of the design problem presented in the item.

As content support, students taking the Grade 11 test are provided with a periodic table reference sheet. No items on the assessment require a calculator or other mathematical tools to answer.

Cognitive Complexity

In addition to being created according to content-area content standards, each item on the NM-MSSA Assessment is assigned a Depth of Knowledge (DOK) level according to the cognitive demand of the item, as influenced by the standard being assessed. DOK is not synonymous with difficulty but rates the complexity of the mental processing a student must use to successfully respond to an item.

The Reading items are mainly categorized as DOK level 2, with a smaller percentage making up DOK levels 1 and 3. The DOK level 1 items generally assess basic comprehension and recall of text. The DOK level 2 items generally assess processing of text using some analysis and low-level inferencing. The DOK level 3 items require a deeper analysis or synthesis of ideas in one or more texts.

The Writing and Language items also mainly fall in DOK level 2, with a smaller percentage making up DOK levels 1 and 3. It is of note that items aligned to Writing standards will not generally be designated as DOK level 1 and items assigned to Language standards will not generally be designated as DOK level 3. The DOK level 1 items require communication of simple ideas and application of basic language conventions. DOK level 2 items generally assess the connection of ideas using a simple organizational structure as well as the application of more complex language conventions. The DOK level 3 items require some higher-level processing skills such as synthesis and analysis, as well as a deeper awareness of audience and purpose, while using complex language conventions to communicate effectively.

In Mathematics, SR and MS items lend themselves best to DOK levels 1 and 2, while CR items may reach the complexity required for DOK level 3 (particularly at the upper grade levels). The DOK level 1 items generally assess basic recall and procedural fluency. The DOK level 2 items generally assess application of skills, modeling, and conceptual understanding. The DOK level 3 items require more strategic thinking and reasoning for more complex problems or questions requiring mathematical justification.

Target percentages for the distribution of operational (core) test points across the cognitive complexity levels (DOK classification) per content area are noted in Tables 4-1 through 4-3.

Table 4-1. Depth of Knowledge Distribution—Reading

DOK	Grade					
	3	4	5	6	7	8
Level 1	0–20%	0–20%	0–20%	0–20%	0–20%	0–20%
Level 2	50–70%	50–70%	50–70%	50–70%	50–70%	50–70%
Level 3	20–40%	20–40%	20–40%	20–40%	20–40%	20–40%
Total	100%	100%	100%	100%	100%	100%

Table 4-2. Depth of Knowledge Distribution—Writing and Language

DOK	Grade					
	3	4	5	6	7	8
Level 1	15–35%	15–35%	15–35%	15–35%	15–35%	15–35%
Level 2	40–60%	40–60%	40–60%	40–60%	40–60%	40–60%
Level 3	15–35%	15–35%	15–35%	15–35%	15–35%	15–35%
Total	100%	100%	100%	100%	100%	100%

Table 4-3. Depth of Knowledge Distribution—Mathematics

DOK	Grade					
	3	4	5	6	7	8
Level 1	5–25%	5–25%	5–25%	5–25%	0–20%	0–30%
Level 2	50–80%	50–80%	50–80%	50–80%	50–80%	50–80%
Level 3	5–30%	5–30%	5–30%	5–30%	5–30%	5–30%
Total	100%	100%	100%	100%	100%	100%

For Science, because the New Mexico STEM Ready! Science Standards are NGSS-aligned, the cognitive complexity of the items is evaluated with a different framework than Depth of Knowledge. This framework, Cognitive Complexity Framework for SSIB, is based on Achieve's *A Framework to Evaluate Cognitive Complexity in Science* (September 2019).

Under the Cognitive Complexity Framework for SSIB, four indicators are used to classify the cognitive complexity of each item: stimulus, science and engineering practice, disciplinary core idea, and crosscutting concept. For each indicator, the classification in terms of high, medium, or low complexity is based on how the students are using the indicator to respond to the item – specifically, to what degree does students' engagement with the indicator contribute to the level of sense-making required by the item.

The evaluation of cognitive complexity is done at the individual item level. For an operational NM-ASR test form, after summing the operational points that reflect cognitive complexity at each complexity level across all four indicators, the target distribution is that at least 10% of the total test points should be high cognitive complexity and no more than 35% of the total test points should be low cognitive complexity.

4.3 Item Writer Training

ELA and Mathematics items on the NM-MSSA Assessment and science items on the NM-ASR Assessment were primarily developed by Cognia content specialists using detailed internal specifications documents as well as item development best practices outlined in section 4.4.1. In addition, Cognia content specialists incorporated the *New Mexico Instructional Scope* and *New Mexico Bias and Sensitivity Guidelines* into their item development and subsequent content reviews.

The writing prompts on the NM-MSSA Assessment were primarily developed by New Mexico educators, who received training as part of an Item-Writing Workshop. In May of 2020, NM PED invited teachers from across New Mexico to participate in an Item-Writing Workshop in which they would help develop stimuli and writing prompts for the NM-MSSA Assessment. Approximately 50 teachers from a diverse range of school districts were able to participate. The New Mexico participants were all licensed educators with a variety of experience and expertise, including language arts teachers, special education teachers, EL teachers, instructional leads and coaches, and educators who worked on the committees to develop the NM Instructional Scope documents. See Appendix F for additional details.

In June, the New Mexico teachers received an initial training session facilitated by Cognia and PED. The training included:

- An overview of the vision and goals associated with New Mexico's Balanced Assessment System
- The purpose of the new writing assessment and its role in the NM-MSSA summative assessment
- The components and structure of the writing prompts
- The specifications associated with the writing-prompt stimuli
- An overview of the writing-prompt rubrics

The participants then worked with Cognia content specialists over several weeks in a series of virtual sessions to study the processes of writing passages and associated writing tasks. Out of precaution during the pandemic all sessions were held virtually. Specific training sessions addressed:

- Understanding the Writing standards and three purposes for writing
- Selecting an appropriate, culturally relevant topic
- Moving from a topic to an outline for a passage(s) and associated writing prompt
- Writing the passage or set of passages
- Developing and finalizing a writing task

Throughout the workshop teachers met with Cognia content specialists to draft, refine, and revise their ideas and writing. They met with each other in small peer groups, as well as with Cognia content specialists, while they developed writing prompts specifically designed for students across the state of New Mexico.

Overall, the teachers developed approximately 90 writing prompts, which included over 100 passages and/or stimuli. A total of 54 writing prompts were field-tested on the NM-MSSA 2020–21 Assessment. Additional writing prompts were field-tested on the NM-MSSA 2021–22 Assessment.

4.4 Item Review Committees and Processes

Items used on NM-MSSA and NM-ASR Assessments are developed to measure achievement on the New Mexico Common Core State Standards and the New Mexico STEM Ready! Science Standards in the assessed content areas, respectively. Cognia content specialists, in collaboration with NM PED, ensure this alignment, and ongoing independent evaluations are held to verify alignment. In addition, independent reviews are scheduled to ensure that items and passages conform to bias and sensitivity guidelines.

4.4.1 Content and Item Reviews

The test developers at Cognia review newly developed items for:

- alignment to the intended content standard;
- item integrity, including content and structure, format, clarity, and possible ambiguity;
- desired correct responses;
- appropriateness and quality of graphics;
- appropriateness of scoring-guide descriptions and distinctions;
- completeness of associated item documentation (e.g., scoring guide, content codes, key, grade level, DOK/cognitive complexity); and
- appropriateness for the designated grade level.

Newly developed passages and items for the NM-MSSA and NM-ASR Assessment also undergo review by nationally representative panels of content and assessment experts, including educators from across the state of New Mexico. The purpose of these reviews is to evaluate items and determine their suitability for assessment by answering the following four questions:

- Does the item align with the assigned content standard(s)?
- Is the content accurate?
- Are the content and context grade-level appropriate?
- Does the item provide maximum accessibility for all students?

(Note that for the newly developed items that were field tested in the 2023 NM-ASR tests, however, the educator committee that previously reviewed the content of the items was comprised completely of NM educators, as Science was not using a national model at that time.)

4.4.2 Bias and Sensitivity Review

Bias and sensitivity review is an essential component of the passage- and item-review process. All Cognia content specialists receive training in bias and sensitivity issues. Controversial and biased topics

are avoided in the test development process. Internal reviews include review of not only content but context, with a particular awareness of bias and sensitivity issues that are specific to New Mexico.

Since no one person is well versed in the full spectrum of possible concerns, the bias and sensitivity review committee helps to ensure that all potential issues are identified. All passages and items undergo bias and sensitivity review prior to field-testing.

The bias and sensitivity review committee comprises a diverse group of people who represent a variety of national and international student subgroups, including New Mexican panelists from diverse backgrounds. The people currently serving on this committee include business professionals, educators, a school administrator, an ESL tutor, graduate school students, and retired professionals. United States racial and ethnic groups represented on this committee include African American, Asian American, Hispanic/Latino/Latina, Native American, and White. These representatives have varied experiences with urban/suburban/rural environments and economically disadvantaged students. International populations represented on this committee currently include South American, Middle Eastern, South Asian, and East Asian. We have summarized in the tables below the specific information we have regarding committee member demographics. See Appendix G for lists of New Mexican committee members.

Table 4-4. Number of Bias & Sensitivity Panelists by Gender

Gender	Number
Male	3
Female	6

Table 4-5. Number of Bias & Sensitivity Panelists by Race/Ethnicity

Race/Ethnicity	Number
American Indian	0
Asian	3
Black/African American	2
Hispanic or Latino	2
White (non-Hispanic)	2

Again, note that for the newly developed items that were field tested in the 2023 NM-ASR tests, however, the committee that reviewed the items for any bias/sensitivity issues was comprised completely of NM representatives, as Science was not using a national model at that time.

4.5 Test Forms Construction

The Cognia content specialists and psychometricians work collaboratively to produce operational test forms using sequential and iterative procedures that support both the content and construct validity of the assessments.

4.5.1 Item and Stimulus Selection

Subsequent to field-testing and item data review, Cognia test developers carefully select the items that will appear in the operational tests. In consultation with Cognia psychometricians, test developers consider the following criteria in selecting sets of items for the operational tests:

- **Content coverage/match to test design and blueprints.** The test designs and blueprints stipulate a specific number of items by item type and content distribution.
- **Item difficulty and complexity.** Item statistics are evaluated to ensure quality psychometric characteristics, as well as similar levels of difficulty and complexity from year to year.
- **"Cueing" items.** Items are reviewed for any information that might "cue" or provide information that would help to answer another item.

Test developers sort and lay out passages and items into test forms. During assembly of the test forms, the following criteria are considered:

- **Key patterns.** The sequence of keys (correct answers) is reviewed to ensure that their order appears random.
- **Option balance.** Selected-response items are balanced across forms so that key options are not markedly disproportionate.
- **Page fit.** Items always appear one per screen for online testing. ELA passages and, when applicable, common Mathematics and Science stimuli always appear to the left of the associated item.
- **Visual appeal.** Every effort is made to make each item as accessible as possible. Each item's presentation may differ slightly depending on the delivery method and size of the screen.
- **Psychometric Targets.** The psychometric properties of the collective items on a form are evaluated against a target Test Characteristic Curve (TCC) to help ensure the overall difficulty of the test form is equivalent (to the extent possible) from year to year.

During operational test construction, content specialists use a psychometric tool in our online test construction system to help ensure the desired test design and psychometric criteria for each form have been met. The Psychometrics team uploads a target Test Characteristic Curve (TCC) and Test Information Function (TIF) into the system, which allows the content specialists to evaluate how closely the proposed test is matching the psychometric targets as the form is built. The Content team is then able to export a file containing the test form metadata, along with the psychometric data, for review and approval by the Psychometric team.

Once the operational form has been constructed, the Psychometric team uses a report generated from our online test construction tool to conduct a review of the TCC and the TIF. The review focuses on a comparison of the TCC and TIF values for the proposed and target tests at cut scores 1, 2, and 3. The Psychometric team provides feedback on whether the proposed test sufficiently matches the target test in terms of expected difficulty (i.e., the TCC) and measurement precision (i.e., the TIF) at each cut score. The content team receives a final recommendation from Psychometrics to either a) move forward with the test form if it meets all psychometric parameters or b) improve the degree to which the proposed TCC and/or TIF matches the target at one or more cut scores.

To the extent possible, the content team will subsequently make changes to the test form to bring it closer to the target by replacing items on the form in accordance with guidance provided by the Psychometrics team. If any changes are made to the test form, the Psychometrics team conducts another review and provides final approval.

A reviewer designated by the NM PED per grade level and content area reviews the test form and, prior to approval, specifically considers the following criteria:

- **Construct validity.** The test content is evaluated to determine the degree to which the test measures what it claims, or purports to be measuring, and items/tasks are aligned to the appropriate indicator/standard/measurable outcome.

- **Key accuracy.** Item keys (and the number of designated keys) are reviewed to ensure accuracy.
- **Positive phrasing in item stems.** Items are checked to ensure that negative words such as "not" and/or "except" are rarely, if ever, used.
- **Specific determiners.** Words such as "always," "never," "totally," and "absolutely" are avoided whenever possible to prevent inadvertent cueing of correct or incorrect answers.
- **Clueing/clanging item associations.** The items on the test are reviewed to ensure that the answer to an item is not given away within another item on the same form (clueing) or that an item's context is too similar to another item on the same form (clanging).
- **Bias/sensitivity concerns.** The test is reviewed by all appropriate stakeholders within the NM PED and assessment bureaus to ensure that the content is appropriate for New Mexico students.
- **Errors or typos.** The test is reviewed to verify that the content and metadata are accurate and there appear to be no obvious human errors.

4.5.2 Selection Specifications to Meet Blueprint Requirements

All NM-MSSA and NM-ASR Assessment items are appropriately field-tested prior to operational use. Once stimuli have been field-tested with a set of items, content specialists evaluate the statistics from the items associated with each stimulus. Often, items associated with a stimulus demonstrate a range of student performance, which is largely dependent upon factors inherent to each item. However, if a circumstance is encountered where many items associated with a stimulus are not performing as expected, this is evaluated carefully. While this scenario does not automatically mean the stimulus contains content that is not comprehensible or accessible, it does signal the need to thoroughly review the stimulus in relation to the item content and reevaluate the acceptability of the stimulus. Cognia assessment content specialists can also review all the aspects of item content, and this is especially important when data indicate that further scrutiny is warranted.

The process for item data includes the following information for all field-tested items:

- classical item difficulty for all items (i.e., *p*-value)
- score distributions for polytomous items
- item option selection distribution for multiple-choice items
- 10 most frequent student responses for multi-select items and technology-enhanced items
- item-total and option-total correlations
- Item Response Theory (IRT) statistics
- Differential item functioning (DIF) using the standardization DIF procedure (Dorans & Kulick, 1986)[1] to produce classifications for female versus male, economically disadvantaged versus non-disadvantaged, Asian versus White, Black versus White, Hispanic versus White, Native American versus White, Multi-racial versus White, and Native Hawaiian/Pacific Islander versus White.

[1] *DIF occurs when an item has difficulty measures that vary across contexts for similarly able subgroups of examinees. DIF procedures are designed to identify items on which the performances of certain subgroups of interest differ from each other after controlling for construct-relevant achievement. In order to ensure meaningful results, DIF statistics are not computed for populations containing less than 200 students in both subgroups. Analysis was conducted using field-test data to detect potential DIF at the item level. The standardizations DIF procedure (Dorans & Kulick, 1986) was employed to evaluate subgroup differences. The computed DIF indices have a theoretical range of -1.0 to 1.0 for multiple-choice items. Critical values are defined as 0.05 and 0.10 and the values are flagged as statistically significant, alpha = 0.05. If the absolute value of standardized DIF is equal to or greater than 0.10, the item is classified "C" DIF; items with absolute values greater than or equal to 0.05 are classified as "B" DIF; otherwise, items are classified as "A" DIF.*

The flags listed in Tables 4-6 and 4-7 are used to identify those items that require an additional level of scrutiny.

Table 4-6. Criteria for Flagged Items Based on Classical Test Theory (CTT) Statistics

Item-Flagging Criteria	Concern
If *p*-value of keyed response < 0.10	Item too difficult
If *p*-value of keyed response > 0.90	Item too easy
If *p*-value of distractor* > *p*-value of keyed response	Possible mis-key
If *p*-value of distractor* > 0.35	Possible second correct option
If item-total correlation < 0.15	Poorly discriminating item
If item-total correlation < 0.00	Non-discriminating or negatively discriminating item
If DIF analysis is B or C	Possible bias in item (B, B-, C, C-)

**Note: These analyses examine item score and item option selection distribution for polytomous and selected-response items, respectively.*

Table 4-7. Criteria for Flagged Items Based on Item Response Theory (IRT) Statistics

Item-Flagging Criteria	Concern
If IRT *a*-parameter < 0.50	Poorly discriminating
If IRT *b*-parameter < -3.00	Easy item
If IRT *b*-parameter > 3.00	Hard item
If IRT *c*-parameter > 0.35	Low-ability students answer correctly (i.e., guessing)
If IRT standard error of estimation > 0.10	Possible issue with item fit

In ELA and Mathematics, the item content of each flagged item is reviewed and discussed by at least two content specialists before a decision is made regarding acceptability of the item. At the end of the process, all field-tested items are designated with a status of “Accept,” “Rework,” or “Reject.” Accepted items become eligible for operational testing. Rework items are eligible to be edited and field-tested again so new item data can be generated. Rejected items are removed from the pool of items eligible for operational testing.

In Science, in addition to the type of content specialist review described above, 2022 field test items were also reviewed by NM PED and a committee of New Mexico educators, to provide additional feedback on the performance of the items relative to their usage on the 2023 operational forms.

Cognia understands that item-level data review must be conducted thoroughly and carefully because of the impacts on test construction, which need to be consistent from administration to administration. Being experts in their respective content areas, Cognia’s content specialists also understand that some assessed standards are typically more challenging for students than others, and the specialists are able to simultaneously make good decisions about both content and data in accepting or rejecting items for operational use based on the item statistics. Finally, Cognia understands that items with DIF statistic flags need to be scrutinized for potential sources of bias. While a flag does not automatically mean the item contains biased content, it does signal the need to thoroughly review the item content and evaluate the ways in which the different focal groups would have access and ability to answer the item to ensure it is fair for all students. For ELA and Science, items with C DIF are avoided for operational use. For Math, items with C DIF may be used on a test form if there is no scenario-based context associated with the item (e.g., it is solely an equation, etc.) There would then be no apparent rationale stemming from the item content as to why the item favors one student subgroup over another.

Chapter 5. Test Administration

Orderly and secure test administrations are necessary to protect secure test content and ensure that test data are validity-interpretable to meet score reporting and accountability reporting requirements.

5.1 Roles and Responsibilities

As indicated in the Test Coordinator's Manual, District Assessment Coordinators are the primary source of assessment information for district staff, school staff, parents, and community. It is the District Test Coordinator's (DTC) responsibility to keep the local educational agency (LEA) informed about current assessment policy and changes, and to provide teachers with available resources for content area assessments. Manuals are used to ensure the uniformity of administration procedures from school to school. These manuals—the Test Coordinator's Manual and the Test Administration Manual—stress the importance of test security and ethical administration while the tests are in the schools and contain explicit directions and scripts for test administrators to read aloud to test-takers. These documents may be accessed on the New Mexico Help and Support Website at: https://newmexico.onlinehelp.cognia.org/archived-test-administration/.

Roles of additional staff are listed below.

5.1.1 Test Administrators

Test Administrators are vital to the success of both the NM-MSSA and NM-ASR Assessments. The Test Administrator (TA) must administer the assessment to students by following the procedures provided in the Test Administration Manual. All TAs involved in test administration, preparation, and security are required to attend training provided by the DTC in accordance with the PED regulation 6.10.7 NMAC. TAs must hold one of the following valid PED licenses from the State of New Mexico:

- school instructor;
- administrator;
- school counselor;
- student success advisor; or
- instructional support providers (e.g., educational diagnostician, psychologist, social worker).

Only long-term substitutes who hold one of the above PED licenses may serve as TAs. Short-term substitutes, educational assistants (EAs), school nurses, and coaches may not serve as TAs unless they also hold one of the valid licenses listed above.

In the event that schools require additional staff to administer either the NM-MSSA or the NM-ASR, other staff members (who have received the necessary training and who have signed the PED Confidentiality Agreement) may be used to provide one-to-one accommodations.

5.1.2 School Test Coordinators

The School Test Coordinator (STC) is appointed at the local level. The STC's point of contact for matters relating to assessment is the DTC. In some smaller districts, the DTC serves as STC for one or more schools in the district.

5.1.3 Proctors

Proctors assist TAs but may not administrate a test without a TA present. Proctors are generally Educational Assistants (EAs) but can be any school employee who does not otherwise hold one of the approved PED licenses. No proctor should assist with a group that includes a child who is a close relative (child, grandchild, niece/nephew, etc.).

5.2 Test Administrator Manual

For Spring 2023, the Test Administrator Manual (TAM) outlined the steps to follow before, during, and after administration of the Spring 2023 New Mexico MSSA and New Mexico Assessment of Science Readiness (NM-ASR) Assessments. Understanding of and compliance with each of these steps is vital for successful administration.

The TAM covers administration policies such as security guidelines and administration information, accessibility features and accommodations including requirements for computer-based tests (CBT) and paper-based tests (PBT), preparing for CBTs and PBTs, administering CBTs and PBTs, directions and scripts for use during CBT and PBT administrations, and what to do at the completion of CBTs and PBTs.

5.3 TA and Proctor Training Requirements and 2023 Test Administrations

All TAs and proctors involved in test administration, preparation, and security are required to attend training provided by the DTC in accordance with the PED regulation 6.10.7 NMAC. Training should include information on test security policies and procedures, test administration procedures, documentation and provision of testing accommodations, and the importance of strictly following all directions in the manuals.

5.4 Testing Irregularity Reports

During the Spring 2023 NM-MSSA and NM-ASR testing administration window, the NM PED received 23 testing irregularity reports. Test administrators and coordinators are trained to report test administration irregularities. The NM PED defines a testing irregularity as any incident in the handling or administration of a test that results in questioning the accuracy of the data or security of the test that may or may not result in an invalidation.

Of the 23 reports, 18 were from the Lordsburg school district. In this district, 18 third and fourth graders were administered the assessment remotely. Remote administration of the statewide summative NM-MSSA is against PED policy but is allowed for the interim assessments. After consultation with Cognia's psychometric team the PED determined that these student scores would be reported.

Five other irregularities were submitted for the following reasons:

- While taking the NM-MSSA Mathematics assessment, a student was provided a submit –and-complete option on the second question of the assessment. The student clicked to submit, and the test was over. It was requested that the test be voided and started again. This request was granted, and the student retested.
- A student clicked through the test and then submitted without answering any questions. This test was invalidated, and the student retested.

- A student did not complete session 1 before starting session 2. Session 2 was then not completed. Eventually another test administrator was able to get the student to finish. The DTC recommended invalidating the scores and this request was approved.
- A student did not click submit at the end of the test. A test administrator brought the student in later to log on and submit the completed test. This successfully submitted this test for scoring.
- A student taking the Reading test had no pencil for scratch-paper during a computer-based test administration. It was reported as an irregularity and the scores were submitted.

5.5 Test Security

The New Mexico Statewide Assessment Program requires that the NM-MSSA and NM-ASR tests be treated with the highest level of test security and accountability. The security of NM-MSSA and NM-ASR materials must be maintained before, during, and after the test administration. TAs, proctors, and school and district test coordinators are required to follow the guidelines in the TAM for distributing, collecting, and returning testing materials. All testing personnel are required to have access to a central, locked storage space for safekeeping of test materials until print materials are returned to Cognia.

To maintain the validity of the tests administered in the statewide assessment program, keeping all test questions secure is absolutely necessary. If security is breached or compromised, the assessment results may not be valid. If one student, school, or district has advantages not awarded to another, the test administration is no longer standardized and loses the important distinction of being appropriate for program accountability.

TAs must follow these security guidelines before, during, and after testing:

- Receive training on test security and administration by the STC or the DTC.
- Complete the New Mexico Confidentiality Agreement and return it to the STC. (The Confidentiality Agreement form is available on the PED website).
- Follow the testing schedule established by the school.
- Ensure TA is not assigned to a classroom in which a relative is being tested.
- Carry out standard examination procedures.
- Ensure secure test materials are secured in a central and locked area when not in use.
- Use the security checklist or a similar tracking tool daily, as provided by the STC, during test administration to check in and check out all test materials.
- Report any possible breaches of security to the STC immediately. Examples of security breaches include but are not limited to:
    - improper handling of test materials, such as
        - someone reproducing any student responses,
        - allowing any unauthorized access to test materials before, during, or after testing, or
        - leaving test materials (including computers being used for CBTs) unsecured when the TA or a proctor is not in the classroom, and
    - improper test administration procedures, such as
        - coaching students during testing,
        - altering student responses in any way, or
        - stray mark cleanup, including but not limited to erasing double-marks, lightly erased, or lightly marked answers.

School and district staff members are prohibited from studying or discussing online test questions in any manner, either among themselves or with students before, during, or after testing.

5.5.1 Prevention and Detection Measures and Procedures

The NM PED has a process in place for on-site technical assistance and monitoring of schools to ensure that proper testing administration procedures are being followed. During monitoring visits, the staff member has a checklist of questions to ask and evidence to gather. The monitoring covers the following key topics:

- Communication: This includes how School Test Coordinators receive information from the PED and vendors about the assessments and how information is shared with others.
- Staff Training: This section includes verification of a process to ensure all required staff have completed training prior to testing.
- Test security: Questions in this section are focused on the storage of materials and accurate administration of the assessments.
- Test environment observations: In this section of the checklist, the observer makes note of how many assessments are being administered and observes at least two rooms to ensure protocols are being followed.
- Participation and verification: This section focuses on determining eligibility of students for ACCESS and DLM testing.
- There is also a section to note any other observations and follow-up needed.

Chapter 6. Scoring: Scope of Work, Processes, and Procedures

6.1 Scope of Work

The 2022–23 New Mexico MSSA, ASR, and SBA consist of operational and field-test, multi-point open-ended response items in ELA, SLA, SBA, mathematics, and science. Table 6-1 outlines the number and type of each item per grade.

Table 6-1. Overview of NM Scope-of-Work 2022–2023

Content Area	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade HS
ELA	OPFT–3 / OE–4, 3	OPFT–3 / OE–4, 3	OPFT–3 / OE–4, 3	OPFT–3 / OE–4, 3	OPFT–3 / OE–4, 3	OPFT–3 / OE–4, 3	--
SLA	OPFT–1 / OE–4, 3	OPFT–1 / OE–4, 3	OPFT–1 / OE–4, 3	OPFT–1 / OE–4, 3	OPFT–1 / OE–4, 3	OPFT–1 / OE–4, 3	--
SBA	--	--	--	--	--	--	OP-3 / OR-4 OP-4 / SR-2
Mathematics	OP–2 / OE–2, 1 OP–2 / OE–4, 2 FT–6 / OE–2, 1 FT–6 / OE–4, 2	OP–2 / OE–2, 1 OP–2 / OE–4, 2 FT–7 / OE–2, 1 FT–5 / OE–4, 2	OP–2 / OE–2, 1 OP–2 / OE–4, 2 FT–7 / OE–2, 1 FT–5 / OE–4, 2	OP–2 / OE–2, 1 OP–2 / OE–4, 2 FT–6 / OE–2, 1 FT–6 / OE–4, 2	OP–2 / OE–2, 1 OP–2 / OE–4, 2 FT–6 / OE–2, 1 FT–6 / OE–4, 2	OP–2 / OE–2, 1 OP–2 / OE–4, 2 FT–6 / OE–2, 1 FT–6 / OE–4, 2	-- -- -- --
Science	--	--	OP-3 / OE 4 FT-5 / OE 4	--	--	OP-3 / OE 4 FT-5 / OE 4	OP-3 / OE 4 FT-5 / OE 4

OP=Operational, FT=Field Test, OPFT= Operational Field Test, OE#, #= multi-point open-ended response item.

6.2 NM-MSSA and ASR Operational Scoring: Processes and Procedures

6.2.1 Score Verification of Multiple-Choice Items

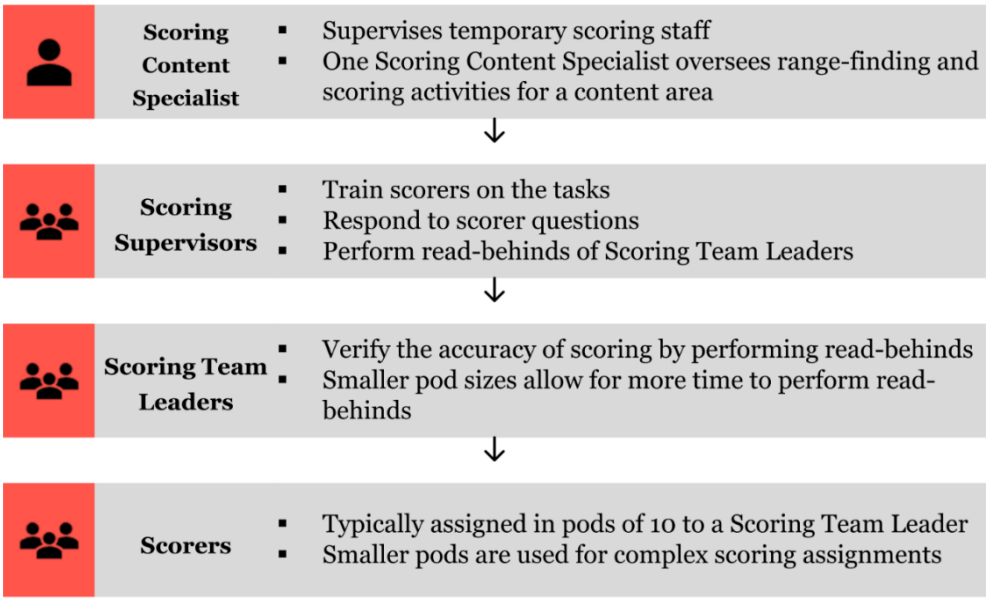
For both computer-based tests (CBT) and paper-based tests (PBTs), responses to multiple-choice items were compared to scoring keys using item analysis software. This robust software compared each student response to multiple-choice items to the respective answer key and assigned a maximum score of 1 point for correct responses and 0 points for incorrect answers. In PBTs, if students filled in multiple bubbles in response to one item, the response was assigned 0 points. At the end of an administration, a second independent validation of all the student responses was conducted to compare and validate results to ensure accurate machine scoring.

6.2.2 Scoring of Open-Ended Response Items

6.2.2.1 Personnel Structure

Cognia's personnel structure for scoring responses consisted of four hierarchical levels as shown in Figure 6-1.

Figure 6-1. Cognia Scoring Staff



All responses were scored by fully vetted scorers who were supervised by Scoring Team Leaders (STLs). The Scoring Supervisors monitored the work of the STLs assigned to them. The Scoring Content Specialist monitored the work of the Scoring Supervisors, STLs, and scorers. Scoring Content Specialists are full-time Cognia staff who report to the Scoring Content Group Manager, who in turn reports to the Director for Content and Quality in the Scoring Services department. This hierarchical structure whereby each level monitors the one below ensures reliable quality and consistency in scoring.

Scoring Content Specialist

The Scoring Content Specialist functioned as the primary lead for his or her designated content area and as a liaison between scoring activities and the Scoring Project Manager to ensure that established quality standards and production schedules were met.

During scoring, the Scoring Content Specialist was responsible for supervising all scoring staff working on the project, including Scoring Supervisors and STLs. The Scoring Content Specialist was also responsible for assuring the consistency and accuracy of scoring work performed by individual scorers and across groups of scorers.

Scoring Supervisors

Scoring Supervisors managed the scorer training and supervised the STLs and scorers working on a designated item and/or content. Scoring Supervisors worked closely with the STLs to ensure consistency and provide counsel and retraining to scorers as necessary. In addition, Scoring Supervisors engaged in supervisory oversight and performed quality-control checks to ensure the consistency and accuracy of the STLs. Scoring Supervisors who were responsible for monitoring training and conducted the retraining of scorers were selected for their ability to instruct and for their level of expertise in their respective disciplines.

Scoring Team Leaders

The STLs were responsible for supervising and monitoring the group of scorers assigned to them. STLs worked closely with their scorers to maintain consistently accurate scoring. They provided quality checks, and they counseled scorers as necessary. STLs were responsible for monitoring and maintaining accurate scoring of their assigned scorers. This included performing read-behinds on scorers and monitoring other quality-control measures. STLs were responsible for arbitrating responses scored by multiple scorers when the assigned scores varied by more than one score point. The arbitration process ensured that such responses received the necessary attention by providing an additional review before assigning a third and final resolution score. In addition to the essential quality control, the arbitration process provided continued opportunities for scorer training.

Because the read-behinds that the STLs performed moderated the scoring process and thus maintained the integrity of the scores, individuals chosen to fill STL positions were selected for their accuracy and content knowledge.

Scorers

Scorers are individuals who evaluate student responses and assign scores.

6.2.2.2 Scorer Recruitment

Cognia actively sought a diverse pool of scorers with a broad range of backgrounds: teachers, scientists, business professionals, graduate school students, retired educators, and the like.

The minimum requirement to assume a position as a scorer or Scoring Team Leader is 48 college credits, which include classes related to the content area being scored. Scoring Supervisors must hold a bachelor's degree with classes related to the content area being scored. In addition, screened bilingual applicants had to demonstrate proficiency in both English and Spanish. Each bilingual applicant must be able to speak, read, write, and translate to and from English and Spanish to carry out their responsibilities in both English and Spanish. All potential scorers and leadership staff submitted documentation (e.g., résumés and/or transcripts) as evidence of meeting the education and experience requirements. Each scorer and leadership staff member signed a binding non-disclosure/confidentiality agreement as well.

6.2.2.3 Scoring Platforms

For the scoring of the 2022-23 test administrations of MSSA and ASR, Cognia engaged two scoring platforms: OSCAR and iScore. iScore was used to score MSSA-ELA, SLA, and SBA and OSCAR was used for the scoring of ASR and MSSA Mathematics due to its capabilities for facilitating online rangefinding. Both systems ensure the security of student responses and test items. During scoring, no student names or schools/districts associated with viewed student work are visible to scorers, and all Scoring Services temporary associates are subject to the same non-disclosure requirements as full-time Cognia staff. Cognia maintained security by using a highly secure server-to-server interface, ensuring that access to all student response images was limited only to scorers, appropriate Cognia staff, and educators participating in the rangefinding activities.

Scorers evaluated most student responses from images rendered by the online testing platforms and a small number of responses from scanned images of paper-based tests. Whether administered in an online or a paper/pencil environment, all responses were scored applying the same scoring criteria.

Prior to the beginning of image scoring, databases were created to receive submitted student responses for each content area and item to be scored. To provide maximum security for all testing and scoring materials, each scorer was required to log on to the scoring systems using a unique combination of an assigned username, a password, and a 6-digit code that was delivered via text or email.

6.2.2.4 Leadership Training

Scoring Supervisors and select STLs were given a separate training session one day before scorer training. Scoring staff, including Scoring Supervisors and STLs, responsible for scoring student responses in iScore were required to achieve the same standard as scorers on item qualification sets: a minimum accuracy scoring rate of 70 percent exact, and 90 percent exact plus adjacent agreement (70/90).

6.2.2.5 Scorer Training

For the scoring of NM-MSSA Mathematics and NM-ASR common operational items, all scorer training was conducted using pre-recorded, interactive training modules.

These modules allowed for self-paced, individual training. Modules were produced by experienced Scoring Supervisors who prepared all training materials for an image slide-show presentation which was overlaid with sound. The format of this training process replicated the traditional face-to-face group training led by a Scoring Supervisor. Each recording started with a discussion of the item and the rubric followed by a detailed discussion of each anchor paper and its rubric-based scoring rationale. After the conclusion of anchor paper training, scorers would gain access to a set of practice papers, to which they would apply the scoring standards as detailed in the rubric and as exemplified in the anchor papers to determine the correct score. For all items that appeared on a prior year's Mathematics or Science test and that had been trained via pre-recorded training modules, the same modules were used to train scorers this year. In doing so, Cognia provides a consistent training experience across the years.

After submitting the score of each practice paper, scorers would get immediate feedback as to whether their score was accurate, and they would receive the justification as to why the response received the score it did. The system is set up such that even if scorers assigned the correct score to the practice paper, they would still receive further explanation of the scoring rationale. For any questions that were not covered by the modules, Scoring Supervisors were available to further elaborate and provide clarification. The modules are designed such that scorers can go back and replay the training on specific papers as needed. This allowed scorers who required more training to review at their own pace, while scorers who were faster in absorbing the scoring standards could move on and proceed with their first attempt to qualify. After module training, scorers continued to have access to electronic versions of the training material in PDF format, so that they could consistently refer to the exemplars during qualification or live scoring.

Scorers were given two opportunities to qualify. If scorers were unable to attain a score match of at least 70 percent exact and 90 percent adjacent agreement on the first qualifying set, they were retrained by discussing the responses contained in the first qualification set with respect to the score-point descriptions of the rubric and by comparing them to the responses of the anchor set. Following this retraining, scoring leadership would administer a second qualification set. If scorers achieved a scoring accuracy rate of at least 70 percent exact and 90 percent adjacent agreement on the second qualification set, then they were allowed to score student responses. Since student responses for Mathematics assessments are assessing two traits (Concepts and Procedures and Mathematical Practices) the minimum threshold of 70% must be achieved on each trait. For ELA, which is also scored on two traits, the 70% threshold applied to both traits combined. Scorers who failed to pass the minimum threshold

were not allowed to score that item. They were either trained on another item or they were dismissed from the project. Appendix H shows the qualification rates for all content areas and grades.

6.2.2.6 Monitoring Scoring Quality

Scorers were required to demonstrate and maintain their ability to score student responses accurately and consistently throughout the scoring process. Both scoring systems enabled scoring leadership to measure and monitor individual and group performance on each scored item in terms of accuracy and consistency, and in terms of read rate (scoring speed) and overall production rate on a constant, real-time basis. Scoring tools employed to measure scoring quality were as follows:

- Read-behind scoring
- Double-blind scoring
- Embedded validity responses
- Recalibration sets

Each scorer's performance on the above procedures was monitored and recorded by the scoring systems and scoring leadership could review data related to the accuracy, consistency, and overall quality of scoring. Scoring leadership was always available to answer scorer questions and counseled and retrained scorers as needed to determine whether a scorer should continue scoring. If a scorer's performance did not meet the prescribed quality standards, scoring leadership initiated a process through which that scorer's work was invalidated and returned to the scoring queue of unscored responses to be re-scored by those scorers who demonstrated scoring accuracy at or above standard.

Read-Behind Scoring

Read-behind scoring allowed scoring leadership to monitor each scorer's scoring performance by way of an immediate real-time snapshot of the scorer's accuracy. The data that were generated by read-behind scoring presented leadership with opportunities to answer questions and to provide counsel to scorers who may have had trouble maintaining the scoring standards.

The number of read-behinds for each scorer varied depending on the accuracy of the scorer. New Mexico scoring specifications for MSSA-ELA, SLA, and SBA require a minimum of two read-behinds per item per hour per scorer, or 10 read-behinds per scorer per full scoring day. For MSSA Math and NM-ASR, the minimum read-behind target was set at 2% of responses scored with an increased goal set if other QC targets were not met. Consistently accurate scorers would only receive the minimum number of read-behinds whereas scorers who exhibited difficulties in maintaining accuracy or consistency received additional read-behinds.

In addition to scorers, scoring leadership was also subject to quality assurance reviews, which were administered by the Scoring Content Specialists. They monitored scoring leadership's accuracy and consistency by reviewing the read-behind results and by performing read-behinds on their STLs.

For the Spanish versions of the NM-MSSA Mathematics and the NM-ASR Assessments, Cognia applied the consensus scoring method. Under this method, two scorers would review student work in tandem and consult with each other on the appropriate score for each student response. This method is particularly effective when the n-count of student work is very low. Scoring accuracy and consistency were maintained via the internal calibration that each scorer provided on the other. Instead of read-behinds, scorers who were selected for consensus scoring the mathematics Spanish responses were constantly monitored by scoring leadership via intermittent participation in the consensus process. This live interaction provided a real-time snapshot of group accuracy.

Double-Blind Scoring and Arbitration Resolution

Double-blind scoring refers to the process of two scorers independently scoring the same response. During this process, neither scorer has any knowledge of the other scorer's score. The double-blind process helps inform scoring leadership about the consistency of scoring among peer scorers who actively score an item. All responses in Mathematics, ELA, and Science Grades 5 & 8 had a minimum of 2 percent of responses double-blind scored. In Science grade 11, the AI engine that is integrated in the scoring platform was employed to support operational scoring. Using a sophisticated algorithm, the AI scoring engine is designed to learn from and mimic human scorers. The model was built using human-scored responses obtained during field-test scoring of these items. Model results were carefully evaluated to determine the appropriate percent of human double-blind scoring. Human second reads were used to verify the reliability of the model in an operational setting. In Science grade 11, AI scoring was applied as the first score with a 10% human double-blind score.

During double-blind scoring, the scoring systems distribute randomly selected responses assigned for double-blind scoring to different scorers without alerting either scorer. Any scoring discrepancies of more than one point between the two scores are routed to an arbitration response queue for resolution by the STL. As described above, the arbitration resolution scoring performed by STLs was blind and did not reveal the previously assigned scorer's scores prior to the STLs entering their score into the system.

The percentage of double-blind responses sent to arbitration by a scorer as a result of a difference in actual scores (i.e., not including blank or unreadable responses) should not have exceeded 10 percent. If a scorer's arbitration percentage exceeded this threshold, scoring leadership counseled, retrained, and/or dismissed the scorer.

Embedded Validity Responses

Validity responses are prescored responses that serve calibration purposes at the onset of scoring an item. The insertion rate for randomly embedded validity responses was 1% per 100 responses scored for NM-ASR and NM-MSSA Math, scorers were not aware when they were scoring an embedded validity response as compared to a live student response. Scorers who demonstrated an accuracy rate of less than 70% exact on each composite score were counseled and the STL increased the number of read-behinds to ensure accuracy.

Recalibration Sets

For NM-MSSA ELA, a set of five calibration papers was administered starting with the second day of scoring an item. This set of five responses, selected by scoring leadership, served as a refresher, and was used to gauge the scorers' ability to maintain accurate scoring of the item on days following their initial item training. Scorers who demonstrated inaccurate scoring on the recalibration set were retrained by the STL or Scoring Supervisor before they could resume live scoring of student responses.

Interrater Reliability

Table 6.5 Summary of Interrater Reliability Statistics for NM-MSSA Math

Grade	Total # of Responses Scored	Total # of Double - Blind Responses Scored	Total % Double-Blind Responses Scored	Score Categories	Score Point Ranges	% Exact	% Adjacent	% Third Reads
3	83,868	1,674	2.0	2	0-2 & 0-4	90.8	6.6	2.6
4	85,853	1,710	2.0	2	0-2 & 0-4	91.6	5.4	3.0
5	87,892	1,754	2.0	2	0-2 & 0-4	93.0	3.7	3.28
6	88,435	1,764	2.0	2	0-2 & 0-4	89.2	0.2	10.6
7	89,942	1,794	2.0	2	0-2 & 0-4	93.3	4.4	2.3
8	95,054	1,896	2.0	2	0-2 & 0-4	89.8	1.7	8.5

Table 6.6 Summary of Interrater Reliability Statistics for NM-ASR

Grade	Total # of Responses Scored	Total # of Double - Blind Responses Scored	Total % Double-Blind Responses Scored	Score Categories	Score Point Ranges	% Exact	% Adjacent	% Third Reads
5	70,614	1,410	2.0	1	0-4	68	30	2.0
8	75,415	1,506	2.0	1	0-4	70	27	3.0
11	66,199	6,439	10.0	1	0-4	55	37	1.8

Table 6.7 Summary of Interrater Reliability Statistics for NM-MSSA ELA

Grade	Total # of Responses Scored	Total # of Double - Blind Responses Scored	Total % Double-Blind Responses Scored	Score Categories	Score Point Ranges	% Exact	% Adjacent	% Third Reads
3	35.038	14,281	40.8	2	1-4 & 1-3	75.3	1.9	22.8
4	36,696	15,442	42.1	2	1-4 & 1-3	77.2	1.1	21.7
5	34,111	12,323	36.1	2	1-4 & 1-3	75.1	1.4	23.5
6	33,753	11,911	35.0	2	1-4 & 1-3	70.1	1.3	28.7
7	31,605	9,308	29.4	2	1-4 & 1-3	71.9	1.3	26.8
8	36,760	13,195	34.8	2	1-4 & 1-3	77.7	1.2	21.1

6.2.2.7 Score-of-Record Rules

Per scoring specifications, the final score-of-record was determined as follows:

- If there was an exact agreement between the scorer and the STL scores, no action was taken—the scorer's original score remained.
- If there was a difference between the scores, either adjacent or discrepant, the STL's score became the score-of-record. Adjacent scores differ by 1 point, while discrepant scores differ by more than 1 point.

6.3 NM-MSSA Mathematics Field Test: Internal Review of Student Work

Due to the low n-count of students who participated in the NM-MSSA Mathematics Field Test, expert scoring staff reviewed student work to determine whether students interacted with the items as expected. Scoring leadership staff provided a written report related to observed trends and student engagement to their colleagues in Content Development to share their observations.

6.4 NM-SBA Operational and NM-ASR Operational Scoring

Due to the low number of students participating in the SBA HS Operational, all scoring was conducted by expert scoring leadership staff applying a consensus scoring approach, whereby two staff members would review student work in tandem and consult with each other on the appropriate score for each student response. Instead of applying the above-mentioned quality control tools used during NM-MSSA Mathematics and ELA and NM-ASR Science operational scoring, scoring accuracy and consistency were maintained via the internal calibration that each staff member provided on the other. The scoring of student work in response to the NM-ASR operational administration followed the procedures as described in section 6.2.

Chapter 7. Classical Item and Test Analysis

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; in particular, racial, ethnic, or gender groups.

Cognia conducts quantitative analyses to help ensure that test items meet these standards. These include statistical evaluations of (1) difficulty indices, (2) item-test correlations, and (3) dimensionality. The details and results for (1) and (2) are presented in this chapter, while the details and results for the dimensionality analyses are presented in section 8.2. All these analyses are based on the administration of NM-MSSA and NM-ASR assessments in spring 2023. Note that the information presented for all these analyses is based on operational items (the items on which student scores are calculated).

7.1 Classical Item Statistics

All operational items were evaluated in terms of classical item difficulty, which under classical test theory practices is defined as the average scored response on an item, divided by the maximum possible score for the item. Although this index is traditionally described as an estimate of item difficulty, it is properly interpreted as an easiness index. The greater in value a classical item difficulty is, the easier the item.

Items that are answered correctly by almost all students provide little information about differences in student abilities, but they 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 they may indicate knowledge or skills that have not yet been mastered by most students. In general, to provide adequate measurement, classical difficulty indices should range from near-chance performance (e.g., 0.25 for four-option multiple-choice items) to 0.90, with a majority of items generally falling around 0.4 to 0.7. However, on standards-referenced assessments such as the NM-MSSA and NM-ASR, it is appropriate to include 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-total correlation is referred to as the item's classical discrimination because it indicates the extent to which successful performance on an item discriminates between high and low scores on the test. Each of the item-total correlations reported here is the Pearson correlation between scored responses on a given item and total raw scores. This Pearson correlation is commonly referred to as the point-biserial correlation (for a dichotomously scored item) and a point-polyserial correlation (for a polytomously scored item). The theoretical range of these correlations is –1.0 to +1.0, with a typical observed range from 0.2 to 0.6. Discrimination indices can be thought of as measures of how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, the

discrimination index can be thought of as a measure of construct consistency. Tables 7-1 and 7-2 list the mean and standard deviation (SD) P-Value and Item-Total Correlations for the operational items in English forms.

A comparison of indices across grade levels is complicated because these indices are population-dependent. Direct comparisons would require that either the items or the students were common across groups. Since that is not the case, it cannot be determined whether differences in these classical indices across grade levels are due to differences in student abilities, differences in item difficulties, or both. Classical item difficulties and item-total correlations are provided in Appendix I.

Table 7-1. Summary Classical Item Statistics for Dichotomous Items

Content Area	Grade	Mean P-Value	SD P-Value	Mean Item-Total Correlation	SD Item-Total Correlation
ELA	3	0.53	0.11	0.46	0.10
	4	0.54	0.12	0.45	0.11
	5	0.52	0.13	0.44	0.10
	6	0.53	0.14	0.43	0.10
	7	0.53	0.15	0.41	0.10
	8	0.53	0.15	0.41	0.10
Mathematics	3	0.52	0.15	0.41	0.11
	4	0.50	0.16	0.41	0.11
	5	0.49	0.16	0.41	0.11
	6	0.48	0.16	0.40	0.11
	7	0.47	0.16	0.40	0.12
	8	0.46	0.17	0.39	0.12
Science	5	0.46	0.17	0.39	0.12
	8	0.45	0.17	0.39	0.12
	11	0.45	0.16	0.39	0.12

Table 7-2. Summary Classical Item Statistics for Polytomous Items

Content Area	Grade	Mean P-Value	SD P-Value	Mean Item-Total Correlation	SD Item-Total Correlation
ELA	3	0.28	0.17	0.58	0.04
	4	0.28	0.17	0.57	0.07
	5	0.29	0.16	0.58	0.08
	6	0.30	0.15	0.58	0.08
	7	0.30	0.15	0.57	0.09
	8	0.30	0.14	0.57	0.09
Mathematics	3	0.30	0.14	0.58	0.09
	4	0.29	0.14	0.59	0.09
	5	0.29	0.13	0.60	0.10
	6	0.28	0.13	0.60	0.10
	7	0.28	0.13	0.60	0.10
	8	0.27	0.14	0.61	0.10
Science	5	0.31	0.16	0.58	0.11
	8	0.33	0.16	0.55	0.12
	11	0.34	0.15	0.53	0.12

7.2 Total Test and Subscore Intercorrelations

When subscores are strongly related to each other, it implies a high internal consistency between subscores. The Pearson correlation matrices among the individual reporting categories (i.e., subscores) are shown in Tables 7-3 and 7-4 for Reading and Mathematics, respectively. The Spring 2023 Writing and Language assessment had only total test scores, without any additional reporting categories. As such, no subscore correlations are reported here for Writing and Language. Results generally indicate that the subscores correlate well with one another and with overall total scores.

Table 7-3. Pearson Correlations of Total Test and Subtest Raw Scores on NM-MSSA ELA, as A Function of Grade

Grade	OP Form	Subtest	Number of Items	Number of Points	Total Test	Subtests						
						1	2	3	4	5	6	7
3	1	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.960	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.923	0.958	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.777	0.817	0.617	1.000	--	--	--	--
		Text Type - Informational Text	11	13	0.919	0.954	0.907	0.793	1.000	--	--	--
		Text Type - Literary Text	9	10	0.873	0.915	0.886	0.728	0.751	1.000	--	--
		Writing & Language	14	20	0.898	0.762	0.739	0.605	0.736	0.685	1.000	--
		Writing Analysis & Language Conventions	12	14	0.897	0.765	0.744	0.602	0.739	0.687	0.976	1.000
	2	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.960	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.923	0.958	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.777	0.817	0.617	1.000	--	--	--	--
		Text Type - Informational Text	11	13	0.919	0.954	0.907	0.793	1.000	--	--	--
		Text Type - Literary Text	9	10	0.873	0.915	0.886	0.728	0.751	1.000	--	--
		Writing & Language	14	20	0.895	0.756	0.734	0.596	0.730	0.678	1.000	--
		Writing Analysis & Language Conventions	12	14	0.897	0.765	0.744	0.602	0.739	0.687	0.971	1.000
	3	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.960	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.923	0.958	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.777	0.817	0.617	1.000	--	--	--	--
		Text Type - Informational Text	11	13	0.919	0.954	0.907	0.793	1.000	--	--	--
		Text Type - Literary Text	9	10	0.873	0.915	0.886	0.728	0.751	1.000	--	--
		Writing & Language	14	20	0.897	0.762	0.741	0.602	0.736	0.685	1.000	--
		Writing Analysis & Language Conventions	12	14	0.897	0.765	0.744	0.602	0.739	0.687	0.979	1.000
4	1	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.949	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.914	0.957	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.775	0.827	0.630	1.000	--	--	--	--
		Text Type - Informational Text	10	13	0.902	0.949	0.912	0.779	1.000	--	--	--
		Text Type - Literary Text	10	10	0.839	0.886	0.843	0.743	0.695	1.000	--	--
		Writing & Language	14	20	0.891	0.727	0.708	0.577	0.692	0.639	1.000	--
		Writing Analysis & Language Conventions	12	14	0.888	0.729	0.712	0.576	0.695	0.642	0.971	1.000
	2	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.949	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.914	0.957	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.775	0.827	0.630	1.000	--	--	--	--
		Text Type - Informational Text	10	13	0.902	0.949	0.912	0.779	1.000	--	--	--
		Text Type - Literary Text	10	10	0.839	0.886	0.843	0.743	0.695	1.000	--	--
		Writing & Language	14	20	0.888	0.726	0.708	0.575	0.691	0.639	1.000	--
		Writing Analysis & Language Conventions	12	14	0.888	0.729	0.712	0.576	0.695	0.642	0.979	1.000
	3	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.949	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.914	0.957	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.775	0.827	0.630	1.000	--	--	--	--

continued

Grade	OP Form	Subtest	Number of Items	Number of Points	Total Test	Subtests 1	2	3	4	5	6	7
4	3	Text Type - Informational Text	10	13	0.902	0.949	0.912	0.779	1.000	--	--	--
		Text Type - Literary Text	10	10	0.839	0.886	0.843	0.743	0.695	1.000	--	--
		Writing & Language	14	20	0.892	0.731	0.712	0.580	0.696	0.643	1.000	--
		Writing Analysis & Language Conventions	12	14	0.888	0.729	0.712	0.576	0.695	0.642	0.975	1.000
5	1	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.943	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.901	0.954	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.795	0.844	0.645	1.000	--	--	--	--
		Text Type - Informational Text	10	12	0.890	0.939	0.911	0.766	1.000	--	--	--
		Text Type - Literary Text	10	11	0.840	0.896	0.836	0.791	0.689	1.000	--	--
		Writing & Language	14	20	0.861	0.683	0.656	0.569	0.651	0.600	1.000	--
		Writing Analysis & Language Conventions	12	14	0.866	0.699	0.673	0.579	0.666	0.613	0.937	1.000
	2	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.943	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.901	0.954	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.795	0.844	0.645	1.000	--	--	--	--
		Text Type - Informational Text	10	12	0.890	0.939	0.911	0.766	1.000	--	--	--
		Text Type - Literary Text	10	11	0.840	0.896	0.836	0.791	0.689	1.000	--	--
		Writing & Language	14	20	0.863	0.696	0.670	0.579	0.664	0.611	1.000	--
		Writing Analysis & Language Conventions	12	14	0.866	0.699	0.673	0.579	0.666	0.613	0.958	1.000
	3	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.943	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	13	15	0.901	0.954	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	7	8	0.795	0.844	0.645	1.000	--	--	--	--
		Text Type - Informational Text	10	12	0.890	0.939	0.911	0.766	1.000	--	--	--
		Text Type - Literary Text	10	11	0.840	0.896	0.836	0.791	0.689	1.000	--	--
		Writing & Language	14	20	0.864	0.694	0.667	0.579	0.660	0.611	1.000	--
		Writing Analysis & Language Conventions	12	14	0.866	0.699	0.673	0.579	0.666	0.613	0.957	1.000
6	1	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.932	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.793	0.846	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.873	0.940	0.615	1.000	--	--	--	--
		Text Type - Informational Text	7	9	0.822	0.881	0.798	0.796	1.000	--	--	--
		Text Type - Literary Text	13	14	0.863	0.927	0.743	0.898	0.639	1.000	--	--
		Writing & Language	14	20	0.852	0.644	0.550	0.603	0.568	0.598	1.000	--
		Writing Analysis & Language Conventions	12	14	0.852	0.649	0.549	0.611	0.570	0.603	0.941	1.000
	2	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.932	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.793	0.846	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.873	0.940	0.615	1.000	--	--	--	--
		Text Type - Informational Text	7	9	0.822	0.881	0.798	0.796	1.000	--	--	--
		Text Type - Literary Text	13	14	0.863	0.927	0.743	0.898	0.639	1.000	--	--
		Writing & Language	14	20	0.846	0.642	0.547	0.602	0.566	0.595	1.000	--
		Writing Analysis & Language Conventions	12	14	0.852	0.649	0.549	0.611	0.570	0.603	0.955	1.000
	3	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.932	1.000	--	--	--	--	--	--

continued

Grade	OP Form	Subtest	Number of Items	Number of Points	Total Test	Subtests						
						1	2	3	4	5	6	7
6	3	Reading Strategy - Analysis and Interpretation	7	8	0.793	0.846	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.873	0.940	0.615	1.000	--	--	--	--
		Text Type - Informational Text	7	9	0.822	0.881	0.798	0.796	1.000	--	--	--
		Text Type - Literary Text	13	14	0.863	0.927	0.743	0.898	0.639	1.000	--	--
		Writing & Language	14	20	0.846	0.645	0.547	0.605	0.567	0.598	1.000	--
		Writing Analysis & Language Conventions	12	14	0.852	0.649	0.549	0.611	0.570	0.603	0.958	1.000
7	1	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.921	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.692	0.765	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.873	0.940	0.500	1.000	--	--	--	--
		Text Type - Informational Text	8	9	0.826	0.863	0.641	0.821	1.000	--	--	--
		Text Type - Literary Text	12	14	0.804	0.902	0.706	0.840	0.561	1.000	--	--
		Writing & Language	14	20	0.838	0.621	0.448	0.598	0.595	0.509	1.000	--
		Writing Analysis & Language Conventions	12	14	0.849	0.630	0.454	0.607	0.606	0.515	0.944	1.000
	2	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.921	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.692	0.765	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.873	0.940	0.500	1.000	--	--	--	--
		Text Type - Informational Text	8	9	0.826	0.863	0.641	0.821	1.000	--	--	--
		Text Type - Literary Text	12	14	0.804	0.902	0.706	0.840	0.561	1.000	--	--
		Writing & Language	14	20	0.839	0.615	0.444	0.592	0.590	0.504	1.000	--
		Writing Analysis & Language Conventions	12	14	0.849	0.630	0.454	0.607	0.606	0.515	0.935	1.000
	3	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.921	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.692	0.765	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.873	0.940	0.500	1.000	--	--	--	--
		Text Type - Informational Text	8	9	0.826	0.863	0.641	0.821	1.000	--	--	--
		Text Type - Literary Text	12	14	0.804	0.902	0.706	0.840	0.561	1.000	--	--
		Writing & Language	14	20	0.844	0.624	0.453	0.600	0.599	0.512	1.000	--
		Writing Analysis & Language Conventions	12	14	0.849	0.630	0.454	0.607	0.606	0.515	0.945	1.000
8	1	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.930	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.738	0.796	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.881	0.946	0.557	1.000	--	--	--	--
		Text Type - Informational Text	9	9	0.841	0.886	0.626	0.880	1.000	--	--	--
		Text Type - Literary Text	11	14	0.844	0.923	0.800	0.839	0.639	1.000	--	--
		Writing & Language	14	20	0.855	0.652	0.507	0.624	0.617	0.570	1.000	--
		Writing Analysis & Language Conventions	12	14	0.866	0.671	0.519	0.642	0.635	0.585	0.944	1.000
	2	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.930	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.738	0.796	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.881	0.946	0.557	1.000	--	--	--	--
		Text Type - Informational Text	9	9	0.841	0.886	0.626	0.880	1.000	--	--	--
		Text Type - Literary Text	11	14	0.844	0.923	0.800	0.839	0.639	1.000	--	--
		Writing & Language	14	20	0.863	0.670	0.525	0.639	0.629	0.589	1.000	--
		Writing Analysis & Language Conventions	12	14	0.866	0.671	0.519	0.642	0.635	0.585	0.959	1.000

continued

Grade	OP Form	Subtest	Number of Items	Number of Points	Total Test	Subtests						
						1	2	3	4	5	6	7
8	3	Total Test	34	43	1.000	--	--	--	--	--	--	--
		Reading	20	23	0.930	1.000	--	--	--	--	--	--
		Reading Strategy - Analysis and Interpretation	7	8	0.738	0.796	1.000	--	--	--	--	--
		Reading Strategy - Comprehension	13	15	0.881	0.946	0.557	1.000	--	--	--	--
		Text Type - Informational Text	9	9	0.841	0.886	0.626	0.880	1.000	--	--	--
		Text Type - Literary Text	11	14	0.844	0.923	0.800	0.839	0.639	1.000	--	--
		Writing & Language	14	20	0.862	0.663	0.519	0.631	0.624	0.581	1.000	--
		Writing Analysis & Language Conventions	12	14	0.866	0.671	0.519	0.642	0.635	0.585	0.947	1.000

Table 7-4. Pearson Correlations of Total Test and Subtest Raw Scores on NM-MSSA Mathematics English Forms Per Grade

Grade		Subtest	Number of Items	Number of Points	Total Test	Subtests				
						1	2	3	4	5
3	1	Total Test	41	51	1.000	--	--	--	--	--
		Measurement & Data/Geometry	12	18	0.940	1.000	--	--	--	--
		Modeling/Structure & Repeated Reasoning	12	13	0.881	0.748	1.000	--	--	--
		Number & Operations in Base Ten/Number & Operations - Fractions	13	14	0.869	0.733	0.688	1.000	--	--
		Operations & Algebraic Thinking	21	22	0.932	0.867	0.831	0.811	1.000	--
		Problem Solving/Reasoning & Argument	14	15	0.917	0.834	0.819	0.841	0.769	1.000
	2	Total Test	40	50	1.000	--	--	--	--	--
		Measurement & Data/Geometry	12	18	0.940	1.000	--	--	--	--
		Modeling/Structure & Repeated Reasoning	12	13	0.881	0.748	1.000	--	--	--
		Number & Operations in Base Ten/Number & Operations - Fractions	12	13	0.855	0.720	0.677	1.000	--	--
		Operations & Algebraic Thinking	20	21	0.928	0.867	0.832	0.788	1.000	--
		Problem Solving/Reasoning & Argument	14	15	0.917	0.834	0.819	0.839	0.767	1.000
4	1	Total Test	41	51	1.000	--	--	--	--	--
		Measurement & Data/Geometry	9	12	0.857	1.000	--	--	--	--
		Modeling/Structure & Repeated Reasoning	19	24	0.955	0.727	1.000	--	--	--
		Number & Operations in Base Ten/Number & Operations - Fractions	9	9	0.725	0.521	0.599	1.000	--	--
		Operations & Algebraic Thinking	20	22	0.932	0.787	0.897	0.684	1.000	--
		Problem Solving/Reasoning & Argument	17	17	0.910	0.778	0.851	0.735	0.754	1.000
5	1	Total Test	41	51	1.000	--	--	--	--	--
		Measurement & Data/Geometry	6	7	0.765	1.000	--	--	--	--
		Modeling/Structure & Repeated Reasoning	18	22	0.941	0.646	1.000	--	--	--
		Number & Operations in Base Ten/Number & Operations - Fractions	13	16	0.904	0.611	0.748	1.000	--	--
		Operations & Algebraic Thinking	21	23	0.954	0.701	0.909	0.854	1.000	--
		Problem Solving/Reasoning & Argument	15	15	0.883	0.695	0.839	0.806	0.760	1.000
6	1	Total Test	44	54	1.000	--	--	--	--	--
		Geometry/Statistics & Probability	9	10	0.818	1.000	--	--	--	--
		Modeling/Structure & Repeated Reasoning	18	24	0.948	0.686	1.000	--	--	--
		Problem Solving/Reasoning & Argument	13	14	0.797	0.538	0.649	1.000	--	--
		Ratios & Proportional Relationships	22	23	0.934	0.784	0.873	0.750	1.000	--
		The Number System/Expressions & Equations	17	18	0.893	0.715	0.845	0.762	0.727	1.000

continued

Grade		Subtest	Number of Items	Number of Points	Total Test	Subtests 1	2	3	4	5
7	1	Total Test	44	54	1.000	--	--	--	--	--
		Geometry/Statistics & Probability	9	12	0.825	1.000	--	--	--	--
		Modeling/Structure & Repeated Reasoning	16	20	0.929	0.676	1.000	--	--	--
		Problem Solving/Reasoning & Argument	15	16	0.894	0.637	0.739	1.000	--	--
		Ratios & Proportional Relationships	25	26	0.943	0.764	0.882	0.858	1.000	--
		The Number System/Expressions & Equations	15	16	0.890	0.749	0.807	0.814	0.742	1.000
8	1	Total Test	45	55	1.000	--	--	--	--	--
		Functions	9	12	0.831	1.000	--	--	--	--
		Geometry/Statistics & Probability	14	17	0.832	0.582	1.000	--	--	--
		Modeling/Structure & Repeated Reasoning	18	20	0.887	0.619	0.597	1.000	--	--
		Problem Solving/Reasoning & Argument	17	17	0.872	0.763	0.747	0.749	1.000	--
		The Number System/Expressions & Equations	22	24	0.919	0.730	0.726	0.865	0.666	1.000

Table 7-5. Pearson Correlations of Total Test and Subtest Raw Scores on NM-Science Grade 5 as a Function of Operational Form

Subtest	Number of Items	Number of Points	Total Test	1	2	3
Operational Form 1						
Total Test	35	64	1.000	--	--	--
1. Earth and Space Sciences	13	24	0.815	1.000	--	--
2.Life Sciences	11	20	0.872	0.726	1.000	--
3.Physical Sciences	11	20	0.858	0.712	0.751	1.000
Operational Form 2						
Total Test	35	64	1.000	--	--	--
1.Earth and Space Sciences	13	24	0.764	1.000	--	--
2.Life Sciences	11	20	0.863	0.739	1.000	--
3.Physical Sciences	11	20	0.857	0.709	0.764	1.000
Operational Form 3						
Total Test	35	64	1.000	--	--	--
1.Earth and Space Sciences	13	24	0.808	1.000	--	--
2.Life Sciences	11	20	0.850	0.719	1.000	--
3.Physical Sciences	11	20	0.809	0.689	0.721	1.000

Table 7-6. Pearson Correlations of Total Test and Subtest Raw Scores on NM-Science Grade 8 as a Function of Operational Form

Subtest	Number of Items	Number of Points	Total Test	1	2	3
			Operational Form 1			
Total Test	35	64	1.000	--	--	--
1. Earth and Space Sciences	12	22	0.816	1.000	--	--
2.Life Sciences	12	22	0.778	0.696	1.000	--
3.Physical Sciences	11	20	0.813	0.693	0.664	1.000
			Operational Form 2			
Total Test	35	64	1.000	--	--	--
1.Earth and Space Sciences	12	22	0.773	1.000	--	--
2.Life Sciences	12	22	0.785	0.736	1.000	--
3.Physical Sciences	11	20	0.803	0.721	0.719	1.000
			Operational Form 3			
Total Test	35	64	1.000	--	--	--
1.Earth and Space Sciences	12	22	0.781	1.000	--	--
2.Life Sciences	12	22	0.728	0.703	1.000	--
3.Physical Sciences	11	20	0.739	0.723	0.684	1.000

Table 7-7. Pearson Correlations of Total Test and Subtest Raw Scores on NM-Science Grade 11 as a Function of Operational Form

Subtest	Number of Items	Number of Points	Total Test	1	2	3
			Operational Form 1			
Total Test	37	68	1.000	--	--	--
1. Earth and Space Sciences	12	22	0.824	1.000	--	--
2.Life Sciences	13	24	0.804	0.719	1.000	--
3.Physical Sciences	12	22	0.809	0.690	0.702	1.000
			Operational Form 2			
Total Test	37	68	1.000	--	--	--
1.Earth and Space Sciences	12	22	0.807	1.000	--	--
2.Life Sciences	13	24	0.794	0.724	1.000	--
3.Physical Sciences	12	22	0.779	0.716	0.729	1.000
			Operational Form 3			
Total Test	37	68	1.000	--	--	--
1.Earth and Space Sciences	12	22	0.761	1.000	--	--
2.Life Sciences	13	24	0.803	0.696	1.000	--
3.Physical Sciences	12	22	0.766	0.674	0.691	1.000

Chapter 8. Psychometrics: Item Response Theory (IRT) Scaling and Equating

This chapter describes the procedures used to scale the NM-MSSA and NM-ASR tests. For the Spring 2023 administration, the NM-MSSA operational tests were (mostly) pre-equated while the NM-ASR operational tests were administered operationally for the first time.

8.1 IRT Models

All NM-MSSA and NM-ASR items were calibrated using item response theory (IRT). IRT uses mathematical models to define a relationship between an unobserved measure of student proficiency, usually referred to as theta (θ), and the probability (p) of getting a dichotomous item correct or of getting a particular score on a polytomous item. 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 & van der Linden, 1997; Hambleton & Swaminathan, 1985). 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. This estimate, $\hat{\theta}$, is considered to be an estimate of the student's performance. It has characteristics that may be preferable to those of raw scores for equating and scaling purposes.

For the NM-MSSA and NM-ASR Assessments, the three-parameter logistic (3PL) model was used for dichotomous (selected-response) items and the Graded-Response Model (GRM) was used for polytomous (constructed-response) items. The 3PL model for dichotomous items can be defined as:

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

Where

U indexes the scored response on an item,

i indexes items,

j indexes students,

a represents item discrimination,

b represents item difficulty,

c is the lower asymptote parameter, and

D is a normalizing constant equal to 1.701.

In the GRM for polytomous items, an item is scored in a k + 1 graded category that can be viewed as a set of k dichotomies. At each point of dichotomization (i.e., at each threshold), a two-parameter model can be used. This implies that a polytomous item with a k + 1 category can be characterized by k Item Category Threshold Curves (ICTCs) of the two-parameter logistic form:

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

Where

U indexes the scored response on an item,

i indexes the items,

j indexes students,

k indexes threshold,

a represents item discrimination,

b represents item difficulty,

d represents item category threshold, and

D is a normalizing constant equal to 1.701.

After computing k ICTCs in the GRM, k + 1 Item Category Characteristic Curves (ICCCs) are derived by subtracting adjacent ICTCs:

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

where P_{ik} represents the probability that the score on item i falls in category k, and

P_{ik}^{*} represents the probability that the score on item i falls at or above the threshold k

Note that P_{i0}^{*} = 1 and $P_{i(m+1)}^{*}$ = 0.

The GRM is also commonly expressed as:

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

The Item Characteristic Curve (ICC) for polytomous items is computed as a weighted sum of ICCCs, where each ICCC is weighted by a score assigned to a corresponding category:

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

See Lord and Novick (1968), Hambleton and Swaminathan (1985), and Baker and Kim (2004) for more information about item calibration and parameter estimation.

8.2 Dimensionality Analysis

Tests are constructed with multiple content-area subcategories and their associated knowledge and skills. Hence, the potential exists for dimensions being invoked beyond the common primary dimension. Generally, the content-area subcategories are highly correlated with each other, and the primary dimension they share typically explains an overwhelming majority of the variance in test scores. 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 scaling and equating of the NM-MSSA and NM-ASR tests.

The purpose of the dimensionality analysis presented in this report is to investigate whether violation of the assumption of test unidimensionality is statistically detectable and, if so, the degree to which unidimensionality is violated. Findings from dimensionality analyses performed on the NM-MSSA operational items for ELA and Mathematics and NM-ASR for science are reported below. (Note: Only operational items were analyzed since they are used for score reporting.)

The dimensionality analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999). Both methods use the estimated average conditional covariances for item pairs as their basic statistical building block. A conditional covariance is the covariance between two items conditioned on 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 indicate 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 display local dependence, conditioned on total score on the non-clustered items. The DIMTEST statistic follows a standard normal distribution under the null hypothesis of unidimensionality.

The DETECT statistic is an effect-size measure 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 multidimensionality; values of 0.4 to 1.0, moderate multidimensionality; and values greater than 1.0, strong multidimensionality (e.g., Roussos & Ozbek, 2006).

DIMTEST and DETECT were separately applied to the NM-MSSA reading, writing and language, and mathematics tests per grade. First, each dataset was split into a training sample and a cross-validation sample.

DIMTEST was then applied to each sample, and the DIMTEST null hypothesis was rejected at a significance level of 0.05 for every grade level per content area. Next, DETECT was used to estimate the effect size for the violations of local independence for all the tests. Table 8-1 displays the multidimensional DETECT effect size estimates, which indicate very weak to weak levels of multidimensionality for every test.

Table 8-1. DETECT Multidimensional Effect Size, as a Function of Content Area and Grade*

Content Area	Grade	OP Form	Multidimensional Effect Size	Interpretation
ELA	3	1	0.282	Small
		2	0.273	Small
		3	0.227	Small
	4	1	0.231	Small
		2	0.237	Small
		3	0.220	Small
	5	1	0.234	Small
		2	0.309	Small
		3	0.252	Small
	6	1	0.237	Small
		2	0.229	Small
		3	0.250	Small
	7	1	0.199	Negligible
		2	0.186	Negligible
		3	0.236	Small
	8	1	0.251	Small
		2	0.239	Small
		3	0.260	Small
Mathematics	3	1	0.200	Negligible
	4	1	0.192	Negligible
	5	1	0.324	Small
	6	1	0.224	Small
	7	1	0.224	Small
	8	1	0.181	Negligible
Science	5	1	0.237	Small
		1	0.236	Small
		1	0.283	Small
	8	2	0.180	Negligible
		2	0.178	Negligible
		2	0.126	Negligible
	11	3	0.083	Negligible
		3	0.109	Negligible
		3	0.089	Negligible

**Calculations based on those students attempting five or more items on the English forms of the given NM-MSSA assessment. Multidimensional effect size < 0.20 interpreted as negligible, 0.20 to 0.40 as small, 0.40 to 1.00 as moderate, and greater than 1.00 as strong.*

8.3 Item Response Theory Results

The tables in Appendix J give the IRT item parameters of all common items on the 2021–22 New Mexico MSSA tests by grade and content area.

Test characteristic curves (TCCs) are based on the IRT item parameters and display the expected (average) raw score associated with each θ_j value between –4.0 and 4.0, or equivalently the expected (average) raw score associated with each observable scale score (see Section 8.4 for details on scale scores). Mathematically, the TCC is computed by summing the ICCs of all items that contribute to the raw score. Using the notation introduced in Section 7.1, the expected raw score at a given value of θ_j is

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

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

j indexes students (here, θ_j runs from –4 to 4), and

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

U indexes the scored response on an item,

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.

Test information functions (TIFs) display the amount of statistical information the test provides at each value of θ_j, or equivalently display the amount of statistical information the test provides at each observable scale score. TIFs depict test score precision across the entire latent trait continuum. There is an inverse relationship between the information from a test and its conditional standard error of measurement (CSEM). The CSEM at a given θ_j [$CSEM(\theta_j)$] is equal to the inverse of the square root of the statistical information at θ_j (e.g., Hambleton, Swaminathan, & Rogers, 1991). That is, the $CSEM(\theta_j)$ is equal to the inverse of the square root of the TIF at a given θ_j [$TIF(\theta_j)$], the expression for which can be written as follows:

$$CSEM(\theta_j) = \frac{1}{\sqrt{TIF(\theta_j)}}$$

Compared to the tails, TIFs are often higher near the middle of the θ distribution, where most students are located and where most items are sensitive by design.

Appendix K contains graphs of the TCC and CSEM, for each content area and grade. Each TCC graph displays the expected raw score (on the vertical axis) for each scale score (on the horizontal axis). Each TCC graph also has a set of vertical lines that indicate the values of the scale score cut scores for the given content area and grade. Each CSEM graph displays the scaled CSEM (see Section 8.4 below) value (on the vertical axis) at each scale score (on the horizontal axis). Each CSEM graph also has a set of vertical lines that indicate the values of the scale score cut scores for the given content area and grade.

8.4 Equating

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year or to equate one year's forms to those given in the previous year.

The NM-MSSA Spring 2023 test forms were pre-equated. The pre-equating process uses item bank values of the IRT item parameters to place the pre-equated test form onto the established IRT scale. Equating ensures that students are not given an unfair advantage or disadvantage because the test form they took is easier or harder than those taken by other students.

8.5 Reported Total Test and Subtest Scale Scores

The *θ* scale used in IRT calibrations is not readily understood by most stakeholders. As such, reporting scales are used for NM-MSSA reporting. The reporting scales are linear transformations of the underlying *θ* scale. To obtain a student's scale score on a given assessment, the student's raw score (i.e., total number of points earned) is translated into a value on the underlying *θ* scale using TCC mapping. The student's *θ* value is translated into a scale score (*SS*) using the following linear equation:

$$SS = \beta_0 + \beta_1\theta$$

where β_0 is an intercept constant and

β_1 is a slope constant,

m is the slope, and

b is the intercept.

The CSEM can also be translated into a scaled CSEM. Whereas values of the CSEM are on the *θ* scale, values of the scaled CSEM are on the reporting scale. The scaled CSEM is obtained via the following equation:

$$Scaled\ CSEM = \beta_1 \times CSEM(\theta)$$

Table 8-2 shows the slope and intercept terms used for the Spring 2023 NM-MSSA and NM-ASR Assessments to calculate the scale scores. See Appendix L for Raw to Scale score Lookup Tables.

Table 8-2. Spring 2023 Scale score Slopes and Intercepts by Content Area and Grade

Content Area	Grade	Slope	Intercept
ELA	3	20.0	352.92840
	4	20.0	457.08760
	5	20.0	556.80480
	6	20.0	654.68740
	7	20.0	755.63620
	8	20.0	857.12120
Mathematics	3	17.5	352.33728
	4	17.5	452.26693
	5	17.5	555.09773
	6	17.5	661.65970
	7	17.5	759.23910
	8	17.5	852.68080
Science	5	12.5	553.56675
	8	10.0	855.10120
	11	7.5	1159.72130

It is important to note that converting from raw scores to *θ* values to scale scores does not change students' achievement-level classifications. Given the relative simplicity of raw scores, it is fair to question why scale scores are reported instead of raw scores. Scale scores make the reporting of grade-level results consistent across test forms and administrations. It is this uniformity across scale scores that facilitates the understanding of student performance. The psychometric advantage of scale scores over raw scores comes from their being linear transformations of *θ*. Since the *θ* scale is used for pre- or post-equating, scale scores are comparable from one year to the next. Raw scores are not.

8.6 Performance Levels

The cut scores used for the Spring 2023 NM-MSSA Assessments are the cut scores that were originally established for the Cognia Interim Assessment, on which New Mexico iMSSA is based (see Appendix M for the NM iMSSA 2021–22 Technical Report Addendum).The interim cut scores were used for the Spring 2023 NM-MSSA Assessments, given the need to report performance level results coupled with the decision to delay setting performance standards for NM-MSSA until 2022. The decision to delay NM-MSSA standard setting was based on the effect of disruption of the COVID-19 pandemic on student instruction and student learning, as well as the impact of the pandemic on Spring 2023 NM-MSSA participation rates.

The cut scores on the theta scale and the reporting scale, used for the Spring 2023 NM-MSSA and NM-ASR Assessments, are presented in Table 8-3.

Table 8-3. Spring 2023 Cutpoints on the Theta Metric and Reporting Scale by Content Area and Grade

		Theta Cut Score			Scale score Cut Score		
Content Area	**Grade**	**1**	**2**	**3**	**1**	**2**	**3**
ELA	3	-0.84070	0.35358	0.87096	336	360	370
	4	-0.84036	0.14562	0.80754	440	460	473
	5	-0.67811	0.15976	0.84779	543	560	573
	6	-1.12288	0.26563	0.95086	632	660	673
	7	-1.20560	0.21819	0.99071	731	760	775
	8	-0.82541	0.14394	0.71275	840	860	871
Mathematics	3	-0.59939	0.43787	1.46087	341	360	377
	4	-0.42244	0.44189	1.61624	444	460	480
	5	-0.38771	0.28013	1.05367	548	560	573
	6	-0.85783	-0.09484	1.00975	646	660	679
	7	-0.59970	0.04348	0.65422	748	760	770
	8	-0.63353	0.41824	1.48261	841	860	878
Science	5	-0.75048	0.51466	1.70117	544	560	574
	8	-0.96101	0.48988	2.73095	845	860	882
	11	-0.76114	0.03716	2.91134	1154	1160	1181

8.6.1 Percentages of Students in Each Performance Level

The performance level distributions for both English and Spanish forms of the Spring 2023 administration of NM-MSSA and NM-ASR Assessments are shown in Table 8-4.

Table 8-4. Performance Level Distribution on NM-MSSA and ASR English Forms, as a Function of Content Area and Grade*

Grade	Number of Students	Frequency of Students				Percentage of Students			
		Novice	Nearing Proficiency	Proficient	Advanced	Novice	Nearing Proficiency	Proficient	Advanced
ELA									
3	20,877	5,462	7,757	3,195	4,463	26.2	37.2	15.3	21.4
4	21,324	5,940	6,540	4,457	4,387	27.9	30.7	20.9	20.6
5	21,870	7,651	6,089	4,925	3,205	35.0	27.8	22.5	14.7
6	22,037	3,971	9,773	4,780	3,513	18.0	44.3	21.7	15.9
7	22,341	4,005	9,279	6,080	2,977	17.9	41.5	27.2	13.3
8	23,594	6,091	8,311	4,767	4,425	25.8	35.2	20.2	18.8
Mathematics									
3	20,884	9,852	6,338	3,851	843	47.2	30.3	18.4	4.0
4	21,365	9,944	6,262	4,219	940	46.5	29.3	19.7	4.4
5	21,872	9,206	5,204	4,628	2,834	42.1	23.8	21.2	13.0
6	22,054	9,738	5,489	5,235	1,592	44.2	24.9	23.7	7.2
7	22,310	12,095	4,894	2,919	2,402	54.2	21.9	13.1	10.8
8	23,601	11,307	7,979	3,644	671	47.9	33.8	15.4	2.8
Science									
5	21,874	5,767	8,900	5,529	1,678	26.4	40.7	25.3	7.7
8	23,595	4,166	12,519	6,694	216	17.7	53.1	28.4	0.9
11	21,157	6,702	6,322	7,980	153	31.7	29.9	37.7	0.7

**Calculations based on those students attempting five or more items on the given NM-MSSA and ASR Assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table 8-5. Performance Level Distribution on NM-MSSA and ASR Spanish Forms, as a Function of Content Areas and Grade*

Grade	Number of Students	Frequency of Students				Percentage of Students			
		Novice	Nearing Proficiency	Proficient	Advanced	Novice	Nearing Proficiency	Proficient	Advanced
SLA									
3	662	257	283	68	54	38.8	42.7	10.3	8.2
4	611	314	202	73	22	51.4	33.1	11.9	3.6
5	272	126	115	29	2	46.3	42.3	10.7	0.7
6	269	116	136	13	4	43.1	50.6	4.8	1.5
7	259	106	121	29	3	40.9	46.7	11.2	1.2
8	278	152	107	15	4	54.7	38.5	5.4	1.4
SMA									
3	680	423	190	60	7	62.2	27.9	8.8	1.0
4	615	398	148	67	2	64.7	24.1	10.9	0.3
5	284	193	59	27	5	68.0	20.8	9.5	1.8
6	272	190	46	33	3	69.9	16.9	12.1	1.1
7	269	225	30	9	5	83.6	11.2	3.3	1.9
8	287	216	64	6	1	75.3	22.3	2.1	0.3
Science									
5	275	126	127	21	1	45.8	46.2	7.6	0.4
8	280	96	165	19	0	34.3	58.9	6.8	0.0
11	299	161	107	31	0	53.8	35.8	10.4	0.0

**Calculations based on those students attempting five or more items on the given NM-MSSA and ASR Assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Chapter 9. Score Reliability

9.1 Classical Reliability Analyses

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 and complement one another. 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: test-retest, alternate forms, split-half, and internal consistency. 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 problem of remembering items 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 forms 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.

Internal consistency reliability reflects the degree to which the items on a test form are related to (or correlate with) each other. Cronbach (1951) provided a statistic, α (coefficient alpha), that estimates

internal consistency reliability. Coefficient alpha is equivalent to the average of all possible split-half reliabilities. The formula for Cronbach's α is as follows:

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

where i indexes the item,

n is the total number of items,

$\sigma^2_{(Y_i)}$ represents an individual item variance, and

σ^2_x represents the total test variance.

Cronbach's α is used to estimate the (unconditional) classical standard error of measurement (SEM), which is given by

$$SEM = \sqrt{\sigma^2_x(1-\alpha)}$$

9.2 IRT Marginal Reliabilities

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

$$\sigma^2_X = \sigma^2_T + \sigma^2_E$$

where σ^2_X is the observed-score variance,

σ^2_T is the true-score variance, and

- σ^2_E is the error variance.

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

$$CTT\ Reliability = 1 - \frac{\sigma^2_E}{\sigma^2_X}.$$

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

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

where $\overline{CSEM(\theta)^2}$ is the mean squared CSEM,

$\overline{CSEM(SS)^2}$ is the mean squared scaled CSEM,

$Var(\hat{\theta})$ is the variance of theta estimates, and

$Var(SS)$is the variance of scale scores.

Using this formula, IRT marginal reliability estimates were calculated for each assessment using the scale scores (and their standard errors).

The reliability of a test can also be evaluated by simply examining directly the CSEMs themselves. CSEMs facilitate the interpretation of individual scale scores. With any given scale-score estimate for a student, the reasonable limits of the true scale score for the student can be calculated by using the CSEM for the scale score.

The tables in Appendix N contain Coefficient α, (classical) SEM, and IRT marginal reliability for the spring 2023 administration of the New Mexico MSSA & ASR tests.

At the total test level and per grade, Coefficient α ranged from 0.84 to 0.90 in ELA, 0.80 to 0.90 in Mathematics, and 0.88 to0.89 in science. Also, at the total test level and per grade, IRT marginal reliability ranged from 0.79 to 0.86 in ELA, 0.61 to 0.78 in Mathematics, and 0.90 for all Science. Note that IRT marginal reliability is partially dependent upon the variance in scale scores. When present, range restriction in smaller samples can reduce the variance in scale scores and therefore reduce the resulting value of IRT marginal reliability.

While subgroup reliability results are included in Appendix N for subgroups with at least 50 students, many of the subgroups have fewer than 100 students per content area and grade. Because the subgroup reliabilities are based on very small samples, no interpretations ought to be made on the adequacy of these subgroup reliabilities.

Given that, the results in Appendix N should be interpreted with appropriate levels of caution. Reliabilities are dependent not only on the measurement properties of a test, but also on the statistical distribution of the studied subgroup. Additionally, reliability estimates can be artificially depressed for subgroups with little variability in test scores (Draper & Smith, 1998).

9.3 Decision Accuracy and Consistency

While related to reliability, the accuracy and consistency of classifying students into achievement categories are even more important statistics in a standards-based reporting framework (Livingston & Lewis, 1995). After the achievement levels were specified and students were classified into those levels, empirical analyses were conducted to estimate the statistical accuracy and consistency of the classifications.

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. Evaluation of decision accuracy is essential, considering all test scores contain measurement error. Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are given to the same group of students. In operational test programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and consistency of classification decisions based on a single administration of a test. The Livingston and Lewis (1995) technique was used to estimate decision accuracy and consistency because the method is easily adaptable to all types of testing formats, including mixed-format tests. The Livingston and Lewis technique uses “true scores” as the term is defined in classical test theory. 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 2021–22 NM-MSSA tests, after various technical adjustments (described in Livingston & Lewis, 1995), a three-by-three contingency table of accuracy was created for each grade and content area, where cell [i, j] represented the estimated proportion of students whose true score fell into performance level i (where i = 1 to 3) and observed score into performance level j (where j = 1 to 3). 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 three-by-three contingency table was created for each grade and content area to show 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 performance level i (where i = 1 to 3) and whose observed score on the second form would fall into performance level j (where j = 1 to 3). 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 κ (kappa; Cohen, 1960), which indicates 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:

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

where $C_{i.}$ is the proportion of students whose observed achievement level would be Level i (where i = 1–3) on the first hypothetical parallel form of the test;

$C_{.i}$ is the proportion of students whose observed achievement level would be Level i (where i = 1–3) on the second hypothetical parallel form of the test; and

C_{ii} is the proportion of students whose observed achievement level would be Level i (where i = 1–3) on both hypothetical parallel forms of the test.

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

The tables in Appendix O contain the decision accuracy and consistency results for the Spring 2023 administration of NM-MSSA. These tables include overall accuracy and consistency indices, kappa, accuracy and consistency values conditional on achievement level, and 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 cutpoint and whose true scores were below the cutpoint. A false negative is the proportion of students whose observed scores were below the cutpoint and whose true scores were above the cutpoint.

For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, if the conditional accuracy value is 0.85 for any achievement level, this figure indicates that among the students whose true scores placed them in this classification, 85 percent would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.80 indicates that 80 percent of students with observed scores in any achievement level would be expected to score in this classification again if a second, parallel test form were used.

Note that, as with other methods of evaluating reliability, accuracy, and consistency, 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 O should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare accuracy and consistency statistics between

grades and content areas. Decision accuracies and consistencies generally ranged from 0.6 to 0.8 at the overall level. At the level of performance level, decision accuracies were stronger for the Needs Support performance level than those for the Near Target and On Target performance levels. This is arguably due to the number of students at each performance level. Fewer students fell in the Near Target and On Target performance levels.

Chapter 10. Score Reporting

10.1 Relationship to SIUs

Score interpretation and use (SIU) statements are claims about how test scores and other performance information can be interpreted and used to guide decisions and actions. We conduct all activities subsequent to development of the MSSA and ASR SIU statements—starting from the performance level descriptors (PLDs) to test design, item development and forms development, and psychometric analyses—to support the SIUs. SIUs also indicate the score reporting elements that we can and should include in score reports.

For example, consider the following NM-MSSA SIU:

> *NM-MSSA scores provide reliable and valid information about important knowledge and skills in grade-level numeracy and literacy that students with the most significant cognitive disabilities are attaining*.

The claims and subclaims in this interpretation statement are that we can report NM-MSSA scores and student proficiency levels because the scores are supported by evidence of score reliability and evidence of validity such as dimensionality and equating studies, thereby supporting the inclusion of student scores and proficiency levels on individual score reports.

10.2 Score Reports

In the Spring of 2022, the Individual Student Reports were redesigned. In the previous year the results for New Mexico MSSA were reported on the interim scale. The reports were produced by subject. In the Spring of 2022, the reports contained the results for the MSSA ELA and math assessments for grades 3-8. These results were combined with science results from the ASR assessment in grades 5 and 8. Grade 11 reports contained the results from the ASR assessments. The report was redesigned to be printed on 11 x 17 paper for students in grades 3-8. The grade 11 reports remained in 8 1/2 x 11 paper. The reports are printed in color. The individual student report contains the following: scale scores, performance levels, standard error, and reporting category performance indicators. Ways to Support text that was added in the Spring of 2022 was removed in the Spring of 2023. Item level reporting was added to the reports in Spring 2023. For additional information concerning the student report, see Appendix P– Reporting Business Requirements.

10.3 Scale Score

A scale score is a numerical value that summarizes student performance. Not all students respond to the same set of test items, so each student's scale score accounts for the slight differences in difficulty among the various forms and administrations of the test. The resulting scale score allows for an appropriate comparison across test forms and administration years within a grade or course and content area. NM-MSSA and NM-ASR reports provide overall scale scores for Reading, Writing and Language, and Mathematics and in grades 5,8 and 11, science, which determine a student's performance level for each content area. Scale-score ranges differ by grade for all tests.

For example, a student who earns an overall scale score of 800 on one form of the grade 8 Mathematics assessment would be expected to earn an overall scale score of 800 on any other form of the grade 8

Mathematics assessment. Furthermore, the student's overall scale score and level of mastery of concepts and skills would be comparable to a student who took the same assessment the previous year or the following year. For cumulative scale-score distributions see Appendix Q; for scale score descriptive statistics, see Appendix R.

A reading scale score and a Writing & Language scale score is also reported for Language Arts. These scores are reported on the same scale as the overall scale score.

10.4 Lexile and Quantile Scores

The NM-MSSA provides additional measures of students' abilities in reading and mathematics. Lexile measures indicate students' reading abilities and Quantile measures indicate students' abilities and readiness for instruction in mathematics.

A Lexile measure indicates the difficulty of materials a student can read and understand. Lexile measures are reported for students who earned an overall Language Arts scale score. The Lexiles are only reported for students who took the Language Arts test in English. A lookup table is used to assign the Lexile score based on the overall Language Arts scale score. Lexile measures can help parents/guardians and educators find reading materials at an appropriate difficulty level for students and monitor a student's growth over time.

A Quantile measure shows what mathematical skills a student has mastered and which skills they may need additional instruction in. Quantile measures are reported for students who earned an overall Mathematics scale score. The Quantiles are only reported for students who took the mathematics test in English. A lookup table is used to assign the Quantile score based on the overall mathematics scale score. Quantile measures can be used by parents/guardians and educators to identify the mathematical skills needed to support students' learning and monitor their growth in mathematics over time.

A QR code is provided in the Individual Student Reports for each subject which links to the website where students and parents/guardians can find out more information about Lexiles and Quantiles.

10.5 Performance Level

Each NM-MSSA/ASR performance level is a broad category that is defined by a student's overall scale score and is used to report overall student performance by describing how well students met the expectations for their grade level/course. There are four performance levels for the Spring 2023 NM-MSSA and NM-ASR Assessments: Advanced, Proficient, Nearing Proficiency, and Novice.

There are Policy Definitions that guide the development of items, test forms, and reporting interpretations for NM-MSSA, SLA, and NM-ASR.

Range PLDs specific to each content area also exist. Range PLDs describe the knowledge and skills that students throughout the range of each proficiency level are expected to be able to demonstrate in each grade and content area. For example, in line with the nature of the science standards, the science range PLDs combine science and engineering practices, disciplinary core ideas, and crosscutting concepts that students in grades 5, 8, and 11 are expected to integrate and demonstrate. The range PLDs appear in Appendix B.

Reporting PLDs summarize the overall expectation of knowledge, skills, and abilities for each performance level. These appear on student reports to support the reporting of NM-MSSA performance in ELA and Mathematics, SLA, and the translated versions of NM-MSSA Mathematics and New Mexico

Assessment of Science Readiness (NM-ASR). The Reporting PLDs appear in the Score Report Interpretation Guide (Appendix S).

10.6 Reporting Category Performance Indicators

Reporting Category performance for NM-MSSA and NM-ASR assessments is indicated by performance indicators that indicate whether the student performed above standard, at/near standard, or below standard in each reporting category. Additional information about reporting category performance indicators is in the Score Report Interpretation Guide, Appendix S in this document.

10.7 Individual Test Questions

On each subject page there is a table containing the following information:

- Report Sequence– this is a number that corresponds to the sequencing of items in an online resource which lists items with item metadata.
- Reporting category abbreviation – These values indicate the reporting category assignment for each item being reported in the table. A dash (-) is used to indicate that the item is not associated with any reporting category.
- Language Arts has a row for Reading Strategy abbreviation. A dash (-) is used to indicate that item does not have a reading strategy assigned.
- Points Earned on each item – The values are formatted so that the points earned on the item over the total points possible for the item e.g., 1/2 indicating the student earned 1 point out of a possible 2 points for the item.
- The writing prompt traits are represented in the table with abbreviations in the report sequence row. PW is used for Production of Writing and UC for Use of Conventions. The numerical value for report sequence is the same for both traits.
- Mathematics has a row for Practice Category abbreviations – A dash (-) indicates that a Practice Category is not assigned to the item.
- A blank space in the points earned column indicates the student did not answer the question.

10.8 Performance Comparison

On the individual student report the student's performance is compared to the performance of students in their school, district, and the state overall. Inclusion rules for the school, district and state aggregations are described in the Reporting Business Requirements attached in Appendix P.

The comparison is made based on the student's earned scale score to the average scale score at the school, district, and state levels.

10.9 Additional Resources

For each content area, additional resources are provided to support families in the development of these skills at home.

Chapter 11. Validity Arguments to Support Intended Score Interpretations and Uses

This chapter presents the primary intended score interpretation and two primary intended score uses. This chapter also presents the claims and subclaims that underlie these three score interpretations and uses (SIUs) and the evidence that supports the claims and subclaims. The New Mexico MSSA and ASR validity argument model is introduced and applied to develop validity arguments to support the four SIUs.

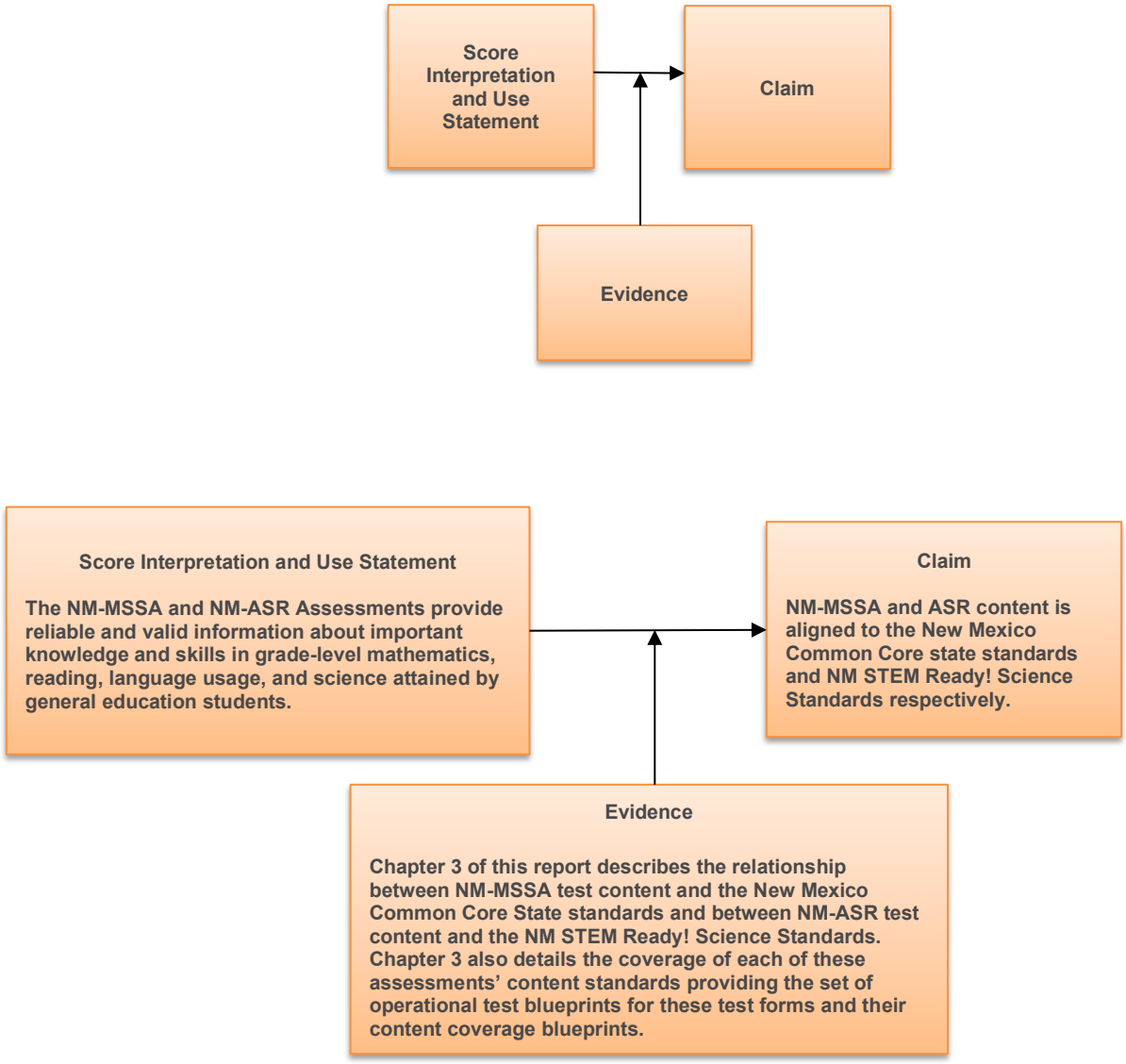
It is important to note that the 2023 NM-MSSA and ASR tests were administered at the end of a school year in which COVID-19 still had a strong impact on instruction and learning. The fact that the 2023 NM-MSSA is pre-equated shields the item parameters, equating results, and psychometric characteristics of the 2023 assessment from deleterious COVID-19 effects. That shielding enables valid interpretations of student performance in 2023, which is likely to reflect whatever deleterious COVID-19 effects there may be, specifically loss of high-quality opportunity to learn and impacts on test performance. The combination of these two facts (pre-equated model and the similarity of student results from past years) indicates that the scores can be interpreted similarly in 2023 and 2019.

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

Further, *Standards* states that “a sound validity argument integrates various strands of evidence into a coherent account of the degree to which existing evidence and theory support the intended interpretations of test scores for specific uses” (p. 21). We use these views in the *Standards,* that evidence must be used to support score interpretation and use claims, as the basis for the NM-MSSA validity argument model, which we describe next.

Emerging practice in state assessment programs is to construct validity arguments based on Toulmin’s model of argumentation (Toulmin. 1958), Chapelle’s proposed practice-oriented adaptation (2021), and Kane’s formulation of validity arguments (2013). A model for NM-MSSA validity arguments, derived from these three conceptualizations, is shown in Figure 11-1. The first panel shows the NM-MSSA model; the second panel is an illustration for an NM-MSSA validity argument for a score interpretation and use statement.

Figure 11-1. NM-MSSA and NM-ASR Validity Argument Model



Adapted from Chapelle (2021) Figures 2.1-2.3, Kane (2013) Figure 1, and Toulmin (1958).

Table 11-1. Relationships Among Score Interpretations and Uses, Claims and Sub-Claims, and Supporting Evidence

Claims	Claims and Subclaims that Support Score Interpretations and Uses
	SIU 1: Primary Intended Score Interpretation
	The NM-MSSA and ASR assessments are designed to measure whether students are on track to be ready for college or career, as defined by the state standards. The NM-MSSA and ASR Assessments provide reliable and valid information about important knowledge and skills in grade-level standards attained by general education students.
Claim 1.1: The content of the tests represents the content of the standards.	
	1.1.1 Assessment content is aligned to the New Mexico Common Core State Standards and New Mexico STEM Ready! Science Standards
	1.1.2 Assessment items are aligned to the New Mexico Common Core State Standards and New Mexico STEM Ready! Science Standards.
Claim 1.2: The test items are construct-relevant.	
	1.2.1. Items require application of the knowledge, skills, and abilities (KSAs) of the targeted construct.
	1.2.2. Items are free of bias and sensitivity issues.
Claim 1.3: Test scores on the NM-MSSA and ASR Assessments provide reliable information about student performance and accurate classifications into performance levels.	
	1.3.1. Test scores and performance level categorizations are adequately reliable for their intended purpose.
	1.3.2. Item characteristics support intended interpretations about all students who take the assessment.
	1.3.3. Test characteristics support intended interpretations about all students who take the assessment.
Claim 1.4: Item and test scoring is implemented accurately.	
	1.4.1. Machine-scored items were scored accurately.
	1.4.2. Constructed-response item scoring training and monitoring procedures met industry standards.
	SIU 2: Intended Score Use for Individual Students
	Performance on the NM-MSSA and ASR indicates a student's progress toward college and career readiness. NM- MSSA and ASR scale scores can be used to compare an individual student's performance to the performance of other students in the school, district, and state.
Claim 2.1: Educators, schools, and districts can use results from the NM-MSSA and ASR Assessments to describe student achievement status with respect to mastery of the content standards.	
	2.1.1. Test scores and performance level categorizations of individual students are adequately reliable and valid measures of student achievement status with respect to mastery of the content standards.
	SIU 3: Intended Score Use for Groups of Students
	SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.
Claim 3.1: Educators can use results from the NM-MSSA and ASR Assessments to support instructional planning for groups of students.	
	3.1.1. Teachers find the performance level descriptors and their students' performance levels useful for planning instruction, especially for students whose test scores fall within performance levels 1 and 2.
	3.1.2. Teachers find their students' scale score information useful for planning instruction, especially for students whose test scores fall within performance levels 1 and 2.
Claim 3.2: Schools and districts can use results from the NM-MSSA and ASR Assessments to make comparisons between organizations (e.g., schools, districts).	
	3.2.1. Test scores and performance levels for groups of students are adequately reliable and valid to enable school, district, and state leaders to monitor changes in means, standard deviations, and performance level percentages for classroom, school, district, and state groups.
	3.2.2. Test scores and proficiency level categorizations of groups of students are adequately reliable and valid to enable monitoring of grade-level performance and student-cohort performance.

Evidence that supports SIUs and claims in NM-MSSA and ASR validity arguments is summarized below, using the rating scale defined in Table 11-2.

Table 11-2. Relevance and Completeness or Completeness of Evidence in Support of SIUs and Claims Underlying Validity Arguments for NM-MSSA and ASR Score Interpretations and Uses

Complete Evidence	When all required pieces of relevant evidence are provided to support a validity argument
Moderate to Substantial Evidence	When several pieces of relevant evidence are provided, but not all required pieces of evidence are provided
Limited Evidence	When only one or two pieces of evidence are provided, where the evidence may be only marginally relevant or where more than 1or 2 pieces of evidence are required
No Evidence	When no relevant evidence exists

11.1 Primary Intended Score Interpretation

The primary intended score interpretation for NM-MSSA and ASR (SIU 1) states that the Assessments provide reliable and valid information about important knowledge and skills in grade-level Reading, Language Usage, Mathematics, and Science attained by general education students.

Claim 1.1. The content of the tests represents the content of the standards.

Items used on NM-MSSA and NM-ASR Assessments are developed to measure achievement on the New Mexico Common Core state standards and New Mexico STEM Ready! Science Standards respectively. Additionally, a third-party independent contractor completed a content alignment study on both the NM-MSSA and NM-ASR Assessments. The results indicate that the content of the assessments represents the New Mexico content standards adopted for both NM-MSSA ELA, and Mathematics and NM-ASR. In addition, independent reviews that involved New Mexico educators were conducted to ensure that items and passages conform to bias and sensitivity guidelines.

Subclaim1.1.1. NM-MSSA Assessment content is aligned to the New Mexico Common Core State Standards and NM-ASR Assessment content is aligned to the New Mexico STEM Ready! Science Standards.

Evidence: Chapter 3 of this report describes the relationship between NM-MSSA and ASR test content and either the New Mexico Common Core State Standards or New Mexico STEM Ready! Science Standards. Chapter 3 also details the coverage of the content standards on NM-MSSA and NM-ASR, providing the set of operational test blueprints for test forms and the content coverage blueprints. Overall, the alignment study indicated there was strong degree of alignment between the NM-MSSA and NM-ASR test forms and the standards / PEs they are intended to measure. Each test form was found to either fully or partially meet the criteria.

Summary of evidence: Complete evidence.

Subclaim 1.1.2. Assessment items are aligned to the New Mexico Common Core State Standards and New Mexico STEM Ready! Science Standards.

Evidence: Chapter 4 describes the item specifications and standardized item writer training in support of new item development. Chapter 4 also details the item review process performed by

item review committees to ensure item content alignment with the intended content standard. The results of the independent alignment study indicate that the assessment content is aligned with New Mexico state content standards. Overall, the study indicated there was strong degree of alignment between the NM-MSSA and NM-ASR test forms and the standards / PEs they are intended to measure. Each test form was found to either fully or partially meet the criteria.

Summary of evidence: Complete evidence.

Claim 1.2. The test items are construct-relevant.

Subclaim 1.2.1. Items require application of the KSAs of the targeted construct.

Evidence: The 2023 operational NM-MSSA and ASR items are aligned to the New Mexico state content standards. The evidence for element 1.2.1 is directly linked to the subclaims 1.1.1 and 1.1.2 above.

Summary of evidence: Complete evidence.

Subclaim 1.2.2. Items are free of bias and sensitivity issues.

Evidence*:* During the item development process, the items followed a rigorous development cycle that includes reviews by New Mexico PED staff and by Item Content and Bias and Sensitivity panelists. The item development process also includes data reviews, during which item-level statistics—including differential item functioning (DIF) statistics—are reviewed. See Chapter 4 for a detailed description of the item review process.

Additionally, Cognia has undertaken an Equity Enhancement Evaluation process in which all steps in the Cognia PADDI process (Principled Assessment Design, Development, and Implementation) are being examined to correct shortcomings in principles and practices related to equitable assessment and opportunities to enhance equity in our assessment practices. One outcome of this process may be the identification of the need for more evidence to support this subclaim.

Summary of evidence*:* Complete evidence, based on current Cognia procedures for the Spring 2023 testing season. Cognia has undertaken an Equity Enhancement Evaluation process in which all steps in the Cognia PADDI process (Principled Assessment Design, Development, and Implementation) are being examined to correct shortcomings in principles and practices related to equitable assessment and opportunities to enhance equity in our assessment practices. One outcome of this process may be the identification of the need for more evidence to support this subclaim.

Claim 1.3: Test scores on the NM-MSSA and ASR Assessments provide reliable information about student performance and accurate classifications into performance levels.

Subclaim 1.3.1. Test scores and performance level categorizations are adequately reliable for their intended purpose.

Evidence:

Score Reliability: Chapter 9 provides a description of both classical and IRT reliability theory and interpretation and a review of the relevant equations. Appendix N contains the reliability results by content area and grade. Appendix N also contains reliability results disaggregated by student subgroups. These reliability estimates are consistent with industry standards, which can be observed in technical reports posted online by other state assessment programs.

Scale score Standard Errors: Chapter 8 provides a description of calculation and interpretation of the scale scores and Chapter 9 provides a description of the calculation of the standard error for a scale score. The average standard error for reported scale scores is reported in Appendix R. The scale score standard error can be compared to the scale score range and the scale score standard deviation to provide some context for interpretation. These standard error estimates are consistent with industry standards, which can be observed in technical reports posted online by other state assessment programs.

Decision Consistency and Accuracy Estimates: Decision accuracy is an estimate of the probability that the observed classification is the true classification. Decision consistency is an estimate of the probability that students would receive the same classification if they tested twice on parallel forms. Chapter 9 describes the theory and equations underlying the estimation of classification accuracy and consistency. Decision accuracy and consistency results are provided in Appendix O. These decision consistency and accuracy estimates are consistent with industry standards, which can be observed in technical reports posted online by other state assessment programs.

Summary of evidence: Complete evidence.

Subclaim 1.3.2. Item characteristics support intended interpretations about all students who take the assessments.

Evidence: The psychometric characteristics most pertinent to evaluating the adequacy of individual items are the estimated item parameters. The item parameter estimates are provided in Appendix J. For dichotomously scored items, the item parameters include the discrimination, difficulty, and lower asymptote parameters. For polytomously scored items, the item parameter estimates include the discrimination, location, and item-category parameters. All items undergo statistical analyses at the time of field-testing, including classical, DIF, and IRT analyses. As stated in Chapter 4, the results of these analyses are reviewed in Data Review meetings with the New Mexico educators and PED staff. After field-testing and prior to operational administration, items from the previous operational administration are reviewed for their item information function (IIF) contributions at the performance level cuts to evaluate and rate the quality of each item. After each operational administration, dimensionality analyses are also conducted to determine how the items correlate with each other in terms of the underlying constructs of the test.

Summary of evidence: Complete evidence.

Subclaim 1.3.3. Test characteristics support intended interpretations about all students who take the assessments.

Evidence:

Model fit analysis verified that the IRT model fits the assessment data for all grades and content areas.

High correlations (e.g., greater than or equal to 0.7) among content area subdomain indicators (e.g., Reading and Language Use in ELA; Operations and Algebraic Thinking in Mathematics) and the relatively low unreliability of these indicators demonstrate that such indicators must be interpreted and used cautiously, and in conjunction with other information about student achievement and learning needs in these areas.

Dimensionality: Dimensionality analysis was conducted on each grade-level test. Chapter 7, section 7.2, provides a detailed description of the dimensionality hypothesis testing and effect-size estimation methods and provides dimensionality results. Minor violations of local independence were noted.

Conditional Standard Errors of Measurement: Chapter 8 provides a detailed description of the psychometric model that was fitted to the data, the test information function (TIF), and the inverse transformation of the TIF into the Conditional Standard Error of Measurement (CSEM). The TIF and CSEM are inverse transformations of each other. Whereas the TIF indicates test score precision, the CSEM indicates the converse, i.e., test score imprecision or measurement error. The TIF and its analogue, the CSEM, are the most pertinent products of the psychometric model in evaluating the adequacy of a test (form). Appendix K shows the CSEMs for each test. By examining the value of CSEM at each of the performance level cut scores, the psychometric appropriateness and accuracy of each test can be evaluated.

Content Coverage: Subclaims 1.1.1, 1.1.2, and 1.2.1 above detail the evidence in support of the content coverage and the alignment of the content to the New Mexico standards.

Scoring: Subclaims 1.4.1 and 1.4.2 detail the evidence in support of accurate item and test scores.

Summary of evidence*:* Complete evidence.

Claim 1.4: Item and test scoring are implemented accurately.

Subclaim 1.4.1. Machine-scored items were scored accurately.

Evidence*:* As described in Section 6.2.1 of Chapter 6 and in Chapter 7, a classical item analysis on the set of machine-scored items is performed prior to scaling and equating. This ensures that for each machine-scored item, the response designated as the correct response was indeed the correct response.

Summary of evidence*:* Complete evidence.

Subclaim 1.4.2. Constructed-response item scoring training and monitoring procedures met industry standards.

Evidence*:* As detailed in Chapter 6, scorer recruitment, training, qualification, and scoring-monitoring procedures follow industry standards. Section 6.2.2, Scoring of Open-Ended Response Items, describes all the procedures that are used to ensure the accuracy of the scoring for the open-ended (constructed) response items, including administrator training and monitoring, benchmarking and identification of scoring materials, scorer recruitment and qualifications, scoring leadership, qualification, specific scoring rules to ensure accuracy, monitoring of quality control, quality reports, and interrater reliability.

Summary of evidence: Complete evidence.

11.2 Primary Intended Score Uses

11.2.1 Intended Score Use for Individual Students

Claim 2.1: Educators and school and district administrators can use results from the NM-MSSA Assessments to describe and monitor student achievement status with respect to mastery of the content standards.

Subclaim 2.1.1. NM-MSSA test scores and performance level categorizations of individual students are adequately reliable and valid measures of student achievement status with respect to mastery of the content standards.

Evidence:

Scale score Standard Errors: Chapter 8 provides a description of calculation and interpretation of the scale scores and Chapter 9 provides a description of the calculation of the standard error for a scale score. The average standard error for reported scale scores is reported in Appendix R. The scale score standard error can be compared to the scale score range and the scale score standard deviation to provide some context for interpretation.

Decision Consistency and Accuracy Estimates: Decision accuracy is an estimate of the probability that the observed classification is the true classification. Decision consistency is an estimate of the probability that students would receive the same classification if they tested twice on parallel forms. Chapter 9 describes the theory and equations underlying the estimation of classification accuracy and consistency. Decision accuracy and consistency results are provided in Appendix O.

Content Coverage: Subclaims 1.1.1, 1.1.2, and 1.2.1 above detail the evidence in support of the content coverage and the alignment of the content to the New Mexico standards.

Scoring: Subclaims 1.4.1 and 1.4.2 detail the evidence in support of accurate item and test scores.

Summary of evidence*:* Complete evidence. Model fit analysis verified that the IRT model fits the assessment data for all grades and content areas.

11.2.2 Intended Score Use for Groups of Students

Claim 3.1: Educators can use results from the NM-MSSA and ASR Assessments to support instructional planning for groups of students.

Subclaim 3.1.1. Teachers find the performance level descriptors and their students' performance levels useful for planning instruction, especially for students whose test scores fall within performance levels 1 and 2.

Evidence: Cognia, in collaboration with PED, has provided multiple professional learning sessions to help New Mexico teachers understand how to use test scores for instructional planning using interims (NM-iMSSA ELA and Math) and formative item sets. While the professional learning session is focused on interim assessments, many parts of the session also covered the balance assessment system, which includes using both the interim and summative results to support New Mexico students.

- Once on-site, Cognia professional learning staff engaged teachers and leaders in Assessment Literacy conversations to understand the importance of the New Mexico Balanced assessment system.
- After New Mexico educators and district leaders have gained an understanding of each type of assessment, the discussion topic transitioned to a deep dive into the interim data sets and the connection between the interim results and summative results from summative assessments.
- The last part of the training is to engage educators and district leaders in how to use the data to drive instructional design/delivery to support students, emphasizing the importance of multiple measures (such as using both interim and summative assessment results). Cognia professional development staff would then spend time looking at the available resources and discussing best practices for using them to support New Mexico students.

As of November 2023, a total of 47 on-site sessions and 12 virtual sessions have been delivered to New Mexico schools, and a total of 822 educators and district/school leaders have participated in the professional learning sessions.

Summary of evidence: Moderate to substantial evidence. Additional evidence may include a teacher survey to understand the degree to which teachers use test scores and other score-based information for instructional planning, especially for low-performing students.

Subclaim 3.1.2. Teachers find their students' scale score information useful for planning instruction, especially for students whose test scores fall within performance levels 1 and 2.

Evidence: Same evidence as subclaim 3.1.2

Summary of evidence: Moderate to substantial evidence. Additional evidence may include a teacher survey to understand the degree to which teachers use test scores and other score-based information for instructional planning, especially for low-performing students.

Claim 3.2: Schools, districts, and state-level stakeholders can use results from the NM-MSSA and ASR Assessments to make comparisons between organizations (e.g., schools, districts).

Subclaim 3.2.1. Test scores and performance levels for groups of students are adequately reliable and valid to enable school, district, and state leaders to monitor changes in means, standard deviations, and performance level percentages for classroom, school, district, and state groups.

Evidence: Evidence for the reliability and validity of the scores and the corresponding scoring processes is presented above under Claim 1.3, which cites Chapter 6 on scoring, Chapter 8 on IRT scaling and equating, and Chapter 9 on classical and IRT reliability and decision accuracy and consistency. The reliability of aggregated scores (e.g., means) is typically as high as or higher than individual score reliabilities (e.g., Brennan, 1995). Appendix N contains the overall and subgroup reliability results. Appendix O contains the decision accuracy and consistency results for the overall test as well as by performance level and by cut score. Subclaims 1.1.1, 1.1.2, and 1.2.1 above detail the evidence in support of the content coverage and the alignment of the content to the New Mexico standards. Subclaims 1.4.1 and 1.4.2 detail the evidence in support of accurate item and test scores. Additionally, model fit analysis verified that the IRT model fits the assessment data for all grades and content areas.

Summary of evidence: Moderate to substantial evidence. Additional evidence may include a district or school leader survey to understand the degree to which teachers use test scores and other score-based information to monitor changes in the aggregated test scores.

Subclaim 3.2.2. Test scores and proficiency level categorizations of groups of students are adequately reliable and valid to enable monitoring of grade-level performance and student-cohort performance.

Evidence: Evidence for the reliability and validity of the scores and the corresponding scoring processes is presented above under Claim 1.3, which cites Chapter 6 on scoring, Chapter 8 on IRT scaling and equating, and Chapter 9 on classical and IRT reliability and decision accuracy and consistency. The reliability of aggregated scores (e.g., means) is typically as high as or higher than individual score reliabilities (e.g., Brennan, 1995). Appendix N contains the overall and subgroup reliability results. Appendix O contains the decision accuracy and consistency results for the overall test as well as by performance level and by cut score. Subclaims 1.1.1, 1.1.2, and 1.2.1 above detail the evidence in support of the content coverage and the alignment of the content to the New Mexico standards. Subclaims 1.4.1 and 1.4.2 detail the evidence in support of accurate item and test scores. Additionally, model fit analysis verified that the IRT model fits the assessment data for all grades and content areas.

Summary of evidence: Moderate to substantial evidence. Additional evidence may include a teacher survey to understand the degree to which teachers use test scores and other score-based information for monitoring grade-level performance and student-cohort performance.

11.3 Conclusions and Next Steps

The majority of the claims and subclaims that support the four claims—that is, the primary intended score interpretations and three intended score uses—are supported by solid evidence. These claims and subclaims and their supporting evidence comprise the validity arguments for NM-MSSA and ASR scores. Table 11-3 summarizes the relevance ratings for each claim and subclaim. Table 11-3 indicates the following:

Primary Score Intended Score Interpretation

Of the four claims and nine subclaims that support the intended score interpretation, all 9 sets of evidence are complete.

Intended Score Use for Individual Students

The one claim that with one supporting subclaim that supports the first intended score use, the evidence for this claim and subclaim is complete.

Intended Score Use for Groups of Students

Of the two claims and four supporting subclaim sets of evidence, all four sets of evidence are moderate to substantial.

Table 11-3. Status of Evidence for All SIUs, Claims, and Subclaims

SIUs, Claims, and Subclaims	Relevance of the Evidence to the Validity Argument			
	No Evidence Exists Currently	Limited	Moderate to Substantial	Complete
SIU 1: Primary Intended Score Interpretation				
The NM-MSSA Assessments provide reliable and valid information about important knowledge and skills in grade-level reading, writing & language usage, and Mathematics attained by general education students.				
1.1.1. NM-MSSA content is aligned to the New Mexico Common Core State Standards.				X
1.1.2. NM-MSSA items are aligned to the New Mexico Common Core State Standards.				X
1.2.1. Items require application of the KSAs of the targeted construct.				X
1.2.2. Items are free of bias and sensitivity issues.				X
1.3.1. NM-MSSA scores and performance level categorizations are adequately reliable for their intended purpose.				X
1.3.2. Item characteristics support intended interpretations about all students who take the NM-MSSA.				X
1.3.3. Test characteristics support intended interpretations about all students who take the NM-MSSA.				X
1.4.1. Machine-scored items were scored accurately.				X
1.4.2. Constructed-response item scoring training and monitoring procedures met industry standards.				X
SIU 2: Intended Score Use for Individual Students				
2.1.1. NM-MSSA test scores and performance level categorizations of individual students are adequately reliable and valid measures of student achievement status with respect to mastery of the content standards.				X
SIU 3: Intended Score Use for Groups of Students				
3.1.1. Teachers find the performance level descriptors and their students' performance levels useful for planning instruction, especially for students whose test scores fall within performance levels 1 and 2.			X	
3.1.2. Teachers find their students' scale score information useful for planning instruction, especially for students whose test scores fall within performance levels 1 and 2.			X	
3.2.1. NM-MSSA scores and performance levels for groups of students are adequately reliable and valid to enable school, district, and state leaders to monitor changes in means, standard deviations, and performance level percentages for classroom, school, district, and state groups.			X	
3.2.2. NM-MSSA scores and proficiency level categorizations of groups of students are adequately reliable and valid to enable monitoring of grade-level performance and student-cohort performance.			X	

11.3.1 Research Agenda

The Score Card ratings provide a road map for a research agenda for the NM-MSSA and NM-ASR programs. Specifically, PED and Cognia can work together to identify the highest priority claims and subclaims for which *No Evidence Exists Currently* and where the evidence is *Limited* and plan studies to gather relevant evidence and strengthen validity arguments. This will be a topic of discussion and planning for more immediate and longer-term efforts during the 2022–2023 school year.

References

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

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

Brennan, R. L. (1995). The conventional wisdom about group mean scores. *Journal of Educational Measurement*, *32*(4), 385 –396.

Chapelle, C. A. (2021). *Argument-based validation in testing and assessment*. Sage.

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., & 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.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer Academic Publishers.

Hambleton, R. K., & van der Linden, W. J. (1997). *Handbook of modern item response theory*. New York, NY: Springer-Verlag.

Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, D.C.: National Council on Measurement in Education.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1–73.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, *32*, 179–197.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of meta test scores*. Reading, MA: Addison-Wesley.

Persuasiveness. (2019). In *Vocabulary.com* Retrieved from https://www.lexico.com/en/definition/persuasive

Plausibility. (2019). In *Lexico.com* Retrieved from https://www.lexico.com/en/definition/plausibility

Relevance. (2019). In *Lexico.com* Retrieved from https://www.lexico.com/en/definition/relevance

Roussos, L. A., & Ozbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement*, *43*, 215–243.

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

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

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

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Zhang, J., & Stout, W. F. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika*, *64*, 213–249.

Appendices

APPENDIX A
LIST OF ACRONYMS

Common Terms and Acronyms Used in Assessment Reports	
3PL	Three-parameter logistic
AERA	American Educational Research Association
APA	American Psychological Association
CBT	Computer-based test
CCSSO	Council of Chief State School Officers
CFT	Census field test
CP	Concepts and procedures standard
CR	constructed response items
CRESST	National Center for Research on Evaluation, Standards, and Student Testing
CSEM	conditional standard error of measurement
CTT	Classical Test Theory
DETECT	Dimensionality Evaluation to Enumerate Contributing Traits
DIF	differential Item functioning
DIMTEST	computer program used by Cognia
DOK	depth of knowledge
DTA	Directions for Test Administration
DTC	District Test Coordinator
EA	Educational assistant
EBSR	evidence-based selected-response item
EL	English learner
ELA	English language arts
ESSA	Every Student Succeeds Act
ETS	Engineering, technology, and applications of science
GRM	graded-response model
HMS	Hobbs Municipal Schools
ICC	item characteristic curve
ICCC	item category characteristic curve
ICTC	item category threshold curve
IIF	Item information function
IRT	Item Response Theory
KSA	knowledge, skills, and abilities
LCPS	Las Cruces Public Schools
LEA	local educational agency
LEP	limited English proficiency
MC	multiple choice items
MP	Mathematical practice standard
MS	machine scored items
NAEP	National Assessment of Educational Progress
NCEO	National Center on Educational Outcomes
NCME	National Council on Measurement in Education

continued

Common Terms and Acronyms Used in Assessment Reports	
NM-ASR	New Mexico Assessment of Science Readiness
NM-MSSA	New Mexico Measures of Student Success and Achievement
NM PED	New Mexico Public Education Department
OE	Open-ended questions
PADDI	Principled Assessment Design, Development, and Implementation
PBT	Paper-based test
PLD	performance level descriptor
REA	Research, Evaluation, and Accountability Bureau
SAFT	Stand-alone field test
SEM	standard error of measurement
SIU	score interpretations and uses
SLA	Spanish language arts
SR	Selected response items
SS	Scaled score
STC	School test coordinator
STL	scoring team leader
TA	test administrators
TAC	Technical Advisory Committee
TAM	test Administration Manual
TAMELA	Transitional assessment for mathematics and ELA
TCC	test characteristic curve
TIF	test information function
WP	writing prompt

APPENDIX B
PLDS, SIUS, AND TEST SPECIFICATIONS

Policy Proficiency Level Descriptors for NM-MSSA and NM-ASR

English and Spanish Versions

Level 4. Advanced

Students demonstrate evidence of **thorough** understanding and use of college and career readiness knowledge, skills, and abilities.

Level 3. Proficient

Students demonstrate evidence of **satisfactory** understanding and use of college and career readiness knowledge, skills, and abilities.

Level 2. Nearing Proficiency

Students demonstrate evidence of **partial** understanding and use of college and career readiness knowledge, skills, and abilities.

Level 1. Novice

Students demonstrate evidence of **emerging** understanding and use of college and career readiness knowledge, skills, and abilities.

STATEMENTS OF SCORE INTERPRETATIONS AND USES (SIUs) FOR THE NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT (NM-MSSA) ASSESSMENT

English Language Arts

Score Interpretation and Use (SIU) Statements for the NM-MSSA and NM-ASR Assessment Programs

The phrase “intended score interpretations for uses” appears several times in the *Standards for Educational and Psychological Testing* and is at the core of the field’s views on validity and validation. It also is central to responding successfully to USDE peer review requirements. For the NM-MSSA English language arts assessments, the phrase refers broadly to **test scores** (i.e., total test scale scores, subdomain indicators), **aggregations of test scores** (e.g., the percentage of students at and above Level 3. Proficient), and **other test performance informational elements** (e.g., the definition of Proficient in the Proficiency Level Descriptors).

SIU Statements for the NM-MSSA: English Language Arts

Using this broad interpretation of the phrase, the intended score interpretations and uses for NM-MSSA are stated below. These statements reflect input from PED’s Technical Advisory Committee and multiple statewide webinars with educator and parent stakeholders on the similar NM-ASR SIU statements. PED and Cognia will use the final, approved SIU statements to guide decisions about test design and score reporting.

NM-MSSA score reports include scale scores for ELA, Reading, and Writing & Language. The ELA scale score includes performance on the Reading, Writing & Language, and Writing sections of the test. The Reading scale score is based only on reading items. The Writing & Language scale score includes both Writing & Language selected response items and the Writing task. Score reports for NM-MSSA Writing include only rubric scores (i.e., no scale scores).

Intended Interpretations and Uses for Individual Students and Groups of Students

Score Interpretation/Use Statement	Explanation/Annotation
NM-MSSA Program Purpose Statements	
Program Purpose Statement, Grades 3–8 NM-MSSA The NM-MSSA grades 3–8 assessments are designed to measure whether students are on track to be ready for college or career, as defined by the state, by showing they have mastered *the New Mexico Common Core State Standards* for English language arts. Results are presented using scale scores and proficiency levels. Proficient performance in each grade indicates mastery of both currently assessed grade level and preceding grades' expectations and progress toward college and career readiness. The English language arts standards require a solid understanding of concepts, a high degree of procedural skill and fluency, and the application of English language arts to solve problems. See details at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/.	NM-MSSA scores should be interpreted in relation to the *New Mexico Common Core State Standards* that are targeted by the assessment. ***College readiness*** indicates that a student is prepared to enter directly into and succeed (i.e., earn a C or better) in entry-level, credit-bearing college and relevant technical courses at two- and four-year public institutions of higher education, without the need for remediation. ***Career readiness*** indicates that students have developed the academic and technical skills (i.e., workplace competencies in one or more of 16 career clusters) necessary to succeed in future careers and to become lifelong learners. ***College and Career Readiness*** is defined by the State and can be found in the following College and Career Readiness Bureau's web page: https://webnew.ped.state.nm.us/bureaus/college-career-readiness/
Program Purpose Statement, Grade 8 NM-MSSA Performance on the grade 8 NM-MSSA indicates mastery of (a) grade 8 content standards, and (b) solid understanding of concepts, a high degree of procedural skill and fluency, and the application of English language arts to solve problems. See details at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/. It also is (c) a predictor of being on track for college and career readiness as defined by the College Board's *College and Career Readiness Benchmarks*: see https://collegereadiness.collegeboard.org/about/scores/benchmarks. Cognia will establish a psychometric link from grade 8 NM-MSSA Reading, Writing & Language, and Mathematics scores to College Board PSAT scores that will enable monitoring of student progress toward the College Board CCR Benchmarks.	The program purpose statements apply in grades 3–8. In addition, performance on the grade 8 NM-MSSA can be interpreted as a predictor of performance on the PSAT 8/9 and PSAT 10, specifically prediction of status in relation to the College Board CCR Benchmarks. The current links are based on a small empirical validation study conducted outside of NM, which is one indicator of college and career readiness. Linking studies conducted specifically for NM students will provide the links necessary for monitoring student progress toward college and career readiness.
Individual Students	
Master Claim Performance on the NM-MSSA indicates a student's progress toward college and career readiness.	College and career readiness requires that students can, without significant scaffolding, comprehend and evaluate grade level problems in English language arts as set forth in the *New Mexico Common Core State Standards* for English language arts.

Score Interpretation/Use Statement	Explanation/Annotation
Interpretations Using Proficiency Level Labels and Proficiency Level Descriptors (PLDs) Student scale scores coincide with one of four levels: Advanced, Proficient, Nearing Proficiency, and Novice.[1] The PLD for each proficiency level describes what students are expected to know and be able to do in relation to the *New Mexico Common Core State Standards* for English language arts in grades 3–8. New Mexico students are expected to perform at the Proficient level to demonstrate mastery of the knowledge and skills needed to indicate college and career readiness.	A student's proficiency level indicates how the student performed in relation to the knowledge and skills assessed in English language arts at that grade level. Proficiency level descriptors indicate the knowledge and skills that students are expected to be able to demonstrate at a level.
Interpretations Using Proficiency Level Descriptors A student's proficiency level indicates that the student can be expected to demonstrate the knowledge and skills described at that level and in the levels below.	The student's proficiency level also indicates that the student has mastered the knowledge and skills of the preceding proficiency levels.
Interpretations Using Scale Scores Scale scores provide a measure of student performance regardless of which form of the NM-MSSA is administered.	Scale scores indicate the student's performance, regardless of which form of the NM-MSSA is taken. The **proposed** scale score reporting scale is under discussion.
Uses of Scale Scores Scale scores can be used to compare an individual student's performance to the performance of other students in the school, district, and state.	Scale scores also indicate a student's performance in relation to the performance of other students. A student's scale score should be interpreted as the range of possible scores within the error band around that score, not only as a single number. (Other terms for "error band" include "margin of error" and "confidence interval.") Differences between scale scores (e.g., for two students or a student's score and a proficiency level cut score) that are within the margin of error should be interpreted as "statistical ties" (i.e., not reliably different).
Interpretation of Reading Scale Scores NM-MSSA Reading scale scores are based on student responses to the Reading items in the larger ELA assessment. These items assess mastery of skills that require students to understand key ideas and details, analyze elements of craft and structure, and integrate knowledge and ideas in informational and literary texts, which is the progression for the next level of the reading curriculum. (See http://www.corestandards.org/ELA-Literacy/CCRA/R/ for details.)	Student performance on Reading and Writing & Language claims indicators are based on the items that target each of those areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.

[1] **NM-MSSA Policy Proficiency Level Descriptors**

Advanced. Students demonstrate evidence of **thorough** understanding and use of college and career readiness knowledge, skills, and abilities.

Proficient. Students demonstrate evidence of **satisfactory** understanding and use of college and career readiness knowledge, skills, and abilities.

Nearing Proficiency. Students demonstrate evidence of **partial** understanding and use of college and career readiness knowledge, skills, and abilities.

Novice. Students demonstrate evidence of **emerging** understanding and use of college and career readiness knowledge, skills, and abilities.

Score Interpretation/Use Statement	Explanation/Annotation
Interpretation of Writing & Language Scale Scores NM-MSSA Writing & Language scale scores are based on student responses to the Writing & Language items and Writing task in the larger ELA assessment. These items assess mastery of skills that require students to communicate clearly to an external, sometimes unfamiliar audience and adapt the form and content of writing to accomplish a particular task and purpose; to determine or clarify the meaning of grade-appropriate words; and to gain control over many conventions of standard English grammar, usage, and mechanics as well as learn other ways to use language to convey meaning effectively, which is the progression for the next level of the writing and language curriculum. (See http://www.corestandards.org/ELA-Literacy/CCRA/R/ for details.)	Student performance on Reading and & Language is based on the items that target each of those areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Item Level Reporting for Individual Students Individual student performance on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of English language arts standards.
Groups of Students	
SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.	
Group Mean Scale Scores Group mean scale scores can be compared to other schools, districts, and the state, and for all students and student subgroups (e.g., gender, English learners, students with disabilities).	Mean (i.e., average) scale scores enable comparison of performance among schools, districts, and other groupings of students. Mean scale scores and percentages of students in a proficiency level for small groups (e.g., fewer than 25 students) are unstable and should be interpreted with caution because of concerns about reliability and stability.
Percentages of Students in Proficiency Levels Percentages of students in the four proficiency levels can be compared to other schools, districts, and the state, and for all students and student subgroups.	These are the percentages of students in each English language arts proficiency level. The PLD for each English language arts proficiency level indicates the degree of mastery of the knowledge and skills needed to indicate college and career readiness in relation to the *New Mexico Common Core State Standards*. The percentages of students in each level indicate the percentage of students who need to reach the next proficiency level. Means and percentages of students in a proficiency level for small groups should be interpreted with caution because of concerns about reliability and stability.
Item Level Reporting for Student Groups Student group performance (e.g., boys, girls, English learners) on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of English language arts standards.

Unintended Score Interpretations and Uses

Until the NM-MSSA is in operational use, we only can speculate on what unintended interpretations and uses of NM-MSSA scores, and other information may arise. Where unintended interpretations and uses may occur, it is the responsibility of that user to provide supporting evidence, and not the responsibility of PED (as specified in the *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-MSSA scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and

uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations, misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following.

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and proficiency level classifications—are estimates accompanied by a standard error. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Interpreting and using NM-MSSA scores correctly requires considering the width of the margin of error around a score. For example, students with a scale score 2 points below the cut score for the Proficient level could, hypothetically, have scored above the Proficient cut score on a different day because the NM-MSSA scale score standard errors are expected to be 2–3 points. Interpretations of NM-MSSA scores should account for the margin of error around each score estimate.

Drawing Conclusions and Making Decisions Based Solely on NM-MSSA Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less than optimal decisions, such as students failing to receive additional instruction based solely on their NM-MSSA score or teacher teams not being eligible for additional professional learning based solely on their students' NM-MSSA scores. Interpretations and uses of NM-MSSA scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators and Item Level Performance Information

Subdomain indicators for English Language Arts are the scale scores for the Reading and Writing & Language assessment sections and the two rubric scores for the Writing section.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. It would be unwise to assume that students who currently are performing at the Proficient and Advanced levels will continue at these levels without sustained effort. Similarly, it would be unwise—and unfair—to assume that students who currently are performing at the Novice and Nearing Proficiency levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.

Misinterpretations about students' current proficiency levels and future performance is not really a misinterpretation of NM-MSSA scores. It is a logical error in concluding that current performance determines future performance.

Overinterpreting NM-MSSA Scores as Indicators of College and Career Readiness

The New Mexico *Common Core State Standards* are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM-MSSA indicates readiness for college and career is supported only by the evidence contained in the English language arts content standards. NM-MSSA scores also can be interpreted as predictors of future performance, in college and career training. However, until empirical prediction studies are completed, this interpretation of NM-MSSA performance should be made with caution and with attention to the strong, but limited, evidence in the content standards.

STATEMENTS OF SCORE INTERPRETATIONS AND USES (SIUs) FOR THE NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT (NM-MSSA) ASSESSMENT

Reading in English and Spanish

Score Interpretation and Use (SIU) Statements for the NM-MSSA and NM-ASR Assessment Programs

The phrase "intended score interpretations for uses" appears several times in the *Standards for Educational and Psychological Testing* and is at the core of the field's views on validity and validation. It also is central to responding successfully to USDE peer review requirements. For the NM-MSSA ELA and mathematics assessments, the phrase refers broadly to **test scores** (i.e., total test scale scores, subdomain indicators), **aggregations of test scores** (e.g., the percentage of students at and above Level 3. Proficient), and **other test performance informational elements** (e.g., the definition of Proficient in the Proficiency Level Descriptors).

SIU Statements for the NM-MSSA: Reading in English and Spanish

Using this broad interpretation of the phrase, the intended score interpretations and uses for NM-MSSA Reading are stated below. These statements reflect input from and PED's Technical Advisory Committee and multiple statewide webinars with educator and parent stakeholders on the similar NM-ASR SIU statements.

PED and Cognia will use the final, approved SIU statements to guide decisions about test design and score reporting.

NM-MSSA score reports include scale scores for ELA, Reading, and Writing & Language. The ELA scale score includes performance on the Reading, Writing & Language, and Writing sections of the test. The Reading scale score is based only on reading items. The Writing & Language scale score includes both Writing & Language selected response items and the Writing task. Score reports for NM-MSSA Writing include only rubric scores (i.e., no scale scores).

Intended Interpretations and Uses for Individual Students and Groups of Students: NM-MSSA Reading in English and Spanish

Score Interpretation/Use Statement	Explanation/Annotation
NM-MSSA Program Purpose Statements	
Program Purpose Statement, Grades 3–8 NM-MSSA The NM-MSSA grades 3–8 assessments are designed to measure whether students are on track to be ready for college or career, as defined by the state, by showing they have mastered *the New Mexico Common Core State Standards* for English language arts. Results are presented using scale scores and proficiency levels. Proficient performance in each grade indicates both mastery of currently assessed grade level and preceding grades' expectations and progress toward college and career readiness. The reading standards require understanding key ideas and details, analyzing elements of craft and structure, and integrating knowledge and ideas in informational and literary texts. See details	NM-MSSA scores should be interpreted in relation to the *New Mexico Common Core State Standards* that are targeted by the assessment. ***College readiness*** indicates that a student is prepared to enter directly into and succeed (i.e., earn a C or better) in entry-level, credit-bearing college and relevant technical courses at two- and four-year public institutions of higher education, without the need for remediation. ***Career readiness*** indicates that students have developed the academic and technical skills (i.e., workplace competencies in one or more of 16 career clusters) necessary to succeed in future careers and to become lifelong learners.

Score Interpretation/Use Statement	Explanation/Annotation
at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/.	***College and Career Readiness*** is defined by the State and can be found in the following College and Career Readiness Bureau's web page: https://webnew.ped.state.nm.us/bureaus/college-career-readiness/
Program Purpose Statement, Grade 8 NM-MSSA Performance on the grade 8 NM-MSSA indicates mastery of (a) grade 8 content standards, and (b) skills that require students to understand key ideas and details, analyze elements of craft and structure, and integrate knowledge and ideas in informational and literary texts, which is the progression for the next level of the reading curriculum. See details at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/. It also is (c) a predictor of being on track for college and career readiness as defined by the College Board's *College and Career Readiness Benchmarks:* see https://collegereadiness.collegeboard.org/about/scores/benchmarks . Cognia will establish a psychometric link from grade 8 NM-MSSA scores to College Board PSAT scores that will enable monitoring of student progress toward the College Board CCR Benchmarks.	The program purpose statements apply in grades 3–8. In addition, performance on the grade 8 NM-MSSA can be interpreted as a predictor of performance on the PSAT 8/9 and PSAT 10, specifically prediction of status in relation to the College Board CCR Benchmarks. The current links are based on a small empirical validation study conducted outside of NM, which is one indicator of college and career readiness. Linking studies conducted specifically for NM students will provide the links necessary for monitoring student progress toward college and career readiness.
Individual Students	
Master Claim Performance on the NM-MSSA indicates a student's progress toward college and career readiness.	College and career readiness requires that students can, without significant scaffolding, comprehend and evaluate complex texts across a range of types and disciplines as set forth in the *New Mexico Common Core State Standards* for English language arts.
Interpretations Using Proficiency Level Labels and Proficiency Level Descriptors Student scores coincide with one of four levels: Advanced, Proficient, Nearing Proficiency, and Novice.[2] The PLD for each proficiency level describes what students are expected to know and be able to do in relation to the *New Mexico Common Core State Standards* for reading in grades 3–8. New Mexico students are expected to perform at the Proficient level to demonstrate mastery of the knowledge and skills needed to indicate college and career readiness.	A student's proficiency level indicates how the student performed in relation to the knowledge and skills assessed in reading at that grade level. Proficiency level descriptors indicate the knowledge and skills that students are expected to be able to demonstrate at a level.
Interpretations Using Proficiency Level Descriptors A student's proficiency level indicates that the student can be expected to demonstrate the knowledge and skills described at that level and in the levels below.	The student's proficiency level also indicates that the student has mastered the knowledge and skills of the preceding proficiency levels.

[2] **NM-ASR Policy Proficiency Level Descriptors**

Advanced. Students demonstrate evidence of **thorough** understanding and use of college and career readiness knowledge, skills, and abilities.

Proficient. Students demonstrate evidence of **satisfactory** understanding and use of college and career readiness knowledge, skills, and abilities.

Nearing Proficiency. Students demonstrate evidence of **partial** understanding and use of college and career readiness knowledge, skills, and abilities.

Novice. Students demonstrate evidence of **emerging** understanding and use of college and career readiness knowledge, skills, and abilities.

Score Interpretation/Use Statement	Explanation/Annotation
Interpretations Using Scale Scores Scale scores provide a measure of student performance regardless of which form of the NM-MSSA is administered.	Scale scores indicate the student's performance, regardless of which form of the NM-MSSA is taken. The **proposed** scale score reporting scale is under discussion.
Uses of Scale Scores Scale scores can be used to compare an individual student's performance to the proficiency levels and to the performance of other students in the school, district, and state.	Scale scores also indicate a student's performance in relation to the performance of other students. A student's scale score should be interpreted as the range of possible scores within the error band around that score, not only as a single number. (Other terms for "error band" include "margin of error" and "confidence interval.") Differences between scale scores (e.g., for two students or a student's score and a proficiency level cut score) that are within the margin of error should be interpreted as "statistical ties" (i.e., not reliably different).
Interpretation of Literary Text Indicators Student performance on Literary Text indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient.[2] Performance on Literary Text items involves reading, comprehending, analyzing, and interpreting grade-level appropriate literary text; themes, text structures, points of view, key ideas and details; new and unfamiliar words; and comparing and contrasting two literary texts.	Student performance on each of the four Reading claims indicators (i.e., Literary Text, Informational Text, Comprehension, and Analysis and Interpretation) is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Informational Text Indicators Student performance on Informational Text indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient.[3] Performance on Informational Text items requires reading, comprehending, analyzing, and interpreting main ideas and details and new and unfamiliar words; and comparing and contrasting information, arguments, and evidence in two texts; all in grade-level appropriate informational text.	Student performance on each of the four Reading claims indicators (i.e., Literary Text, Informational Text, Comprehension, and Analysis and Interpretation) is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Comprehension Indicators Student performance on Comprehension indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [2] Performance on Comprehension items requires applying reading skills and strategies to comprehend central ideas and themes, identify supporting details, and determine meaning of words and phrases in grade-level appropriate literary and informational text.	Student performance on each of the four Reading claims indicators (i.e., Literary Text, Informational Text, Comprehension, and Analysis and Interpretation) is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Analysis and Interpretation Indicators Student performance on Analysis and Interpretation indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [2]	Student performance on each of the four Reading claims indicators (i.e., Literary Text, Informational Text, Comprehension, and Analysis and Interpretation) is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge

[3] NM-MSSA subdomain indicators are reported as Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. These subdomain indicators are calculated by comparing a student's subdomain performance to the subdomain performance distribution of students who are just barely Nearing Proficient on the total test and using the standard deviation of that distribution to determine the Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient indicators.

Score Interpretation/Use Statement	Explanation/Annotation
Performance on Analysis and Interpretation items requires applying reading skills and strategies to grade-level appropriate literary and informational text to analyze ideas, events, and characters and examine relationships among text elements; interpret themes, purposes, claims, and evidence; determine and evaluate points of view; analyze the impact of literal and figurative language; analyze word choice; compare and contrast information and author methods in two texts; and make inferences and draw conclusions using evidence from the texts to support analyses and interpretations.	(DOK) level that categorizes the cognitive demand of the item content.
Item Level Reporting for Individual Students Individual student performance on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of Reading standards.
Groups of Students	
SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., gender, English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.	
Group Mean Scale Scores Group mean scale scores can be compared to other schools, districts, and the state, and for all students and student subgroups (e.g., gender, English learners, students with disabilities, racial/ethnic subgroups).	Mean (i.e., average) scale scores enable comparison of performance among schools, districts, and other groupings of students. Mean scale scores and percentages of students in a proficiency level for small groups (e.g., fewer than 25 students) are unstable and should be interpreted with caution.
Percentages of Students in Proficiency Levels Percentages of students in proficiency levels 1–4 can be compared to other schools, districts, and the state, and for all students and student subgroups.	These are the percentages of students in each reading proficiency level. The PLD for each reading proficiency level indicates the degree of mastery of the knowledge and skills needed to indicate college and career readiness in relation to the *New Mexico Common Core State Standards*. The percentages of each students in each level indicate the percentage of students who need to reach the next proficiency level. Means and percentages of students in a proficiency level for small groups should be interpreted with caution because of concerns about reliability and stability.
Item Level Reporting for Student Groups Student group performance (e.g., boys, girls, English learners) on individual test items or groups of items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of Reading standards.

Unintended Score Interpretations and Uses

Until the NM-MSSA is in operational use, we only can speculate on what unintended interpretations and uses of NM-MSSA scores and other information may arise. Where unintended interpretations and uses may occur, it is the responsibility of that user to provide supporting evidence, and not the responsibility of PED (as specified in the *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-MSSA scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations, misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following.

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and proficiency level classifications—are estimates accompanied by a standard error. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Interpreting and using NM-MSSA scores correctly requires considering the width of the margin of error around a score. For example, students with a scale score 2 points below the cut score for the Proficient level could, hypothetically, have scored above the Proficient cut score on a different day because the NM-MSSA scale score standard errors are expected to be 2–3 points. Interpretations of NM-MSSA scores should account for the margin of error around each score estimate.

Drawing Conclusions and Making Decisions Based Solely on NM-MSSA Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less than optimal decisions, such as students failing to receive additional instruction based solely on their NM-MSSA score or teacher teams not being eligible for additional professional learning based solely on their students' NM-MSSA scores. Interpretations and uses of NM-MSSA scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators and Item Level Performance Information

Subdomain indicators (e.g., Literary Text, Comprehension) are based on fewer items than are NM-MSSA total test scores. As a result, they are less stable estimates of student achievement and learning needs in that subdomain. In addition, because the performance indicators for the reading subdomains are highly correlated, differences in those performance indicators may be smaller than the proficiency level labels may suggest. Interpretations and uses of indicator scores should be supplemented with additional information.

Overinterpreting NM-MSSA Scores as Indicators of College and Career Readiness

The New Mexico *Common Core State Standards* are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM-MSSA indicates readiness for college and career is supported only by the evidence contained in the English language arts content standards. NM-MSSA scores also can be interpreted as predictors of future performance, in college and career training. However, until empirical prediction studies are completed, this interpretation of NM-MSSA performance should be made with caution and with attention to the strong, but limited, evidence in the content standards.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. It would be unwise to assume that students who currently are performing at the Proficient and Advanced levels will continue at these levels without sustained effort. Similarly, it would be unwise—and unfair—to assume that students who currently are performing at the Novice and Nearing Proficiency levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.

This misinterpretation about students and the future is not really a misinterpretation of NM-MSSA scores. It is a logical error in concluding that current performance determines future performance.

STATEMENTS OF SCORE INTERPRETATIONS AND USES (SIUs) FOR THE NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT (NM-MSSA) ASSESSMENT

Writing & Language in English

Score Interpretation and Use (SIU) Statements for the NM-MSSA and NM-ASR Assessment Programs

The phrase "intended score interpretations for uses" appears several times in the *Standards for Educational and Psychological Testing* and is at the core of the field's views on validity and validation. It also is central to responding successfully to USDE peer review requirements. For the NM-MSSA ELA and mathematics assessments, the phrase refers broadly to **test scores** (i.e., total test scale scores, subdomain indicators), **aggregations of test scores** (e.g., the percentage of students at and above Level 3. Proficient), and **other test performance informational elements** (e.g., the definition of Proficient in the Proficiency Level Descriptors).

SIU Statements for the NM-MSSA: Writing & Language in English

Using this broad interpretation of the phrase, the intended score interpretations and uses for NM-MSSA are stated below. These statements reflect input from PED's Technical Advisory Committee and multiple statewide webinars with educator and parent stakeholders on the similar NM-ASR SIU statements. PED and Cognia will use the final, approved SIU statements to guide decisions about test design and score reporting.

NM-MSSA score reports include scale scores for ELA, Reading, and Writing & Language. The ELA scale score includes performance on the Reading, Writing & Language, and Writing sections of the test. The Reading scale score is based only on reading items. The Writing & Language scale score includes both Writing & Language selected response items and the Writing task. Score reports for NM-MSSA Writing include only rubric scores (i.e., no scale scores).

Intended Interpretations and Uses for Individual Students and Groups of Students: NM-MSSA Writing & Language in English

Score Interpretation/Use Statement	Explanation/Annotation
NM-MSSA Program Purpose Statements	
Program Purpose Statement, Grades 3–8 NM-MSSA The NM-MSSA grades 3–8 assessments are designed to measure whether students are on track to be ready for college or career, as defined by the state, by showing they have mastered *the New Mexico Common Core State Standards* for Writing & Language. Results are presented using scale scores and proficiency levels. Proficient performance in each grade indicates both mastery of currently assessed grade level and preceding grades' expectations and progress toward college and career readiness. The Writing and Language standards require students to analyze the use of introductions and conclusions; develop writing topics;	NM-MSSA scores should be interpreted in relation to the *New Mexico Common Core State Standards* that are targeted by the assessment. ***College readiness*** indicates that a student is prepared to enter directly into and succeed (i.e., earn a C or better) in entry-level, credit-bearing college and relevant technical courses at two- and four-year public institutions of higher education, without the need for remediation. ***Career readiness*** indicates that students have developed the academic and technical skills (i.e., workplace competencies in one

Score Interpretation/Use Statement	Explanation/Annotation
use precise and effective language (grades 4–8); use transition words within pieces of writing; use correct conventions of standard English; use knowledge of language; and acquire and use vocabulary. See details at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/	or more of 16 career clusters) necessary to succeed in future careers and to become lifelong learners. ***College and Career Readiness*** is defined by the State and can be found in the following College and Career Readiness Bureau's web page: https://webnew.ped.state.nm.us/bureaus/college-career-readiness/
Program Purpose Statement, Grade 8 NM-MSSA Performance on the grade 8 NM-MSSA indicates mastery of (a) grade 8 content standards, and (b) skills that require analysis and production of different modes of writing, as well as analysis and use of English conventions and vocabulary. It also is (c) a predictor of being on track for college and career readiness as defined by the College Board's *College and Career Readiness Benchmarks*: see https://collegereadiness.collegeboard.org/about/scores/benchmarks . Cognia will establish a psychometric link from grade 8 NM-MSSA scores to College Board PSAT scores that will enable monitoring of student progress toward the College Board CCR Benchmarks.	The program purpose statements apply in grades 3–8. In addition, performance on the grade 8 NM-MSSA can be interpreted as a predictor of performance on the PSAT 8/9 and PSAT 10, specifically prediction of status in relation to the College Board CCR Benchmarks. The current links are based on a small empirical validation study conducted outside of NM, which is one indicator of college and career readiness. Linking studies conducted specifically for NM students will provide the links necessary for monitoring student progress toward college and career readiness.
Individual Students	
Master Claim Performance on the NM-MSSA indicates a student's progress toward college and career readiness.	College and career readiness requires students to use correct conventions of standard English, use knowledge of language, and acquire and use vocabulary, as set forth in the *New Mexico Common Core State Standards* or English Language Arts.
Interpretations Using Proficiency Level Labels and Proficiency Level Descriptors (PLDs)[4] Student scores coincide with one of four levels: Advanced, Proficient, Nearing Proficiency, and Novice. The PLD for each proficiency level describes what students are expected to know and be able to do in relation to the *New Mexico Common Core State Standards* for writing and language in grades 3–8. New Mexico students are expected to perform at the Proficient level to demonstrate mastery of the knowledge and skills needed to indicate college and career readiness.	A student's proficiency level indicates how the student performed in relation to the knowledge and skills assessed in writing and language at that grade level. Proficiency level descriptors indicate the knowledge and skills that students are expected to be able to demonstrate at a level.
Interpretations Using Proficiency Level Descriptors (PLDs) A student's proficiency level indicates that the student can be expected to demonstrate the knowledge and skills described at that level and in the levels below.	The student's proficiency level also indicates that the student has mastered the knowledge and skills of the preceding proficiency levels.

[4] **Policy Proficiency Level Descriptors for NM-MSSA and NM-ASR**

Advanced. Students demonstrate evidence of thorough understanding and use of college and career readiness knowledge, skills, and abilities.

Proficient. Students demonstrate evidence of satisfactory understanding and use of college and career readiness knowledge, skills, and abilities.

Nearing Proficiency. Students demonstrate evidence of partial understanding and use of college and career readiness knowledge, skills, and abilities.

Novice. Students demonstrate evidence of emerging understanding and use of college and career readiness knowledge, skills, and abilities.

Score Interpretation/Use Statement	Explanation/Annotation
Interpretations Using Scale Scores Scale scores can be used to compare an individual student's performance to the proficiency levels and to the performance of other students in the school, district, and state.	Scale scores indicate the student's performance, regardless of which form of the NM-MSSA is taken. The **proposed** scale score reporting scale is under discussion.
Uses of Scale Scores Scale scores can be used to compare an individual student's performance to the proficiency levels and to the performance of other students in the school, district, and state.	Scale scores also indicate a student's performance in relation to the performance of other students. A student's scale score should be interpreted as the range of possible scores within the error band around that score, not only as a single number. (Other terms for "error band" include "margin of error" and "confidence interval.") Differences between scale scores (e.g., for two students or a student's score and a proficiency level cut score) that are within the margin of error should be interpreted as "statistical ties" (i.e., not reliably different).
Interpretation of the Writing Analysis and English Language Conventions Subdomain Indicator[5] Student performance on Language & Writing and Writing Analysis indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. Performance on Writing Analysis items involves analyzing the use of introductions and conclusions, developing writing topics, using precise and effective language (grades 4–8), and using transition words within pieces of writing. Performance on English Language Conventions items involves using correct conventions of standard English, using knowledge of language, and acquiring and using vocabulary.	Student performance on the Language & Writing and Writing Analysis claims indicator is based on the items that target these claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Item Level Reporting for Individual Students Individual student performance on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of Writing & Language standards.
Groups of Students	
SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., gender, English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.	
Mean Scale Scores Group mean scale scores can be compared to other classrooms, schools, districts, and the state, and for all students and student subgroups (e.g., gender, English learners, students with disabilities, racial/ethnic subgroups).	Mean (i.e., average) scale scores enable comparison of performance among schools, districts, and other groupings of students. Mean scale scores and percentages of students in a proficiency level for small groups (e.g., fewer than 25 students) are unstable and should be interpreted with caution.
Percentages of Students in Proficiency Levels Percentages of students in proficiency levels 1–4 can be compared to other classrooms, schools, districts, and the state, and for all students and student subgroups.	These are the percentages of students in each writing and language proficiency level. The PLD for each writing and language proficiency level indicates the degree of mastery of the knowledge and skills needed to indicate college and career readiness in relation to the *New*

[5] NM-MSSA subdomain indicators are reported as Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. These subdomain indicators are calculated by comparing a student's subdomain performance to the subdomain performance distribution of students who are just barely Nearing Proficient on the total test and using the standard deviation of that distribution to determine the Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient indicators.

Score Interpretation/Use Statement	Explanation/Annotation
	Mexico Common Core State Standards. The percentages of each students in each level indicate the percentage of students who need to reach the next proficiency level. Means and percentages of students in a proficiency level for small groups should be interpreted with caution because of concerns about reliability and stability.
Item Level Reporting for Student Groups Student group performance (e.g., boys, girls, English learners) on individual test items or groups of items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of Writing & Language standards.

Unintended Score Interpretations and Uses

Until the NM-MSSA is in operational use, we only can speculate on what unintended interpretations and uses of NM-MSSA scores and other information may arise. Where unintended interpretations and uses may occur, it is the responsibility of that user to provide supporting evidence, and not the responsibility of PED (as specified in the *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-MSSA scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations, misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following.

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and proficiency level classifications—are estimates accompanied by a standard error. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Interpreting and using NM-MSSA scores correctly requires considering the width of the margin of error around a score. For example, students with a scale score 2 points below the cut score for the Proficient level could, hypothetically, have scored above the Proficient cut score on a different day because the NM-MSSA scale score standard errors are expected to be 2–3 points. Interpretations of NM-MSSA scores should account for the margin of error around each score estimate.

Drawing Conclusions and Making Decisions Based Solely on NM-MSSA Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less than optimal decisions, such as students failing to receive additional instruction based solely on their NM-MSSA score or teacher teams not being eligible for additional professional learning based solely on their students' NM-MSSA scores. Interpretations and uses of NM-MSSA scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators and Item Level Performance Information

Subdomain indicators (i.e., Writing Analysis and English Language Conventions) are based on fewer items than are NM-MSSA total test scores. As a result, they are less stable estimates of student achievement and learning needs in that subdomain. In addition, because the performance indicators for the Writing & Language subdomains are highly correlated, differences in those performance indicators may be smaller than the proficiency level labels may suggest. Interpretations and uses of indicator scores should be supplemented with additional information.

Overinterpreting NM-MSSA Scores as Indicators of College and Career Readiness

The New Mexico *Common Core State Standards* are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM-MSSA indicates readiness for college and career is supported only by the evidence contained in the language arts content standards. NM-MSSA scores also can be interpreted as predictors of future performance, in college and career training. However, until empirical prediction studies are completed, this interpretation of NM-MSSA performance should be made with caution and with attention to the strong, but limited, evidence in the content standards.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. It would be unwise to assume that students who currently are performing at the Proficient and Advanced levels will continue at these levels without sustained effort. Similarly, it would be unwise—and unfair—to assume that students who currently are performing at the Novice and Nearing Proficiency levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.

This misinterpretation about students and the future is not really a misinterpretation of NM-MSSA scores. It is a logical error in concluding that current performance determines future performance.

STATEMENTS OF SCORE INTERPRETATIONS AND USES (SIUs) FOR THE NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT (NM-MSSA) ASSESSMENT

Writing

Score Interpretation and Use (SIU) Statements for the NM-MSSA and NM-ASR Assessment Programs

The phrase "intended score interpretations for uses" appears several times in the *Standards for Educational and Psychological Testing* and is at the core of the field's views on validity and validation. It also is central to responding successfully to USDE peer review requirements. For the NM-MSSA ELA and mathematics assessments and NM-ASR, the phrase refers broadly to **test scores** (i.e., total test scale scores, subdomain indicators), **aggregations of test scores** (e.g., the percentage of students at and above Level 3. Proficient), and **other test performance informational elements** (e.g., the definition of Proficient in the Proficiency Level Descriptors).

SIU Statements for the NM-MSSA: Writing

Using this broad interpretation of the phrase, the intended score interpretations and uses for NM-MSSA are stated below. These statements reflect input from PED's Technical Advisory Committee and multiple statewide webinars with educator and parent stakeholders on the similar NM-ASR SIU statements.
PED and Cognia will use the final, approved SIU statements to guide decisions about test design and score reporting.
NM-MSSA score reports include scale scores for ELA, Reading, and Writing & Language. The ELA scale score includes performance on the Reading, Writing & Language, and Writing sections of the test. Score reports for NM-MSSA Writing include only rubric scores (i.e., no scale scores).

Intended Interpretations and Uses for Individual Students and Groups of Students

Score Interpretation/Use Statement	Explanation/Annotation
NM-MSSA Program Purpose Statements	
Program Purpose Statement, Grades 3–8 NM-MSSA The NM-MSSA grades 3–8 assessments are designed to measure whether students are on track to be ready for college or career, as defined by the state, by showing they have mastered *the New Mexico Common Core State Standards* (NMCCSS) for ELA, Reading, Writing & Language, and Writing. Writing results are presented using rubric scores only. In grades 3–8, these rubric scores are Production of Writing (1–4 points) and Use of Conventions (1–3 points). Proficient writing performance in each grade indicates mastery of both current grade level and preceding grade level expectations and progress toward college and career readiness. Proficient writing performance is inherent in the description in each rubric's score level 3. The Writing standards require students to (a) learn to use writing as a way of offering and supporting opinions, demonstrating	NM-MSSA scores should be interpreted in relation to the *New Mexico Common Core State Standards* that are targeted by the assessment. ***College readiness*** indicates that a student is prepared to enter directly into and succeed (i.e., earning a C or better) in entry-level, credit-bearing college and relevant technical courses at two- and four-year public institutions of higher education, without the need for remediation. ***Career readiness*** indicates that students have developed the academic and technical skills (i.e., workplace competencies in one or more of 16 career clusters) necessary to succeed in future careers and to become lifelong learners. ***College and Career Readiness*** is defined by the State and can be found in the following College and Career Readiness Bureau's web page: https://webnew.ped.state.nm.us/bureaus/college-career-readiness/

Score Interpretation/Use Statement	Explanation/Annotation
understanding of the subjects they are studying, and conveying real and imagined experiences and events; (b) communicate clearly to an external, sometimes unfamiliar audience; and (c) adapt the form and content of their writing to accomplish a particular task and purpose. See http://www.corestandards.org/ELA-Literacy/CCRA/W/	
Individual Students	
Master Claim Performance on the NM-MSSA indicates a student's progress toward college and career readiness.	College and career readiness requires that students can, without significant scaffolding, (a) write to support opinions, demonstrate understanding of the subjects they are studying, and convey real and imagined experiences and events; (b) communicate clearly to an external audience; and (c) adapt the form and content of their writing to accomplish a particular task and purpose as set forth in the *New Mexico Common Core State Standards* for English Language Arts.
Interpretation of Production of Writing Rubric Scores In grades 3–8, student writing is scored for Production of Writing using a 1–4 point rubric for each of three different purposes for writing: opinion/argument, informative/explanatory, and narrative. Writing that is scored as a 3 indicates that the student (a) presents ideas that generally address the task; (b) generally develops the topic as appropriate for the intended purpose for writing with mostly pertinent facts, definitions, details, examples, and other information from relevant sources, or descriptive detail and somewhat varied narrative techniques; (c) establishes and generally maintains an organizational plan or sequence of events with an introduction and generally related conclusion; and (d) generally makes language choices to create cohesion or transitions and to generally contribute to a style and tone appropriate to the purpose for writing.	Student performance is reported, using this rubric, on one of four score levels. Score point 3 indicates Proficient performance on this element of writing. The scoring rubrics and other information about scoring student writing are available at https://newmexico.onlinehelp.cognia.org/wp-content/uploads/sites/10/2020/11/NM-MSSA_Writing-Assessment-Resources-for-Teachers.pdf
Interpretation of Use of Conventions Rubric Scores In grades 3–8, student writing is scored for Use of Conventions using a 1–3 point rubric. Writing that is scored as a 3 indicates that the student (a) demonstrates general command of standard English grammar and usage, (b) demonstrates general command of standard English conventions relative to the length and complexity of text, (c) produces writing with infrequent errors that do not interfere with meaning or confuse the reader.	Student performance is reported, using this rubric, on one of four score levels. Score point 3 indicates Proficient performance on this element of writing. The scoring rubrics and other information about scoring student writing are available at https://newmexico.onlinehelp.cognia.org/wp-content/uploads/sites/10/2020/11/NM-MSSA_Writing-Assessment-Resources-for-Teachers.pdf
Groups of Students	
SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.	
Group Mean Rubric Scores Group mean rubric scores can be compared to other schools, districts, and the state, and for all students and student subgroups (e.g., gender, English learners, students with disabilities).	Mean (i.e., average) rubric scores enable comparison of performance among schools, districts, and other groupings of students. Mean rubric scores and percentages of students in each rubric score level for small groups (e.g., fewer than 25 students) are unstable and should be interpreted with caution because of concerns about reliability and stability.
Percentages of Students in Rubric Score Levels Percentages of students in the three or four rubric score levels can be compared to other schools, districts, and the state, and for all students and student subgroups.	These are the percentages of students in each Writing rubric score level. The score for each Writing rubric score level indicates the degree of mastery of the knowledge and skills needed to indicate college and

Score Interpretation/Use Statement	Explanation/Annotation
	career readiness in relation to the *New Mexico Common Core State Standards*. The percentages of students in each score level indicate the percentage of students who need to reach the next rubric score level. Means and percentages of students in a rubric score level for small groups should be interpreted with caution because of concerns about reliability and stability.

Unintended Score Interpretations and Uses

Until the NM-MSSA is in operational use, we only can speculate on what unintended interpretations and uses of NM-MSSA scores and other information may arise. Where unintended interpretations and uses occur, it is the responsibility of that user to provide supporting evidence, and not the responsibility of PED (as specified in the *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-MSSA scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations, misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following.

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and rubric score levels—are estimates. Scale scores are accompanied by standard errors. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Writing rubric scores typically are not accompanied by standard errors. Instead, rater disagreement may indicate an error in scoring or other interpretations. (See "Overinterpreting Subdomain Indicators and Writing Rubric Scores" below for additional details.)

Drawing Conclusions and Making Decisions Based Solely on NM-MSSA Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less than optimal decisions, such as students failing to receive additional instruction based solely on their NM-MSSA score or teacher teams not being eligible for additional professional learning based solely on their students' NM-MSSA scores. Interpretations and uses of NM-MSSA scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators: The Two Writing Rubric Scores

Subdomain indicators in the overall ELA assessment account for performance on the Writing assessment. The overall ELA assessment and the Reading and Writing & Language assessments are reported using scale scores. Scale scores are accompanied by standard errors.
In contrast, the two writing rubric scores are the subdomain indicators for the Writing assessment. Each student essay is scored by two raters on two rubrics (i.e., Production of Writing, 1–4 points, and Use of Conventions, 1–3 points). The student's essay score is the sum of the two rater scores. One-point differences (i.e., "disagreement") between rater scores are summed. Two- or three-point differences between rater scores are resolved by a third rater. Disagreements between rater scores may be interpreted as (a) scoring errors, or (b) indicators that the essay reflects features of two adjacent rubric score levels. In addition, because the two Writing rubric scores are correlated, differences across rubric

scores may be smaller than the rubric descriptions may suggest. Interpretations and uses of rubric scores should be supplemented with additional information about students' writing experiences and performance.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. It would be unwise to assume that students who currently are performing at the Production of Writing rubric score 3 (out of a total possible 4 points) will continue to write at this level without sustained effort every school year. Similarly, it would be unwise—and unfair—to assume that students who currently at lower rubric score levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.
Misinterpretations about students' current proficiency levels and future performance is not really a misinterpretation of NM-MSSA scores. It is a logical error in concluding that current performance determines future performance.

Overinterpreting NM-MSSA Scores as Indicators of College and Career Readiness

The New Mexico *Common Core State Standards* are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM-MSSA indicates readiness for college and career is supported only by the evidence contained in the English Language Arts content standards. NM-MSSA scores also can be interpreted as predictors of future performance, in college and career training. However, until empirical prediction studies are completed, this interpretation of NM-MSSA performance should be made with caution and with attention to the strong, but limited, evidence in the content standards.

STATEMENTS OF SCORE INTERPRETATIONS AND USES (SIUs) FOR THE NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT (NM-MSSA) ASSESSMENT

Mathematics

Score Interpretation and Use (SIU) Statements for the NM-MSSA and NM-ASR Assessment Programs

The phrase "intended score interpretations for uses" appears several times in the *Standards for Educational and Psychological Testing* and is at the core of the field's views on validity and validation. It also is central to responding successfully to USDE peer review requirements. For the NM-MSSA ELA and mathematics assessments, the phrase refers broadly to **test scores** (i.e., total test scale scores, subtest indicators), **aggregations of test scores** (e.g., the percentage of students at and above Level 3: Proficient), and **other test performance informational elements** (e.g., the definition of Proficient in the Proficiency Level Descriptors).

SIU Statements for the NM-MSSA: Mathematics

Using this broad interpretation of the phrase, the intended score interpretations and uses for NM-MSSA are stated below. These statements reflect input from multiple statewide webinars with educator and parent stakeholders and PED's Technical Advisory Committee.
PED and Cognia will use the final, approved SIU statements to guide decisions about test design and score reporting.

Intended Interpretations and Uses for Individual Students and Groups of Students

Score Interpretation/Use Statement	Explanation/Annotation
NM-MSSA Program Purpose Statements	
Program Purpose Statement, Grades 3–8 NM-MSSA The NM-MSSA grades 3–8 assessments are designed to measure whether students are on track to be ready for college or career, as defined by the state, by showing they have mastered *the New Mexico Common Core State Standards* for mathematics. Results are presented using scale scores and proficiency levels. Proficient performance in each grade indicates both mastery of currently assessed grade level and preceding grades' expectations and progress toward college and career readiness. The mathematics standards require a solid understanding of concepts, a high degree of procedural skill and fluency, and the application of mathematics to solve problems. See details at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/.	NM-MSSA scores should be interpreted in relation to the *New Mexico Common Core State Standards* that are targeted by the assessment. ***College readiness*** indicates that a student is prepared to enter directly into and succeed (i.e., earn a C or better) in entry-level, credit-bearing college and relevant technical courses at two- and four-year public institutions of higher education, without the need for remediation. ***Career readiness*** indicates that students have developed the academic and technical skills (i.e., workplace competencies in one or more of 16 career clusters) necessary to succeed in future careers and to become lifelong learners. ***College and Career Readiness*** is defined by the State and can be found in the following College and Career Readiness Bureau's web page: https://webnew.ped.state.nm.us/bureaus/college-career-readiness/

Score Interpretation/Use Statement	Explanation/Annotation
Program Purpose Statement, Grade 8 NM-MSSA Performance on the grade 8 NM-MSSA indicates mastery of (a) grade 8 content standards, and (b) solid understanding of concepts, a high degree of procedural skill and fluency, and the application of mathematics to solve problems. See details at https://webnew.ped.state.nm.us/bureaus/assessment-3/nm-mssa/. It also is (c) a predictor of being on track for college and career readiness as defined by the College Board's *College and Career Readiness Benchmarks*: see https://collegereadiness.collegeboard.org/about/scores/benchmarks . Cognia will establish a psychometric link from grade 8 NM-MSSA scores to College Board PSAT scores that will enable monitoring of student progress toward the College Board CCR Benchmarks.	The program purpose statements apply in grades 3–8. In addition, performance on the grade 8 NM-MSSA can be interpreted as a predictor of performance on the PSAT 8/9 and PSAT 10, specifically prediction of status in relation to the College Board CCR Benchmarks. The current links are based on a small empirical validation study conducted outside of NM, which is one indicator of college and career readiness. Linking studies conducted specifically for NM students will provide the links necessary for monitoring student progress toward college and career readiness.
Individual Students	
Master Claim Performance on the NM-MSSA indicates a student's progress toward college and career readiness.	College and career readiness requires that students can, without significant scaffolding, comprehend and evaluate grade-level problems in mathematics as set forth in the *New Mexico Common Core State Standards* for mathematics with connections to the Standards for Mathematical Practices.
Interpretations Using Proficiency Level Labels and Proficiency Level Descriptors (PLDs) Student scores coincide with one of four levels: Advanced, Proficient, Nearing Proficiency, and Novice.[6] The PLD for each proficiency level describes what students are expected to know and be able to do in relation to the *New Mexico Common Core State Standards* for mathematics in grades 3–8. New Mexico students are expected to perform at the Proficient level to demonstrate mastery of the knowledge and skills needed to indicate college and career readiness.	A student's proficiency level indicates how the student performed in relation to the knowledge and skills assessed in mathematics at that grade level. Proficiency level descriptors indicate the knowledge and skills that students are expected to be able to demonstrate at a level.
Interpretations Using Proficiency Level Descriptors A student's proficiency level indicates that the student can be expected to demonstrate the knowledge and skills described at that level and in the levels below.	The student's proficiency level also indicates that the student has mastered the knowledge and skills of the preceding proficiency levels.
Interpretations Using Scale Scores Scale scores provide a measure of student performance regardless of which form of the NM-MSSA is administered.	Scale scores indicate the student's performance, regardless of which form of the NM-MSSA is taken. The **proposed** scale score reporting scale is under discussion.

[6] **NM-MSSA Policy Proficiency Level Descriptors**

Advanced. Students demonstrate evidence of **thorough** understanding and use of college and career readiness knowledge, skills, and abilities.

Proficient. Students demonstrate evidence of **satisfactory** understanding and use of college and career readiness knowledge, skills, and abilities.

Nearing Proficiency. Students demonstrate evidence of **partial** understanding and use of college and career readiness knowledge, skills, and abilities.

Novice. Students demonstrate evidence of **emerging** understanding and use of college and career readiness knowledge, skills, and abilities.

Score Interpretation/Use Statement	Explanation/Annotation
Uses of Scale Scores Scale scores can be used to compare an individual student's performance to the performance of other students in the school, district, and state.	Scale scores also indicate a student's performance in relation to the performance of other students. A student's scale score should be interpreted as the range of possible scores within the error band around that score, not only as a single number. (Other terms for "error band" include "margin of error" and "confidence interval.") Differences between scale scores (e.g., for two students or a student's score and a proficiency level cut score) that are within the margin of error should be interpreted as "statistical ties" (i.e., not reliably different).
Interpretation of Operations & Algebraic Thinking Grades 3–5 Indicators Student performance in the Operations & Algebraic Thinking indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient.[7] Performance on Operations & Algebraic Thinking items reflects application of mathematical operations; using algebraic representations to solve problems involving whole numbers; and identifying, explaining, and extending arithmetic patterns.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Number & Operations in Base Ten and Number & Operations – Fractions Grades 3–5 Indicators Student performance on Number & Operations in Base Ten and Number & Operations – Fractions indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7] Performance on Number & Operations in Base Ten items reflect understanding and using whole number place values to represent and interpret numbers, understanding the concept of fractions, representing fractions and decimal fractions, and comparing the sizes of whole numbers and fractions.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Measurement & Data and Geometry Grades 3–5 Indicators Student performance on Measurement & Data and Geometry indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7] Performance on Measurement & Data and Geometry items reflects understanding measurement principles and applying them to solve problems, representing and analyzing data in simple graphs, and understanding geometric principles and using them to describe objects and solve problems.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Ratios & Proportional Relationships Grades 6 and 7 Indicators Student performance on Ratios & Proportional Relationships indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7]	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.

[7] NM-MSSA subscore indicators are reported as Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. These subscore indicators are calculated using the IRT item parameters only for the items that are tagged for each subscore indicator. They are based on the cut score for Nearing Proficient, means, and conditional standard errors.

Score Interpretation/Use Statement	Explanation/Annotation
Performance on Ratios & Proportional Relationships items reflects understanding, representing, and interpreting ratios and proportional relationships between variables to solve problems.	
Interpretation of Functions Grade 8 Indicators Student performance on Functions indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7] Performance on Functions items reflects understanding the concept of functions and representing linear functions in equations, tables, and graphs; comparing properties of two functions; and interpreting linear and nonlinear functions presented in a variety of forms.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of The Number System and Expressions & Equations Grades 6–8 Indicators Student performance on The Number System and Expressions & Equations indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7] Performance on The Number System items in grades 6 and 7 reflects understanding, representing, and computing with rational numbers. Performance on The Number System items in grade 8 reflects understanding and comparing rational and irrational numbers. Performance on Expressions & Equations items reflects using expressions, equations, and inequalities to represent and solve mathematical and real-world problems.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Geometry and Statistics & Probability Grades 6–8 Indicators Student performance on Geometry and Statistics & Probability indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7] Performance on Geometry and Statistics & Probability items reflects understanding and applying geometric properties related to area, surface area, volume, and angles to solve real-world and mathematical problems; representing and analyzing data in a variety of plots and graphs; and summarizing and describing distributions using multiple measures.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Interpretation of Problem Solving, Reasoning, and Argument Indicators Student performance on Problem Solving, Reasoning, and Argument indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, Did Not Meet Proficient. [7] Performance on Problem Solving, Reasoning, and Argument items reflects applying grade-level appropriate mathematical concepts and procedures and quantitative and logical reasoning to solve standard and nonstandard real-world and mathematical problems; and constructing viable arguments and critiquing the reasoning of others.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.

Score Interpretation/Use Statement	Explanation/Annotation
Interpretation of Modeling, Patterns, and Structure Indicators Student performance on Modeling, Patterns, and Structure indicators is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [7] Performance on Modeling, Patterns, and Structure items reflects using grade-level appropriate quantitative reasoning to interpret mathematical representations, representing real-world mathematical situations using mathematical models, using mathematical models to solve real-world and mathematical problems, and looking for and making use of structure and repeated reasoning.	Student performance on each of the mathematics claims indicators is based on the items that target each of the claim areas. These items also explicitly target a depth of knowledge (DOK) level that categorizes the cognitive demand of the item content.
Item Level Reporting for Individual Students Individual student performance on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of mathematics standards.
Groups of Students	
SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.	
Group Mean Scale Scores Group mean scale scores can be compared to other schools and districts, to the state, and for all students and student subgroups (e.g., gender, English learners, students with disabilities).	Mean (i.e., average) scale scores enable comparison of performance among schools, districts, and other groupings of students. Mean scale scores and percentages of students in a proficiency level for small groups (e.g., fewer than 25 students) are unstable and should be interpreted with caution because of concerns about reliability and stability.
Percentages of Students in Proficiency Levels Percentages of students in the four proficiency levels can be compared to other schools, districts, and the state, and for all students and student subgroups.	These are the percentages of students in each mathematics proficiency level. The PLD for each mathematics proficiency level indicates the degree of mastery of the knowledge and skills needed to indicate college and career readiness in relation to the *New Mexico Common Core State Standards*. The percentages of students in each level indicate the percentage of students who need to reach the next proficiency level. Means and percentages of students in a proficiency level for small groups should be interpreted with caution because of concerns about reliability and stability.
Item Level Reporting for Student Groups Student group performance (e.g., boys, girls, English learners) on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of mathematics standards.

Unintended Score Interpretations and Uses

Until the NM-MSSA is in operational use, we can only speculate on what unintended interpretations and uses of NM-MSSA scores and other information may arise. Where unintended interpretations and uses may occur, it is the responsibility of that user to provide supporting evidence, and not the responsibility of PED (as specified in the *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-MSSA scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations,

misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following:

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and proficiency level classifications—are estimates accompanied by a standard error. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Interpreting and using NM-MSSA scores correctly requires considering the width of the margin of error around a score. For example, students with a scale score 2 points below the cut score for the Proficient level could, hypothetically, have scored above the Proficient cut score on a different day because the NM-MSSA scale score standard errors are expected to be 2–3 points. Interpretations of NM-MSSA scores should account for the margin of error around each score estimate.

Drawing Conclusions and Making Decisions Based Solely on NM-MSSA Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less than optimal decisions, such as students failing to receive additional instruction based solely on their NM-MSSA score or teacher teams not being eligible for additional professional learning based solely on their students' NM-MSSA scores. Interpretations and uses of NM-MSSA scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators and Item Level Performance Information

Subdomain indicators (e.g., Operations and Algebraic Thinking, Measurement and Data) are based on fewer items than are NM-MSSA total test scores. As a result, they are less stable estimates of student achievement and learning needs in that subdomain. In addition, because the performance indicators for the mathematics subdomains are highly correlated, differences in those performance indicators may be smaller than the proficiency level labels may suggest. Interpretations and uses of indicator scores should be supplemented with additional information.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. It would be unwise to assume that students who currently are performing at the Proficient and Advanced levels will continue at these levels without sustained effort. Similarly, it would be unwise—and unfair—to assume that students who currently are performing at the Novice and Nearing Proficiency levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.

Misinterpretations about students' current proficiency levels and future performance is not really a misinterpretation of NM-MSSA scores. It is a logical error in concluding that current performance determines future performance.

Overinterpreting NM-MSSA Scores as Indicators of College and Career Readiness

The New Mexico *Common Core State Standards* are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM-MSSA indicates readiness for college and career is supported only by the evidence contained in the mathematics content standards. NM-MSSA scores also can be interpreted as predictors of future performance in college and career training. However, until empirical prediction studies are completed, this interpretation of NM-MSSA performance should be made with caution and with attention to the strong, but limited, evidence in the content standards.

STATEMENTS OF SCORE INTERPRETATIONS AND USES (SIUs) FOR THE NEW MEXICO ASSESSMENT OF SCIENCE READINESS (NM-ASR)

Score Interpretation and Use (SIU) Statements for the NM-MSSA and NM-ASR Assessment Programs

The phrase "intended score interpretations for uses" appears several times in the Standards for Educational and Psychological Testing and is the core of the field's views on validity and validation. It also is central to responding successfully to USDE peer review requirements. For the NM-ASR, the phrase refers broadly to test scores (i.e., total test scale scores, subdomain scores), aggregations of test scores (e.g., the percentage of students at and above Level 3: Proficient), and other test performance informational elements (e.g., the definition of Proficient in the Proficiency Level Descriptors).

SIU Statements for the NM-ASR

Using this broad interpretation of the phrase, the intended score interpretations and uses for NM-ASR are stated below. These statements reflect input from multiple statewide webinars with educator and parent stakeholders and PED's Technical Advisory Committee.
PED and Cognia will use the final, approved SIU statements to guide decisions about test design and score reporting.
NM-MSSA score reports include scale scores for ELA, Reading, and Writing & Language. The ELA scale score includes performance on the Reading, Writing & Language, and Writing sections of the test. Score reports for NM-MSSA Writing include only rubric scores (i.e., no scale scores).

Intended Score Interpretations and Uses for Individual Students and Groups of Students

Score Interpretation/Use Statement	Explanation/Annotation
NM-ASR Program Purpose Statements	
Program Purpose Statement, Grade 11 NM-ASR The grade 11 NM-ASR is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the New Mexico *STEM Ready! Science Standards*, which require integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts to explain phenomena and solve problems. Results are presented using scale scores and proficiency levels. Proficient performance in grade 11 indicates both mastery of currently assessed grade level and preceding grades' expectations and progress toward college and career readiness.	NM-ASR scores should be interpreted in relation to the New Mexico *STEM Ready! Science Standards* that are targeted by the assessment. ***College readiness*** indicates that a student is prepared to enter directly into and succeed (i.e., earn a C or better) in entry-level, credit-bearing college and relevant technical courses at two- and four- year public institutions of higher education, without the need for remediation. ***Career readiness*** indicates that students have developed the academic and technical skills (i.e., workplace competencies in one or more of 16 career clusters) necessary to succeed in future careers and to become lifelong learners. ***College and Career Readiness*** is defined by the State and can be found in the following College and Career Readiness Bureau's web page: https://webnew.ped.state.nm.us/bureaus/coll ege-career-readiness/ Evidence to support this NM-ASR college and career readiness claim

	is in the New Mexico *STEM Ready! Science Standards*, which are based on the Next Generation Science Standards (NGSS). The NGSS "constructed each performance expectation by linking concepts and practices that build coherently over time throughout K–12, thereby helping to ensure that students who meet the NGSS will be prepared to succeed in science courses in both 2- and 4-year institutions" (see NGSS *Appendix C – College and Career Readiness* at https://www.nextgenscience.org/sites/default/files/resource/files/NGSS%20Appendix%20C%20Final%20072613.pdf).
Program Purpose Statement, Grades 5 and 8 NM-ASR Performance on the grade 8 NM-ASR indicates student mastery of grade levels 3–5 and 6–8 expectations for integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts as presented in the standards, which is the progression for the next level of science curriculum and is a predictor of being on track for college and career readiness.	The four explanations at grade 11 (above) apply in grades 5 and 8. In addition, performance on the grade 8 NM- ASR can be interpreted as a potential predictor of performance on the grade 11 NM- ASR (pending empirical validation), which is one indicator of college and career readiness. Performance on the grade 5 NM-ASR can be interpreted as a potential predictor of performance on the grade 8 NM-ASR (pending
Score Interpretation/Use Statement	**Explanation/Annotation**
Proficient performance in grades 5 and 8 indicates both mastery of currently assessed grade level and preceding grades' expectations and progress toward college and career readiness.	empirical validation), which is a predictor of college and career readiness at grade 11.
Individual Students	
Master Claim Performance on the NM-ASR indicates a student's progress toward college and career readiness.	College and career readiness requires that students can make sense of phenomena and solve real-world problems by applying and interconnecting scientific knowledge and skills as set forth in the New Mexico *STEM Ready! Science Standards*.
Interpretations Using Proficiency Level Labels and Proficiency Level Descriptors (PLDs) Student scores coincide with one of four levels: Novice, Nearing Proficiency, Proficient, and Advanced.[8] The PLD for each proficiency level describes what students can be expected to know and be able to do in relation to the New Mexico *STEM Ready! Science Standards* in grades 5, 8, and 11. New Mexico students are expected to perform at the Proficient level to demonstrate sufficient mastery of knowledge and skills needed to indicate college and career readiness.	A student's proficiency level indicates how the student performed in relation to the knowledge and skills assessed in science at that grade level. Proficiency level descriptors indicate the knowledge and skills that students are expected to be able to demonstrate at a level.

[8] **NM-ASR Policy Proficiency Level Descriptors**

Advanced. Students demonstrate evidence of **thorough** understanding and use of college and career readiness knowledge, skills, and abilities.

Proficient. Students demonstrate evidence of satisfactory understanding and use of college and career readiness knowledge, skills, and abilities.

Nearing Proficiency. Students demonstrate evidence of partial understanding and use of college and career readiness knowledge, skills, and abilities.

Novice. Students demonstrate evidence of emerging understanding and use of college and career readiness knowledge, skills, and abilities.

Interpretations Using Proficiency Level Descriptors A student's proficiency level indicates that the student can demonstrate the knowledge and skills described at that level and in the levels below.	The student's proficiency level also indicates that the student has mastered the knowledge and skills of the preceding proficiency levels.
Interpretations Using Scale Scores Scale scores provide a measure of student performance regardless of which form of the NM- ASR is administered.	Scale scores indicate the student's performance, regardless of which form of the NM-ASR is taken. The **proposed** scale score reporting scale is under discussion.
Score Interpretation/Use Statement	**Explanation/Annotation**
Uses of Scale Scores Scale scores can be used to compare an individual student's performance to the performance of other students in the school, district, and state.	Scale scores also indicate a student's performance in relation to the performance of other students. A student's scale score should be interpreted as the range of possible scores within the error band around that score, not only as a single number. (Other terms for "error band" include "margin of error" and "confidence interval.") Differences between scale scores (e.g., for two students or a student's score and a proficiency level cut score) that are within the margin of error should be interpreted as "statistical ties" (i.e., not reliably different).
Interpretations of Practices and Crosscutting Concepts in Physical Sciences Student performance on this science subdomain is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient.[9]	Student performance in this science subdomain is based on items that target Disciplinary Core Ideas in Physical Sciences plus Science and Engineering Practices and/or Crosscutting Concepts. Because indicators for Physical, Life, and Earth and Space Sciences are likely to be highly correlated and will have non-trivial standard errors, proficiency levels for most students are likely to be identical in all three science subdomains.
Interpretations of Practices and Crosscutting Concepts in Life Sciences Student performance on this science subdomain is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [9]	Student performance in this science subdomain is based on items that target Disciplinary Core Ideas in Life Sciences plus Science and Engineering Practices and/or Crosscutting Concepts. Because indicators for Physical, Life, and Earth and Space Sciences are likely to be highly correlated and will have non-trivial standard errors, proficiency levels for most students are likely to be identical in all three science subdomains.
Interpretations of Practices and Crosscutting Concepts in Earth and Space Sciences Student performance on this science subdomain is reported in three levels: Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. [9]	Student performance in this science subdomain is based on items that target Disciplinary Core Ideas in Earth and Space Sciences plus Science and Engineering Practices and/or Crosscutting Concepts. Because indicators for Physical, Life, and Earth and Space Sciences are likely to be highly correlated and will have non-trivial standard errors, proficiency levels for most students are likely to be identical in all three science subdomains.

[9] NM-ASR subdomain indicators are reported as Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient. These subdomain indicators are calculated by comparing a student's subdomain performance to the subdomain performance distribution of students who are just barely Nearing Proficient on the total test, and by using the standard deviation of that distribution to determine the Met/Exceeded Proficient, Nearing Proficient, and Did Not Meet Proficient indicators.

Score Interpretation/Use Statement	Explanation/Annotation
Item Level Reporting for Individual Students Individual student performance on individual test items may suggest potential areas of strength and learning needs.	Caveat: Students may perform differently on items from other test forms that target the same subset of science standards.
Groups of Students	
SIU statements for groups of students are applicable to aggregate reporting of school, district, and state performance and student subgroups (e.g., English learners, students with disabilities, racial/ethnic subgroups) within those levels of aggregation.	
Group Mean Scale Scores Group mean scale scores can be compared to other schools, districts, and the state, and for all students and student subgroups (e.g., gender, English learners, students with disabilities, racial/ethnic subgroups).	Mean (i.e., average) scale scores enable comparison of performance among schools, districts, and other groupings of students. Mean scale scores and percentages of students in a proficiency level for small groups (e.g., fewer than 25 students) are unstable and should be interpreted with caution because of concerns about reliability and stability.
Percentages of Students in Proficiency Levels Percentages of students in the four proficiency levels can be compared to other schools, districts, and the state, and for all students and student subgroups.	These are the percentages of students in each science proficiency level. The PLD for each science proficiency level indicates the degree of mastery of the knowledge and skills needed to indicate college and career readiness in relation to the *New Mexico STEM Ready! Science Standards*. The percentages of students in each level indicate the percentage of students who need to reach the next proficiency level. Means and percentages of students in a proficiency level for small groups should be interpreted with caution because of concerns about reliability and stability.
Item Level Reporting for Student Groups Student group performance (e.g., boys, girls, English learners) on individual test items or groups of items may suggest potential areas of strength and learning needs—with the caution that a student group may perform differently on other items that target the same Disciplinary Core Ideas, Science and Engineering Practices, and Crosscutting Concepts.	Caveat: Students may perform differently on items from other test forms that target the same subset of science standards.

Unintended Score Interpretations and Uses

Until the NM-ASR is in operational use, we can only speculate on what unintended interpretations and uses of NM-ASR scores and other information may arise. Where unintended interpretations and uses may be in use, it is the responsibility of that user to provide supporting evidence, and not the responsibility of PED (as specified in the *Standards for Educational and Psychological Testing*, 2014). The main concern for misinterpreting or misusing NM-ASR scores is the potential negative consequences for individual students, subgroups of students, and schools, districts, and the state. If unintended interpretations and uses with potential negative consequences arise, PED will take steps to ameliorate the misinterpretations, misuses, and negative consequences. Some common misinterpretations and misuses that can arise include the following.

Interpreting Test Scores as 100% Accurate Indicators of Test Performance

All measurements in the real world, including test scores, are estimates. Test scores—for example, scale scores and proficiency level classifications—are estimates accompanied by a standard error. Standard errors are often referred to as the "margin of error" (e.g., in political polling). Interpreting and using NM-ASR scores correctly requires considering the width of the margin of error around a score. For example, students with a scale score 2 points below the cut score for the Proficient level could, hypothetically, have scored above the Proficient cut score on a different day because the NM-ASR scale score standard errors are expected to be 2–3 points. Interpretations of NM-ASR scores should account for the margin of error around each score estimate.

Drawing Conclusions and Making Decisions Based Solely on NM-ASR Scores

There is wide agreement that conclusions and decisions based on a single piece of evidence can be risky. The risk is that the single piece of evidence can lead to less than optimal decisions, such as students

failing to receive additional instruction based solely on their NM-ASR score or teacher teams not being eligible for additional science professional learning based solely on their students' NM-ASR scores. Interpretations and uses of NM-ASR scores should be supplemented with additional information.

Overinterpreting Subdomain Indicators and Item Level Performance Information

Subdomain indicators (e.g., Interpretations of Practices and Crosscutting Concepts in Life Sciences) are based on fewer items than are NM-ASR total test scores. As a result, they are less stable estimates of student achievement and learning needs in that subdomain. In addition, because the performance indicators for the three science subdomains are highly correlated, differences in those performance indicators may be smaller than the proficiency level labels may suggest. Interpretations and uses of indicator scores should be supplemented with additional information.

Misinterpreting Current Performance as the Most Likely Predictor for Future Performance

A goal of education is to improve students' current achievement—that is, to bend their performance trajectory upward. We assume that students who currently are performing at the Proficient and Advanced levels will continue at these levels only with sustained effort and support. It would be unwise—and unfair—to assume that students who currently are performing at the Novice and Nearing Proficiency levels will perform at these levels in the future. In fact, our duty as educators is to help these students learn more and achieve higher.

Misinterpretations about students' current proficiency levels and future performance is not really a misinterpretation of NM-ASR scores. It is a logical error in concluding that current performance determines future performance.

Overinterpreting NM-ASR Scores as Indicators of College and Career Readiness

The New Mexico *STEM Ready! Science Standards* are designed to prepare students to be able to benefit from college study and postsecondary training. The claim that performance on NM-ASR indicates readiness for college and career is supported only by the evidence contained in the science content standards. NM-ASR scores also can be interpreted as predictors of future performance in college and career training. However, until empirical prediction studies are completed, this interpretation of NM- ASR performance should be made with caution and with attention to the strong, but limited, evidence in the content standards.

Test Specifications

New Mexico Measures of Student Success and Achievement (NM-MSSA)

English Language Arts

Purpose

- Part of a Balanced Assessment System
- Claims/Score Interpretation and Use Statements

Test Specifications

- Reporting Categories and Test Blueprint
- Cognitive Complexity
- Test Design
- Fairness
- Additional NM-MSSA Considerations

Stimulus Specifications

- Specifications for Reading Stimuli
- Specifications for Language Usage Stimuli
- Specifications for Writing Prompt Stimuli

Item Specifications

- Alignment
- Item Types
- Sample Items

Purpose

Part of a Balanced Assessment System

The New Mexico Measures of Student Success and Achievement (NM-MSSA) is New Mexico's statewide summative assessment for Mathematics and English Language Arts, administered at the end of grades 3-8. As the NM-MSSA is a single measure at the end of a grade band, interpretations and uses of NM-MSSA scores should be supplemented with additional measures, including information from classroom summative and formative assessments in mathematics and English Language Arts as well as interim assessments.

Formative assessment may include the use of the Cognia Formative Assessments, which is a collection of formative assessment materials for grades 3–8 being provided by Cognia during the term of the contract with the state to administer the NM-MSSA. The materials are aligned to the CCSS and therefore to the New Mexico *Common Core State Standards*.

Claims/Score Interpretation and Use Statements

The NM-MSSA is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the New Mexico Commons Core State Standards.
In addition to overall ELA scale score, student performance on Reading and Language & Writing and seven subdomains is reported as noted in the reporting categories subsection in the following pages. See the Score Interpretation and Use (SIU) documents for the complete set of NM-MSSA score interpretation and use statements. These documents can be accessed at https://newmexico.onlinehelp.cognia.org/resources-nm-mssa/ .

Test Specifications – Reporting Categories and Test Blueprint

Reporting Categories

The reporting categories for NM-MSSA are based on the clusters of standards found in the New Mexico *Common Core State Standards* for both content and practices as noted below.

Reporting Categories-Reading Grades 3-8

- Reading
- Text Type
  - Literary Text
  - Informational Text
- Reading Strategy
  - Comprehension
  - Analysis and Interpretation Reporting Categories-Writing & Language Grades 3-8
- Writing & Language

- Writing Analysis & Language Conventions
- Production of Writing (rubric score)
- Use of Conventions (rubric score)

Percentages for the distribution of operational (core) test points for each of the reporting categories reflect the distribution in the standards, so as not to over- or underrepresent content. The internal test blueprint has specifications for inclusion on the core forms.

Core Test Blueprints

Specifications for the full test blueprints for the construction of the core forms reflect the reporting category specifications, as well as percentage requirements for each cluster. These constructs represent key aspects of the standards to which items are aligned; as such, the percentage of operational (core) test points for each should be maintained from year to year.

Note that percentages in reading for (a) text type and (b) reading strategy are calculated independently. An individual item may contribute to multiple parts of the blueprint.

All NM-MSSA ELA items are based on passages, either literary or informational (e.g., expository, argument, opinion, memoir, autobiography, etc.). The configuration of texts on the ELA assessment seeks to balance national high-quality assessment guidance (e.g., NAEP, CCSSO, etc.) as well as considerations around test length. The ELA test design incorporates as much of a 50/50 split of literary and informational texts as possible in the elementary grades while still maintaining a limited summative test footprint. Beginning at grade 6, there is a shift in emphasis to informational texts at the upper grade band.

Operational Test Blueprint

	English Language Arts		Grade 3-5		Grade 6-8	
			Ideal # of Core Pts	Ideal % of Core Pts	Ideal # of Core Pts	Ideal % of Core Pts
Reading	Text Type	Literary Text	15	65%	8	35%
		Informational Text	8	35%	15	65%
	Reading Strategy	Comprehension	12-14	52-60%	9-12	39-52%
		Analysis and Interpretation	9-11	39-48%	13-16	56-70%
	Cluster	Key Ideas and Details	9-11	39-48%	7-11	30-48%
		Craft and Structure	7-9	30-39%	6-10	26-43%
		Integration of Knowledge and Ideas	4-6	17-26%	4-6	17-26%
	Total		**23***	**100%***	**23***	**100%***
Writing & Language	Language & Writing Passage Sets	Writing Analysis	9-11	42-52%	9-11	42-52%
		English Language Conventions	3-5	14-23%	3-5	14-23%
	Writing Prompt*	Production of Writing	4	19%	4	19%
		Use of Conventions	3	14%	3	14%
	Total		**21**	**100%**	**21**	**100%**

*All items align to a text type (Literary, Informational), reading strategy (Comprehension, Analysis and Interpretation) and a cluster (Key Ideas and Details, Craft and Structure, and Integration of Knowledge and Ideas).

Test Specifications – Cognitive Complexity

Percentages for the distribution of operational (core) test points across the cognitive complexity levels (DOK classification) are noted in the table below.

Depth of Knowledge Distribution - Reading

DOK	Grade					
	3	4	5	6	7	8
Level 1	0-20%	0-20%	0-20%	0-20%	0-20%	0-20%
Level 2	50-70%	50-70%	50-70%	50-70%	50-70%	50-70%
Level 3	20-40%	20-40%	20-40%	20-40%	20-40%	20-40%
Total	100%	100%	100%	100%	100%	100%

Depth of Knowledge Distribution – Writing & Language

DOK	Grade					
	3	4	5	6	7	8
Level 1	15-35%	15-35%	15-35%	15-35%	15-35%	15-35%
Level 2	40-60%	40-60%	40-60%	40-60%	40-60%	40-60%
Level 3	15-35%	15-35%	15-35%	15-35%	15-35%	15-35%
Total	100%	100%	100%	100%	100%	100%

Test Specifications – Test Design

Assessable Standards

The NM-MSSA assesses the New Mexico *Common Core State Standards*. The following standards are eligible to be included in the NM-MSSA.

- Grade 3-5 test: Reading: Literature (RL.1-7, RL.9), Reading: Informational Text (RI.1- RI.9), Writing, and Language standards.
- Grade 6-8 test: Reading: Literature (RL.1-7, RL.9), Reading: Informational Text (RI.1- RI.9), Writing, and Language standards.

Test Design - Student Experience

The *Student Testing Experience* tables on the following pages provide a summary of the number of items and points by item type, usage (i.e., operational items or field test items), and estimated testing time for each grade's NM-MSSA assessment.

As shown in the test design tables, the types of items on the NM-MSSA are 1-point machine- scored

items (MS-1), 2-point machine-scored items (MS-2), and 7-point writing prompts (WP). Additional item type descriptions and sample items can be found in the item specifications section on page 16.

Each NM-MSSA test is administered in two sessions. The test form contains both core operational items and matrix field test items. The core operational items are items administered to all students taking that core form, and that count toward a student score. Matrix field test items are items administered to subsets of students to "try out" performance (with different students receiving different field test items), and therefore do not count toward student score.

Student Testing Experience (Full Form)

ELA Grades 3-8 (Spring 2021)	Passage-Based Items				Total Items	Total Points	
	Passage Sets	MS-1	MS-2	WP		Min	Max
Core Operational Items	6	32	6	0	38	44	44
Matrix Operational Items	0	0	0	0	0	0	0
Matrix Field Test Items	2	5	1	1	7	14	14
Total Student Experience	**8**	**37**	**7**	**1**	**45**	**58**	**58**
				Estimated Test Time (min)			**150 (60/90)**

ELA Grades 3-8 (Spring 2022 and beyond)	Passage-Based Items				Total Items	Total Points	
	Passage Sets	MS-1	MS-2	WP		Min	Max
Core Operational Items	6	27	5	0	32	37	37
Matrix Operational Items	0	0	0	1	1	7	7
Matrix Field Test Items	2	10	2	0	12	14	14
Total Student Experience	**8**	**37**	**7**	**1**	**45**	**58**	**58**
				Estimated Test Time (min)			**150 (60/90)**

Practice Test

A full-length practice test mirroring the operational test design is available beginning in the 2021-2022 school year. The practice tests and supporting materials can be accessed at https://newmexico.onlinehelp.cognia.org/practice-tests-nm-mssa/.

Test Specifications – Fairness

Fairness is defined as the extent to which the test scores are valid for different groups of test takers. Consideration of universal design, bias, and sensitivity guidelines support the construction of fair, valid assessments.

Universal Design for Assessments

The concept of Universal Design for Assessments focuses on developing content and assessments that reach the widest population of students possible. Stimuli and items on the NM-MSSA are designed to simply and clearly present tasks and to provide maximum readability, comprehensibility, and legibility. The seven elements of Universal Design for Assessments are based on the original UDL guiding principles:

Universal Design for Assessments

Principle	Explanation
Inclusive Assessment Population	Tests designed for state, district, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field-testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct-irrelevant cognitive, sensory, emotional, and physical barriers are removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	Test design facilitates the use of needed accommodations (e.g., all items can be translated to braille).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and presented in understandable language.
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words kept to a minimum) for readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, tables, figures, and illustrations, and to response formats.

Bias

The concept of Bias is defined as the presence of some characteristic of an item that results in differential performance for two individuals of the same ability but from different ethnic, sex, cultural, or religious groups.

Bias can occur whenever content offends or disadvantages a student or group of students due to gender, race, regional background, socioeconomic status, or any other such classification.

Test developers take care to craft content in a way that does not misrepresent specific groups or rest on assumptions made about specific groups, that in turn could negatively impact how students interpret content.

- Stimulus and item content on the NM-MSSA must not present stereotypes or unfair representations of gender, race, ethnicity, disability, culture, or religion.
- Stimulus and item content on the NM-MSSA should not depend on overly-experiential information such as knowledge of technology, consumer goods, pop culture, geographic locations, or sports

and extracurricular activities. While these topics are not completely excluded from use, care must be taken to ensure that the items are presented in a way that does not require a level of knowledge that would not be held by all students.

Sensitivity

Sensitivity refers to the presence of content that is contrary to the acceptable norms of the students, educators, parents, or other members of the community that may interact with the assessment. Sensitive subject matter can impact student performance or attitudes toward testing, and hence, their test scores.

Consideration of bias and sensitivity issues is very important when developing content for an assessment. Test developers must ensure that stimuli and items are free of content that will negatively affect a student's performance not because of what the student knows and can do but because the content evokes an emotional response from that student (or is in some other way distracting to the student).

Subjects/contexts that are likely to prompt emotional distress on the part of students cannot be used on the NM-MSSA (e.g., war, violence, human death or debilitating disease, animal-based medical research). Careful judgment should be applied to standards that cover topics that may be considered controversial by some groups (e.g., evolution examples, population dynamics including death/extinction, environmental impact). Those standards represent content knowledge to be assessed, but the assessment must be done in a sensitive, unbiased way.

Stimulus Specifications

Reading

All NM-MSSA reading items are based on permissioned passages, either literary or informational (e.g., expository, argument, opinion, memoire, autobiography, etc.). The configuration of authentic texts on the assessment seeks to balance national high-quality assessment guidance (e.g., NAEP, CCSSO, etc.) as well as considerations around test length. For grades 3–5, item sets are based on single literary passages, paired literary passages, and paired informational passages. For grades 6–8, item sets are based on paired literary passages, single informational passages, and paired informational passages. In an effort to reflect a 50/50 split of literary and informational texts in the elementary grades, the assessment at grades 3-5 does contain a balance of paired literary and paired informational texts. Beginning at grade 6, there is a shift in emphasis to informational texts at the upper grade band.

Literary passages should include experiences with which students are familiar or provide enough context that a student unfamiliar with the experience is not disadvantaged. Passages should be excellent models of exemplary writing in including such literary elements as character development, a well-crafted plot, a text structure that supports the meaning of the passage, and the development of a point of view. Passages should include text structures and literary devices (figurative language, irony, etc.) appropriate to the grade-level being assessed.

Informational passages must include current and accurate information. In addition, care must be taken to avoid topics and details that could quickly become dated (news stories, technologies, discoveries, etc.). Text content should be consistent with the current best thinking in the various fields represented on the test and be neither overly speculative nor highly tentative, given the likelihood of such texts becoming dated over the long course of test development and administration.

Informational passages must be able to stand on their own and should not require any outside knowledge (other than common knowledge) to understand the topic or author's position.
Informative/explanatory texts and arguments should be accurate, well-reasoned, and logically organized, reflecting a variety of logical text structures including, but not limited to, compare/contrast, cause/effect, order of importance, sequence/steps in a process, problem/solution, description and explanation, question and answer, and cyclical structures. It should also be noted that domain-specific vocabulary must be easily understood via context clues, minimal footnotes, and/or authorial explanation.

Quantitative measures are certainly helpful in situating a passage within a grade band; they will also help to establish a passage as appropriate (or not) for a particular grade. However, quantitative complexity evaluations must be coupled with thorough qualitative review in order to make an informed grade assignment. The qualitative measures of text complexity are evaluated using the qualitative dimensions of text complexity found in Appendix A of the CCSS, as well as the CCSSO Text Complexity Qualitative Measures Rubrics for Informational and Literary Texts.

The quantitative measures used for NM-MSSA reading passages are shown below. It should be noted that passages may sometimes exceed these specifications if the passage content is deemed appropriate in light of qualitative complexity measures.

Word Count (within the prescribed ranges)

Grade	Passage Type	Word Count
3	Single	300 – 800
	Paired	600 – 1,000
4	Single	300 – 800
	Paired	600 – 1,000
5	Single	300 – 800
	Paired	600 – 1,000
6	Single	500 – 1,000
	Paired	800 – 1,200
7	Single	500 – 1,000
	Paired	800 – 1,200
8	Single	500 – 1,000
	Paired	800 – 1,200

Readabilities (Lexile and TextEvaluator also within the prescribed ranges)

Grade Level	College and Career Ready "Stretch" Lexile Bands
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830 L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11 and 12	1185L to 1385L

Grade Level	TextEvaluator℠
3	310 – 590
4	405 – 655
5	480 – 720
6	550 – 790
7	615 – 860
8	685 – 940

Language Usage

All NM-MSSA language usage passages will be commissioned texts which contain embedded errors. A commissioned passage is an original text written for a specific purpose (e.g., editing task). Passages should demonstrate accurate and engaging expository writing, effective argumentation, and vivid, intelligible narrative writing. The essential elements of each of these genres are derived from the applicable Common Core State Standards in Writing. (e.g., per the standards, argumentation in G7 and G8 should develop a claim in part through engagement with a counterclaim; narratives in all grades should be developed through dialogue, thoughts, and description; etc.)

Expository passages are expected to meet the highest standards of factual accuracy, syntactical and grammatical proficiency, reader engagement, and originality. Generally, expository passages will represent one of three subject areas: Social Studies/History; Science/Social Science/Technical Subjects; and, to a lesser extent, the Humanities. Although written with the general reader in mind, passages nevertheless strive to present compelling information that responds to relevant issues in each field—a new interpretation of an event or phenomenon; an examination of an overlooked (or misunderstood) movement, moment, or figure; an introduction to foundational knowledge in any of the three disciplines, etc.

Passages should assume no content background or expertise on the part of readers, but writers should invest their work with the precision and novelty that rewards attentive reading. This principle extends to technical or discipline-specific language, which should appear where necessary and natural, and never gratuitously or without supporting context. The NM-MSSA writing and language assessment is not intended to evaluate students' reading comprehension abilities. If students struggle simply to read a passage, they will surely be unable to demonstrate a meaningful understanding of how to edit or revise it.

Argument/Opinion passages maintain the same high standards of accuracy, syntax, engagement, and originality expected of expository passages. They also demonstrate cogent argumentation. That is, the writer must establish a position; provide claims, supported by evidence, that develop that position; introduce and rebut a counterclaim (in grade 7 and 8); and, throughout, use rhetorical techniques (persuasive transitions, rhetorical questions, appeals to reason or personal experience, etc.) to advance the position.

Argument/Opinion passages will tend to be informed by issues in the social sciences or current events. Successfully realized topics are not inherently controversial, and the writer's position is not diffident or universalist—that is, the argument is primarily positive in tone and modest in scope. Nonetheless, the position is clear.

Narrative passages succinctly and lucidly describe a fictional event. Although these passages will feature many or all the hallmarks of the narrative form—plot/conflict, climax/epiphany, conclusion, dialogue, characters' thoughts, action, description—they must be coherent in spite of their brevity. They avoid long

stretches of dialogue; flashbacks and other jarring time-shifts; florid, clichéd, or self-consciously "literary" language; and more than three speaking characters.

Determining the complexity of a text requires a multifaceted approach. In addition to evaluating text based on quantitative data (readability statistics), the writers must also rely on qualitative measures to help determine whether a passage being developed is appropriate for the target audience. The qualitative measures of text complexity are evaluated using the qualitative dimensions of text complexity found in Appendix A of the CCSS, as well as the CCSSO Text Complexity Qualitative Measures Rubrics for Informational and Literary Texts.

The quantitative measures used for NM-MSSA language usage passages are shown below. It should be noted that passages may sometimes exceed these specifications if the passage content is deemed appropriate in light of qualitative complexity measures.

Word Count (within the prescribed ranges)

Grade	Word Count
3	250–400
4	275–425
5	300–450
6	325–550
7	350–625
8	350–625

Readabilities (Lexile and TextEvaluator also within the prescribed ranges)

Grade Level	College and Career Ready "Stretch" Lexile Bands
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830 L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11 and 12	1185L to 1385L

Grade Level	TextEvaluator℠
3	310 – 590
4	405 – 655
5	480 – 720
6	550 – 790
7	615 – 860
8	685 – 940

Writing Prompts

All NM-MSSA writing prompts will be partnered with one to three brief text stimuli. These may be intact (whole) passages or excerpts (only a part) of a more extended text. Some possible text types include:

- Story
- Memoir
- Biography
- Poem
- Article
- Essay
- Review (book, music, performance, etc.)
- Editorial
- Website
- Letter
- Journal (diary, log) entry
- Instructions
- Advertisement
- Brochure
- Memo
- Script
- Transcript
- Display text (e.g., to accompany an artifact in a museum)

The number of text stimuli will vary depending on the purpose for writing. Narrative prompts will be associated with 1-2 text stimuli, while informative/explanatory and opinion/argument prompts will be associated with 2-3 text stimuli. The passages may be either permissioned or commissioned.

Word Count (maximum **total** word count 800-1000)

Grade	Word Count
3	800
4	850
5	900
6	950
7-8	1,000

Readability (Lexile within the prescribed ranges)

Grade Level	College and Career Ready "Stretch" Lexile Bands
1	190L to 530L
2	420L to 650L
3	520L to 820L
4	740L to 940L
5	830 L to 1010L
6	925L to 1070L
7	970L to 1120L
8	1010L to 1185L
9	1050L to 1260L
10	1080L to 1335L
11 and 12	1185L to 1385L

Item Specifications

Alignment

The items on the NM-MSSA are aligned to the New Mexico *Common Core State Standards*. Each item is aligned to one of the content standards in NMCCSS.

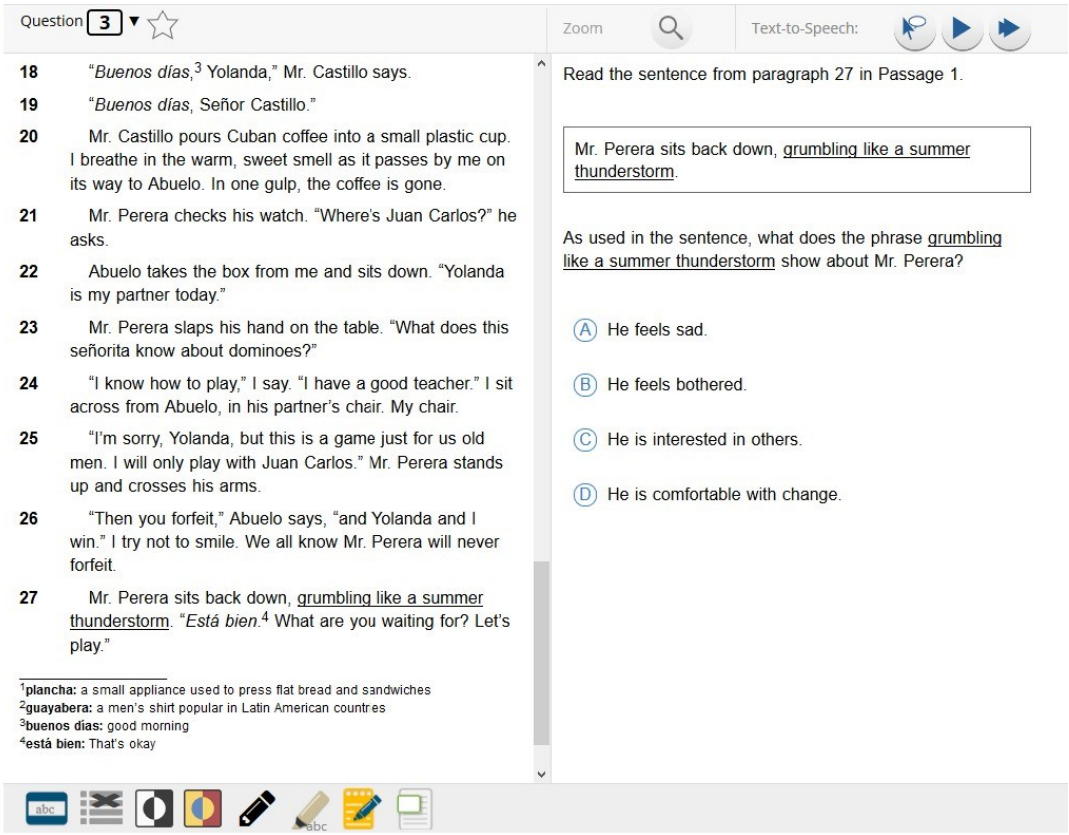
Item Types

The types of items on the NM-MSSA are 1-point machine-scored items (MS-1), 2-point machine-scored items (MS-2), and 7-point writing prompt (WP):

- MS-1 items are worth 1 point. These items may be multiple choice or multiple select.
- MS-2 items are worth 2 points. In ELA, these items are also known as evidence-based selected response items (EBSR). These items have two parts (Part a and Part b) for students to answer, and 0, 1, or 2 points total can be earned across Part a and Part b. Each part provides choices from which to select.
- Writing prompts (WPs) are worth 7 points. These items require students to write an extended response to a single prompt. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The WPs items are evaluated using a "Production of Writing" rubric on a scale from 1–4 and a "Use of Conventions" rubric on a scale from 1-3.

Sample Items

- MS-1 items are worth 1 point. These items may be multiple choice or multiple select.



Question 3 ▾ ☆ Zoom Text-to-Speech:

18 "*Buenos días*,[3] Yolanda," Mr. Castillo says.

19 "*Buenos días*, Señor Castillo."

20 Mr. Castillo pours Cuban coffee into a small plastic cup. I breathe in the warm, sweet smell as it passes by me on its way to Abuelo. In one gulp, the coffee is gone.

21 Mr. Perera checks his watch. "Where's Juan Carlos?" he asks.

22 Abuelo takes the box from me and sits down. "Yolanda is my partner today."

23 Mr. Perera slaps his hand on the table. "What does this señorita know about dominoes?"

24 "I know how to play," I say. "I have a good teacher." I sit across from Abuelo, in his partner's chair. My chair.

25 "I'm sorry, Yolanda, but this is a game just for us old men. I will only play with Juan Carlos." Mr. Perera stands up and crosses his arms.

26 "Then you forfeit," Abuelo says, "and Yolanda and I win." I try not to smile. We all know Mr. Perera will never forfeit.

27 Mr. Perera sits back down, grumbling like a summer thunderstorm. "*Está bien*.[4] What are you waiting for? Let's play."

[1]**plancha:** a small appliance used to press flat bread and sandwiches
[2]**guayabera:** a men's shirt popular in Latin American countries
[3]**buenos días:** good morning
[4]**está bien:** That's okay

Read the sentence from paragraph 27 in Passage 1.

Mr. Perera sits back down, grumbling like a summer thunderstorm.

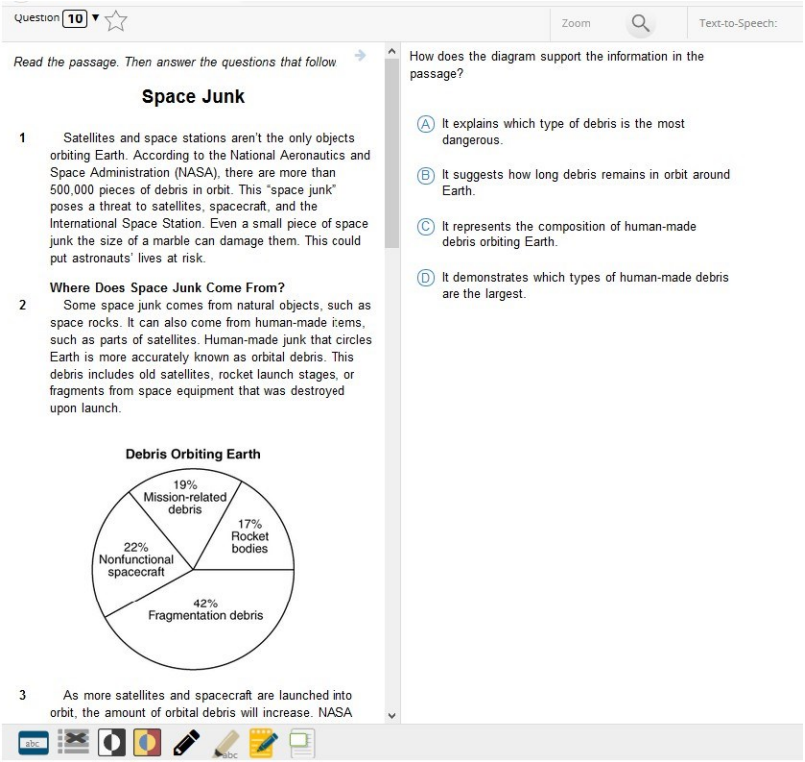
As used in the sentence, what does the phrase grumbling like a summer thunderstorm show about Mr. Perera?

Ⓐ He feels sad.

Ⓑ He feels bothered.

Ⓒ He is interested in others.

Ⓓ He is comfortable with change.



Question 10 ▾ ☆ Zoom Text-to-Speech:

Read the passage. Then answer the questions that follow.

Space Junk

1 Satellites and space stations aren't the only objects orbiting Earth. According to the National Aeronautics and Space Administration (NASA), there are more than 500,000 pieces of debris in orbit. This "space junk" poses a threat to satellites, spacecraft, and the International Space Station. Even a small piece of space junk the size of a marble can damage them. This could put astronauts' lives at risk.

Where Does Space Junk Come From?

2 Some space junk comes from natural objects, such as space rocks. It can also come from human-made items, such as parts of satellites. Human-made junk that circles Earth is more accurately known as orbital debris. This debris includes old satellites, rocket launch stages, or fragments from space equipment that was destroyed upon launch.

Debris Orbiting Earth

3 As more satellites and spacecraft are launched into orbit, the amount of orbital debris will increase. NASA

How does the diagram support the information in the passage?

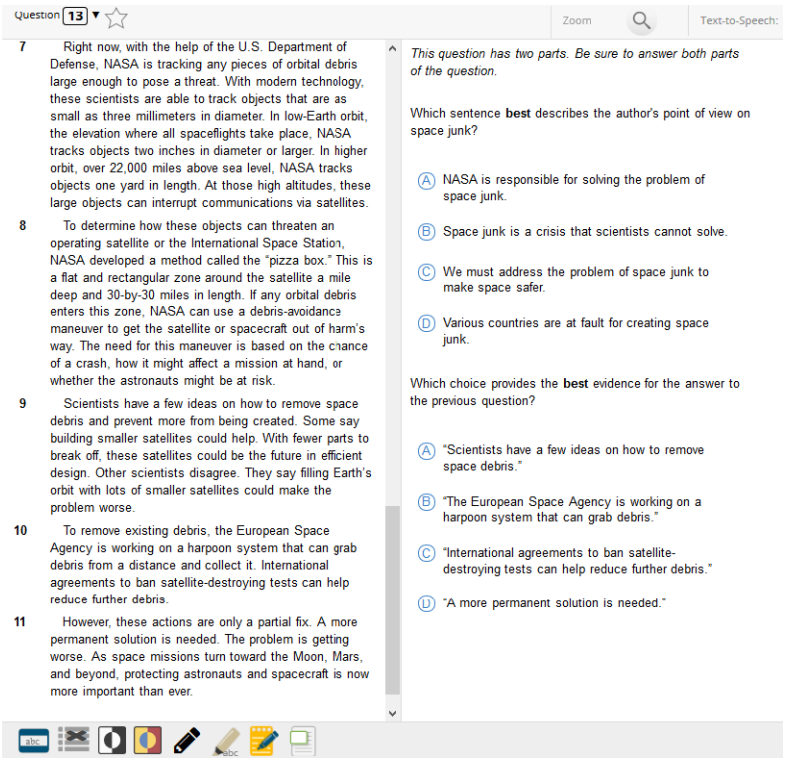
Ⓐ It explains which type of debris is the most dangerous.

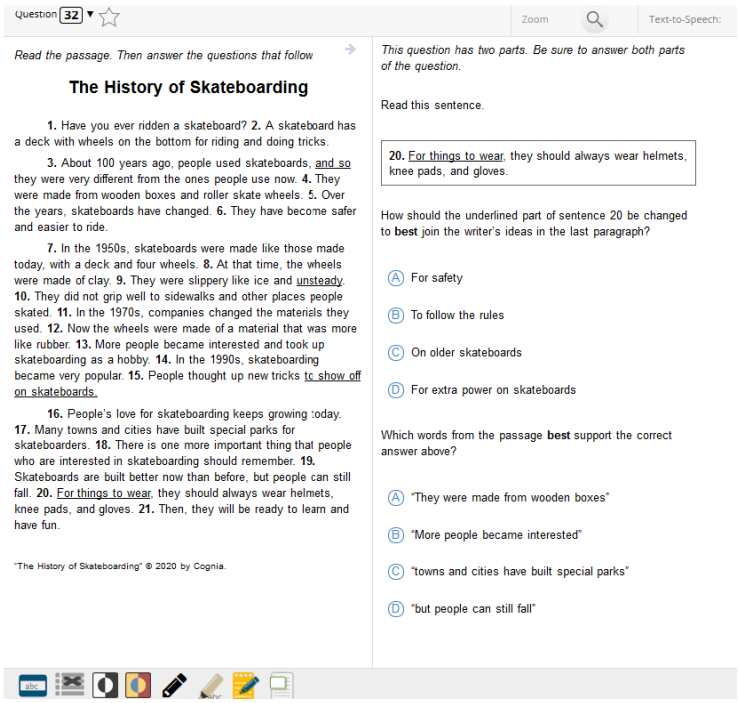
Ⓑ It suggests how long debris remains in orbit around Earth.

Ⓒ It represents the composition of human-made debris orbiting Earth.

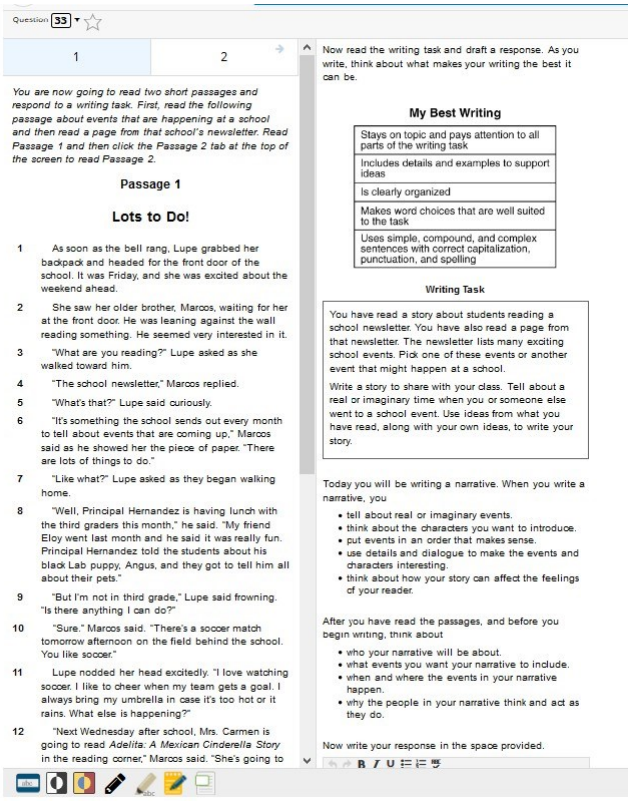
Ⓓ It demonstrates which types of human-made debris are the largest.

- MS-2 items are worth 2 points. In ELA, these items are also known as evidence-based selected response items (EBSR). These items have two parts (Part a and Part b) for students to answer, and 0, 1, or 2 points total can be earned across Part a and Part b. Each part provides choices from which to select.





- Writing prompts (WPs) are worth 7 points. These items require students to write an extended response to a single prompt. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The WPs items are evaluated using a “Production of Writing” rubric on a scale from 1–4 and a “Use of Conventions” rubric on a scale from 1-3.



Test Specifications

New Mexico Measures of Student Success and Achievement (NM-MSSA)

Mathematics

Purpose

- Part of a Balanced Assessment System
- Claims/Score Interpretation and Use Statements

Test Specifications

- Reporting Categories and Test Blueprint
- Cognitive Complexity
- Test Design
- Fairness
- Additional NM-MSSA Considerations

Item Specifications

- Alignment
- Item Types
- Sample Items

Purpose

Part of a Balanced Assessment System

The NM-MSSA is New Mexico's statewide summative assessment for Mathematics and English Language Arts, administered at the end of grades 3-8. As the NM-MSSA is a single measure at the end of a grade band, interpretations and uses of NM-MSSA scores should be supplemented with additional measures, including information from classroom summative and formative assessments in mathematics and English Language Arts as well as interim assessments.
Formative assessment may include the use of the Cognia Formative Assessments, which is a collection of formative assessment materials for grades 3–8 being provided by Cognia during the term of the contract with the state to administer the NM-MSSA. The materials are aligned to the CCSS and therefore to the New Mexico Common Core State Standards.

Claims/Score Interpretation and Use Statements

The NM-MSSA is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the New Mexico Common Core State Standards. In addition to the overall scale score, student performance on three mathematical content subdomains and two mathematical practices subdomains is reported as noted in the reporting categories subsection in the following pages.
See the Score Interpretation and Use (SIU) documents for the complete set of NM-MSSA score interpretation and use statements. These documents can be accessed at https://newmexico.onlinehelp.cognia.org/resources-nm-mssa/.

Test Specifications – Reporting Categories and Test Blueprint

Reporting Categories

The reporting categories for the NM-MSSA are based on the clusters of standards found in the New Mexico *Common Core State Standards* for both content and practices as noted below.

Reporting Categories

Grades 3-5

- Operations and Algebraic Thinking
- Number & Operations in Base 10/Number & Operations -Fractions
- Measurement and Data/Geometry

Grade 6-7

- Ratios & Proportional Relationships
- The Number System/Expressions & Equations
- Geometry/Statistics & Probability

Grade 8

- Functions
- The Number System/Expressions & Equations
- Geometry/Statistics & Probability

Grades 3-8

- Problem Solving (MP1)/Reasoning & Argument (MP 2/3)
- Modeling (MP4)/Structure & Repeated Reasoning (MP 7/8)

Percentages for the distribution of operational (core) test points for each of the reporting categories reflect the distribution in the standards, so as not to over- or underrepresent content. The internal test blueprint has specifications for inclusion on the core forms.

Core Test Blueprints

Specifications for the full test blueprints for the construction of the core forms reflect the reporting category specifications, as well as percentage requirements for each cluster. These constructs represent key aspects of the standards to which items are aligned; as such, the percentage of operational (core) test points for each should be maintained from year to year.

Note that percentages for (a) content clusters and (b) mathematical practices are calculated independently. An individual item may contribute to multiple parts of the blueprint.

Operational Test Blueprint

		Grade 3		Grade 4		Grade 5	
	Mathematics Grade 3, 4, 5	**# of Core Pts**	**% of Core Pts**	**# of Core Pts**	**% of Core Pts**	**# of Core Pts**	**% of Core Pts**
Concepts & Procedures	Operations and Algebraic Thinking	12-18	27-40%	10-16	22-36%	7-11	16-24%
	Number & Operations in Base 10	5-7	11-16%	8-10	17-22%	7-13	16-29%
	Number & Operations-Fractions	8-10	17-22%	10-16	22-36%	11-15	24-33%
	Measurement and Data	11-15	24-33%	6-10	13-22%	10-14	22-31%
	Geometry	3-5	7-11%	3-5	7-11%	4-8	9-18%
	Subtotal	**45**	**100%**	**45**	**100%**	**45**	**100%**
Practices	Problem Solving*	≥ 8	≥ 16%	≥ 8	≥ 16%	≥ 8	≥ 16%
	Reasoning* & Argument						
	Modeling*	≥ 8	≥ 16%	≥ 8	≥ 16%	≥ 8	≥ 16%
	Structure & Repeated Reasoning*						
	Total	**51****		**51****		**51****	

		Grade 6		Grade 7			Grade 8	
	Mathematics Grades 6 & 7	**# of Core Pts**	**% of Core Pts**	**# of Core Pts**	**% of Core Pts**	**Mathematics Grade 8**	**# of Core Pts**	**% of Core Pts**
Concepts & Procedures	Ratios & Proportional Relationships	8-12	17-25%	8-12	17-25%	Functions	10-16	20-33%
	The Number System	8-12	17-25%	6	13%	The Number System	4	8%
	Expressions & Equations	8-12	17-25%	8-16	17-33%	Expressions & Equations	11-17	22-35%
	Geometry	6-10	13-21%	6-10	13-21%	Geometry	10-16	20-33%
	Statistics & Probability	6-10	13-21%	10-12	21-25%	Statistics & Probability	10-12	20-24%
	Subtotal	**48**	**100%**	**48**	**100%**	**Subtotal**	**49**	**100%**
Practices	Problem Solving*	≥ 8	≥ 15%	≥ 8	≥ 15%	Problem Solving*	≥ 8	≥ 15%
	Reasoning* & Argument					Reasoning* & Argument		
	Modeling*	≥ 8	≥ 15%	≥ 8	≥ 15%	Modeling*	≥ 8	≥ 15%
	Structure & Repeated Reasoning*					Structure & Repeated Reasoning*		
	Total	**54****		**54****		**Total**	**55****	

*All or most items dually coded to Concepts and Procedures and Mathematical Practices standards

**Constructed response items are scored for both Concepts & Procedures and Mathematical Practices. A total of six points from the Mathematical Practices rubric contribute to a student's overall score.

Test Specifications – Cognitive Complexity

Percentages for the distribution of operational (core) test points across the cognitive complexity levels (DOK classification) are noted in the table below.

Depth of Knowledge Distribution

	Grade					
DOK	3	4	5	6	7	8
Level 1	5-25%	5-25%	5-25%	5-25%	0-20%	0-30%
Level 2	50-80%	50-80%	50-80%	50-80%	50-80%	50-80%
Level 3	5-30%	5-30%	5-30%	5-30%	5-30%	5-30%
Total	100%	100%	100%	100%	100%	100%

Test Specifications – Test Design

Assessable Standards

The NM-MSSA assesses the New Mexico *Common Core State Standards* as follows:

- Grade 3 test: All standards in grade 3.
- Grade 4 test: All standards in grade 4.
- Grade 5 test: All standards in grade 5.
- Grade 6 test: All standards in grade 6.
- Grade 7 test: All standards in grade 7.
- Grade 8 test: All standards in grade 8.

Test Design - Student Experience

The *Student Testing Experience* tables on the following pages provide a summary of the number of items and points by item type, usage (i.e., operational items or field test items), and estimated testing time for each grade's NM-MSSA assessment.

As shown in the test design tables, the types of items on the NM-MSSA are 1-point machine-scored items (MS-1), 3-point constructed response items (CR-3), and 6-point constructed response items (CR-6). Additional item type descriptions and sample items can be found in the item specifications section on page 16.

Each NM-MSSA test is administered in two sessions. The test form contains both core operational items and matrix field test items. The core operational items are items administered to all students taking that core form, and that count toward student score. Matrix field test items are items administered to subsets of students to "try out" performance (with different students receiving different field test items), and therefore do not count toward student score.

Student Testing Experience (Full Form)

Mathematics Grade 3,4,5	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	33	2	2	37	51	51
Matrix Operational Items	0	0	0	0	0	0
Matrix Field Test Items	5	1		6	8	11
Total Student Experience	**38**	**5**		**43**	**59**	**62**
				Estimated Test Time (min)		**120**

Mathematics Grade 6,7	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	36	2	2	40	54	54
Matrix Operational Items	0	0	0	0	0	0
Matrix Field Test Items	5	1		6	8	11
Total Student Experience	**41**	**5**		**46**	**62**	**65**
				Estimated Test Time (min)		**120**

Mathematics Grade 8	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	37	2	2	41	55	55
Matrix Operational Items	0	0	0	0	0	0
Matrix Field Test Items	5	1		6	8	11
Total Student Experience	**42**	**5**		**47**	**63**	**66**
				Estimated Test Time (min)		**120**

Test Specifications – Fairness

Fairness is defined as the extent to which the test scores are valid for different groups of test takers. Consideration of universal design, bias, and sensitivity guidelines support the construction of fair, valid assessments.

Universal Design for Assessments

The concept of Universal Design for Assessments focuses on developing content and assessments that reach the widest population of students possible. Stimuli and items on the NM-MSSA are designed to simply and clearly present tasks and to provide maximum readability, comprehensibility, and legibility. The seven elements of Universal Design for Assessments are based on the original UDL guiding principles:

Universal Design for Assessments

Principle	Explanation
Inclusive Assessment Population	Tests designed for state, district, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field-testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct-irrelevant cognitive, sensory, emotional, and physical barriers are removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	Test design facilitates the use of needed accommodations (e.g., all items can be brailled).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and presented in understandable language.
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words kept to a minimum) for readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, tables, figures, and illustrations, and to response formats.

Bias

The concept of Bias is defined as the presence of some characteristic of an item that results in differential performance for two individuals of the same ability but from different ethnic, sex, cultural, or religious groups.

Bias can occur whenever content offends or disadvantages a student or group of students due to gender, race, regional background, socioeconomic status, or any other such classification.

Test developers take care to craft content in a way that does not misrepresent specific groups or rest on assumptions made about specific groups, that in turn could negatively impact how students interpret content.

- Stimulus and item content on the NM-MSSA must not present stereotypes or unfair representations of gender, race, ethnicity, disability, culture, or religion.
- Stimulus and item content on the NM-MSSA should not depend on overly-experiential information such as knowledge of technology, consumer goods, pop culture, geographic locations, or sports and extracurricular activities. While these topics are not completely excluded from use, care must be taken to ensure that the items are presented in a way that does not require a level of knowledge that would not be held by all students.

Sensitivity

Sensitivity refers to the presence of content that is contrary to the acceptable norms of the students, educators, parents, or other members of the community that may interact with the assessment. Sensitive subject matter can impact student performance or attitudes toward testing, and hence, their test scores.

Consideration of bias and sensitivity issues is very important when developing content for an assessment. Test developers must ensure that stimuli and items are free of content that will negatively affect a student's performance not because of what the student knows and can do but because the content evokes an emotional response from that student (or is in some other way distracting to the student).

Subjects/contexts that are likely to prompt emotional distress on the part of students cannot be used on the NM-MSSA (e.g., war, violence, human death or debilitating disease, animal-based medical research).

Careful judgment should be applied to standards that cover topics that may be considered controversial by some groups (e.g., evolution examples, population dynamics including death/extinction, environmental impact). Those standards represent content knowledge to be assessed, but the assessment must be done in a sensitive, unbiased way.

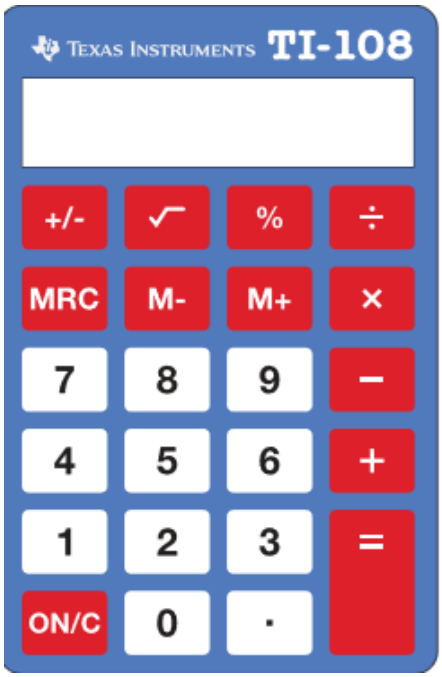
Test Specifications – Additional NM-MSSA Considerations

Calculator Use

Students in grades 3–6 who are taking the paper-pencil test can use their own four-function calculator with a square root key during Session 2. Students in grades 7–8 who are taking the paper-pencil test can use their own scientific calculator during Session 2. The memory on any hand-held calculator must be cleared before the test begins.

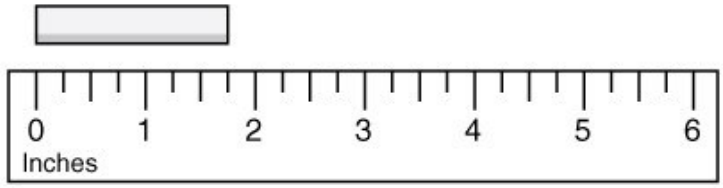
Students may not use their own calculators for the online test unless it is an approved accommodation. The images below show the calculators that are available for use in the online testing platform.

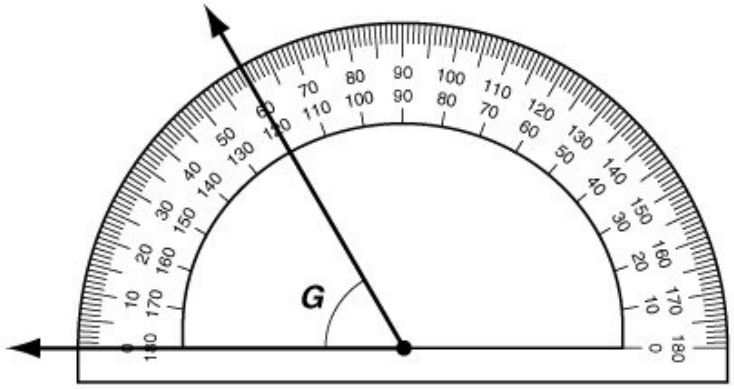
NM-MSSA Calculator Use in Mathematics		
Grade	**Session 1**	**Session 2**
3	None	Basic
4	None	Basic
5	None	Basic
6	None	Basic
7	None	Scientific
8	None	Scientific

Basic	Scientific
	

Rulers and Protractors

For Spring 2021, a ruler or protractor will be embedded within a graphic for items that require students to measure lengths of objects or angles. Below are some examples of the embedded tools that could be found within the items.





Reference Sheets

Reference sheets are not included. Depending on the targeted rigor and complexity of an item, a formula may be embedded within an item.

Practice Test

A full-length practice test mirroring the operational test design is available beginning in the 2020-2021 school year. The practice tests and supporting materials can be accessed at https://newmexico.onlinehelp.cognia.org/practice-tests-nm-mssa/.

Item Specifications

Alignment

The items on the NM-MSSA are aligned to the New Mexico Common Core State Standards.
Each item is aligned to one of the content standards in NMCCSS. Also, the majority of the items are also aligned to a mathematical practices standard in the NMCCSS.

Item Types

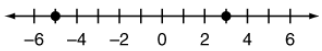
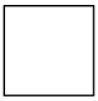
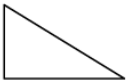
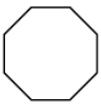
The types of items on the NM-MSSA are 1-point machine-scored items (MS-1), 3-point constructed response items (CR-3), and 4-point constructed response items (CR-4):

- MS-1 items are worth 1 point. These items may be multiple choice or multiple select.

- Constructed response items (CRs) are worth either 3 or 6 points. These items require students to write an extended response to a prompt. The prompt may be a single prompt, or more typically, the items are written with multiple, scaffolded parts for students to respond to. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The CR-3 items are evaluated using a content rubric on a scale from 0–2 and a practices rubric on a scale from 0-1. The CR-6 items are evaluated using a content rubric on a scale from 0–4 and a practices rubric on a scale from 0-2.

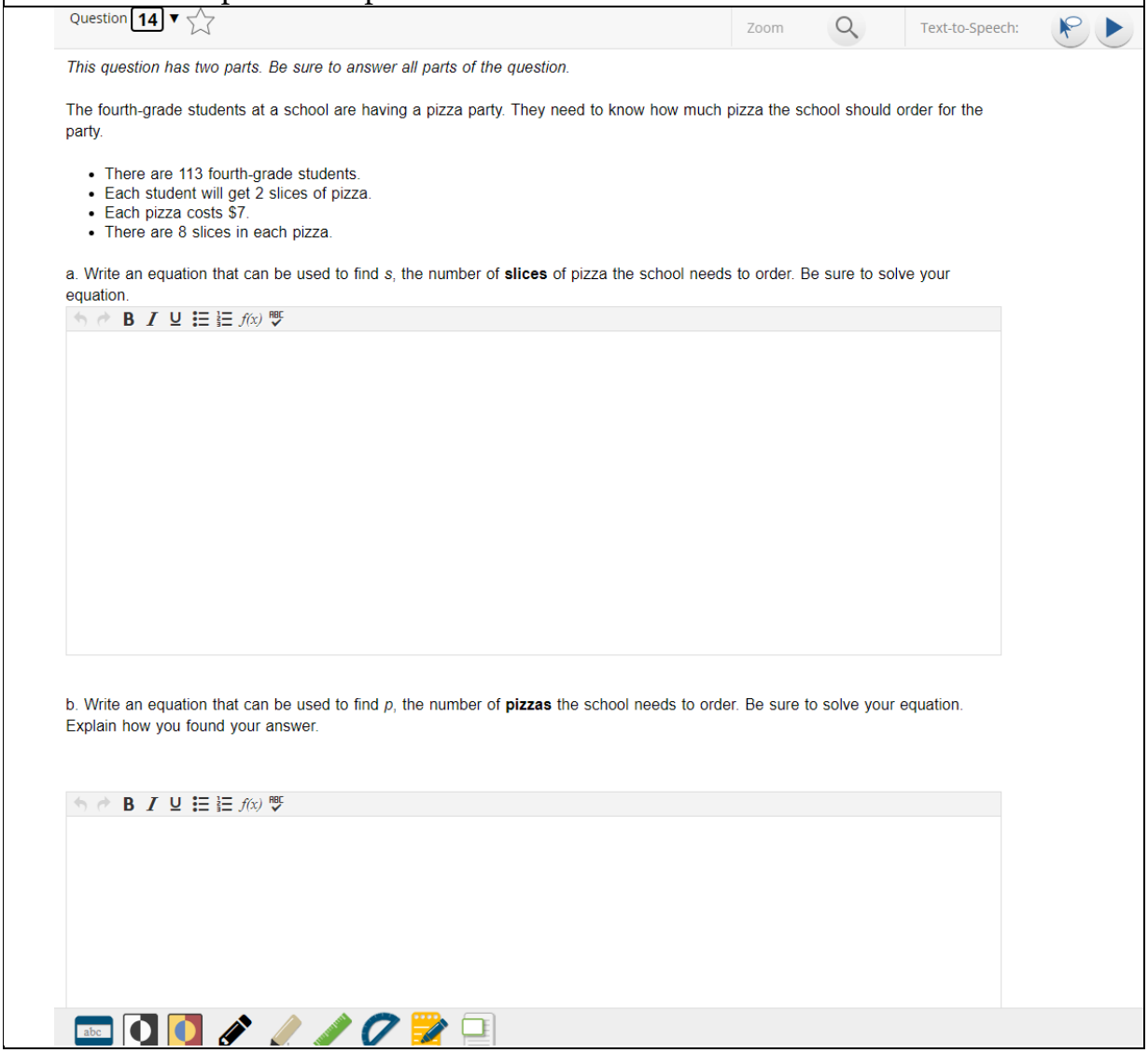
Sample Items

- MS-1 items are worth 1 point. These items may be multiple choice or multiple select. For multiple select items in grades 3-5, the number of choices to select will be given in the item stem. For multiple select items in grades 6-8, the item stem will direct students to “select all that apply.”

Multiple Choice	Multiple Select
Question 8 ▼ ☆ Two points are shown on this number line.  Which expression represents the distance between the two points? Ⓐ $\lvert -5\rvert - \lvert 3\rvert$ Ⓑ $\lvert 3\rvert - \lvert -5\rvert$ Ⓒ $\lvert -5 - 3\rvert$ Ⓓ $\lvert 3 - 5\rvert$	Question 24 ▼ ☆ Select the **two** shapes that have the same number of lines of symmetry. Ⓐ  Ⓑ  Ⓒ  Ⓓ  Ⓔ 

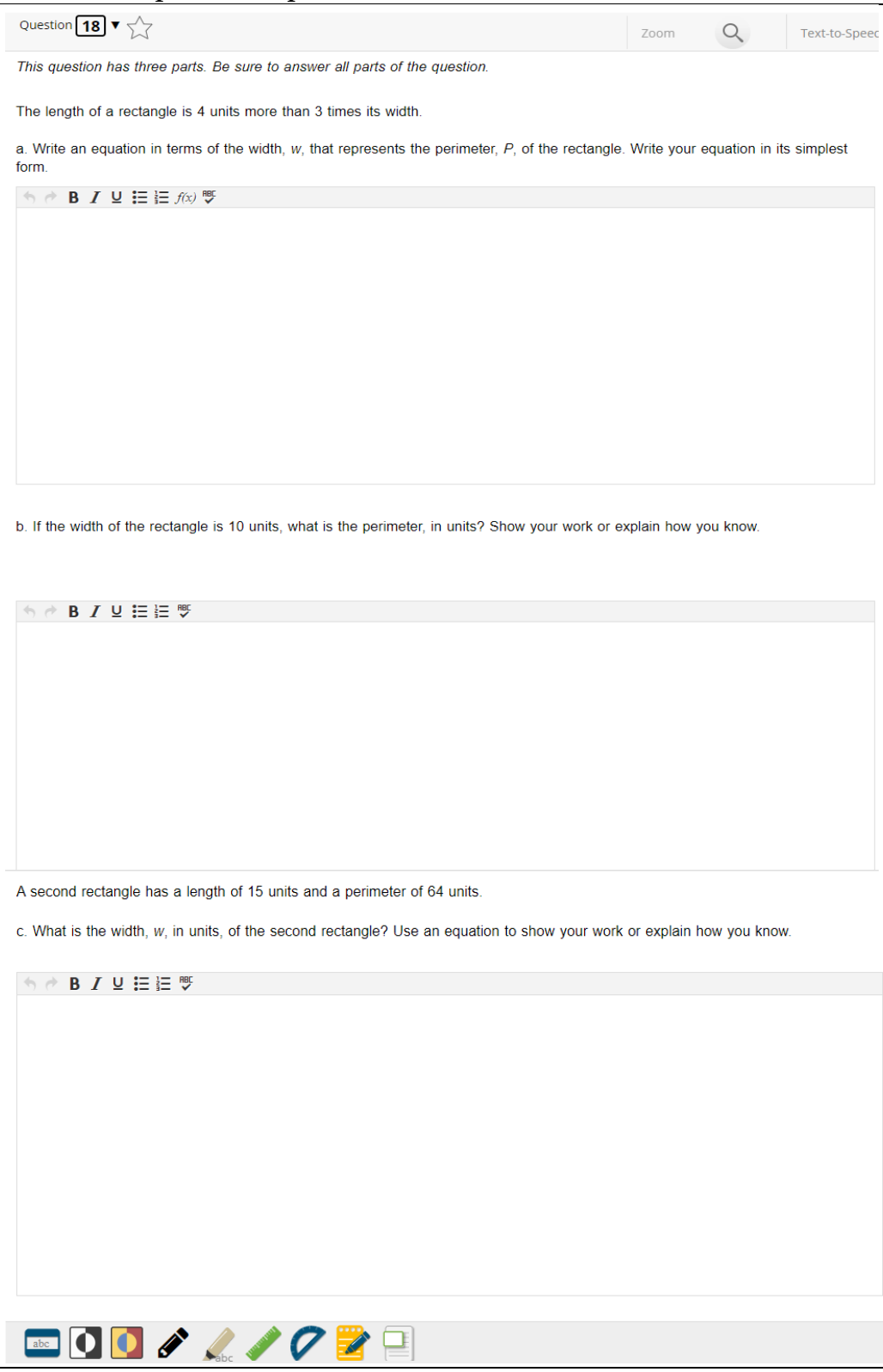
- CR-3 items are worth a total of 3 points. These items require students to write an extended response to a prompt. The prompt may be a single prompt, or more typically, the items are written with multiple, scaffolded parts for students to respond to. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The CR-3 items are evaluated using a content rubric on a scale from 0–2 and a practices rubric on a scale from 0-1.

Constructed Response Sample



- CR-6 items are worth a total of 6 points. These items require students to write an extended response to a prompt. The prompt may be a single prompt, or more typically, the items are written with multiple, scaffolded parts for students to respond to. These items are hand-scored, with scorers using a multi-trait rubric and scoring notes to evaluate responses. The CR-6 items are evaluated using a content rubric on a scale from 0–4 and a practices rubric on a scale from 0-2.

Constructed Response Sample



Appendix A – Mathematical Practices

Cognia Assessments Mathematics

Mathematical Practices, Focus statements, and clarifying bullets

1 Make sense of problems and persevere in solving them.

Focus 1A: Entry into a problem

Determines the form (quantity or measure) of the solution to the mathematical or real-life problem, analyzes the givens to extract only the pertinent information needed to solve the mathematical or real-life problem, and analyzes the givens to identify missing information needed to solve the mathematical or real-life problem.

- Explain the meaning of a problem
- Look for entry points to its solution
- Analyze givens, constraints, relationships, and goals for extraneous or missing information
- Make conjectures about the form and meaning of the solution

Focus 1B: Solution path

Determines and uses an appropriate solution path including the identification and appropriate use of tools to solve a well-posed mathematical or real-life problem.

- Plan a solution pathway rather than simply jumping into a solution attempt
- Consider analogous problems
- Try special cases and simpler forms of the original problem in order to gain insight into its solution
- Monitor and evaluate progress and change course if necessary
- Transform representations to get the information they need
- Understand the approaches of others to solving complex problems and identify correspondences between different approaches

Focus 1C: Appropriate solution to a problem

Determines whether the solution to a well-posed mathematical or real-life problem is accurate and/or realistic.

- Check solutions to problems using a different method
- Ensure the solution makes sense
- Verify the necessary precision is used in the solution
- Analyze the problem to ensure the proper units are specified in the solution

2 Reason abstractly and quantitatively.

Focus 2A: Decontextualize and manipulate

Abstracts a given situation, represents it symbolically, and manipulates the symbols.

- Decontextualize—
  - To abstract a given situation

- To represent it symbolically
- To manipulate the representing symbols as if they have a life of their own, without necessarily attending to their referents
- Create a coherent representation of the problem at hand

Focus 2B: Contextualize

Shows understanding of the referents for the symbols involved in a mathematical sentence representing a situation.

- Contextualize—
  - To pause as needed during the manipulation process in order to probe into the referents for the symbols involved
- Consider the units involved

Focus 2C: Quantitative reasoning

Knows and uses different properties of operations and objects and shows understanding of the meaning of the quantities.

- Make sense of quantities and their relationships in problem situations
- Attend to the meaning of quantities, not just how to compute them
- Know and flexibly use different properties of operations and objects

3 Construct viable arguments and critique the reasoning of others.

Focus 3A: Construct arguments

Makes and defends arguments.

- Understand and use stated assumptions, definitions, and previously established results in constructing arguments
- Communicate arguments properly to others
- Reason inductively about data, making plausible arguments that take into account the context from which the data arose
- Elementary students can construct arguments using concrete referents such as objects, drawings, diagrams, and actions whereas middle school and high school students learn to determine domains to which an argument applies

Focus 3B: Evaluate arguments

Evaluates arguments.

- Respond to the arguments of others
- Compare the effectiveness of two plausible arguments
- Distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is
- Listen to or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments

4 Model with mathematics.

Focus 4A: Create models.

Creates an appropriate model.

- Identify important quantities in a practical situation and map their relationships using such tools as diagrams, two-way tables, graphs, flowcharts, and formulas

Focus 4B: Interpret and analyze models.

Interprets and analyzes models.

- Analyze relationships mathematically to draw conclusions
- Improve models if they have not served its purpose
- Interpret their mathematical results in the context of the situation
- Explain correspondences between equations, verbal descriptions, tables, and graphs

Focus 4C: Use the model to solve problems.

Uses a model to solve a problem.

- Apply the mathematics they know to solve problems arising in everyday life, society, and the workplace
- Use models to develop designs, predict outcomes, describe phenomena, solve problems, and explain causes and effects
- Make assumptions and approximations to simplify a complicated situation
- Reflect on whether the results make sense
- Draw diagrams of important features and relationships, graph data, and search for regularity or trends
- Rely on using models to help conceptualize and solve a problem

7 Look for and make use of structure.

Focus 7/8A: Structure

Look for and make use of structure.

- Use the structure of an expression to rewrite it in another form
- Step back for an overview and shift perspective
- See complicated things as single objects or as being composed of several objects
- Use the geometric attributes of figures to classify/sort
- Use mathematical properties of numbers, operations, and equality to explain and analyze mathematical or real-world problems
- Recognize the elements of effective data representation for a data set
- Use familiar/known structures to see something in a different way
- Use structure of numbers and shapes to identify relationships between them

8 Look for and express regularity in repeated reasoning.

Focus 7/8B: Repeated Reasoning

- Look for and express regularity in repeated reasoning.
- Discern a pattern (e.g. the less you subtract, the greater the difference)
- Notice if calculations result in repeated values (e.g., notice when there is a repeating decimal)
- Look both for general methods and for shortcuts
- Abstract general principles from repeated phenomena (e.g., slope, formulas for area or perimeter, correlation)

Test Specifications

New Mexico Assessment of Science Readiness (NM-ASR)

Contents

Purpose of the NM-ASR

Part of a Balanced Assessment System

The NM-ASR is New Mexico's statewide summative assessment for science, administered at the end of grades 5, 8, and 11. As the NM-ASR is a single measure at the end of a grade band, interpretations and uses of NM-ASR scores should be supplemented with additional measures, including information from classroom summative and formative assessments in science.
Formative assessment may include the use of STEM Gauge, which is a collection of formative assessment materials for grades K–8 being provided by Cognia during the term of their contract with the state to administer the NM-ASR. The materials are aligned to the NGSS and therefore to the *New Mexico STEM Ready! Science Standards*. The materials for STEM Gauge may be accessed at the following site: http://go.cognia.org/instructional-support-materials-for-new-mexico-science-educators.
Claims/Score Interpretation and Use Statements
The NM-ASR is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the *New Mexico STEM Ready! Science Standards*. The standards require integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts to explain phenomena and solve problems.

In addition to overall scale score, student performance on three science subdomains is reported:

- Practices and Crosscutting Concepts in Physical Sciences
- Practices and Crosscutting Concepts in Life Sciences
- Practices and Crosscutting Concepts in Earth and Space Sciences

Test Specifications – Test Design

Assessable Standards

The NM-ASR assesses the New Mexico STEM Ready! Science Standards as follows:

- Grade 5 test: All standards in grades 3, 4, and 5, *except* 5-SS-1 NM.
- Grade 8 test: All standards in the middle school grade band (6–8), **including** MS-ESS3-3 NM.
- Grade 11 test: All standards in the high school grade band (9–12), *except* HS-LS2-7 NM and HS-SS-1 NM (but **including** HS-SS-2 NM).

Test Design

The NM-ASR test is administered in three sessions. The test is administered online as a computer-based test (CBT).

Online accommodations are available for the CBT. Paper, large-print, and Braille test forms, as well as computer- and print-based Spanish test forms, are also provided.

No calculator is provided for the NM-ASR, as no items require calculator use. A periodic table will be provided as a reference for high school (grade 11).

The NM-ASR contains both machine-scored (MS) items and hand-scored open-ended (OE) items. Many of the items are organized in item clusters (CL), and there are some standalone items, both 2-point machine-scored items (MS-2) and 4-point open-ended items (OE). Additional item type descriptions and sample items can be found in the item specifications section on page 14.

Both core operational items (which count for a student's score) and matrix field test items (which are try-out items that do not count for a student's score) are included on the NM-ASR test.

The total number of test items, points, and estimated testing time for the NM-ASR are shown in the following tables.

How to read the student testing experience tables:

As a reminder,

- MS-1 items are worth 1 point.
- MS-2 items are worth 2 points.
- OE items are worth 4 points.

An example breakdown of items and points is shown for the first row of the grade 5 table.

Grade 5	Cluster/Passage Items			Standalone Items		Total Number of Items	Total Number of Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	**6 psgs** x 0 points = *0 points*	**12 items** x 1 point = *12 points*	**12 items** x 2 points = *24 points*	**8 items** x 2 points = *16 points*	**3 items** x 4 points = *12 points*	**12 + 12 + 8 + 3 = 35 items**	12 + 24 + 16 + 12 = *64 points*

Table 1: Grade 5 NM-ASR Student Testing Experience

Grade 5	Cluster/Passage Items			Standalone Items		Total Items	Total Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	8	3	35	64
Matrix Field Test Items	2	4	4	4	1	13	24
Total Student Experience	8	16	16	12	4	48	88
					Estimated Testing Time (min)		150

Table 2: Grade 8 NM-ASR Student Testing Experience

Grade 8	Cluster/Passage Items			Standalone Items		Total Items	Total Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	8	3	35	64
Matrix Field Test Items	2	4	4	4	1	13	24
Total Student Experience	8	16	16	12	4	48	88
					Estimated Testing Time (min)		150

Table 3: Grade 11 NM-ASR Student Testing Experience

Grade 11	Cluster/Passage Items			Standalone Items		Total Items	Total Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	10	3	37	68
Matrix Field Test Items	2	4	4	5	1	14	26
Total Student Experience	8	16	16	15	4	51	94
					Estimated Testing Time (min)		165

Practice Test

Full-length practice tests mirroring the operational test design and supporting materials can be accessed at https://newmexico.onlinehelp.cognia.org/practice-tests-nm-asr/.

Test Specifications – Reporting Categories

The reporting categories for NM-ASR are based on the three content domains. Percentages for the distribution of operational (core) test points for each of the reporting categories reflect the distribution in the standards, so as not to over- or underrepresent content.

Based on this representativeness, the fourth content domain of Engineering, Technology, and Applications of Science as well as the NM-specific content domain of Science and Society are not reported as a subscore (as there are very few standards out of the total in each grade band). Items coded to these standards **do** count toward total test score.

Table 4: Grade 5 NM-ASR Reporting Categories

Reporting Category	Typical Number of Clusters	Typical Number of Standalone MS-2	Typical Number of Standalone OE	Number of Core Points	Percent of Core Points (+/-4%)
Practices and Crosscutting Concepts in Physical Sciences	2	4–6	1	24–28	40%
Practices and Crosscutting Concepts in Life Sciences	2	1–3	1	18–22	30%
Practices and Crosscutting Concepts in Earth and Space Sciences	2	1–3	1	18–22	30%

Table 5: Grade 8 NM-ASR Reporting Categories

Reporting Category	Typical Number of Clusters	Typical Number of Standalone MS-2	Typical Number of Standalone OE	Number of Core Points	Percent of Core Points (+/-4%)
Practices and Crosscutting Concepts in Physical Sciences	2	2–4	1	20–24	35%
Practices and Crosscutting Concepts in Life Sciences	2	2–4	1	20–24	35%
Practices and Crosscutting Concepts in Earth and Space Sciences	2	1–3	1	18–22	30%

Table 6: Grade 11 NM-ASR Reporting Categories

Reporting Category	Typical Number of Clusters	Typical Number of Standalone MS-2	Typical Number of Standalone OE	Number of Core Points	Percent of Core Points (+/-4%)
Practices and Crosscutting Concepts in Physical Sciences	2	3–5	1	22–26	35%
Practices and Crosscutting Concepts in Life Sciences	2	3–5	1	22–26	35%
Practices and Crosscutting Concepts in Earth and Space Sciences	2	1–3	1	18–22	30%

Test Specifications – Cognitive Complexity

Because the *New Mexico STEM Ready! Science Standards* are NGSS-aligned, the cognitive complexity of items on the NM-ASR is evaluated with a different framework than Depth of Knowledge.

For the items on the NM-ASR, four indicators are used to classify the cognitive complexity of each item: stimulus, science and engineering practice, disciplinary core idea, and crosscutting concept. For each indicator, the classification in terms of high, medium, or low complexity is based on how the students are

using the indicator to respond to the item—specifically, to what degree does students' engagement with the indicator contribute to the level of sense-making required by the item.

On the NM-ASR, after summing the operational (core) test points at each cognitive complexity level across all four indicators, **at least 10% of the points should be high cognitive complexity** and **no more than 35% of the points should be low cognitive complexity**.

The descriptors for each indicator at the three complexity levels (high, medium, low) are presented in the following tables.

Table 7: Descriptors for Stimulus at Three Complexity Levels

LEVEL	STIMULUS
High	- Phenomenon is novel, complex, and/or unfamiliar to students - Students must synthesize multiple pieces of information and do a significant amount of "figuring out" to make sense of the phenomenon
Medium	- Phenomenon is somewhat novel, but may be analogous to what many students are familiar with - Students must use multiple pieces of information and do an intermediate amount of "figuring out" to make sense of the phenomenon
Low	- Phenomenon is familiar and/or more straightforward for students - Students only need to use simple/straightforward information, and/or a single piece of information, and do a minimal amount of "figuring out" to answer the question or contribute to making sense of the phenomenon

Table 8: Descriptors for Science and Engineering Practice (SEP) at Three Complexity Levels

LEVEL	SCIENCE AND ENGINEERING PRACTICE (SEP)
High	- Students must apply the SEP, or multiple SEPs, in a sophisticated way to make sense of the phenomenon (e.g., synthesis to perform more connections, steps, combination of SEP elements, such as having to combine data, produce a new graph or model as evidence, etc.) - Often little to no scaffolding that helps students apply the SEP
Medium	- Students must apply the SEP to make sense of the phenomenon - Typically, some scaffolding that helps students apply the SEP
Low	- Students only need to use the SEP in a simple, mechanical way to answer the question or contribute to making sense of the phenomenon - Often a large amount of scaffolding that helps students apply the SEP

Table 9: Descriptors for Disciplinary Core Idea (DCI) at Three Complexity Levels

LEVEL	DISCIPLINARY CORE IDEA (DCI)
High	- Students must apply and connect DCIs in a sophisticated way to make sense of the phenomenon, i.e., • application of science ideas (often multiple, grade-band appropriate ideas) in unique ways or new combinations • knowledge transfer to construct new understanding, make sense of novel phenomena - Often little to no scaffolding that helps students apply the DCI
Medium	- Students must apply or reason with the DCI(s) to make sense of the phenomenon - Typically, some scaffolding that helps students apply the DCI
Low	- Students use the DCI in a simple, straightforward way (i.e., little to no application or reasoning) to answer the question or contribute to making sense of the phenomenon - Often a large amount of scaffolding that helps students apply the DCI

Table 10: Descriptors for Crosscutting Concept (CCC) at Three Complexity Levels

LEVEL	CROSSCUTTING CONCEPT (CCC)
High	- Students must apply the CCC in an in-depth way to expand thinking and make non-typical connections to make sense of the phenomenon
Medium	- Students must use the CCC as specified by the CCC sub-bullet detail to make sense of the phenomenon
Low	- Students only use the CCC in a general way to answer the question or contribute to making sense of the phenomenon

Test Specifications – Fairness

Fairness is defined as the extent to which the test scores are valid for different groups of test takers. Consideration of universal design, bias, and sensitivity guidelines support the construction of fair, valid assessments.

Universal Design for Assessments

The concept of Universal Design for Assessments focuses on developing content and assessments that reach the widest population of students possible. Stimuli and items on the NM-ASR are designed to simply and clearly present tasks and to provide maximum readability, comprehensibility, and legibility. The seven elements of Universal Design for Assessments are based on the original UDL guiding principles:

Table 11: Elements of Universal Design for Assessments

Universal Design for Assessments	
Principle	**Explanation**
Inclusive Assessment Population	Tests designed for state, district, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field-testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct-irrelevant cognitive, sensory, emotional, and physical barriers are removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	Test design facilitates the use of needed accommodations (e.g., all items can be brailled).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and presented in understandable language.
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words kept to a minimum) for readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, tables, figures, and illustrations, and to response formats.

Bias

The concept of bias is defined as the presence of some characteristic of an item that results in differential performance for two individuals of the same ability but from different ethnic, sex, cultural, or religious groups.

Bias can occur whenever content offends or disadvantages a student or group of students due to gender, race, regional background, socioeconomic status, or any other such classification.

Test developers take care to craft content in a way that does not misrepresent specific groups or rest on assumptions made about specific groups that in turn could negatively impact how students interpret content.

- Stimulus and item content on the NM-ASR must not present stereotypes or unfair representations of gender, race, ethnicity, disability, culture, or religion.
- Stimulus and item content on the NM-ASR should not depend on overly experiential information such as knowledge of technology, consumer goods, pop culture, geographic locations, or sports and extracurricular activities. While these topics are not completely excluded from use, care must

be taken to ensure that the items are presented in a way that does not require a level of knowledge that would not be held by all students.

Sensitivity

Sensitivity refers to the presence of content that is contrary to the acceptable norms of the students, educators, parents, or other members of the community that may interact with the assessment. Sensitive subject matter can impact students' performance or attitudes toward testing, and hence, their test scores.

Consideration of bias and sensitivity issues is very important when developing content for an assessment. Test developers must ensure that stimuli and items are free of content that will negatively affect a student's performance not because of what the student knows and can do but because the content evokes an emotional response from that student (or is in some other way distracting to the student).
Subjects/contexts that are likely to prompt emotional distress on the part of students cannot be used on the NM-ASR (e.g., war, violence, human death or debilitating disease, animal-based medical research). Careful judgment should be applied to standards/performance expectations (PEs) that cover topics that may be considered controversial by some groups (e.g., evolution examples, population dynamics including death/extinction, environmental impact). Those PEs represent content knowledge to be assessed, but the assessment must be done in a sensitive, unbiased way.

Stimulus Specifications

All items for the NM-ASR have a stimulus. For clusters, all items in the set are associated with a common stimulus that presents a science phenomenon or engineering design problem. For standalone items (MS-2, OE), the item includes a lead stimulus that provides a specific science phenomenon or engineering design problem, or context thereof. Phenomenon refers to something observable that happens in the real world, whether natural or man-made. Engineering design problem refers to a personal or societal need or want.

Specifications for Cluster Stimuli

1. The stimulus must present a single, rich science phenomenon or engineering design problem aligned to the standards/performance expectations (PEs).
2. The stimulus may present any variety of elements to provide the necessary information to support sense-making (via the items) around the phenomenon or problem: text paragraphs, passages, graphs, data tables, models, drawings, etc.
3. The stimulus must be rich enough to support the development of enough items for the cluster, in the context of a storyline (sequence of sense-making) around the phenomenon or problem using the DCIs, SEPs, and CCCs of the targeted PEs.
4. All information in the stimulus should be necessary, but not conceptually sufficient, for students to respond (i.e., students must also use their own knowledge of the constructs in the PE(s) to answer the items, rather than simply identify given information).
5. The stimulus phenomenon or problem must be grade-appropriate, engaging, and relevant for students at that grade level.
6. The stimulus should adhere to the specifications in the following table regarding length, wording, and complexity. *

Table 12: Stimulus Specifications by Grade Band

Stimulus Characteristic	Elementary School (Grades 3–5)	Middle School (Grades 6–8)	High School (Grades 9–12)
Text Word Count*	100–300 words	100–400 words	100–400 words
Vocabulary Level *(excluding science content vocabulary)*	Grade 3	Grade 5 maximum	Grade 8 maximum
Readability/Lexile Maximum	820L (Gr 3)	1010L (Gr 5)	1185L (Gr 8)
Qualitative Text Characteristics	Simple sentence structures, clear/uncomplicated graphics, lower vocabulary demands, use of only essential science vocabulary	Slight mix of simple and more complex phrasing and sentence structure, average to moderately complex graphics, average vocabulary demands	Mix of simple and more complex phrasing and sentence structure, average to moderately complex graphics, average vocabulary demands

**Count should balance text and graphic load—in a stimulus with more and/or complex graphics, the word count should be lower; in a stimulus with few and/or very simple graphics, the word count could, if needed, be at the higher end of range.*

Items aligned to the NM-Specific Standards may sometimes exceed these specifications, especially word count, because of the detailed NM-specific contexts that must be provided.

Specifications for Standalone Item Stimuli

1. **MS-2 items:** The stimulus must present a hook or driving reason for the question being asked, and it must set a phenomenon- or problem-based context, aligned to the PE, for the item. The stimulus will typically not be as extensive as a stimulus for an item cluster.
2. **OE items:** The stimulus must present a hook or driving reason for the question being asked, and it must include a phenomena or problem, aligned to the PE, to drive the item. The stimulus for open-ended items will typically be more concise than for item clusters but more detailed than for MS-2 standalone items.

Item Specifications

Alignment

The items on the NM-ASR are aligned to the *New Mexico STEM Ready! Science Standards*, including both the NGSS and the NM-Specific Standards. Each item is aligned to a performance expectation (PE) as well as dimensions of the PE. All items must have either 2-dimensional or 3-dimensional alignment.

Item Types

The types of items on the NM-ASR are item clusters, 2-point machine-scored standalone items (MS-2), and 4-point open-ended standalone items (OE):

- An item cluster is a set of items all associated with a common stimulus. Clusters contain four items. These items may be multiple choice, multiple select, or technology-enhanced, with two of the items being worth 1 point and two of the items being worth 2 points. The clusters typically align to two PEs, and all clusters measure all three dimensions of the PEs being assessed.
- Standalone MS-2 items are worth 2 points. These items have two parts (Part a and Part b) for students to answer, and 0, 1, or 2 points total can be earned across Part a and Part b. These items may be multiple choice, multiple select, or technology-enhanced (e.g., drag-and-drop, hot spot, drop-down selections).
- Open-ended items are worth 4 points. These items require students to write an extended response to a prompt. The prompt may be a single prompt, or more typically, the items are written with multiple, scaffolded parts for students to respond to. These items are hand-scored, with scorers using a rubric and scoring notes to evaluate responses on a scale from 0–4.

Samples of each of these item types are included on the following pages.

Table 13: Cluster Item Type Sample

Clusters	A set of 4 items all associated with an introductory passage, or "stimulus." The entire cluster is worth 6 points. See Figure 4 to see the structure of the cluster.
Stimulus	A stimulus typically contains both text and graphics such as diagrams, tables, or graphs. An example stimulus from the NM-ASR grade 5 practice test is shown in ***Figure 1*** The items associated with the cluster assess two Physical Sciences PEs: • 5-PS1-3: Make observations and measurements to identify materials based on their properties. • SEP: Planning and Carrying Out Investigations • DCI: PS1.A: Structure and Properties of Matter • CCC: Scale, Proportion, and Quantity • 5-PS1-4: Conduct an investigation to determine whether the mixing of two or more substances results in new substances. • SEP: Planning and Carrying Out Investigations • DCI: PS1.B: Chemical Reactions • CCC: Cause and Effect

Machine-Scored 1-Point Item (MS-1)	Two of the items in the cluster are machine-scored items worth 1 point each.
	• These items may be:
	  ○ multiple-choice,
	  ○ multi-select, or
	  ○ technology-enhanced items (e.g., drag-and-drop, hot spot, drop-down selections).
	• See ***Figure 2*** to view sample.

Machine-Scored 2-Point Item (MS-2)	The other two items in the cluster are machine-scored items worth 2 points each.
	• These items have two parts, with
	  ○ Part a worth 1 point and
	  ○ Part b also worth 1 point.
	• Each part of the item may be presented as
	  ○ multiple-choice,
	  ○ multi-select, or
	  ○ technology-enhanced (e.g., drag-and-drop, hot spot, drop-down selections).
	See ***Figure 3*** to view sample.

Figure 1: Cluster Stimulus Sample – Grade 5 NM-ASR Practice Test

Read the information. Then answer the questions that follow.

Investigating Gas Production

In class, a teacher demonstrates a chemical reaction by mixing vinegar and baking soda to produce bubbles of gas. Eliana wonders whether mixing other substances could also produce a gas. She decides to test different combinations of sugar, water, vinegar, and baking soda.

Some properties of these substances are shown in the table.

Properties of Substances

Substance	Color	Solid or Liquid	Attracted to a Magnet	Conducts Electricity
Sugar	White	Solid	No	No
Water	Clear	Liquid	No	Yes
Vinegar	Clear	Liquid	No	Yes
Baking soda	White	Solid	No	No

Investigation 1
Eliana mixes a small amount of each liquid and solid in a bowl and observes whether bubbles of gas are produced. Her observations are shown in the table.

Investigation 1 Observations

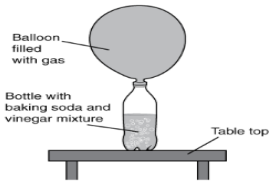
Liquid Used	Solid Used	Gas Produced
Water	Sugar	No
Water	Baking soda	No
Vinegar	Sugar	No
Vinegar	Baking soda	Yes

Investigation 2
Next, Eliana wonders whether changing the amount of baking soda would change the amount of gas produced. To investigate, she follows these steps:

1. Record the mass of a balloon.
2. Pour 50 milliliters of vinegar into a bottle.
3. Put 5 milliliters of baking soda inside the balloon. Hold the balloon so that the baking soda stays inside the balloon and attach the open end of the balloon to the top of the bottle.
4. Lift the balloon so that the baking soda falls into the bottle with vinegar.
5. Wait one minute.
6. Carefully remove the balloon from the bottle without allowing any gas to escape.
7. Measure the mass of the balloon filled with gas.
8. Calculate the mass of gas produced by subtracting the mass of the balloon from the mass of the balloon filled with gas.
9. Repeat steps 1–8 until three trials have been completed.
10. Repeat steps 1–9 with 10 milliliters and 15 milliliters of baking soda.

The results of one trial are shown in the diagram.

Gas in a Balloon

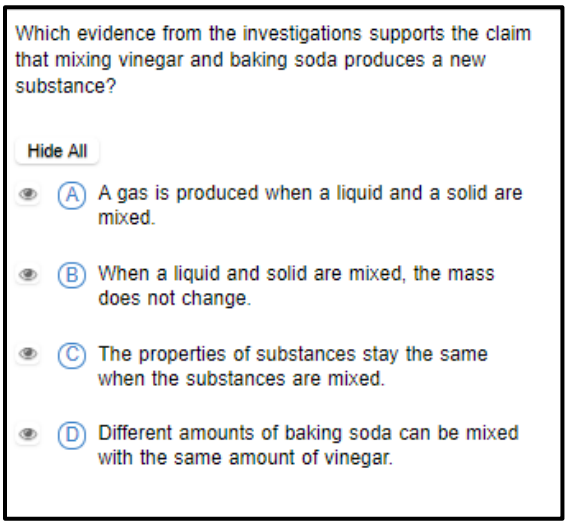


Eliana's data are shown in the table.

Investigation 2 Data

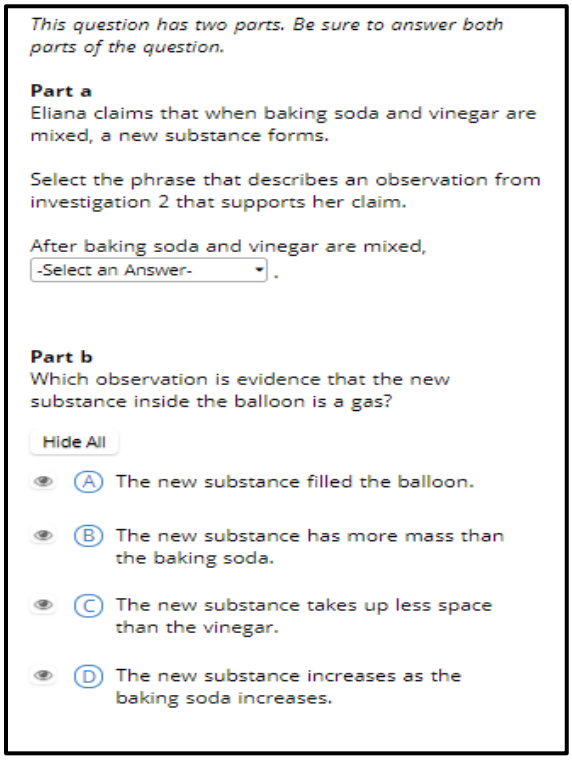
Amount of Baking Soda (milliliters)	Mass of Gas Produced (grams)			Average Mass of Gas Produced (grams)
	Trial 1	Trial 2	Trial 3	
5	1.0	0.8	1.2	1.0
10	1.5	1.9	1.4	1.6
15	2.4	1.9	2.6	2.3

Figure 2: Machine-Scored 1-Point (MS-1) Cluster Item – Grade 5 NM-ASR Practice Test



MS-1 cluster item, grade 5 practice test, aligned to PE 5-PS1-4: Conduct an investigation to determine whether the mixing of two or more substances results in new substances.
The dimensions for the PE are SEP: Planning and Carrying Out Investigations; DCI: PS1.B: Chemical Reactions; CCC: Cause and Effect. This particular MS-1 item in the cluster assesses the DCI and CCC dimensions.

Figure 3: Machine-Scored 2-Point (MS-2) Cluster Item Sample – Grade 5 NM-ASR Practice Test



MS-2 cluster item, grade 5 practice test, aligned to PE 5-PS1-4: Conduct an investigation to determine

whether the mixing of two or more substances results in new substances.
The dimensions for the PE are SEP: Planning and Carrying Out Investigations; DCI: PS1.B: Chemical Reactions; CCC: Cause and Effect. This particular MS-2 item in the cluster assesses the DCI and CCC dimensions.

Figure 4: Cluster Structure

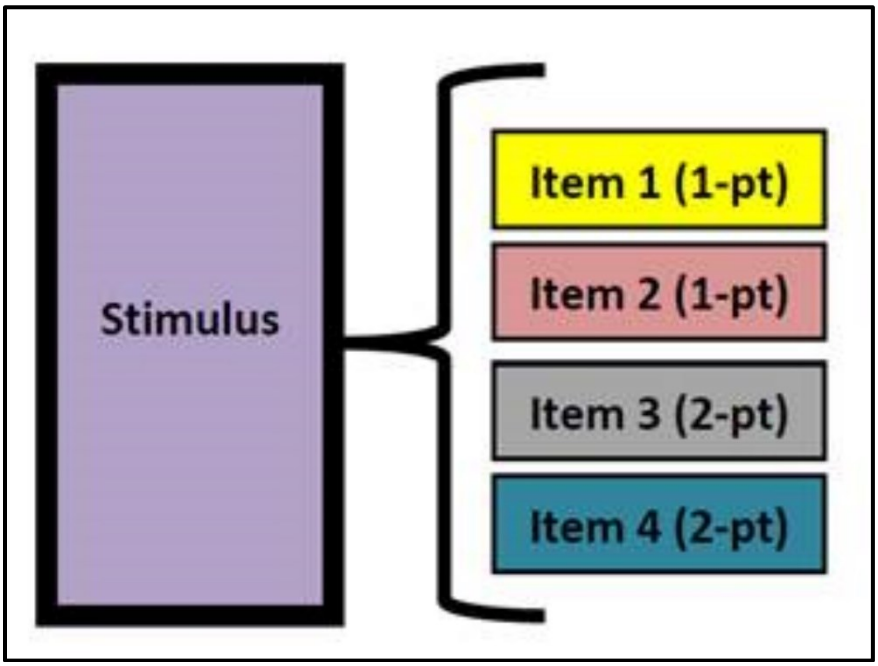


Table 14: MS-2 Item Type Sample

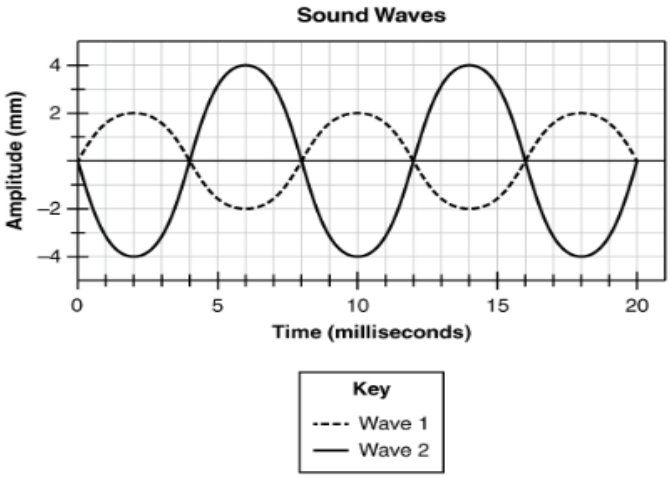
MS-2 Items	As in the cluster, the standalone MS-2 items are worth 2 points and have two parts, with ○ Part a worth 1 point and ○ Part b also worth 1 point.
	Each part of the item may be presented as ○ multiple-choice, ○ multi-select, or ○ technology-enhanced (e.g., drag-and-drop, hot spot, drop-down selections).
	See ***Figure 5*** to view sample.

Some MS-2 items are standalone, or individual, machine-scored items worth two points.

Figure 5: Machine-Scored 2-Point Item Sample – Grade 8 NM-ASR Practice Test

This question has two parts. Be sure to answer both parts of the question.

Elaine is researching the characteristics of sound waves as she listens to music. She finds a diagram that shows two sound waves.



Part a

Select the number that describes the sound waves in the diagram.

Wave 1 has -Select an Answer- the energy of Wave 2.

Part b

Elaine predicts that the energy in each sound wave would double if the frequency of the wave doubled.

Which statement describes her prediction?

Hide All

(A) Her prediction is correct because the energy of a wave is proportional to the wave's frequency.

(B) Her prediction is incorrect because the energy of a wave is proportional to the square of the wave's frequency.

(C) Her prediction is correct because the energy of a wave is proportional to speed and a wave with twice the frequency has twice the speed.

(D) Her prediction is incorrect because the energy of a wave is proportional to the wavelength and a wave with twice the frequency has half the wavelength.

MS-2 item, grade 8 practice test, aligned to PE MS-PS4-1: Use mathematical representations to describe a simple model for waves that includes how the amplitude of a wave is related to the energy in a wave. The dimensions for the PE are SEP: Using Mathematics and Computational Thinking; DCI: PS4.A: Wave Properties; CCC: Patterns. This particular MS-2 standalone item assesses the SEP, DCI, and CCC dimensions.

Table 15: OE Item Type Sample

OE Items	The prompt or question may be • a single prompt, or • more typically, the item will be written with multiple, scaffolded parts for students to answer.
	The items are worth 4 points each.
	These items are hand-scored for: ○ 4, ○ 3, ○ 2, ○ 1, or ○ 0 points. The items are scored by trained scorers using a rubric and scoring notes.
	See ***Figure 6*** to view sample.

OE, or open-ended, items are standalone items that require students to provide a written response to a prompt or question.

Figure 6: Open-Ended 4-Point Item Sample – Grade 11 NM-ASR Practice Test

This question has two parts. Be sure to answer both parts of the question.

Some students work at a local aquarium. One of their tasks is to care for mollusks and corals in ocean water in a tank at the aquarium. The students need to make sure that the ocean water has the right balance of calcium ions (Ca^{2+}) and carbonate ions (CO_3^{2-}) that the mollusks and corals need to build their shells and skeletons.

To do this, the students need to ensure that calcium and carbonate ions are continuously added to the ocean water in the tank. The students know that ocean water contains calcium carbonate, which naturally breaks down into calcium and carbonate ions. The equilibrium relationship between the components in the water is shown in the equation.

Equilibrium Equation

$$CO_2 + H_2O + CaCO_3 \rightleftharpoons Ca^{2+} + 2H^+ + 2CO_3^{2-}$$

The students decide to test the equilibrium relationships in the equation. With ocean water as an input, the students remove calcium ions (Ca^{2+}) as the ions form in the water in the tank. The students observe that as they remove calcium ions, more calcium ions form in the tank. They realize that this is an example of Le Chatelier's principle that describes the equilibrium relationships in the water.

The people who work at the aquarium tell the students that ocean water contains carbon dioxide (CO_2) and that increasing amounts of CO_2 in ocean water can cause some of the calcium carbonate ($CaCO_3$) in the shells and skeletons of ocean organisms to dissolve.

The students want to solve this problem by decreasing the amount of carbon dioxide in ocean water.

a. Describe one way students could decrease the amount of CO_2 in ocean water by applying Le Chatelier's principle.

b. Describe one constraint on implementing the change described in Part (a).

OE item, grade 11 practice test, aligned to PE HS-PS1-6: Refine the design of a chemical system by specifying a change in conditions that would produce increased amounts of products at equilibrium. The dimensions for the PE are SEP: Constructing Explanations and Designing Solutions; DCI: PS1.B: Chemical Reactions and ETS1.C: Optimizing the Design Solution; CCC: Stability and Change. This particular OE item assesses the SEP, DCI, and CCC dimensions.

APPENDIX C
PARTICIPATION RATES

Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA & ASR assessments.

Table C-1. Participation Rates on NM-MSSA ELA, as a Function of Subgroup and Grade*

Group	Subgroup	Grade 3	4	5	6	7	8
Overall		20,881	21,363	21,870	22,037	22,391	23,648
Gender	Female	10,221	10,540	10,690	10,879	10,974	11,584
	Male	10,652	10,811	11,177	11,151	11,397	12,046
	Unknown	8	12	3	7	20	18
Ethnicity	African American or Black	700	648	668	697	654	717
	American Indian or Alaska Native	2,556	2,621	2,557	2,546	2,741	2,762
	Asian	350	396	419	377	381	395
	Caucasian	16,854	17,249	17,794	17,992	18,187	19,366
	Hawaiian Native or Other Pacific Islander	68	75	67	94	117	108
	Multi	341	359	357	316	284	287
	Unknown	12	15	8	15	27	13
Hispanic	Yes	13,027	13,199	13,751	13,892	13,973	14,956
	No	7,842	8,149	8,111	8,130	8,391	8,679
	Unknown	0	0	0	0	0	0
Econ. Dis.	Yes	10,233	10,938	11,274	10,810	10,639	11,163
	No	8,290	8,703	8,847	9,632	10,073	10,758
	Unknown	2,358	1,722	1,749	1,595	1,679	1,727
English Learners	Yes	3,074	3,613	4,102	3,846	4,174	4,172
	No	17,807	17,750	17,768	18,191	18,217	19,476
Foster Care	Yes	31	21	24	31	14	19
	No	9,763	9,993	10,235	9,964	9,544	10,149
	Unknown	11,087	11,349	11,611	12,042	12,833	13,480
Homeless	Yes	344	361	349	394	323	346
	No	17,741	18,834	19,307	19,697	20,064	21,202
	Unknown	2,796	2,168	2,214	1,946	2,004	2,100
Homeschool	Yes	0	0	2	1	1	2
	No	20,881	21,363	21,868	22,036	22,390	23,646
	Unknown	0	0	0	0	0	0
Migrant	Yes	9	12	11	17	6	11
	No	13,273	13,583	13,988	15,355	15,551	16,366
	Unknown	7,599	7,768	7,871	6,665	6,834	7,271
Military	Yes	208	177	200	148	187	208
	No	12,002	12,289	12,459	13,839	14,043	14,726
	Unknown	8,671	8,897	9,211	8,050	8,161	8,714
Special Ed	Yes	3,537	3,774	3,936	3,999	3,799	4,088
	No	16,848	17,102	17,484	17,588	18,139	19,084
	Unknown	496	487	450	450	453	476

**Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA assessment.*

Table C-2. Participation Rates on NM-MSSA Mathematics, as a Function of Subgroup and Grade*

Group	Subgroup	Grade 3	4	5	6	7	8
Overall		20,884	21,368	21,872	22,054	22,359	23,634
Gender	Female	10,230	10,543	10,685	10,885	10,954	11,566
	Male	10,646	10,813	11,184	11,162	11,385	12,050
	Unknown	8	12	3	7	20	18
Ethnicity	African American or Black	702	647	668	702	655	717
	American Indian or Alaska Native	2,555	2,619	2,558	2,547	2,733	2,764
	Asian	355	400	423	381	384	396
	Caucasian	16,851	17,251	17,792	17,998	18,160	19,350
	Hawaiian Native or Other Pacific Islander	68	75	67	94	116	109
	Multi	341	361	356	317	284	285
	Unknown	12	15	8	15	27	13
Hispanic	Yes	13,022	13,204	13,751	13,893	13,939	14,933
	No	7,850	8,149	8,113	8,146	8,393	8,688
	Unknown	0	0	0	0	0	0
Econ. Dis.	Yes	10,240	10,928	11,278	10,815	10,625	11,156
	No	8,290	8,718	8,842	9,644	10,059	10,762
	Unknown	2,354	1,722	1,752	1,595	1,675	1,716
English Learners	Yes	3,079	3,626	4,108	3,853	4,176	4,165
	No	17,805	17,742	17,764	18,201	18,183	19,469
Foster Care	Yes	31	21	25	32	14	18
	No	9,771	9,991	10,239	9,976	9,530	10,141
	Unknown	11,082	11,356	11,608	12,046	12,815	13,475
Homeless	Yes	346	361	346	393	321	349
	No	17,746	18,839	19,311	19,716	20,038	21,197
	Unknown	2,792	2,168	2,215	1,945	2,000	2,088
Homeschool	Yes	0	0	2	1	1	2
	No	20,884	21,368	21,870	22,053	22,358	23,632
	Unknown	0	0	0	0	0	0
Migrant	Yes	9	12	11	17	6	11
	No	13,275	13,587	13,983	15,366	15,533	16,373
	Unknown	7,600	7,769	7,878	6,671	6,820	7,250
Military	Yes	207	177	199	148	187	206
	No	12,001	12,291	12,455	13,850	14,027	14,732
	Unknown	8,676	8,900	9,218	8,056	8,145	8,696
Special Ed	Yes	3,541	3,776	3,936	3,993	3,793	4,090
	No	16,847	17,105	17,487	17,611	18,113	19,066
	Unknown	496	487	449	450	453	478

**Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA assessment.*

Table C-3. Participation Rates on NM-ASR Science, as a Function of Subgroup and Grade*

Group	Subgroup	Grade 5	8	11
Overall		21,888	23,668	21,202
Gender	Female	10,689	11,590	10,710
	Male	11,196	12,061	10,483
	Unknown	3	17	9
Ethnicity	African American or Black	670	716	582
	American Indian or Alaska Native	2,553	2,763	2,697
	Asian	421	396	414
	Caucasian	17,810	19,387	17,233
	Hawaiian Native or Other Pacific Islander	67	108	87
	Multi	359	286	188
	Unknown	8	12	1
Hispanic	Yes	13,759	14,963	13,006
	No	8,121	8,693	8,195
	Unknown	0	0	0
Econ. Dis.	Yes	11,286	11,163	8,044
	No	8,851	10,772	12,209
	Unknown	1,751	1,733	949
English Learners	Yes	4,114	4,179	2,872
	No	17,774	19,489	18,330
Foster Care	Yes	25	18	4
	No	10,241	10,156	5,670
	Unknown	11,622	13,494	15,528
Homeless	Yes	345	351	278
	No	19,326	21,217	19,973
	Unknown	2,217	2,100	951
Homeschool	Yes	2	2	1
	No	21,886	23,666	21,201
	Unknown	0	0	0
Migrant	Yes	11	12	1
	No	14,001	16,367	16,059
	Unknown	7,876	7,289	5,142
Military	Yes	200	207	216
	No	12,474	14,727	15,151
	Unknown	9,214	8,734	5,835
Special Ed	Yes	3,934	4,097	2,819
	No	17,505	19,093	18,381
	Unknown	449	478	2

**Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA assessment.*

Table C-4. Participation Rates on NM-MSSA Spanish Language Arts (SLA), as a Function of Subgroup and Grade*

Group	Subgroup	Grade					
		3	4	5	6	7	8
Overall		662	611	272	269	259	278
Gender	Female	329	308	140	134	130	130
	Male	333	303	132	135	129	148
	Unknown	0	0	0	0	0	0
Ethnicity	African American or Black	8	3	4	2	4	2
	American Indian or Alaska Native	0	0	0	3	3	2
	Asian	3	3	2	1	4	2
	Caucasian	645	599	261	258	243	268
	Hawaiian Native or Other Pacific Islander	4	3	1	2	2	4
	Multi	2	2	4	3	3	0
	Unknown	0	1	0	0	0	0
Hispanic	Yes	651	599	265	264	253	277
	No	11	11	7	5	6	1
	Unknown	0	0	0	0	0	0
Econ. Dis.	Yes	484	448	144	139	141	147
	No	137	138	114	113	103	119
	Unknown	41	25	14	17	15	12
English Learners	Yes	605	561	229	221	215	248
	No	57	50	43	48	44	30
Foster Care	Yes	0	0	0	2	0	0
	No	354	352	154	150	124	151
	Unknown	308	259	118	117	135	127
Homeless	Yes	17	16	19	13	16	18
	No	542	504	224	227	206	233
	Unknown	103	91	29	29	37	27
Homeschool	Yes	0	0	0	0	0	0
	No	662	611	272	269	259	278
	Unknown	0	0	0	0	0	0
Migrant	Yes	5	3	8	0	4	2
	No	467	413	193	168	176	195
	Unknown	190	195	71	101	79	81
Military	Yes	9	1	0	2	1	1
	No	359	366	174	146	147	162
	Unknown	294	244	98	121	111	115
Special Ed	Yes	70	56	7	12	3	3
	No	590	545	260	256	255	274
	Unknown	2	10	5	1	1	1

**Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA assessment.*

Table C-5. Participation Rates on NM-MSSA Mathematics (Spanish Transadapted), as a Function of Subgroup and Grade*

Group	Subgroup	Grade					
		3	4	5	6	7	8
Overall		680	615	284	272	269	287
Gender	Female	335	308	146	136	134	135
	Male	345	307	138	136	135	152
	Unknown	0	0	0	0	0	0
Ethnicity	African American or Black	8	3	4	2	4	2
	American Indian or Alaska Native	0	0	0	3	4	2
	Asian	3	3	2	1	4	2
	Caucasian	662	604	272	261	252	277
	Hawaiian Native or Other Pacific Islander	5	3	2	2	2	4
	Multi	2	1	4	3	3	0
	Unknown	0	1	0	0	0	0
Hispanic	Yes	666	602	277	267	263	286
	No	14	12	7	5	6	1
	Unknown	0	0	0	0	0	0
Econ. Dis.	Yes	493	447	152	140	145	150
	No	147	142	118	116	107	123
	Unknown	40	26	14	16	17	14
English Learners	Yes	619	559	236	226	223	255
	No	61	56	48	46	46	32
Foster Care	Yes	0	0	0	1	0	0
	No	372	358	165	152	131	157
	Unknown	308	257	119	119	138	130
Homeless	Yes	17	16	21	12	17	18
	No	560	507	234	232	213	241
	Unknown	103	92	29	28	39	28
Homeschool	Yes	0	0	0	0	0	0
	No	680	615	284	272	269	287
	Unknown	0	0	0	0	0	0
Migrant	Yes	5	2	8	1	4	2
	No	481	424	203	172	183	204
	Unknown	194	189	73	99	82	81
Military	Yes	9	1	0	2	1	1
	No	374	376	184	151	154	171
	Unknown	297	238	100	119	114	115
Special Ed	Yes	70	55	7	11	3	2
	No	608	550	272	260	265	284
	Unknown	2	10	5	1	1	1

**Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA assessment.*

Table C-6. Participation Rates on NM-ASR Science (Spanish Transadapted), as a Function of Subgroup and Grade*

Group	Subgroup	Grade		
		5	8	11
Overall		275	280	299
Gender	Female	143	132	162
	Male	132	148	137
	Unknown	0	0	0
Ethnicity	African American or Black	4	2	3
	American Indian or Alaska Native	0	2	5
	Asian	2	2	0
	Caucasian	263	270	287
	Hawaiian Native or Other Pacific Islander	2	4	4
	Multi	4	0	0
	Unknown	0	0	0
Hispanic	Yes	268	279	293
	No	7	1	6
	Unknown	0	0	0
Econ. Dis.	Yes	147	150	169
	No	114	117	130
	Unknown	14	13	0
English Learners	Yes	229	247	240
	No	46	33	59
Foster Care	Yes	0	0	0
	No	161	156	132
	Unknown	114	124	167
Homeless	Yes	20	18	23
	No	227	233	273
	Unknown	28	29	3
Homeschool	Yes	0	0	0
	No	275	280	299
	Unknown	0	0	0
Migrant	Yes	8	2	1
	No	195	197	202
	Unknown	72	81	96
Military	Yes	0	1	1
	No	177	165	173
	Unknown	98	114	125
Special Ed	Yes	6	2	2
	No	265	277	296
	Unknown	4	1	1

**Participation is defined as those students who took and attempted at least 5 items on the given NM-MSSA assessment.*

APPENDIX D
ACCOMMODATION FREQUENCIES

Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.

Table D-1. Number of Students Taking NM-MSSA ELA, as a Function of Accommodation or Accessibility Feature and Grade*

Accommodation/Accessibility Feature	Grades					
	3	4	5	6	7	8
EL: Commercial Word-to-Word Dictionary	49	67	76	127	146	124
EL: Customized Dual Language Glossary	5	11	0	0	1	0
EL: Directions in Native Language	34	45	72	75	36	30
EL: Picture Dictionary	6	10	3	5	9	4
EL: Pocket Word-to-Word Translator	1	1	4	3	1	1
IEP/504: Allow Accessibility Mode Testing	24	42	50	24	40	53
IEP/504: Assistive Technology Devices Presentation	3	5	4	11	8	4
IEP/504: Assistive Technology Devices Responses	4	5	6	0	1	2
IEP/504: Braille	2	1	0	0	0	1
IEP/504: Constructed Response Human Scribe	7	3	4	1	1	1
IEP/504: Human Reader English	61	72	124	58	38	29
IEP/504: Human Signer	10	5	3	2	4	3
IEP/504: Large-print	1	0	2	0	1	2
IEP/504: Read Aloud to Self	149	198	231	132	87	94
IEP/504: Selected Response Human Scribe	5	3	2	1	1	1
Online test only: Braille Notetaker	0	0	1	0	0	0
Online test only: Braille Writer	0	0	1	0	0	0
Online Test only: Color Contrast	1,168	1,153	1,228	842	929	1,016
Online test only: ELA ASL Video	17	10	8	10	10	10
Online test only: ELA Text-to-Speech English	528	607	619	620	508	505
Online test only: Headphones/Noise Buffer	557	564	642	555	431	506
Online test only: Human Signer for Test Directions	12	5	9	10	14	9
Online test only: Refreshable Braille	0	0	1	0	0	0
Online test only: Screen Reader	8	19	18	13	11	5
Online test only: Speech-to-Text	114	131	154	168	98	121
Online test only: Test was marked for Masking Answer	1,363	1,380	1,325	949	971	1,053
Online test only: Test was marked for Masking Custom	1,173	1,090	1,135	808	880	974
Online test only: Test was marked for Reverse Contrast	1,153	1,093	1,157	784	866	882
Online test only: Test was marked for Tactile Graphics	7	7	7	10	9	2
Online test only: Word Prediction	36	63	80	61	33	26
Online test only: Word Prediction (embedded)	202	238	330	261	197	187
Online test only: Human Scribe	38	47	44	30	28	13

**Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.*

Table D-2. Number of Students Taking NM-MSSA Mathematics, as a Function of Accommodation or Accessibility Feature and Grade*

Accommodation/Accessibility Feature	Grades					
	3	4	5	6	7	8
EL: Commercial Word-to-Word Dictionary	50	67	76	127	144	122
EL: Customized Dual Language Glossary	5	11	1	0	1	0
EL: Directions in Native Language	36	45	74	76	34	27
EL: Picture Dictionary	5	10	4	5	9	4
EL: Pocket Word-to-Word Translator	1	1	4	3	1	1
IEP/504: Allow Accessibility Mode Testing	24	42	50	24	40	53
IEP/504: Assistive Technology Devices Presentation	4	5	4	11	8	4
IEP/504: Assistive Technology Devices Responses	4	5	6	0	2	3
IEP/504: Braille	2	2	0	0	0	1
IEP/504: Constructed Response Human Scribe	8	5	4	1	1	1
IEP/504: Human Reader English	69	98	144	100	70	43
IEP/504: Human Signer	11	5	3	2	4	3
IEP/504: Large-print	3	2	3	0	1	3
IEP/504: Read Aloud to Self	142	198	227	123	83	89
IEP/504: Selected Response Human Scribe	7	5	3	1	1	1
Online test only: Basic Calculator on non-Calculator section of Math	122	178	227	436	489	531
Online test only: Braille Notetaker	0	0	1	0	0	0
Online test only: Braille Writer	0	0	1	0	0	0
Online Test only: Color Contrast	1,168	1,152	1,229	844	924	1,008
Online test only: Headphones/Noise Buffer	558	564	642	554	430	501
Online test only: Human Signer for Test Directions	13	5	9	10	14	9
Online test only: Math ASL Video	17	10	9	10	12	12
Online test only: Mathematics Text-to-Speech English	5,378	5,677	5,760	4,058	3,766	3,911
Online test only: Mathematics Tools	623	693	754	602	472	463
Online test only: Refreshable Braille	0	0	1	0	0	0
Online test only: Scientific Calculator on non-Calculator section of Math	25	40	49	134	280	384
Online test only: Screen Reader	8	19	18	12	10	5
Online test only: Speech-to-Text	113	130	154	167	98	120
Online test only: Test was marked for Masking Answer	1,364	1,381	1,328	949	962	1,045
Online test only: Test was marked for Masking Custom	1,173	1,089	1,137	806	876	967
Online test only: Test was marked for Reverse Contrast	1,153	1,092	1,158	786	862	876
Online test only: Test was marked for Tactile Graphics	7	7	7	10	9	2
Online test only: Word Prediction	37	63	80	61	33	26
Online test only: Word Prediction (embedded)	203	236	329	262	195	185
Online test only: Human Scribe	37	47	44	29	28	14

**Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.*

Table D-3. Number of Students Taking NM-ASR Science, as a Function of Accommodation or Accessibility Feature and Grade*

Accommodation/Accessibility Feature	Grades		
	5	8	11
EL: Commercial Word-to-Word Dictionary	76	124	11
EL: Customized Dual Language Glossary	1	2	0
EL: Directions in Native Language	74	30	4
EL: Picture Dictionary	4	4	0
EL: Pocket Word-to-Word Translator	4	1	4
IEP/504: Allow Accessibility Mode Testing	50	52	9
IEP/504: Assistive Technology Devices Presentation	4	4	2
IEP/504: Assistive Technology Devices Responses	6	3	3
IEP/504: Braille	0	0	0
IEP/504: Constructed Response Human Scribe	2	1	0
IEP/504: Human Reader English	139	40	4
IEP/504: Human Signer	3	3	0
IEP/504: Large-print	2	3	0
IEP/504: Read Aloud to Self	225	91	30
IEP/504: Selected Response Human Scribe	3	1	0
Online test only: Braille Notetaker	1	0	1
Online test only: Braille Writer	1	0	1
Online Test only: Color Contrast	1,229	1,020	940
Online test only: Headphones/Noise Buffer	640	503	129
Online test only: Human Signer for Test Directions	9	9	6
Online test only: Refreshable Braille	1	0	0
Online test only: Science Text-to-Speech English	4,643	3,452	1,598
Online test only: Screen Reader	17	5	2
Online test only: Speech-to-Text	152	121	26
Online test only: Test was marked for Masking Answer	1,329	1,058	867
Online test only: Test was marked for Masking Custom	1,137	980	865
Online test only: Test was marked for Reverse Contrast	1,159	887	904
Online test only: Test was marked for Tactile Graphics	6	2	1
Online test only: Word Prediction	79	26	3
Online test only: Word Prediction (embedded)	329	185	4
Online test only: Human Scribe	43	14	3

**Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.*

Table D-4. Number of Students Taking NM-MSSA SLA, as a Function of Accommodation or Accessibility Feature and Grade*

Accommodation/Accessibility Feature	Grades					
	3	4	5	6	7	8
EL: Commercial Word-to-Word Dictionary	11	25	13	26	17	19
EL: Customized Dual Language Glossary	1	1	1	17	8	7
EL: Directions in Native Language	58	66	37	75	36	47
EL: Picture Dictionary	9	3	2	14	6	2
EL: Pocket Word-to-Word Translator	0	15	1	13	5	2
IEP/504: Allow Accessibility Mode Testing	0	0	0	0	1	0
IEP/504: Assistive Technology Devices Presentation	0	0	0	0	1	0
IEP/504: Assistive Technology Devices Responses	0	0	0	0	0	0
IEP/504: Braille	0	0	0	0	0	0
IEP/504: Constructed Response Human Scribe	0	0	0	0	0	0
IEP/504: Human Reader Spanish	0	1	0	0	0	0
IEP/504: Human Signer	0	0	0	0	0	0
IEP/504: Large-print	0	0	0	0	0	0
IEP/504: Read Aloud to Self	4	7	0	1	0	0
IEP/504: Selected Response Human Scribe	0	0	0	0	0	0
Online test only: Braille Notetaker	0	0	0	0	0	0
Online test only: Braille Writer	0	0	0	0	0	0
Online Test only: Color Contrast	4	5	0	0	1	0
Online test only: Headphones/Noise Buffer	18	28	0	3	1	1
Online test only: Human Signer for Test Directions	0	0	0	0	0	0
Online test only: Refreshable Braille	0	0	0	0	0	0
Online test only: Screen Reader	0	0	0	0	0	0
Online test only: SLA Text-to-Speech Spanish	13	10	7	7	2	6
Online test only: Speech-to-Text	2	2	1	0	0	0
Online test only: Test was marked for Masking Answer	5	11	2	1	2	2
Online test only: Test was marked for Masking Custom	6	6	0	0	1	0
Online test only: Test was marked for Reverse Contrast	3	4	0	0	0	0
Online test only: Test was marked for Tactile Graphics	0	0	0	0	0	0
Online test only: Word Prediction	0	4	3	0	0	0
Online test only: Word Prediction (embedded)	4	2	4	0	0	0
Online test only: Human Scribe	0	3	0	0	0	0

**Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.*

Table D-5. Number of Students Taking NM-MSSA Mathematics (Spanish Transadapted), as a Function of Accommodation or Accessibility Feature and Grade*

Accommodation/Accessibility Feature	Grades					
	3	4	5	6	7	8
EL: Commercial Word-to-Word Dictionary	11	24	13	27	19	25
EL: Customized Dual Language Glossary	1	1	1	17	8	7
EL: Directions in Native Language	59	67	37	75	40	53
EL: Picture Dictionary	10	3	4	14	6	2
EL: Pocket Word-to-Word Translator	0	15	1	13	5	2
IEP/504: Allow Accessibility Mode Testing	0	0	0	0	1	0
IEP/504: Assistive Technology Devices Presentation	0	0	0	0	1	0
IEP/504: Assistive Technology Devices Responses	0	0	0	0	0	0
IEP/504: Braille	0	0	0	0	0	0
IEP/504: Constructed Response Human Scribe	0	0	0	0	0	0
IEP/504: Human Reader Spanish	0	1	0	0	0	0
IEP/504: Human Signer	0	0	0	0	0	0
IEP/504: Large-print	0	0	0	0	0	0
IEP/504: Read Aloud to Self	4	6	0	1	0	0
IEP/504: Selected Response Human Scribe	0	0	0	0	0	0
Online test only: Basic Calculator on non-Calculator section of Math	1	2	0	0	0	0
Online test only: Braille Notetaker	0	0	0	0	0	0
Online test only: Braille Writer	0	0	0	0	0	0
Online Test only: Color Contrast	4	5	0	0	1	1
Online test only: Headphones/Noise Buffer	18	27	1	3	1	2
Online test only: Human Signer for Test Directions	0	0	0	0	0	0
Online test only: Math ASL Video	0	0	0	0	0	0
Online test only: Mathematics Text-to-Speech Spanish	140	117	37	46	29	43
Online test only: Mathematics Tools	8	8	1	4	2	0
Online test only: Refreshable Braille	0	0	0	0	0	0
Online test only: Scientific Calculator on non-Calculator section of Math	0	0	0	0	0	0
Online test only: Screen Reader	0	0	0	0	0	0
Online test only: Speech-to-Text	2	2	1	0	0	0
Online test only: Test was marked for Masking Answer	6	11	3	2	3	3
Online test only: Test was marked for Masking Custom	6	6	0	0	1	0
Online test only: Test was marked for Reverse Contrast	4	4	1	0	0	0
Online test only: Test was marked for Tactile Graphics	0	0	0	0	0	0
Online test only: Word Prediction	0	4	3	0	0	0
Online test only: Word Prediction (embedded)	4	2	4	0	0	0
Online test only: Human Scribe	0	3	0	0	0	0

**Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.*

Table D-6. Number of Students Taking NM-ASR Science (Spanish Transadapted), as a Function of Accommodation or Accessibility Feature and Grade*

Accommodation/Accessibility Feature	Grades		
	5	8	11
EL: Commercial Word-to-Word Dictionary	13	24	13
EL: Customized Dual Language Glossary	1	5	0
EL: Directions in Native Language	37	50	9
EL: Picture Dictionary	4	2	2
EL: Pocket Word-to-Word Translator	1	2	6
IEP/504: Allow Accessibility Mode Testing	0	0	0
IEP/504: Assistive Technology Devices Presentation	0	0	0
IEP/504: Assistive Technology Devices Responses	0	0	0
IEP/504: Braille	0	0	0
IEP/504: Constructed Response Human Scribe	0	0	0
IEP/504: Human Reader Spanish	0	0	0
IEP/504: Human Signer	0	0	0
IEP/504: Large-print	0	0	0
IEP/504: Read Aloud to Self	0	0	0
IEP/504: Selected Response Human Scribe	0	0	0
Online test only: Braille Notetaker	0	0	0
Online test only: Braille Writer	0	0	0
Online Test only: Color Contrast	0	1	1
Online test only: Headphones/Noise Buffer	1	2	1
Online test only: Human Signer for Test Directions	0	0	0
Online test only: Refreshable Braille	0	0	0
Online test only: Science Text-to-Speech Spanish	34	40	8
Online test only: Screen Reader	0	0	0
Online test only: Speech-to-Text	0	0	0
Online test only: Test was marked for Masking Answer	3	3	0
Online test only: Test was marked for Masking Custom	0	0	0
Online test only: Test was marked for Reverse Contrast	1	0	1
Online test only: Test was marked for Tactile Graphics	0	0	0
Online test only: Word Prediction	3	0	0
Online test only: Word Prediction (embedded)	4	0	0
Online test only: Human Scribe	0	0	0

**Only students who met the attemptedness rule (i.e., attempted 5 or more items) contributed to the frequencies in these tables.*

Appendix E
2022–23 Assessment Accommodations & Accessibility Manual

NEW MEXICO
Public Education Department

AFAM

NM-ASR and NM-MSSA Spring 2023

Accessibility Features and Accommodations Manual

Guidance for Districts and Decision-Making Teams to Ensure that Spring 2023 Science, Mathematics, and ELA Summative Assessments Produce Valid Results for All Students

NM-ASR
NEW MEXICO ASSESSMENT OF SCIENCE READINESS

NM-MSSA
NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT

Available online at: newmexico.onlinehelp.cognia.org/accommodations-information

Table of Contents

Background

Audience and Purpose

The NM Accessibility Features and Accommodations Manual is a comprehensive policy document that provides guidance to districts and decision making teams to ensure the New Mexico Assessment of Science Readiness (NM-ASR) and the New Mexico Measures of Student Success and Achievement (NM-MSSA) summative assessments provide valid results for all participating students.

Introduction

New Mexico Public Education Department (PED) regards assessments as tools for enhancing teaching and learning. PED is committed to providing all students with equitable access to high-quality, 21st century assessments. By applying principles of universal design, using technology, embedding accessibility features, and allowing a broad range of accommodations, PED provides opportunities for the largest possible number of students to demonstrate their knowledge and skills. PED sets and maintains high expectations that all students will have access to the full range of grade-level and course content standards. For additional PED guidance concerning accommodations on the required summative assessments, please refer to the *2022–23 Assessment Accommodations & Accessibility Manual* available at https://webnew.ped.state.nm.us/wp-content/uploads/2022/09/2022_23_Accommodations_Manual.pdf.

PED's goals for promoting student access include:

- Applying principles of universal design for accessible assessments during every stage of the development of the assessment items and performance tasks;
- Minimizing/eliminating features of the assessment that are irrelevant to what is being measured so that all students can more accurately demonstrate their knowledge and skills;
- Measuring the full range of complexity of the standards;
- Using technology for the accessible delivery of the assessments;
- Building accessibility throughout the test without sacrificing assessment validity;
- Using a combination of accessible authoring and accessible technologies from the inception of items and tasks; and
- Engaging state and national experts throughout the development process through item review, bias and sensitivity review, policy development and review, and research.

This manual provides information on the accessibility features and accommodations that will be available during the 2023 Science, Math, and ELA assessments, based on careful review and inclusion of the following:

- Current and field test research on effective practices for assessing student groups, (including students with disabilities and ELs) and backgrounds (cultural, regional, linguistic, dialect, and socio-economic);
- Feedback from state leads and state experts on students with disabilities and ELs;
- Feedback from the content experts.

Participation Guidelines for Paper-Based 2023 Science/Math/ELA Assessments

Although 2023 Science/Math/ELA assessments are computer-based, using an online testing platform, there may be specific instances which require a student to take a paper-based assessment instead. The following conditions may result in a school choosing to administer a paper-based assessment:

- Condition #1: A student is unable to use a computer due to the impact of his or her disability. The student's inability to participate in computer-based assessments should be documented in an Individualized Education Program (IEP) or 504 plan. Examples may include:
    - A student with a disability who cannot participate in the online assessment due to a health-related disability, neurological disorder, or other complex disability, and/or cannot meet the demands of a computer-based test administration;
    - A student with an emotional, behavioral, or other disability who is unable to maintain sufficient concentration to participate in a computer-based test administration, even with test accommodations;
    - A student with a disability who requires assistive technology that is not compatible with the testing platform.
- Condition #2: A student who recently entered the school and has had very little or no prior experience or familiarity with technology.
- Condition #3: A student who is unable to access an online assessment due to religious beliefs.

General Administrative Considerations, Universal Tools, and Accessibility Features

Administrative Considerations for All Students

Detailed guidelines on the administration of the 2023 Science, Math, and ELA assessments will be included in the *Test Administrator's Manual* and the *Test Coordinator's Manual.*

Although students are generally tested in their regular classroom and follow the standard test administration schedule for the grade and content area being assessed, the principal or the test coordinator have the authority to schedule students in testing spaces other than regular classrooms, and at different scheduled times, as long as all requirements for testing conditions and test security are met as set forth in the *Test Administrator's Manual* and *Test Coordinator's Manual*. Decisions may be considered, for example, that benefit students who are easily distracted in large group settings by testing them in a small group or individual setting. In general, changes to the timing, setting, or conditions of testing are left to the discretion of the principal or test coordinator.

In accordance with principles of universal design for assessment, PED is providing the following administrative guidance regarding the timing and scheduling of assessments, and setting/locations for testing. These administrative considerations are available to all students. The principal may determine that any student can receive one or more of the following test administration considerations, regardless of the student's status as a student with a disability or EL.

Table 1: Administrative Considerations for All Students

Consideration	Description
Small Group Testing	Student is tested in a separate location with a small group of students with matching accessibility features, accommodations, or testing needs as appropriate. Small group size can vary from 5–10 students depending on the circumstance and student needs as outlined in the IEP.
Time of Day	Student is tested during a specific time of day based on their individual needs (e.g., ELA in the morning; no testing after lunch).
Separate or Alternate Location	Student is tested in a specifically assigned location.
Specified Area or Setting	Student is tested in a specialized area or setting (e.g., front of the classroom, seat near the door, library, etc.).
Adaptive and specialized equipment or furniture	Student is provided specialized equipment or furniture needed for a successful testing environment (e.g., low lighting; adaptive seat).
Frequent breaks	Guidance on logistics for administrating the 2023 Science, Math, and ELA assessments with frequent breaks: • *Medical Breaks*: Student takes a break due to pre-existing or sudden onset of a temporary or long-term medical condition. Student's testing time stops. • *Individual Bathroom Breaks*: Student requests a bathroom break within their overall allotted testing time. Student's testing time does not stop. • *In-Chair Stretch Break:* Student pauses and stretches. Student's testing time does not stop. • *Other Frequent Breaks,* according to PED policy.

Universal Tools Available to All Students

Table 2 lists the tools available to all students through the computer-based testing platform as well as the equivalent resources for paper-based testing. The universal tools do not need to be assigned in iTester prior to testing. Students should be familiar with using these tools prior to testing through classroom instruction or practice testing.

Table 2: Universal Tools for Computer-Based Testing and Their Paper-Based Testing Equivalents

<table>
<tr><th>CBT Tool and Guidelines</th><th>PBT Equivalent and Guidelines</th></tr>
<tr><td>Answer Eliminator

Before Testing:
• Assignment in iTester: not assigned prior to testing; available to all students

During Testing: The Answer Eliminator tool allows students to eliminate response option(s) by placing a strike though over the option.</td><td>Removable Markers

Before Testing:
• Materials: Test Administrator provides students with blank masking cards/markers

During Testing: The student may cover or uncover answer options with external blank masking cards as needed).</td></tr>
<tr><td colspan="2">Blank Scratch Paper

Before Testing:
• Assignment in iTester: not assigned prior to testing; available to all students
• Materials: Test Administrators must supply at least one page of blank scratch paper (i.e., either unlined, lined, or graph) per student, per unit. If graph paper is used during mathematics instruction, it is recommended that schools provide graph paper as scratch paper for mathematics units. Students with visual impairments may also use braille paper, raised line paper, bold line paper, raised line graph paper, bold line graph paper, abacus, or Math Window.

During Testing: The student uses blank scratch paper (lined, un-lined, or graph) to take notes and/or work through items during testing. Additional pages may be provided as needed. Students are not required to write their names on scratch paper.

After Testing: Test Administrators are responsible for collecting ALL scratch paper after testing is completed to be securely destroyed. Scratch paper must be securely shredded if it has been used. Schools may reuse unused scratch paper only if paper is completely blank.</td></tr>
<tr><td>Bookmark

Before Testing:
• Assignment in iTester: not assigned prior to testing; available to all students

During Testing: Students can bookmark or save items to come back to later.</td><td>Place Markers

Before Testing:
• Materials: Test Administrator provides student with place markers.

During Testing: The student uses non-sticky place markers to “bookmark” items to review later. All place markers must be removed before test booklet or answer document is submitted for scoring.</td></tr>
</table>

CBT Tool and Guidelines	PBT Equivalent and Guidelines
Calculator – Mathematics (Calculator Sessions) **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** An embedded calculator is available to students taking calculator sessions of the mathematics tests. See the TAM for more information on the calculators available for each grade.	Same as CBT **Before Testing:** • Materials: Test Administrator provides students with handheld calculators for the appropriate grades/sessions. See the TAM for more information on the calculators available for each grade. **During Testing:** Students use handheld calculators on the calculator sessions of the mathematics tests.
Expand Passage **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** Stimulus passages can be expanded.	n/a
Line Reader **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** The Line Reader tool can be used to assist in reading by raising and lowering the tool for each line of text onscreen. It is resizable and draggable.	***Straight Edge*** **Before Testing:** • Materials: Test Administrator provides student with blank straight edge. **During Testing:** The student uses a blank straight edge as he or she reads and follows along with the text
Note Pad **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** The Notepad tool can be used to type notes for each separate test question. The Notepad can be moved around on the screen and resized as desired.	see ***Blank Scratch Paper***
Pop-up Glossary **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** Students can view definitions of pre-selected words by selecting words with a book icon to launch a pop-up screen with the word's definition.	***Glossary in Footnotes*** **During Testing:** The student refers to a glossary of pre- selected, construct-irrelevant words in the footnotes of the paper-based test.

CBT Tool and Guidelines	PBT Equivalent and Guidelines
Reference Sheet **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** Available for grade 11 science (English) and grades 5, 8, and 11 (Spanish) only. The reference sheet contains the Periodic Table for grade 11 science tests. An additional reference sheet for grades 5, 8, and 11 Spanish language science tests contains grade-appropriate Spanish-English glossaries of science terms. Students can use the information in the reference sheet to help answer some test questions.	same as CBT **Before Testing:** • Materials: Test Administrator provides printed reference sheets to students taking grade 11 English language science tests or grade 5, 8, or 11 Spanish language science tests. The reference sheet provides a periodic table for students taking grade 11 tests. Additional printed reference sheets containing grade-appropriate Spanish-English glossaries of science terms are provided to students taking Spanish language science tests. **During Testing:** Students can use the information in the reference sheet to help answer some test questions.
Sketch Tool (Not available on constructed response items) **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** The Sketch tool can be used to sketch, highlight, or underline text on the screen. This tool will only appear on items that do not have a sketchpad widget or rich text editor response option.	see ***Blank Scratch Paper***
Text Highlight **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** The Text Highlight tool can be used to select text and highlight the selection.	***Highlighter*** **Before Testing:** • Materials: Test Administrator provides student with highlighter(s). **During Testing:** The student highlights text as needed to recall and/or emphasize.
Zoom View (magnifier) **Before Testing:** • Assignment in iTester: not assigned prior to testing; available to all students **During Testing:** Students can magnify the entire screen in four increments: 100%, 150%, 200%, and 300%.	***Magnification/Enlargement Device*** **Before Testing:** • Materials: Test Administrator provides student with magnification/enlargement device. **During Testing:** The student uses external magnification or enlargement devices to increase the font or graphic size (e.g., projector, CCTV, eye-glass mounted or hand-held magnifiers, electronic magnification systems, etc.).

Accessibility Features

Table 3 lists the accessibility features available to students through the computer-based testing platform as well as the equivalent resources for paper-based testing. **For students taking computer-based tests, accessibility features must be assigned in iTester prior to testing.**

Table 3: Accessibility Features for Computer-Based and Paper-Based Testing

CBT Features and Guidelines	PBT Equivalent and Guidelines
Answer Masking **Before Testing:** • Assignment in iTester: must be assigned prior to testing **During Testing:** The Answer Masking tool allows students to hide answer options to help narrow down the correct answer.	***Removable Markers*** **Before Testing:** • Materials: Test Administrator provides students with blank masking cards/markers **During Testing:** The student may cover or uncover answer options with external blank masking cards as needed).
Color Contrast **Before Testing:** • Assignment in iTester: must be assigned prior to testing **During Testing:** Students can choose a text and background color from a set of 12 predefined color combinations.	***Colored Overlays*** **Before Testing:** • Materials: Test Administrator provides students with colored overlays. **During Testing:** The student uses colored overlays when taking the assessment. The color should match what is currently used during instruction.
Custom Masking **Before Testing:** • Assignment in iTester: must be assigned prior to testing **During Testing:** Provides the ability to mask certain parts of the test interface or question.	***Removable Markers*** **Before Testing:** • Materials: Test Administrator provides students with blank masking cards/markers **During Testing:** The student may cover or uncover answer options with external blank masking cards as needed).
Text-to-Speech (English or Spanish) **Before Testing:** • Assignment in iTester: must be assigned prior to testing **During Testing:** Students can play, pause, skip, or stop audio. They can select specific text for on-demand audio, and the Gear icon allows students to change the volume or speed of the text being read aloud.	***Human Reader (English or Spanish)*** **Before Testing:** • Materials: Human Reader Kits, which include one copy of the student test booklet (and answer document for grades 4–8) and an extra test booklet for Test Administrators. • Test Administrator Training: Test Administrators providing these accommodations must review the following, as applicable: • Human Reader Kits at least two school days prior to paper-based testing, with

CBT Features and Guidelines	PBT Equivalent and Guidelines
	kits provided to schools for this purpose. Review of Human Reader Kits must occur in a SECURE ENVIRONMENT. • Appendix A: Test Administration Protocol for the Human Reader Accommodation for English Language Arts (ELA) Assessments, and the Human Reader Accessibility Feature for Mathematics Assessments. • Appendix I: The 2023 Math and ELA Assessments for Students with Visual Impairments, Including Blindness. **During Testing:** A student receives an audio representation of the mathematics assessment through a human reader.
Reverse Contrast **Before Testing:** • Assignment in iTester: must be assigned prior to testing **During Testing:** Inverts color values on the screen.	n/a

Accommodations for Students with Disabilities and English Learners

Overview

It is important to ensure that performance in the classroom and on assessments is influenced minimally, if at all, by a student's disability or linguistic/cultural characteristics that are unrelated to the content being assessed. For the 2023 Science, Math, and ELA assessments, accommodations are considered to be adjustments to the testing conditions, test format, or test administration that provide equitable access during assessments for students with disabilities and students who are ELs. In general, only accommodations that are used in daily instruction should be assigned on a summative assessment. Under no circumstance should the assessment be the first time a student is exposed to a particular accommodation. In addition, Test Administrators administering the assessment or providing accommodations should be an education professional who is familiar with the student, and who is typically responsible for providing the accommodation in the classroom. To the extent possible, accommodations should:

- Provide equitable access during instruction and assessments;
- Mitigate the effects of a student's disability;
- Not reduce learning or performance expectations;
- Not change the construct being assessed; and
- Not compromise the integrity or validity of the assessment.

Accommodations are intended to reduce and/or eliminate the effects of a student's disability and/ or English language proficiency level; however, **accommodations should never reduce learning expectations by reducing the scope, complexity, or rigor of an assessment.** Moreover, accommodations provided to a student on the 2023 Science, Math, and ELA assessments must be generally consistent with those provided for classroom instruction and classroom assessments. There are some accommodations that may be used for instruction or for formative assessments but are not allowed for the summative assessment because they impact the validity of the assessment results – for example, allowing a student to use a thesaurus or access the internet during a 2023 Science, Math, and ELA assessment. There may be consequences (e.g., invalidating a student's test score) for the use of non-allowable accommodations during the 2023 Science, Math, and ELA assessments. It is important for educators to become familiar with policies regarding accommodations used for the 2023 Science, Math, and ELA assessments.

The guidelines provided in this manual are intended to ensure that valid and reliable scores are produced on the 2023 Science, Math, and ELA assessments, and that an unfair advantage is not given to students who receive accommodations. Outside of the guidance provided in this manual, changes to an accommodation or the conditions in which it is provided may change what the assessment is measuring, and will likely call into question the reliability and validity of the results regarding what a student knows and is able to do as measured by the assessment.

To the extent possible, accommodations should adhere to the following principles:

- Accommodations enable students to participate more fully and fairly in instruction and assessments and to demonstrate their knowledge and skills.
- Accommodations should be based upon an individual student's needs rather than on the category of a student's disability, level of English language proficiency alone, level of or access to grade-level instruction, amount of time spent in a general classroom, current program setting, or availability of staff.
- Accommodations should be based on a documented need in the instruction/assessment setting and should not be provided for the purpose of giving the student an enhancement that could be viewed as an unfair advantage.
- Accommodations for students with disabilities should be described and documented in the student's appropriate plan (i.e., either the IEP or 504 plan).
- Accommodations for ELs should be described and documented.

- Students who are ELs with disabilities qualify to receive accommodations for both students with disabilities and ELs.
- Accommodations should become part of the student's program of daily instruction as soon as possible after completion and approval of the appropriate plan.
- Accommodations should not be introduced for the first time during the testing of a student.
- Accommodations should be monitored for effectiveness.
- Accommodations used for instruction should also be used, if allowable, on local district assessments and state assessments.

Providing an accommodation feature that is not documented in a student's IEP or administering an assessment without documented accommodations will result in the invalidation of a test session.

TAs and proctors should be familiar with individual student needs regarding appropriate accommodations before a student begins and completes a test session in order to mitigate the need for testing invalidation.

If a student completes less than three questions in a test session with or without the appropriate accommodations, the DTC (or STC) should stop the session, complete a testing irregularity report in the DTC portal, request a test invalidation, and reschedule the session under the appropriate conditions. If more than three questions are completed, PED authorization of the invalidation will be required to re-schedule the session and re-test the student.

Scoring and Reporting

Summative assessment scores for students who receive any of the accommodations listed in this manual will be aggregated with the scores of other students and those of relevant groups, and can be included for accountability purposes.

Unique Accommodations

PED has developed a comprehensive list of accessibility features and accommodations that are designed to increase access to the 2023 Science, Math, and ELA assessments and will result in valid, comparable assessment scores. However, students with disabilities or ELs may require additional accommodations that are not found in this manual. PED will individually review requests for unique accommodations in their respective state on an individual basis and will provide approval after determining whether the accommodation would result in a valid score for the student. Refer to Appendix D: Unique and Emergency Accommodations.

Emergency Accommodations

An emergency accommodation may be appropriate for a student who incurs a temporary disabling condition that interferes with test performance shortly before or during the assessment window. A student who does not have an IEP or 504 plan may require an accommodation as a result of a recently- occurring accident or illness. Cases include students who have a recently-fractured limb (e.g., arm, wrist, shoulder); whose only pair of eyeglasses has broken; or a student returning after a serious or prolonged illness or injury. An emergency accommodation should be given only if the accommodation will result in a valid score for the student (i.e., does not change the construct being measured by the test[s]). If the principal (or designee) determines that a student requires an emergency accommodation on the 2023 Science/Math/ELA assessment, a Nonstandard Accommodation Request Form must be completed within the DTC portal. If approved, the form must be kept on file. Requests for emergency accommodations will be approved after it is determined that use of the accommodation would result in a valid score for the student. Refer to Appendix D: Unique and Emergency Accommodations.

Student Refusal Form

If a student refuses an accommodation listed in his or her IEP, 504 plan, or if required, an EL plan, the school should document in writing that the student refused the accommodation, and the accommodation must be offered and remain available to the student during testing. This form must be completed and placed in the student's file and a copy must be sent to the parent on the day of refusal. Principals (or designee) should work with Test Administrators to determine who, if any others, should be informed when a student refuses an accommodation documented in an IEP, 504 plan, or if required, an EL plan. Refer to Appendix E: Student Accommodation Refusal Form.

Ongoing Research and Data Collection on Use of Accommodations

PED will continue to research the effectiveness, validity, differential impact, relevance, and feasibility of the accommodations, and revise as needed.

Accommodations for Students with Disabilities

Table 4 lists the ACCOMMODATIONS for students with disabilities that describe changes in the assessment format and method in which the assessment is administered. The table also outlines the before, during, and after testing activities necessary to successfully administer these accommodations. Accommodations for students with disabilities must be assigned to the student in the iTester portal before testing. This information is included in the "before testing" guidance.

Table 4: Accommodations for Students with Disabilities (IEP, 504)

<table>
<tr><th>CBT Accommodation and Guidelines</th><th>PBT Accommodation and Guidelines</th></tr>
<tr><td colspan="2">Allow Accessibility Mode Testing

(See Assistive Technology Device Presentation [Non-Screen Reader], Assistive Technology Device Responses)</td></tr>
<tr><td colspan="2">Assistive Technology Device Presentation (Non-Screen Reader),
Assistive Technology Device Responses

Before Testing:
• Assignment in iTester: must be assigned prior to testing
  • Note: Test coordinators should ensure the Allow Accessibility Mode (AAM) accommodation is turned on for all students who will require Windows-based third-party accessibility software.
• Testing: Assistive technology should be tested using a practice test to determine whether the assistive technology will interact with iTester and can be used successfully during computer-based testing. For more information, refer to the Testing With Third Party Assistive Technology guidelines available here: newmexico.onlinehelp.cognia.org/cbt-guides/.

During Testing: Students may use a range of assistive technologies on the 2023 Science/Math/ELA assessments, including devices that are compatible with the online testing platform, and those that are used externally on a separate computer.

After Testing: Test Administrators are responsible for collecting all nonscorable student work created from assistive technology devices. Content must be cleared off all devices. Paper nonscorable student work must be securely shredded.

For PBT administration, responses must be transcribed verbatim by a test administrator in a standard student test booklet or answer document. Only transcribed responses will be scored. Refer to Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses.</td></tr>
<tr><td>ASL Videos (Mathematics or ELA)

(see also Presentation Options for ELA)

Before Testing:
• Assignment in iTester:
  ○ must be assigned prior to testing
  ○ If a student does not use ASL, a human interpreter and separate test setting will be required.</td><td>see Human Signer</td></tr>
</table>

<table>
<tr><th>CBT Accommodation and Guidelines</th><th>PBT Accommodation and Guidelines</th></tr>
<tr><td>During Testing: The student views an embedded video of a human interpreter. The student may pause and resume the video but cannot adjust the pace.</td><td></td></tr>
<tr><td colspan="2">Basic/Scientific Calculator on Non-Calculator Sections of the Mathematics Test

(See also Mathematics Tools [Non-Calculator Sections])
Before Testing:
• Assignment in iTester: must be assigned prior to testing
• Materials: for PBT administration, the TA provides students with handheld calculators for the appropriate grades/sections, as follows:
  • grades 3–5, all sessions: a four-function calculator with square root and percentage functions
  • grades 6–7, Session 1, Section A: a four-function calculator with square root and percentage functions
  • grade 8, Session 1: a scientific calculator

During Testing:
• For CBT administration, the student has access to the embedded basic or scientific calculator (depending on grade) while taking the non-calculator section(s) of the computer-based test mathematics test.
• For PBT administration, the student uses an appropriate handheld calculator.</td></tr>
<tr><td colspan="2">Braille Notetaker,
Braille Writer

Before Testing:
• Assignment in iTester: must be assigned prior to testing

During Testing: A student who is blind or has a visual impairment may use an electronic braille notetaker or braille writer. The grammar checker, internet, and stored file functionalities must be turned off. For students using braille forms, the Test Administrator directions for filling in a circle, making marks, and erasing do not apply. Students should number their responses to be sure that their answers can be transcribed accurately into a scorable test booklet, answer document, or iTester.

After Testing:
• Student responses generated using an electronic braille notetaker or braille writer must be transcribed verbatim by a Test Administrator into the student's standard test booklet, answer document, or iTester. Only transcribed responses will be scored. Responses must be transcribed by the teacher of the student with visual impairment or a Test Administrator supervised by the teacher of the student with visual impairment.
• Refer to Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses.
Test Administrators are responsible for collecting all nonscorable student work created using assistive technology devices. Test-related content must be deleted from all devices. Nonscorable student work must be securely shredded</td></tr>
<tr><td>ELA Text-to-Speech English

(see Presentation Options for ELA)</td><td>See Human Reader, Human Signer</td></tr>
</table>

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

Headphones as Noise Buffer

Before Testing:

- Assignment in iTester: must be assigned prior to testing
- Materials: Test Administrator provides student with headphones.

During Testing: The student uses headphones or noise buffers to minimize distraction or filter external noise during testing. If headphones are used only as noise buffers, they should not be plugged into the student's device.

Human Reader (English or Spanish)

(see also Presentation Options for ELA)

Before Testing:

- Assignment in iTester: must be assigned prior to testing
- Materials: Human Reader Kits, which include one copy of the student test booklet (and answer document for grades 4–8) and an extra test booklet for Test Administrators.
- Test Administrator Training: Test Administrators providing these accommodations must do the following, as applicable:
  - Review Human Reader Kits at least two school days prior to paper-based testing, with kits provided to schools for this purpose. Review of Human Reader Kits must occur in a SECURE ENVIRONMENT.
  - Review Appendix A: Test Administration Protocol for the Human Reader Accommodation for English Language Arts (ELA) Assessments, and the Human Reader Accessibility Feature for Mathematics Assessments.
  - Review Appendix I: The 2023 Math and ELA Assessments for Students with Visual Impairments, Including Blindness.

During Testing: A human reader will read the test to a student. The student may either be tested in a small group or a separate setting based on the student's experiences during classroom assessments.

Human Scribe

(see Response Options)

Human Signer

(see also Presentation Options for ELA)

Before Testing:

- Assignment in iTester: must be assigned prior to testing
- Test Administrator Training: Human Signers must review:
  - Test administration scripts included in the *Test Administrator's Manual.*
  - Appendix H: Human Signer Guidelines (signers only).

During Testing: A human signer will sign the test to a student. The student may either be tested in a small group or a separate setting based on the student's experiences during classroom assessments.

Human Signer for Test Directions

Before Testing:

- Assignment in iTester: must be assigned prior to testing
- Test Administrator Training: Human Signers must review:
  - Test Administrator Scripts included in the *Test Administrator's Manual.*
  - Appendix H: Human Signer Guidelines (signers only).

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

During Testing: A human signer will sign the test directions to a student. The student may either be tested in a small group or a separate setting based on the student's experiences during classroom assessments.

Mathematics Tools (Non-Calculator Sections)

Before Testing:

- Purpose: The purpose of the mathematics tools on the non-calculator sections accommodation is to provide access for students with a disability that *severely limits or prevents* their ability to perform basic calculations (i.e., student is unable to perform single-digit addition, subtraction, multiplication, or division). For these students, a calculation device may be used on the non-calculator AND calculator sections of the mathematics assessments. The IEP or 504 plan must specify which device(s) or manipulatives.
- Assignment in iTester: must be assigned prior to testing
- Materials:
    - Allowable mathematics tools include:
        - Arithmetic tables (e.g., addition charts, subtraction charts, multiplication charts; division charts).
        - Two-color chips (e.g., single-sided or double- sided).
        - Counters and counting chips.
        - Square tiles.
        - Base 10 blocks.
        - 100s chart.

A student with a visual impairment may need other mathematics tools, such as a large print ruler (embedded ruler is designed in 18 point font), braille ruler, tactile compass, or braille protractor.

During Testing: A student uses a calculation device (e.g., four-function calculator, large key, or other adapted calculator), arithmetic table (including addition/subtraction and/or multiplication/division charts), and/or manipulatives (IEP or 504 plan must specify which device or manipulative) on the NON-CALCULATOR SECTIONS of the mathematics assessments. If a talking calculator is used, the student must use headphones or be tested in a separate setting.

Important Guidelines for identifying students to receive this accommodation: IEP teams and 504 Plan Coordinators should carefully review the following guidelines before identifying students to receive this accommodation. If all guidelines are NOT met, and the student is given Calculation Device and Mathematics Tools without proper documentation, the student's assessment score may be invalidated and the score would not be counted in the overall assessment results (i.e., the student would be considered a "non-participant" for the mathematics assessment.)

In making decisions whether to provide the student with this accommodation, IEP teams and 504 Plan Coordinators should consider whether the student has:

- A disability that *severely limits or prevents* the student's ability to perform basic calculations (i.e., single-digit addition, subtraction, multiplication, or division), even after varied and repeated attempts to teach the student to do so.

Before listing the accommodation in the student's IEP/504 plan, teams should also consider whether:

- The student is unable to perform calculations without the use of a calculation device, arithmetic table, or manipulative during routine instruction.
- The student's inability to perform mathematical calculations is documented in evaluation summaries from locally-administered diagnostic assessments.
- The student receives ongoing, intensive instruction and/or interventions to learn to calculate without using a calculation device, in order to ensure that the student continues to learn basic calculation and fluency.

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

For a student who receives this accommodation, no claims should be inferred regarding the student's ability to perform basic mathematical calculations without the use of a calculator.

Paper-Based Edition

Before Testing:

- Assignment in iTester: not assigned/documented in iTester
- Materials: Paper-Based Edition of the assessment

During Testing: For schools administering the computer-based assessments, a paper-based assessment is available for students who (1) are unable to take a computer-based assessment due to a disability; (2) recently entered the school and has very little or no prior experience or familiarity with technology; (3) attend a school providing paper-based assessments as the primary mode; or (4) are unable to access an online assessment due to religion or beliefs.

Paper-Based Edition Braille

Before Testing:

- Assignment in iTester: not assigned/documented in iTester
- Materials: Braille Kits are required for administration. Braille Kits include Test Administrator Braille Scripts, one copy of the student's paper Braille Assessment, and a standard test booklet or answer document for transcription.
- Test Administrator Training: Test Administrators of students with visual impairments must review:
    - Braille Kits, which will be provided to schools at least two full school days prior to testing in a SECURE ENVIRONMENT for the Test Administrator to verify that the braille code is accurate on the test booklet cover and review the braille test administration scripts, including information specific to administering paper-based braille. **Important:** Reading, viewing, copying, or reproducing passages or test items is prohibited.
    - Appendix I: The 2023 Math and ELA Assessments for Students with Visual Impairments, Including Blindness.
    - If needed by the student, braille test booklets or answer documents may be disassembled for testing (but must be reassembled for return). It is critical that Test Administrators count the number of pages in the test booklet or answer document prior to disassembling the test booklets or answer documents to help ensure that all pages are returned.

During Testing: A student who is blind or has a visual impairment and is unable to take the computer-based test with a refreshable braille display may take the ELA and mathematics assessments using the hard-copy contracted braille edition. Tactile graphics are already embedded in the hard copy braille edition. For students using braille forms, the Test Administrator directions for filling in a circle, making marks, and erasing do not apply. Students should number their responses to be sure that their answers can be transcribed accurately into a scorable test booklet or answer document.

After Testing:

- Responses must be transcribed verbatim by a Test Administrator in a standard student test booklet or answer document, which is included in the Braille Test Kit. Only transcribed responses will be scored.
- Refer to Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses for protocol.
- Test Administrators are responsible for collecting all nonscorable student work created from assistive technology devices. Content must be deleted off all devices. Nonscorable student work must be securely shredded.

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

- If the braille test booklet or answer document was disassembled, it must be reassembled for return. To reassemble test booklets or answer documents, the Test Administrator may staple or binder clip all pages for return. Failure to return all pages will be considered a breach of security.

Paper-Based Edition Large Print

Before Testing:

- Assignment in iTester: not assigned/documented iTester
- Materials: Large Print Test Kits include a large print test booklet and a standard test booklet or answer document for transcription.
- Test Administrator Training: Test Administrators of students with visual impairments must review:
  - Appendix I: The Spring 2023 Math and ELA Assessments for Students with Visual Impairment, Including Blindness.

During Testing: A large print paper-based form of each assessment is available for a student with a visual impairment who is unable to take a computer-based assessment. The font size for the large print edition will be 18 point on paper sized 11" x 17". Students will not record their answers in standard print test booklets or answer documents. Instead, students will circle their answers in a large print test booklet. For constructed response items, students will write their answers on the lines provided in their large print test booklets. In mathematics, students will need to write their answers in boxes at the top of the answer grids, but they do not need to bubble in their answers. Test Administrators should refer to the TAM Scripts for instances where they should demonstrate an activity or display information. Demonstrations should be conducted where they are visible for each student (e.g., on the board, near the student).

After Testing:

- Responses must be transcribed verbatim by a Test Administrator in a standard student test booklet or answer document, which is included in the Large Print Test Kit. Only transcribed responses will be scored. At least two persons must be present during transcription of student responses (one transcriber and one observer confirming accuracy). It is recommended that one of the individuals be a District Test Coordinator or School Test Coordinator. Refer to Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses.

Presentation Options for ELA

- ***ELA Text-to-Speech English***
- ***ASL Video (ELA)***
- ***Human Reader***
- ***Human Signer***

Before Testing:

- Purpose: The purpose of the text-to-speech, ASL video, Human Reader, and Human Signer accommodations for the ELA assessment is to provide access to printed or written texts on the ELA assessments for a very small number of students with print-related disabilities who would otherwise be unable to participate in the assessment because their disability s*everely limits or prevents* their ability to access printed text by decoding. This accommodation is not intended for students reading somewhat (i.e., only moderately) below grade level.
- Assignment in iTester: must be assigned prior to testing
- Tools for Identification: IEP teams/504 Plan Coordinators should use the decision-making tool available in Appendix C: Text-to-Speech, ASL Video, or Human Reader/Human Signer Guidance for English Language Arts (ELA) Assessments to inform their decision-making.
- Materials: Human Reader Kits, which include one copy of the student test booklet (and answer document for grades 4-8) and an extra test booklet for Test Administrators (Human Reader/Signer).

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

- Test Administrator Training: Test Administrators providing these accommodations must review the following, as applicable:
  - Human Reader Kits at least two school days prior to paper-based testing, with kits provided to schools for this purpose. Review of Human Reader Kits must occur in a SECURE ENVIRONMENT.
  - Appendix A: Test Administration Protocol for the Human Reader Accommodation for English Language Arts (ELA) Assessments, and the Human Reader Accessibility Feature for Mathematics Assessments.
  - Appendix F: ELA Audio Guidelines.
  - Appendix H: Human Signer Guidelines (signers only).
  - Appendix I: The 2023 Math and ELA Assessments for Students with Visual Impairments, Including Blindness.
  - The *Kiosk User Guide*, available at newmexico.onlinehelp.cognia.org/cbt-guides/, for Text-to-Speech functionality

During Testing: A student receives an audio representation of the ELA assessment either through embedded text-to-speech, embedded ASL video, or a Human Reader/Signer. For Human Reader, the Test Administrator will need to reference Appendix F: ELA Audio Guidelines. **Note:** If headphones are *not* used for text-to-speech, or the student has a Human Reader or Signer, the student must be tested in a separate setting.

Important Guidelines on identifying students for these accommodations: IEP teams and 504 Plan Coordinators should carefully review the following guidelines before identifying students to receive these accommodations on the ELA assessments. If all guidelines are NOT met, and the student is given the text-to-speech, ASL video, or Human Reader/Human Signer accommodation on an English language arts (ELA) assessment, the student's assessment score may be invalidated and the score would not be counted in the overall assessment results (i.e., the student would be considered a "non-participant" for the English language arts (ELA) assessment.)

In making decisions on whether to provide a student with this accommodation, IEP teams and 504 Plan Coordinators should consider whether the student has:

- Blindness or a visual impairment and has not learned (or is unable to use) braille;
  OR
- A disability that *severely limits or prevents* him/her from accessing printed text, even after varied and repeated attempts to teach the student to do so (e.g., student is unable to decode printed text);
  OR
- Deafness or a hearing impairment and is severely limited or prevented from decoding text due to a documented history of early and prolonged language deprivation.

Before listing the accommodation in the student's IEP or 504 plan, teams/ coordinators should consider whether:

- The student has access to printed text during routine instruction through a reader, other spoken-text audio format, or signer;
- The student's inability to decode printed text or read braille is documented in evaluation summaries from locally-administered diagnostic assessments; and the student receives ongoing, intensive instruction and/or interventions in the foundational reading skills to continue to attain the important college and career-ready skill of independent reading.

Decisions about who receives this accommodation will be made by IEP teams and 504 Plan Coordinators. For a student who receives one of these accommodations, no claims should be inferred regarding the student's ability to demonstrate foundational reading skills (i.e., decoding).

<table>
<tr><th>CBT Accommodation and Guidelines</th><th>PBT Accommodation and Guidelines</th></tr>
<tr><td colspan="2">Read Aloud to Self

Before Testing:
• Assignment in iTester: must be assigned prior to testing

During Testing: The student reads aloud the assessment to themselves. Students may use an external device such as a whisper phone. The student must be tested in a separate setting.</td></tr>
<tr><td>Refreshable Braille Display with Screen Reader

Before Testing:
• Assignment in iTester: must be assigned prior to testing
• Materials and Equipment: iTester screen reader compatibility has been tested with JAWS 19 and 20; for optimal screen reader usage, PED recommends using JAWS 19 or 20. A braille testing kit is required for test administration.
• Screen Reader Testing: Screen reader software SHOULD be tested using a practice test to determine whether the assistive technology will interact with iTester and can be used successfully during computer-based testing. For more information, refer to the Testing With Third Party Assistive Technology guidelines available here: newmexico.onlinehelp.cognia.org/cbt-guides/.
• Test Administrator Training: Test Administrators should review Appendix I: The 2023 Science, Math, and ELA Assessments for Students with Visual Impairments, Including Blindness.

During Testing: A student who is blind or has a visual impairment takes the Mathematics or ELA assessments using his or her preferred screen reader software with a refreshable braille display. A student who uses a screen reader with refreshable braille will also need a tactile graphics booklet, which contains only the graphics portion of test questions and visual descriptions of pictures and multimedia where applicable. If the student is not using headphones, the student must be tested in a separate setting.

After Testing: Tactile graphics booklets contain secure item content and should be handled as secure test materials. Test Administrators should</td><td>See Paper form Braille</td></tr>
</table>

<table>
<tr><th>CBT Accommodation and Guidelines</th><th>PBT Accommodation and Guidelines</th></tr>
<tr><td>return tactile graphics to Test Coordinators. Test Coordinators must return tactile graphics with the nonscorable materials.</td><td></td></tr>
<tr><td colspan="2">Response Options
Speech-to-Text
Human Scribe
Before Testing:
Assignment in iTester:
must be assigned prior to testing
If a student is using an allowable 3rd party external Assistive Technology that provides speech-to-text functionality that will interact with iTester, see Assistive Technology Device Responses for additional information.
Materials: External device provided by the student, if needed. If the student uses speech-to-text software, such as Dragon® Naturally Speaking, then a separate computer must be provided; one to run the assessment on iTester and a second computer to run the software. iTester does not contain embedded speech-to-text software.
Test Administrator Training: Test Administrators providing the scribe accommodation must review:
Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses.
During Testing: Student dictates responses either verbally, using an external speech-to-text device, an augmentative/assistive communication device (e.g., picture/word board), or by dictating, signing, gesturing, pointing, or eye-gazing. The student must be tested in a separate setting. The student must be familiar with any assistive technology external device used for test administration. Note: iTester does not have embedded Speech-to-Text functionality—students must use allowable Assistive Technology or an external third party device (responses must be transcribed).

After Testing:
Responses must be transcribed exactly as dictated/signed (e.g., the human scribe may not change, embellish, or interpret a student's responses when transcribing) into the student's standard test booklet or answer document. Only transcribed responses will be scored.
Refer to Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses.
Test Administrators are responsible for collecting all paper nonscorable student work created using assistive technology devices. Test-related content must be deleted from all devices. Nonscorable student work must be securely shredded.</td></tr>
<tr><td>Screen Reader

Before Testing:
Assignment in iTester:
must be assigned in prior to testing
For ELA, the student does not use a refreshable braille display or hard copy braille edition because they have either not yet learned, or are unable to use, braille.</td><td>See Paper form Braille</td></tr>
</table>

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines
• Materials and Equipment: iTester screen reader compatibility has been tested with JAWS 19 and 20; for optimal screen reader usage, PED recommends using JAWS 19 or 20. A braille testing kit is required for test administration. • Screen Reader Testing: Screen reader software SHOULD be tested using a practice test to determine whether the assistive technology will interact with iTester and can be used successfully during computer-based testing. For more information, refer to the *Testing With Third Party Assistive Technology* guidelines available here: newmexico.onlinehelp.cognia.org/cbt-guides/. • Test Administrator Training: Test Administrators should review Appendix I: The 2023 Science, Math, and ELA Assessments for Students with Visual Impairments, Including Blindness. **During Testing:** A student who is blind or has a visual impairment takes the assessments using his or her preferred screen reader software. A student who uses a screen reader will also need a tactile graphics booklet, which contains only the graphics portion of test questions and visual descriptions of pictures and multimedia, where applicable. If the student is not using headphones, the student must be tested in a separate setting. **After Testing:** Tactile graphics booklets contain secure item content and should be handled as secure test materials. Test Administrators should return tactile graphics to Test Coordinators. Test Coordinators must return tactile graphics with the nonscorable materials.	
Speech-to-Text (see Response Options)	See ***Human Scribe, Human Signer***
Tactile Graphics **Before Testing:** • Assignment in iTester: must be assigned prior to testing • See **Screen Reader** for additional information. **During Testing:** • A student who is blind or has a visual impairment who uses a screen reader or refreshable braille will also need a braille kit in order to access tactile graphics.	

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

- Tactile graphics will be embedded in the braille Paper Form assessments, when needed.

After Testing: Braille booklets contain secure item content and should be handled as secure test materials. Test Administrators should return braille materials to Test Coordinators. Test Coordinators must return braille materials with the nonscorable materials.

Word Prediction (external)

Before Testing:

- Assignment in iTester:
  - must be assigned prior to testing
  - If a student is using an allowable 3rd party external Assistive Technology that provides speech-to-text functionality that will interact with iTester, see **Assistive Technology Device Responses** for additional information.
- Materials: External Word Prediction Device.

During Testing: The student uses an external word prediction device that provides a bank of frequently- or recently-used words on-screen after the student enters the first few letters of a word. The student must be familiar with the use of the external device prior to assessment administration. The device may not connect to the internet or save information.

After Testing:

- Student responses generated using the External Word Prediction Device software must be transcribed verbatim by a Test Administrator into iTester. Only transcribed responses submitted in iTester will be scored. **Note:** If the student is writing his/her responses directly into iTester through the external software for word prediction, then transcribing is not necessary.
- Refer to Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses.
- Test Administrators are responsible for collecting all nonscorable student work created using external word prediction device software. Test-related content must be deleted from all devices. Nonscorable student work must be securely shredded.

Important Guidelines for identifying students to receive this accommodation: IEP teams and 504 Plan Coordinators should carefully review the following guidelines before identifying a student to receive this accommodation.

In making decisions whether to provide the student with this accommodation, IEP teams and 504 Plan Coordinators are instructed to consider whether the student has:

- A physical disability that *severely limits or prevents* the student from writing or keyboarding responses;
  OR
- A disability that *severely limits or prevents* the student from recalling, processing, and expressing written language, even after varied and repeated attempts to teach the student to do so.

Before listing the accommodation in the student's IEP/504 plan, teams/ coordinators are instructed to consider whether:

- The student's inability to express in writing is documented in evaluation summaries from locally administered diagnostic assessments;
- The student routinely uses a word-prediction device or software during classroom writing assignments; and

The student receives ongoing, intensive instruction, and/or intervention in language processing and writing, as deemed appropriate by the IEP team/504 Plan Coordinator.

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines

Word Prediction (Embedded)
 	
Before Testing:
• Assignment in iTester:
  ○ must be assigned prior to testing
  ○ available on English language and Spanish language tests
  ○ available to users on Chromebook, Mac, and Windows
• This accommodation requires extra files to be downloaded to the student's workstation when they log into their test. Therefore it is recommended that students with this accommodation log in a few minutes before or after other students in the test group to minimize the download time.
 	
During Testing: Students will have access to the CoWriter word prediction application in any open-ended items. It does not require a current CoWriter account.
	See ***Word Prediction (external)***

Accommodations for English Learners

Table 5 lists the ACCOMMODATIONS for EL students that describe changes in the assessment format and method in which the assessment is administered. The table also outlines the before, during, and after testing activities necessary to successfully administer these accommodations. Accommodations for students with disabilities must be assigned to the student in the iTester portal before testing. This information is included in the "before testing" guidance.

Table 5: Accommodations for English Learners (EL)

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines
Commercial Word-to-Word Dictionary **Before Testing:** • Assignment in iTester: must be assigned prior to testing • Materials: Word-to-word dictionaries are provided to students by their school, based on those used by the student for routine classroom instruction. **During Testing:** The student uses a published bilingual, word-to-word dictionary that does not include definitions, pronunciation, phrases, sentences, or pictures. The student should be familiar with the dictionary they will use during testing. Students should be given ample time to complete the test using the accommodation. If no printed word-to-word dictionary can be found for a specific language, an electronic translator may be used. The device may not connect to the internet or store information, and therefore, web-based translators are not allowed	
Customized Dual Language Glossary **Before Testing:** • Assignment in iTester: must be assigned prior to testing	

Directions in Native Language
 	
Before Testing:
• Assignment in iTester: must be assigned prior to testing
• Materials:
  ○ The 2023 Science/Math/ELA assessments provide written test administration directions in Spanish
  ○ If written general test administration directions are not available in the student's native language, a local translator fluent both in English and the student's native language may translate and read the directions in the language of the student.
• Test Administrator Training: Test Administrators, or other qualified interpreters, providing the general administration directions in languages other than English must review the directions in advance in order to provide consistent transadaptations. Test Administrators providing this accommodation will ideally be literate and fluent in English, as well as in the student's native language; or may collaborate with a local translator, if available.
 	
During Testing: The Test Administrator, or other qualified interpreter, reads aloud the general administration instructions in the student's native language. The student may request that directions be repeated. The student must be tested in a separate setting.	

Picture Dictionary **Before Testing:** • Assignment in iTester: must be assigned prior to testing	

CBT Accommodation and Guidelines	PBT Accommodation and Guidelines
Pocket Word-to-Word Translator **Before Testing:** • Assignment in iTester: must be assigned prior to testing	
Spanish Language Version **Before Testing:** • Assignment in iTester: o must be assigned prior to testing • Students must be placed in separate iTester class and that class must be assigned the Spanish version of the test when scheduling that class for a test session. Students must change the kiosk to the Spanish version before logging in. • Test Administer Training: Test Administrators providing this accommodation should ideally be literate and fluent in English and Spanish, or may be assisted by a translator, if available, since test administration directions will be read to the student in Spanish. **During Testing:** A student takes the science, mathematics, or English Language Arts assessment with content presented in Spanish **Note:** If the student is also receiving a Human Reader or Text-to-Speech accessibility feature, the test can be read aloud in Spanish only (i.e., the test cannot be read aloud in English in addition to Spanish).	

Table 6 lists the accommodations on 2023 Science/Math/ELA assessments that are available to ELs, cross-referenced with recommendations regarding the effectiveness of the accommodation based on the English Language Proficiency (ELP) level of the student.

Table 6: Guidance on Selection of Accommodations for English Learners on 2023 Science, Math, and ELA Assessments

Accommodations	Most likely to benefit ELs at this ELP Level		
	Beginning	Intermediate	Advanced
Commercial Word-to-Word Dictionary	◉	●	●
Speech-to-Text **Human Scribe**	●	◉	○
Directions in Native Language	●	◉	○
Spanish Language Version	●	◉	○
Paper-Based Edition of the Assessment in Spanish	●	◉	○
Large Print Edition of the Assessment in Spanish	●	◉	○
Text-to-Speech in Spanish **Human Reader Spanish**	●	◉	○

KEY for Table 6:

● Highly recommended for use by ELs at this ELP level
◉ Recommended for use by ELs at this ELP level
○ May not be appropriate for students at this ELP level

Appendix A: Test Administration Protocol for the Human Reader Accommodation for English Language Arts (ELA) Assessments, and the Human Reader Accessibility Feature for Mathematics Assessments

In cases where a student requires a text-to-speech accommodation on the English language arts (ELA) and/or a text-to-speech accessibility feature on the mathematics assessments, but cannot participate in the computer-based assessment and takes the paper-based assessment instead, a Human Reader must provide the accommodation to the student. Human Readers who provide the accommodation to a student on the English language arts (ELA) or the accessibility feature on the mathematics assessments must follow these procedures during testing to ensure the standardization of the oral presentation of the assessments.

Procedures for Human Readers Providing the Human Reader Accommodation for ELA Assessments or the Human Reader Accessibility Feature for the Mathematics Assessments

1. Readers must be trained locally to administer each assessment, as indicated in the *Test Administrator Manual (TAM). Readers must sign the Staff Confidentiality Agreement available at* **webnew.ped.state.nm.us/bureaus/assessment-3/district-test-coordinator/.**
2. Readers must speak in a clear and consistent voice throughout the test administration, using correct pronunciation, and without vocal inflections that may provide clues to, or mislead, a student.
3. Readers should be provided a Human Reader Kit (which includes a copy of the test and the test administrator's directions) two school days prior to the start of testing, in order to become familiar with the words, terms, symbols, signs, and/or graphics that will be read aloud to the student. Readers must also refer to Appendix F: ELA Audio Guidelines and/or Appendix G: Mathematics Audio Guidelines to ensure consistency in how items are read. **Note: Review of Human Reader Kits must occur in a SECURE ENVIRONMENT**.
4. Readers must read verbatim (word for word) only what is printed in the test book (or in rare cases, on the computer screen) without changing, emphasizing, or adding words. Readers may not clarify (except for test directions), provide additional information, assist, or influence the student's selection of a response in any way.
5. Readers should emphasize only the words printed in boldface, italics, or capital letters and inform the student that the words are printed that way. No other emphasis or vocal inflection is permitted.
6. Readers may repeat passages, test items, and response options, as requested, according to the needs of the student. Readers should not rush through the test and should ask the student if they are ready to move to the next item.
7. Readers may not attempt to solve mathematics problems, or determine the correct answer to a test item while reading, as this may result in pauses or changes in inflection which may mislead the student.
8. Readers must attempt to maintain a neutral facial expression, neither smiling nor frowning during the test, which may be interpreted by the student as approval or disapproval of the student's answers.
9. Readers must be familiar with the student's IEP or 504 plan, and should know in advance which accommodations are required by the student, and for which test (ELA and/or Mathematics) the student is designated to receive a Human Reader.
10. Readers must be aware of whether a student requires additional tools, devices, or adaptive equipment that has been approved for use during the test, such as a magnifier, closed circuit television (CCTV), abacus, brailler, slate and stylus, etc.

11. If a reader is unsure how to pronounce an unfamiliar word, advise the student of the uncertainty and spell the word.
12. When reading a word that is pronounced like another word with a different spelling, the reader may spell the word after pronouncing it, if there is any doubt about which word is intended.
13. Readers must spell any words requested by the student.
14. When reading passages, readers must be aware of punctuation marks. Readers may read the passage, or selected lines a second time, with all punctuation marks indicated.
15. When test items refer to a particular line, or lines, of a passage, reread the lines before reading the question and answer choices. For example, the reader should say, "Question X refers to the following lines…," then read the lines to the student, followed by question X and the response options.
16. When reading selected response items, readers must be careful to give equal stress to each response option and to read all of them before waiting for a response.
17. If a reader is also scribing the student's responses, or if another adult will scribe, and the student designates a response choice by letter only ("D," for example), the reader must ask the student if he/she would like the response to be reread before the answer is recorded in the answer booklet.
18. If the student chooses an answer before the reader has read all the answer choices, the Human Reader must ask if the student wants the other response options to be read.
19. After the reader finishes reading a test item and all response options, the reader must allow the student to pause before responding. If the pause has been lengthy, say: "Do you want me to read the question or any part of it again?" When rereading questions, readers must avoid emphasis on words not bolded, italicized, or capitalized.

Procedures for Providing the Human Reader Accommodation for ELA Assessments or the Human Reader Accessibility Feature for the Mathematics Assessments to a Small Group of Students

Human Readers may read the test aloud to a small group of students, rather than individually, provided that each student has the Human Reader accommodation/accessibility feature listed in an IEP or 504 Plan.

The following procedures must be followed:

- Check individual state policies on the maximum allowable number of students in a Human Reader group.
- Students with the Human Reader accessibility feature for mathematics or Human Reader accommodation for ELA that need to be grouped together must be taking the same test form, since test questions will differ on each form of the test.
- Students not receiving the Human Reader accessibility feature for mathematics or the Human Reader accommodation for ELA may not be tested in the same location as students who are receiving the human accessibility feature for mathematics or Human Reader accommodation for ELA.

Appendix B: Protocol for the Use of the Scribe Accommodation and for Transcribing Student Responses

Scribing a student's responses by an adult Test Administrator is a response accommodation that allows students to provide test responses to an adult Test Administrator who writes or types the responses directly onto the assessment for the student. Students receiving the scribe accommodation may respond to assessment items either:

- verbally,
- using a speech-to-text device or other augmentative/assistive communication device (e.g., picture/ word board),
- signing (e.g., American Sign Language, signed English, Cued Speech),
- gesturing,
- pointing, or
- eye-gazing

Note: Scribing may include "dragging and dropping" selected response items, as appropriate.

The scribe accommodation is appropriate for students with a physical disability that *severely limits or prevents* the student's motor process of writing, typing, or recording responses during testing. This includes students with reduced ability to record responses due to pain, fracture, paralysis, loss of function, or loss of endurance, as well as students whose handwriting is indecipherable or illegible. Scribes are also an appropriate accommodation for students who have a documented disability in the area of written expression which results in significant interference in their ability to express their knowledge in writing/keyboarding, even after varied and repeated attempts to teach the student to do so.

If a student requires a scribe due to a recently-occurring, though temporary, illness or injury, a Nonstandard Accommodations Request Form (see Appendix D) must be completed and kept on file at the school.

If a student requires a scribe due to an ongoing inability to express his or her responses through writing/ keyboarding, this should be documented in evaluation summaries from locally-administered diagnostic assessments, and must be listed in the student's IEP or 504 plan. The student should be receiving ongoing, intensive instruction and/or interventions to learn written expression, as deemed appropriate by the IEP team or 504 Plan Coordinator.

The use of a scribe is permitted in the following 2023 Science, Math, and ELA assessments:

- Science
- Mathematics
- English Language Arts (ELA) assessments for Evidence Based Selected Response, and Technology Enhanced Constructed Response items
- English Language Arts (ELA) assessments for Prose Constructed Responses. **Note:** For this accommodation, refer to selection and administration guidelines in the *Accessibility Features and Accommodations Manual*

Qualifications of the Scribe

Individuals who provide the scribe accommodation to a student must:

- be trained by the school or district, as indicated in the Test Administrator Manuals;
- sign a Confidentiality Agreement Form; and
- be fluent in receptive and expressive American Sign Language (ASL), signed English, or other sign system, for students who are deaf or hard of hearing.

Preferably, the scribe will already be familiar with and have experience scribing for the student. If the scribe is unfamiliar with the student, then scribe and student should have the opportunity to practice the scribing process together prior to taking the assessment.

Administering the Scribe Accommodation

- A scribe may administer the scribe accommodation only to one student at a time during a test session. The student must be tested in a separate setting.
- The scribe must write legibly, if transcribing a student's response into a test book.
- The scribe must transcribe responses verbatim from the student, and may not prompt or question the student, or correct a student's responses. The scribe may ask the student to restate (or sign) words or parts, as needed.
- A student using a scribe must be given the same opportunity as other students to plan and draft a constructed response. The scribe may write an outline, plan, or draft exactly as directed by the student without any cueing and guidance to the student.
- The scribe should be informed of the preferred method or format for recording the student's response before the date of the assessment. During testing, the student may dictate constructed responses either:
    1. Directly to a human scribe who records the responses at the time they are given (computer- and paper-based testing)
    2. Into a speech-to-text converter (e.g., voice recognition software), augmentative communication device, or assistive technology device to be transcribed by the scribe at a later time into the online testing platform or unto a paper-based book/answer document). A student must be given the opportunity to review and edit his or her responses before they are finalized into the online testing platform or paper-based test book/answer document.
- When using a speech-to-text converter, augmentative communication device, or other assistive technology device, hard copies of the student's response must be printed out for transcription purposes unless the device being used does not have the capability to print. In cases where printing a response is not possible, scribing must take place as the student dictates or otherwise produces the response. All electronic files must be deleted immediately after the testing session.
- The scribe must allow the student to review the scribed response in order to make edits. If requested by the student, the scribe may read the scribed response back to the student. The student may dictate changes or edits to the scribe, and the scribe must make those changes exactly as dictated by student, even if a change is incorrect. All changes must be made during the test session.

Additional Guidelines for the English language arts (ELA) Assessment–Prose Constructed Responses

Capitalization and Punctuation

For the English language arts (ELA) Assessment—Prose Constructed Responses only, the student is responsible for all capitalization and punctuation. This can be accomplished either after testing or during testing using one or more of the following Rules for Punctuation:

1. After dictation: The student can dictate the entire response at one time. The scribe will write/type the response without capitalization and punctuation. When the student is finished dictating, the scribe will show the response to the student. The student will tell the scribe which letters are to be capitalized and where punctuation should be added.

2. During dictation: The student may add capitalization and punctuate as he/she dictates.
    a. For example, when stating the sentence "The fox ran." The student will say, "Capital T, the fox ran, period"
    b. If a sentence includes other punctuation, for example a comma, the student must indicate the comma. For example, when stating, "The boy bought apples, oranges, and bananas." The student will say, "Capital T, the boy bought apples, comma, oranges, comma, and bananas, period"

Students must be given the opportunity to proofread their responses, even if they provide capitalization and punctuation during dictation.

Rules for Capitalization

The *scribe can automatically capitalize in these cases:*

1. The scribe should capitalize the first letter of a sentence if the student has indicated the punctuation in the previous sentence. For example, if the student said, "Capital T, the fox ran, period. The fox jumped, period." The scribe would write "The fox ran. The fox jumped."
2. The first word in a new paragraph when students have indicated for the scribe to begin a new paragraph.

The *student must specify capitalization in the following cases:*

1. The first letter of a sentence, if the student has not indicated punctuation in the previous sentence. For example, if the student said, "Capital T, the fox ran, the fox jumped, period." The scribe would write "The fox ran the fox jumped."
2. Other capitalization (e.g., capitalization of proper nouns, acronyms, etc.)

Scribe Parameters during the Assessment

The following scribing practices are acceptable:

- The scribe may ask "Are you finished?" Or "Is there anything you want to add or delete?"
- The scribe may respond to procedural questions asked by the student such as, "Do I have to use the entire space to answer the question?" The scribe may indicate "no."
- If the student requests that the scribe read a response that was already dictated, the scribe must read what the student dictated previously in an even voice, being careful not to cue the student to errors.

The following scribing practices are unacceptable:

- The scribe cannot influence the student's response in any way.
- The scribe cannot give the student specific directions, clues, or prompts; e.g., "First, set the equations equal to one another;" or "Make sure that the equation is set equal to zero."
- The scribe cannot tell the student if his/her answer is correct or incorrect.
- The scribe cannot answer a student's questions related to the content; e.g., "Is this the right way to set up the problem?" Or "Can you tell me what this word means?"
- The scribe cannot alert the student to mistakes he/she made during testing.

Special Considerations When Scribing for a Student Who Uses Sign Language or Cued Speech

- The scribe for a student who signs their responses must be fluent in ASL, signed English, or other sign systems the student uses.
- When responses are dictated by a student using American Sign Language (or other signed system), the scribe may ask clarifying questions regarding the use of classifiers. Classifiers give descriptive information about a noun or verb such as location and kind.

- The scribe will write the student's responses in English. The transcription of ASL will not be done in a word-to-word format, but instead will be written in English without changing or enhancing the meaning of the content, adding information, or explaining concepts unknown to the student (e.g., student signs "HOUSE WHITE LIVE THERE ME." Scribe writes "I live in the white house.")
- Scribe must follow all other acceptable scribing practices.

Use of Speech-to-Text/Voice-Recognition Software/Devices

Speech-to-text conversion, or voice recognition, software allows students to dictate responses into their computer microphone and have the responses converted to printed text. For this accommodation, students will use their own assistive technology devices at a separate computer station equipped with speech-to-text/voice recognition software in order to respond to multiple-choice, open-ended items, and extended responses on the 2023 Science, Math, and ELA assessments. Students who use voice recognition software routinely, and for whom this accommodation is listed in their IEP, may use speech-to-text/ voice recognition software as an accommodation on the 2023 Science, Math, and ELA assessments. Students must become familiar with the software and must have opportunities to practice using it prior to testing. It is also important that students who use speech-to-text devices be given the opportunity to develop planning notes using speech-to-text, and to view what they produce via speech-to-text.

Upon completion of a test, the student's responses should be printed out and the guidelines for transcribing student responses followed.

Guidelines for Transcribing Student Responses (Paper-based testing only)

Certain situations involving scribing of responses during administration of 2023 Science, Math, and ELA assessments may require a Test Administrator to transcribe a student's response in a standard, scorable test booklet or answer document. These situations may include:

- Answers were recorded in the wrong section of a Test Booklet or Answer Document, or in an incorrect Test Booklet or Answer Document.
- A student takes the test using a special test format that requires answers to be transcribed (e.g., large print).
- A student uses a speech-to-text converter, augmentative communication device, or assistive technology device to be transcribed by the scribe at a later time.
- As an accommodation, a student records answers in a test booklet, answer document, or on blank paper, instead of in the required Test Booklet or Answer Document.
- A Test Booklet or Answer Document becomes unusable (e.g., torn, wrinkled).

If a student's responses must be transcribed after test administration is completed, the following steps must be followed:

- At least two persons must be present during any transcription of student responses. One of these persons will be the transcriber, and the other will be an observer confirming the accuracy of the transcription. It is highly recommended that one of the individuals be an authorized District Test Coordinator or School Test Coordinator.
- The student's response must be transcribed verbatim into the Answer Document or Test Booklet. The student's original response in an Answer Document/Test Booklet should be returned with secure test materials. The District Test Coordinator or School Test Coordinator should write "DO NOT SCORE" or draw an "X" in large font on the front of the original Answer Document/Test Booklet. Do not cover the barcode. Return them with nonscorable test materials.

- Braille transcription: Only an eligible Test Administrator who is a certified Teacher of Students with Visual Impairment, including Blindness, or someone working under the direct supervision of an eligible Test Administrator who is a certified Teacher of Students with Visual Impairment, including Blindness may transcribe the student's responses onto the paper form of the 2023 Science, Math, and ELA assessments.
- Any original student responses that were printed from an assistive technology device or recorded separately on blank paper (or on other external devices) must be securely shredded.

Procedures for Transcribing Student Responses for Computer-Based Testing

Selected Response and Technology Enhanced Items
For selected response and technology enhanced items, student responses must be entered into iTester during the test session by the Test Administrator. Once the student reaches the end of the test with all Selected Response and Technology Enhanced Items completed, the Test Administrator should have the student EXIT the test but not submit the test.

Constructed Response Items
During administration of computer-based 2023 Science, Math, and ELA assessments, students who require use of a speech-to-text converter, augmentative communication device, or assistive technology device will need constructed responses transcribed into iTester by a Test Administrator before the online testing window closes. In these situations, the following steps must be followed.

- As the student encounters constructed responses, he/she should use his/her device to respond to the questions. The student will then continue testing in iTester, leaving these items unanswered in iTester.
- Once the student reaches the item they should click "Finish" to take them to the test review screen. On the test review screen confirm all answers to be transcribed appear as "unanswered".
- Click on "Exit" *NOT* "Turn-In" to exit the testing kiosk.
    - Note: if a student clicks "Turn-In" in error, contact the support desk. The support desk can reactivate the student's test session which will allow the transcriber to log back into the test session that has been turned-in.
- When ready to transcribe responses into the test, log into the test using the students log in credentials, session access code and proctor password, if needed.
- Navigate to the unanswered items left for transcription and transcribe student's answers.
- At least two persons must be present during any transcription of student responses. One of the individuals must be an authorized Test Administrator.
- The student's responses must be transcribed verbatim into iTester. (See note above about scribing signed responses in English).
- Once all items have been transcribed, the Test Administrator will submit the test by clicking "Turn-In" on the test review screen.
- After transcription is complete, all original student responses that were printed from an assistive technology device must be securely shredded.

Appendix C: Text-to-Speech, ASL Video, or Human Reader/Human Signer Guidance for English Language Arts (ELA) Assessments

Individualized Education Program (IEP) or 504 Plan Decision-Making Tool

Directions: This is an optional tool that has been developed to assist IEP teams and 504 Plan Coordinators in identifying students who may be appropriate candidates to receive the accommodation for text-to-speech (computer-based), ASL video (computer-based), or Human Reader/Human Signer (paper-based) for the ELA summative.

Student's Name:_____ **D.O.B.:**_____ **Grade:** _____

School/Program:_____ **State ID #/Local ID#:** _____

District:_____ **State:** _____

IEP Team Members or 504 Plan Coordinator/Staff		
Title	**Name**	**Date**
IEP team Chairperson or 504 Coordinator:		
Special Education Teacher(s):		
General Education Teacher(s):		
IEP team member(s) qualified to interpret reading evaluation results:		
Parent(s)/Guardian:*		
Student (if a team participant):		
Other IEP team member(s):		
Verification of Parent/Guardian Notification (optional):* _____ *(Parent/Guardian Initials)* *I have been informed by my child's school that my child will receive a text-to-speech, ASL video or Human Reader/Human Signer accommodation for an English language arts (ELA) assessment.*		

* If the parent/guardian does not initial this form, the school should attach documentation of notification to the parent and date of notification to this form regarding the decision to provide the text-to-speech, ASL video, or Human Reader/Human Signer accommodation to the student, and keep this form with the student's records.

If all guidelines listed are met, and the student is given the **text-to-speech, ASL video, or Human Reader/Human Signer accommodation** for the English language arts (ELA) assessment, he/she will receive a valid score on the assessment. If all guidelines are not met, and the student is given the **text-to-speech, ASL video, or Human Reader/Human Signer** accommodation on an English language arts (ELA) assessment, the student's assessment score may be invalidated and the score would not be counted in the overall assessment results; i.e., the student would be considered a "non-participant" for the English language arts (ELA) assessment.

Guidelines for IEP Team or 504 Plan Consideration	Additional Guidance	Agree/ Disagree
The student has an Individualized Education Program (IEP) or 504 plan.	Student has an approved IEP or current 504 plan.	❑ Agree ❑ Disagree
In making decisions on whether to provide the student with this accommodation, IEP teams and 504 Plan Coordinators are instructed to consider whether the student has: • Blindness or a visual impairment and has not yet learned (or is unable to use) braille; OR • A disability that *severely limits or prevents* him/her from accessing printed text, even after varied and repeated attempts to teach the student to do so (e.g., student is unable to decode printed text); OR • Deafness or a hearing impairment and is severely limited or prevented from decoding text due to a documented history of early and prolonged language deprivation.	For the screen reader accommodation, the IEP team or 504 Plan Coordinator must determine whether the student is blind or has a visual impairment and has not yet learned (or is unable to use) braille. For the **text-to-speech, ASL video, or Human Reader/Human Signer accommodation**, the IEP team or 504 Plan Coordinator must determine whether the student has a disability that *severely limits or prevents* him or her from decoding text. This accommodation is not intended for a student reading somewhat (i.e., moderately) below grade level. The IEP or 504 plan must document objective evidence from a variety of sources (including state assessments, district assessments, AND one or more locally-administered diagnostic assessments or other evaluation) that indicate that the student's ability to decode text is severely limited or prevented or that the student is blind or visually impaired and has not yet learned (or is unable to use) braille. States may provide additional guidance for their respective states based on PED policy or practice.	❑ Agree ❑ Disagree
Before listing the accommodation in the student's IEP or 504 plan, teams and plan coordinators should also consider whether: • The student has access to printed text during routine instruction through a reader or other spoken-text audio format, or interpreter; • The student's inability to decode printed text or read braille is documented in evaluation summaries from locally-administered diagnostic assessments; or • The student receives ongoing, intensive instruction and/or interventions in the foundational reading skills to continue to attain the important college and career-ready skill of independent reading.	States may provide additional guidance for their respective states in order to define intensive instruction and interventions based on PED policy or practice.	❑ Agree ❑ Disagree

List the data and/or evaluation sources that were used to document the decision to give the text-to-speech, ASL video, or Human Reader/Human Signer accommodation to the student on the English language arts (ELA) assessment(s):

1) Name of Diagnostic Evaluation or Educational Assessment: _____

_____

Name and Title of Test Administrator: _____
Most Recent Testing Date: _____
Score(s): _____
Provide a Summary of the Results: _____

_____

_____

_____

2) Name of Diagnostic Evaluation or Educational Assessment: _____

_____

Name and Title of Test Administrator: _____
Most Recent Testing Date: _____
Score(s): _____
Provide a Summary of the Results: _____

3) List any additional assessment data, scores, and/or evaluation results that were used to guide the decision-making process for IEP teams or 504 Plan Coordinators regarding the text-to-speech, ASL video, or Human Reader/Human Signer accommodation for the English language arts (ELA) assessment(s):

_____

_____

_____

List the instructional interventions and supports specifically related to reading that are currently provided through daily instruction to the student:

- Intensive reading interventions have been provided to the student for _____ years.
- List the specific school years and frequency _____
- Describe and list the specific reading intervention(s) provided to the student: _____

  _____

  _____

List any additional relevant information regarding the student:

_____

_____

Appendix D: Unique and Emergency Accommodations

Directions: The form on the following page should be used for students with unique or emergency accommodations. If a student with a disability or an EL requires an accommodation (i.e., a "unique accommodation") that is not listed in the *Accessibility Features and Accommodations Manual,* and does not change the construct being measured by the test, the DTC may request the use of an accommodation not currently listed in this manual by using this form. This form is also appropriate in cases where a student needs a new accommodation immediately prior to the assessment due to unforeseen circumstances and there is not sufficient time for a 504 plan to be developed with appropriate accommodations. Cases could include students who have a recently-fractured limb (e.g., fingers, hand, arm, wrist, or shoulder); whose only pair of eyeglasses has broken; or a student returning from a serious or prolonged illness or injury. If the principal or School Test Coordinator determines that a student requires an emergency accommodation on the day of the test, this form must be completed and submitted to the District Test Coordinator. The DTC will submit to PED for approval.

NEW MEXICO
Public Education Department

Request for Nonstandard Assessment Accommodation 2022–23

Purpose of Form: This form is to request a unique accommodation (e.g., testing at home, use of electronic devices for medical monitoring) that is not identified in the accommodations manual and is to be used on a state-required assessment. The New Mexico Student Assessment Accommodations Manual can be found on the DTC Resources web page.

Procedure for Requesting Accommodation:

- The nonstandard accommodation request must be documented in a student's IEP, 504, or EL Plan, and the district or charter school must retain the form for a period of five (5) years from the date of the test.
- This form should be submitted through the DTC portal.
- PED will review the request and provide a response within five (5) business days.

Student Information

Student State Identification (SSID) Number (9 digits):	
Student Initials Only:	Student Date of Birth:

District/Charter/BIE School Contact Information

District Name:	School Name:
Name of District Test Coordinator:	Email:
Name of person requesting accommodation:	Email:

For which assessment(s) is this accommodation requested?

Assessment(s):
Accommodation Requested:
Justification for Request:

Is the requested non-standard accommodation documented in the IEP, 504 Plan, EL Plan, or Individualized Healthcare Plan? ☐**Yes** ☐**No**

Is the requested accommodation routinely used in classroom instruction/testing? If no, explain in Justification for Request section. ☐**Yes** ☐**No**

For PED Use Only	
Assigned to:	Date:
☐ Reason for Denial:	
☐Approved for:	
Date District Notified:	

Appendix E: Student Accommodation Refusal Form

Directions: If a student refuses an accommodation listed in his or her Individualized Education Program (IEP), 504 plan, or an EL plan, the school should document in writing that the student refused the accommodation, and the accommodation must be offered and remain available to the student during testing. This form must be completed and placed in the student's file and a copy sent to the parent on the day of refusal. Principals (or designee) should work with Test Administrators to determine who, if any others, should be informed when a student refuses an accommodation documented in an IEP, 504 plan or an EL plan.

Student Name:_____ Date: _____

Grade:_____ Student ID#: _____

School Name: _____

School District: _____

Assessment Type: _____

Test Administrator: _____

Accommodation(s) refused: _____

Reason for refusal: _____

_____

_____

Comments: _____

_____

_____

Student's Signature (optional): _____

Signature of Test Administrator: _____

Keep this form on file at the school.
A copy must be sent home to the parent.

Appendix F: ELA Audio Guidelines

Version 3.0

Table of Contents

Visuals

Guidelines for Text-to-Speech Descriptions

Use these guidelines to describe visuals for text-to-speech scripts:

- Read the title.
- Provide a general overview of the image (i.e., A map of South America, a graphic organizer with a center circle and four circles radiating outward).
- Begin with the main section of the image.
- Describe the details in a succinct manner using grade-level appropriate vocabulary.
- Omit minor details that are irrelevant (a box to the left of the person).
- If facial expressions or body language are important, do not assume a blind student can interpret them. For example, it is better to describe a person as worried than to state that the person has furrowed brows.
- When describing several people in an image, label each one clearly so they are not mixed up (i.e., tall man, elderly man, little boy).
- Describe only what is seen in the image. Do not provide interpretation or additional information.

Classifications for Embed Coding Scheme for Text Descriptions

An embed code within the alt text will be included for all test items with visual elements. The embed code will be classified as a 1, 2 or 3. The description of each level is listed below:

[1] is not construct-relevant and can be eliminated (e.g., it is only there for engagement purposes). For example, a picture of an elephant added purely for engagement would has alt text that reads “elephant [1]” or “picture of elephant [1].”

[2] is construct-relevant and can be represented using accompanying textual description. Example of text where reading the graph is construct-relevant: The graph title is Roller Rink costs. Key, dashed line represents Roller Rink A, solid line represents Roller Rink B. The x-axis is labeled number of people. The y-axis is labeled cost in dollars. The dashed arrow starts at zero people, sixty dollars and points to a little less than sixteen people, midway between one hundred and one hundred ten dollars. The solid arrow starts at zero people, a little less than ten dollars and points to a little more than fourteen people, a little less than one hundred ten dollars. [2]

[3] is construct-relevant and can be represented using accompanying textual description together with a tactile representation or physical manipulative. Example of text where reading the graph is construct-relevant: The graph title is Roller Rink costs. Key, dashed line represents Roller Rink A, solid line represents Roller Rink B. The x-axis is labeled number of people. The y-axis is labeled cost in dollars. [3]

Ellipses

Example

22. Which statement best represents a turning point in the story?

    A. "Suddenly he seemed to know that if he were to survive, he must learn how to fly . . . "

    B. "Albert jumped up and down and screeched for them to rescue him, but they could do nothing."

    C. "When he tried to climb the rocks to the ridge top, he slid backward on his rear."

    D. "Albert watched as his brother pumped his wings wildly and zigzagged far above the ground."

Audio Guideline

Text Only/Text and Graphics

When an ellipsis is used to signify missing text in a sentence, read as "pause 'dot, dot, dot' pause."

Note: Pauses in each application of the audio guidelines in this document are represented by an En Dash with a space on either side of the En Dash.

Application of Audio Guideline

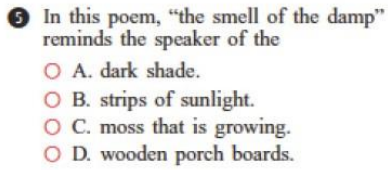
Example

Which statement best represents a turning point in the story?

A: Suddenly he seemed to know that if he were to survive, he must learn how to fly – dot – dot – dot –

Quotations and Quotation Marks

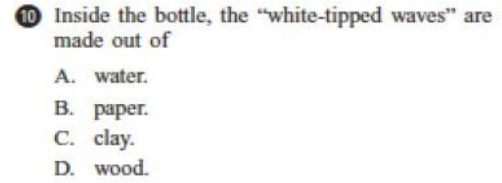
Example 1



5 In this poem, "the smell of the damp" reminds the speaker of the

- A. dark shade.
- B. strips of sunlight.
- C. moss that is growing.
- D. wooden porch boards.

Example 2



10 Inside the bottle, the "white-tipped waves" are made out of

A. water.
B. paper.
C. clay.
D. wood.

Example 3
Mill argues against using St. Paul's epistles as a means for discrimination against women because "The powers that be are ordained of God' gives his sanction to military despotism to that alone, as the Christian form of political government, or commands passive obedience to it."

Audio Guideline
Text Only/Text and Graphics

a. Quotation marks should be read as "quote" before the text and "end quote" after the text.
b. If the quotes surround the title of a work, do not say, "quote."
c. If both single and double quotes occur in a single passage, item, or paragraph, specify with "single quote," "end single quote," "double quote," and "end double quote."

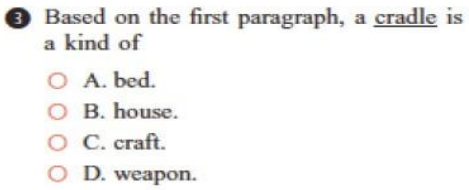
Application of Audio Guideline
Example 1:
In this poem – quote – the smell of the damp – end quote – reminds the speaker of
A dark shade.
B strips of sunlight.
C moss that is growing.
D wooden porch boards.

Example 2:
Inside the bottle, the – quote – white-tipped waves – end quote – are made out of
A water.
B paper.
C clay.
D wood.

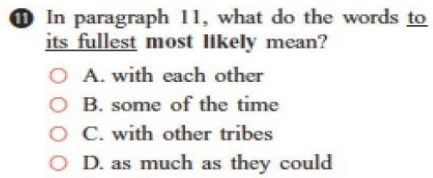
Example 3
Mill argues against using St. Paul's epistles as a means for discrimination against women because – double quote – single quote – the powers that be are ordained of God – end single quote – gives his sanction to military despotism to that alone, as the Christian form of political government, or commands passive obedience to it – end double quotes –

Emphasis for Underline, Bold, Italics, Capitalization

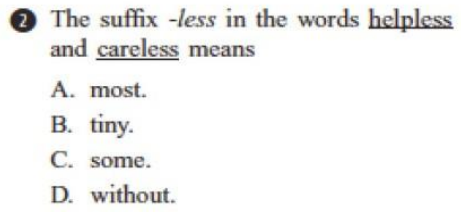
Example 1



Example 2



Example 3



Audio Guideline
Text Only/Text and Graphics

Emphasize words that are underlined, bolded, italicized, or capitalized.

Pause before and after the emphasized word(s) to differentiate between emphasis and normal formatting.

Do not read differently or pause for italics, underline, or bold if they are being used for the directions before a passage or item and are not part of the prompt, question, or answers.

Application of Audio Guideline

Example 1
Based on the first paragraph, a – cradle – is a kind of
A: bed.
B: house.
C: craft.
D: weapon.

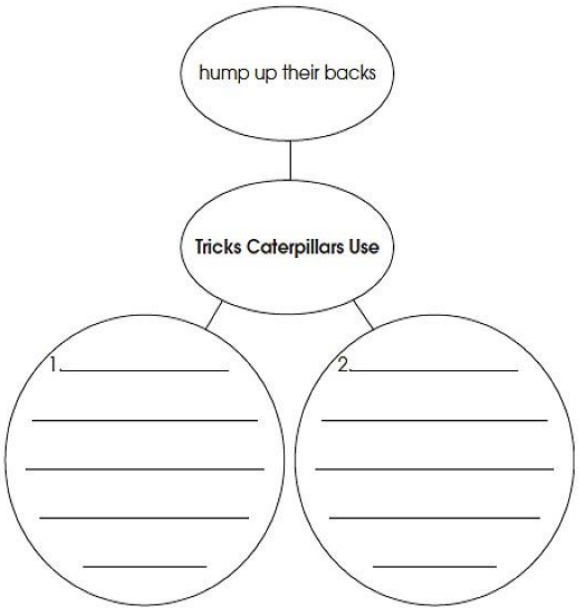
Example 2
In paragraph eleven, what do the words – to its fullest – most likely – mean?
A: with each other
B: some of the time
C: with other tribes
D: as much as they could

Example 3
The suffix – less – in the words – helpless – and – careless – means
A: most.
B: tiny.
C: some.
D: without.

Word Webs

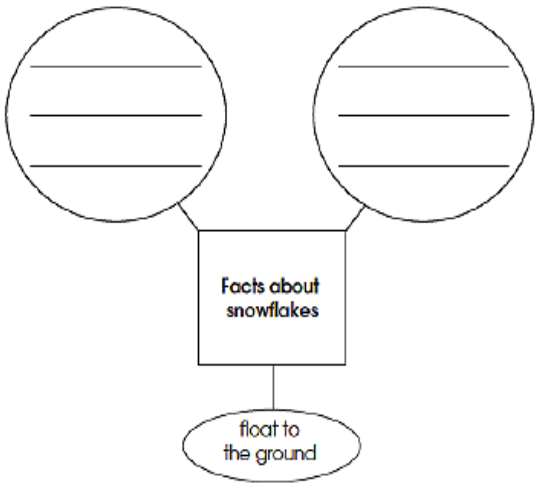
Example 1

13. Using the reading selection, write two other tricks caterpillars use to try to get away from their enemies.

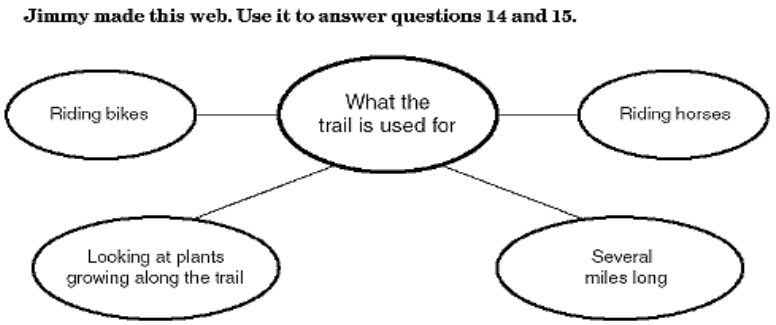


Example 2

10. Use details from the reading selection to complete the web below.



Example 3



Audio Guideline

Text Only
Read the title of the word web, if available, before reading the rest of the text in the word web.

Text and Graphics
Begin by giving a very brief orientation that includes

- that it is a word web
- the attributes of the word web (number of cells, rows, etc.)

Read the word web in a logical manner that helps the student easily navigate the information. While many word webs can be read left to right, top to bottom, some word webs are better read bottom to top or from the middle.

Use common language throughout the item and the test when referring to word webs and their attributes (labels, blank cells, stems, etc.).

Application of Audio Guideline

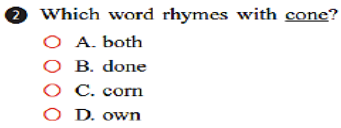
Example 1
A word web containing four cells. The center cell is labeled “Tricks Caterpillars Use.” A cell connecting to the center cell is labeled “hump up their backs.” The two other cells connecting to the center cell contain space to write two other tricks caterpillars use.

Example 2
A word web containing four cells. The center cell is labeled “Facts about snowflakes.” A cell connecting to the center cell is labeled “float to the ground.” The two other cells connecting to the center contain space to write.

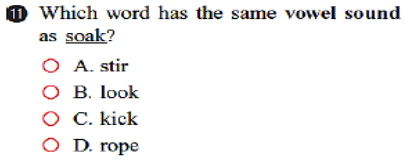
Example 3
A web containing five cells. The center cell is labeled “What the trail is used for.” The four cells connecting to the center cell are labeled “Riding bikes,” “Riding horses,” “Looking at plants growing along the trail,” and “Several miles long.”

Pronunciation

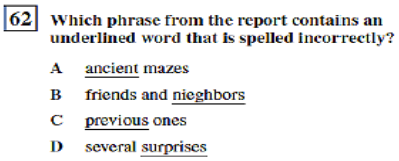
Example 1



Example 2



Example 3



Audio Guideline

Text Only
If the question or stem has the word that rhymes or has a specific sound, read that word, but do not read the answers.

Do not try and read aloud misspelled words as pronunciation is somewhat subjective.

Text and Graphics
When an item is measuring rhyming of words or sounds of words, speak the individual letters in the word instead of speaking the word. If the question or stem has the word that rhymes or has a specific sound, read that word and spell out the answer options.

For questions containing intentionally misspelled words, spell out any word for which the student needs to consider spelling correctness/incorrectness.

Do not try and read aloud misspelled words as pronunciation is somewhat subjective.

Application of Audio Guideline

Example 1

Text Only
Which word rhymes with cone?
A: A
B: B
C: C
D: D

Text and Graphics
Which word rhymes with – cone?
A: B – O – T – H
B: D – O – N – E
C: C – O – R – N
D: O – W – N

Example 2

Text Only
Which word has the same vowel sound as soak?
A: A
B: B
C: C
D: D

Text and Graphics
Which word has the same vowel sounds as – soak?
A: S – T – I – R
B: L – O – O – K
C: K – I – C – K
D: R – O – P – E

Example 3

Text Only
Which phrase from the report contains an underlined word that is spelled incorrectly?
A: A
B: B
C: C
D: D

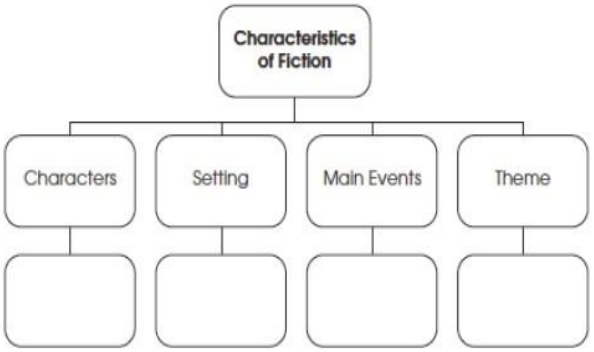
Text and Graphics
Which phrase from the report contains an underlined word that is spelled incorrectly?
A: A – N – C – I – E – N – T mazes
B: friends and N – I – E – G – H – B – O – R – S
C: P – R – E – V – I – O – U – S ones
D: several S – U – R – P – R – I – S – E – S

Graphic Organizers

Example 1

38. "We put the crushed cocoa beans into a chocolate pot."

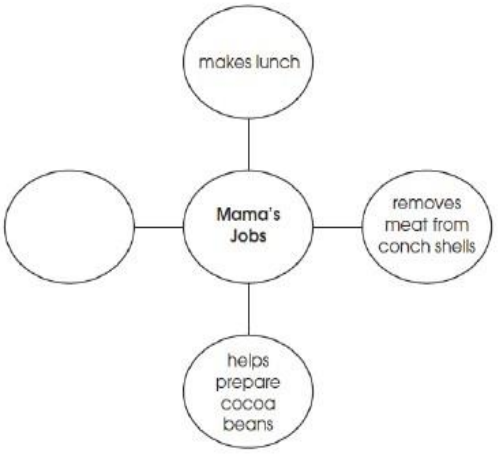
Which column in the graphic organizer below would include this detail?



A. Characters
B. Setting
C. Main Events
D. Theme

Example 2

41.



According to information in the selection, which phrase should be added to the graphic organizer above?

A. makes chowder from conchs
B. hollows a log to make a canoe
C. plants cacao trees in the shade
D. crushes cocoa beans in a mortar

Audio Guideline

Text Only
Read the title of the graphic organizer, if available, before reading the rest of the text in the graphic organizer.

Text and Graphics

If the organizer is structured like a table or has a structure similar to a table, refer to the Math Audio Guidelines document.

If the organizer is structured like a word web, follow the rules in this document for word webs.

Application of Audio Guideline
Example 1

Graphic organizer with a cell labeled "Characteristics of Fiction" at the top. Below the top cell there are four columns and two rows. The first row has columns labeled "Characters," "Setting," "Main Events," and "Theme." Below each labeled cell is a blank cell.

Example 2

Center cell, Mama's Jobs; connecting cells, read clockwise from the top, makes lunch, removes meat from conch shells, helps prepare cocoa beans, blank.

Different Types of Text

Play, Example 1

Setting: Deep in the forest. Tall stool is center, shorter stool is left.
At Rise: Leopard is seated on tall stool, beating drum. Turtle enters left and slowly moves to center and sits on smaller stool.
Leopard (pounding drum and chanting): The forest is mine all night and all day. . .
Turtle (shouting over drum): Good morning, Leopard. I've been listening to your music. You have a fine sounding drum and a fine voice as well.
(Leopard stops pounding drum and looks up.)

Play, Example 2

Jay: Who's that? (Turning the flashlight on the man)
Louie: Get that light outta my face and go back to sleep, Kid.
Jay: There's nothing here to steal, Mister. I swear.
Louie: Is that you, Jay?
Jay: Yeah, who are you?
Louie: It's Uncle Louie.
Jay: Uncle Louie? No kidding? . . . Arty! It's Uncle Louie.

Application of Audio Guideline

Example 1

Setting: – (Voice 1) – Deep in the forest. Tall stool is center, shorter stool is left.
At Rise: – (Voice 1) – Leopard is seated on tall stool, beating drum. Turtle enters left and slowly moves to center and sits on smaller stool.
Leopard – (Voice 1) – pounding drum and chanting: – (Voice 2) – The forest is mine all night and all day – dot – dot – dot -
Turtle – (Voice 1) – shouting over drum: – (Voice 2) – Good morning, Leopard. I've been listening to your music. You have a fine sounding drum and a fine voice as well. – (Voice 1) – Leopard stops pounding drum and looks up.

Example 2

Jay – (Voice 1) – Who's that? – (Voice 2) – Turning the flashlight on the man.
Louie – (Voice 1) – Get that light outta my face and go back to sleep, Kid.
Jay – (Voice 1) – There's nothing here to steal, Mister. I swear.
Louie – (Voice 1) – Is that you, Jay?
Jay – (Voice 1) – Yeah, who are you?
Louie – (Voice 1) – It's Uncle Louie.
Jay – (Voice 1) – Uncle Louie? No kidding? – dot – dot – dot – Arty! It's Uncle Louie.

Poem, Example 1

Carrying the Snake to the Garden
In the cellar
was the smallest snake
I have ever seen.
It coiled itself
in a corner
and watched me
with eyes
like two little stars
set into coal,
and a tail
that quivered.
One step
of my foot
and it fled
like a running shoelace,
but a scoop of the wrist
and I had it
in my hand.
I was sorry
for the fear,
so I hurried
upstairs and out the kitchen door
to the warm grass
and the sunlight
and the garden.
It turned and turned
in my hand
but when I put it down
it didn't move.
I thought
it was going to flow
up my leg
and into my pocket.
I thought, for a moment,
as it lifted its face,
it was going to sing.
And then it was gone.
—Mary Oliver

Poem Example 2

Sheepdog

In the green field stand the scattered sheep,
pretending innocence,
and the Shepherd standing
just beyond the field—
and at the Shepherd's feet, poised,
the rough-coat collie dog, with one thought only.
It is the woolies.
Her eyes, one blue, one brown
never leave them.

When the Shepherd's whistle
releases her,
she's off, like an arrow, running east,
her bared teeth showing
the wolf that still lives in her.
She circles wide, closing in,
a black and white blur at
the edge of a sheep's bad dream.
But the Shepherd whistles, twice for *right*
and once for *left,*
and the dog holds back,
bringing order out of her own wildness,
serving the man's need.

By sundown,
the circle is complete.
The sheep are penned.
The tired Shepherd, the panting dog
head for home, each
more than they would be alone,
the ring the dog marked, running,
symbol of their union.

Audio Guideline

Text Only
Read the poem paying attention to the layout of the stanzas. Do not reference given line numbers. Use extended pauses for the start of a new stanza.

Text and Graphics
Read the poem paying attention to the layout of the stanzas. Reference the line numbers associated with the first and last line of a stanza. For example, say, "Start of stanza line 12 . . . End of stanza line 18." Use extended pauses for the start of a new stanza or reference the new stanza if deemed necessary. Use the above rules for emphasis.

Application of Audio Guideline
Example 1
Read the poem as is line by line.

Example 2
In the green field stand the scattered sheep,
pretending innocence,
and the Shepherd standing
just beyond the field
and at the Shepherd's feet, poised,
the rough–coat collie dog, with one thought only.
– *It is the woolies.* –
Her eyes, one blue, one brown
never leave them. – End of stanza – line 9
Start of stanza – line 10 – When the Shepherd's whistle
releases her,
she's off, like an arrow, running east,
her bared teeth showing
the wolf that still lives in her.
She circles wide, closing in,
a black and white blur at
the edge of a sheep's bad dream.
But the Shepherd whistles, twice for – right
and once for – left,
and the dog holds back,
bringing order out of her own wildness,
serving the man's need. – end of stanza – line 22
start of stanza – line 23 – By sundown,
the circle is complete.
The sheep are penned.
The tired Shepherd, the panting dog
head for home, each
more than they would be alone,
the ring the dog marked, running,
symbol of their union. – end of stanza – line 30 –

Political Cartoons

Example

Look at the cartoon below. Then answer the following.

"I'M SORRY, KID, BUT IT REALLY HURTS ME MORE THAN IT HURTS YOU"

According to the cartoon, what is a criticism of the juvenile justice system?

A. The system gives judges little choice in punishment.
B. The juvenile justice system wastes too much money.
C. The government has too much control over the lives of juveniles.
D. The courts make the community responsible for juveniles' actions.

Audio Guideline

Text Only
Read the title of the political cartoon, if available, before reading the rest of the text in the political cartoon.

Text and Graphics
Start by stating that it is a political cartoon.
Pay special attention to any writing in the cartoon (labels, titles, signs, etc.).
Read the caption of the cartoon.

Application of Audio Guideline

Example
A political cartoon showing an officer standing behind a boy who is standing before a judge. The judge has an open book that is titled "Comprehensive guidelines for sentencing juvenile offenders." The caption of the cartoon is I'm sorry, kid, but it really hurts me more than it hurts you.

Maps

Example
(Part of a passage and section on Machu Picchu that references many of the countries, cities, and geographical features labeled)

Audio Guideline

Text Only
Read the title of the map if available, then read the key, compass rose, and map from top to bottom, left to right as much as possible.

Text and Graphics
Read the title of the map if available, then read the key, compass rose, and map from top to bottom, left to right as much as possible.
For maps, a few words can be used to describe the map unless the item requires the student to use the map to answer the question.

Application of Audio Guideline

Example
A map showing a portion of South America: Ecuador; Amazon River; Urbamba River; Peru; Vilcabamba, Brazil; Machu Picchu; Andes Mountains; Cuzco, Bolivia; Atacama Desert; Chile; Argentina.

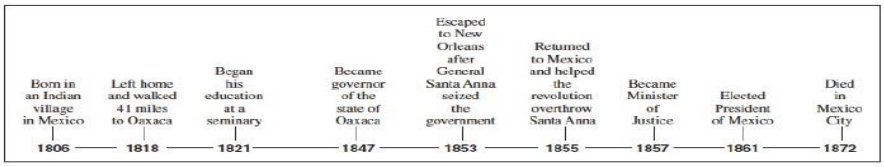
Timelines

Example 1

Timeline

Edmund Halley is born	Halley observes the comet for the first time	Halley visits Isaac Newton to discuss the laws of gravity	Halley focuses on the study of comets	Halley dies	The comet returns to view as Halley predicted
1656	1682	1684	1704	1742	1759

Example 2



Audio Guideline

Text Only
Read the title of the timeline and text from top to bottom, column to column.

Text and Graphics
State that it is a timeline and read the title first or any brief note of what the timeline represents.

State the direction of the timeline and direction of reading.

Read the timeline in chronological order, keeping text with the corresponding date.

Read the date first, followed by the corresponding text that accompanies it.

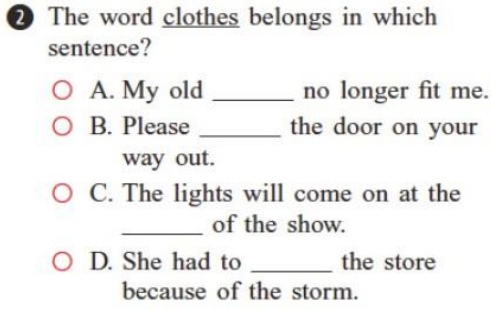
Application of Audio Guideline

Example 1
A timeline of Edmund Halley's life. From left to right, the timeline reads, sixteen fifty-six, Edmund Halley is born; sixteen eighty-two, Halley observes the comet for the first time; sixteen eighty-four, Halley visits Isaac Newton to discuss the laws of gravity; seventeen oh-four, Halley focuses on the study of comets; seventeen forty-two, Halley dies; seventeen fifty-nine, The comet returns to view as Halley predicted.

Example 2
A timeline of Benito Juarez's life. From left to right the timeline reads, eighteen oh-six, Born in an Indian village in Mexico; eighteen eighteen, Left home and walked forty-one miles to Oaxaca; eighteen twenty-one, Began his education at a seminary; eighteen forty-seven, Became governor of the state of Oaxaca; eighteen fifty-three, Escaped to New Orleans after General Santa Anna seized the government; eighteen fifty-five, Returned to Mexico and helped the revolution overthrow Santa Anna; eighteen fifty-seven, Became Minister of Justice; eighteen sixty-one, Elected President of Mexico; eighteen seventy-two, Died in Mexico City.

Fill in the Blank

Example



Audio Guideline

Text Only
Read the blank element with a pause, then “blank” followed by a pause.

Text and Graphics
Read the blank element with a pause, then “blank” followed by a pause.

If the space to be filled in has a question mark, read it as “unknown x” where x is the line, box, bubble, cell, etc.

For technology enhanced items where the blank is in the shape of a box, read the blank box with a pause, then “blank box” followed by a pause.

Application of Audio Guideline
Example
Text Only; Text and Graphics
A: My old – blank – no longer fit me.
B.: Please – blank – the door on your way out.
C: The lights will come on at the – blank – of the show.
D: She had to – blank – the store because of the storm.

Pictures

Example 1

An Air Force LC-130, equipped with skis for landing, transports scientists and workers to Antarctica and back.

American scientists and their helpers who are traveling to the interior of Antarctica fly from Christchurch, New Zealand, on U.S. Air Force planes, operated by the 109th Airlift Wing of the New York Air National Guard. These LC-130s are outfitted with skis instead of wheels for landing on the ice runways.

The flight from Christchurch to McMurdo Station, the biggest American base in Antarctica, takes eight hours. Boomerang flights—ones that turn around midway—are common. The planes can't carry enough fuel to fly to Antarctica and back again to New Zealand. They must refuel in Antarctica. But when there's a blizzard on the ice, the pilots can't land to refuel. So at the midway point, the pilot always radios ahead. If there's a chance of a storm, the plane turns around and flies back to New Zealand. One third of all flights headed for Antarctica are forced to turn around midway. This midway point is called the point of no return.

Example 2

Read the following two selections. Think about how they are alike and how they are different.

Reminiscing

by Ralph Cortez

Watermelons were so much sweeter then,
When boys were the stuff of super men,
And summers seemed so much longer too,
With nothing pending and nothing due.
We were swordsmen—swashbuckling heroes,
Eternal victors—never zeroes;
Second basemen and clean-up hitters;
Forever winners, never quitters.
Play was a ritual in those days,
To go on magical mind forays,
To play the game with aplomb and ease,
To venture forth when and where we'd please.
We would feign death, and then rise up again.
Watermelons were so much sweeter then.

Piano

by D. H. Lawrence

Softly, in the dusk, a woman is singing to me:
Taking me back down the vista of years, till I see
A child sitting under the piano, in the boom of the tingling strings
And pressing the small, poised feet of a mother who smiles as she sings.

In spite of myself, the insidious mastery of song
Betrays me back, till the heart of me weeps to belong
To the old Sunday evenings at home, winter outside
And hymns in the cozy parlor, the tinkling piano our guide.

So now it is vain for the singer to burst into clamor
With the great black piano appassionato. The glamour
Of childish days is upon me, my manhood is cast
Down in the flood of remembrance, I weep like a child for the past.

[Public Domain]

CS21908

Example 3

Whites and African Americans participated and sometimes worked together. Many of the African Americans were escaped slaves themselves, but they continued to risk their lives to help others. There were ordinary farmers, ministers, and housewives. Many well-known political and religious leaders from the black and white communities were also active supporters. In 1859, a congressman named Owen Lovejoy gave a speech in which he announced that he worked with the Underground Railroad. In the speech, he boldly said: *"Owen Lovejoy . . . aids every fugitive that comes to his door and asks it. Proclaim it then from the housetops. Write it on every leaf that trembles in the forest, make it blaze from the sun at high noon."*

Audio Guideline

Text Only
After the paragraph that refers to the picture, read the title, if available. Read embedded text and/or caption, and then read text.

Text and Graphics
Before describing the picture, it should be determined whether the details of the picture are necessary to understanding and responding to the item(s). In many cases, the picture will be used to accompany a passage or reading excerpt as a piece of visual interest that is not essential in responding to the item. In this case, a very brief description may suffice.

In other cases, the caption or embedded text will describe the picture and only limited additional information is necessary.

In general, read the title of the picture or caption (if it is meant to serve as a title) if there is one.

Application of Audio Guideline

Example 1
A picture showing an airplane.

American scientists and their helpers who are traveling to the interior of Antarctica fly from Christchurch, New Zealand, on U.S. Air Force planes, operated by the 109th Airlift Wing of the New York Air National Guard. These LC-130s are outfitted with skis instead of wheels for landing on the ice runways.

The flight from Christchurch to McMurdo Station, the biggest American base in Antarctica, takes eight hours. Boomerang flights—ones that turn around midway—are common. The planes can't carry enough fuel to fly to Antarctica and back again to New Zealand. They must refuel in Antarctica. But when there's a blizzard on the ice, the pilots can't land to refuel. So at the midway point, the pilot always radios ahead. If there's a chance of a storm, the plane turns around and flies back to New Zealand. One third of all flights headed for Antarctica are forced to turn around midway. This midway point is called the point of no return.

Example 2
A picture of a sliced watermelon.

A picture of a piano with musical notes coming from it.

Example 3
A picture of a slave with chains on his hands and feet. The caption reads "Am I Not a Man and a Brother?"

Boxed Sentences or Paragraphs

Example 1

"This is your last chance to change your mind" said the operator.

What does the sentence suggest about a ride on the Space Shot?

Example 2

Nothing was different except the warm glow that was in my belly and my arms and my legs and my head and wouldn't go away.

Which of the following words is an adjective as it is used in the sentence?

Audio Guideline

Text Only
Read the boxed sentence/word as is with a pause before and after to reflect a return to normal formatting.

Text and Graphics
Preface the boxed sentence/word by saying "boxed x" (x being sentence, word, etc.).

Pause after reading the information in the box to indicate a return to normal formatting.

Application of Audio Guideline

Example 1
What does the sentence suggest about a ride on the Space Shot?

Boxed sentence, – This is your last chance to change your mind, – said the operator. –
(Answer options are read.)

Example 2

Which of the following words is an adjective as it is used in the sentence?

Boxed sentence, – Nothing was different except the warm glow that was in my belly and my arms and my legs and my head and wouldn’t go away. –

(Answer options are read.)

Appendix G: Mathematics Audio Guidelines

Version 3.3

Table of Contents

Visuals

Guidelines for Text-to-Speech Descriptions

Use these guidelines to describe visuals for text-to-speech scripts:

Read the title.

Provide a general overview of the image. (i.e., A map of South America, a graphic organizer with a center circle and four circles radiating outward)

Begin with the main section of the image.

Describe the details in a succinct manner using grade-level appropriate vocabulary.

Omit minor details that are irrelevant (a box to the left of the person).

If facial expressions or body language are important, do not assume a blind student can interpret them. For example, it is better to describe a person as worried than to state they have furrowed brows.

When describing several people in an image, label each one clearly so they are not mixed up. (i.e., tall man, elderly man, little boy)

Describe only what is seen in the image, do not provide interpretation or additional information.

Reading Inline Choice Items

Test Nav 8.4 does not yet have the capability to read the options in an inline choice item, therefore, follow these directions for providing phonetic markup.

Use the drop-down menus to complete the sentence.

Example Stem:

A twenty-three point six K-G grocery cart is pushed away from and then rolls back toward a cart rack. Use the graph to complete the sentence describing the motion of the grocery cart.

Example of Inline Choice

The graph shows that the cart travels (Inline Choice dropdown menu) meters between zero and five seconds.

When accessing the dropdown menu, the following answer options are available.

Two point zero
Three point zero
Four point zero
Five point zero
Six point zero

Example of Phonetic Markup

The graph shows that the cart travels - blank - meters between zero and five seconds. The answer choices are: two point zero, three point zero, four point zero, five point zero, six point zero.

Classifications for Embed Coding Scheme for Text Descriptions

An embed code within the alt text will be included for all test items with visual elements. The embed code will be classified as a 1, 2 or 3. The description of each level is listed below:

[1] is not construct-relevant and can be eliminated (e.g., it is only there for engagement purposes). For example, a picture of an elephant added purely for engagement would has alt text that reads "elephant [1]" or "picture of elephant [1]."

[2] is construct-relevant and can be represented using accompanying textual description. Example of text where reading the graph is construct-relevant: The graph title is Roller Rink costs. Key, dashed line represents Roller Rink A, solid line represents Roller Rink B. The x-axis is labeled number of people. The y-axis is labeled cost in dollars. The dashed arrow starts at zero people, sixty dollars and points to a little less than sixteen people, midway between one hundred and one hundred ten dollars. The solid arrow starts at zero people, a little less than ten dollars and points to a little more than fourteen people, a little less than one hundred ten dollars. [2]

[3] is construct-relevant and can be represented using accompanying textual description together with a tactile representation or physical manipulative. Example of text where reading the graph is construct-relevant: The graph title is Roller Rink costs. Key, dashed line represents Roller Rink A, solid line represents Roller Rink B. The x-axis is labeled number of people. The y-axis is labeled cost in dollars. [3]

Accessibility experts will be trained on this embedded coding scheme during the item tagging phase of item development.

Symbols

Money ($)

Example 1
$4.35

Example 2
$2.50

Example 3
$5,390

Audio Guideline

Read dollars and cents if there is a decimal point.

Do not read shortcuts for numbers. For instance $.25 and $1.50 should be read as twenty-five cents instead of a quarter. This will allow a more standardized presentation of monetary quantities.

If the amount is less than one dollar, read “X cents” and do not read the zero ($0.35 is “thirty-five cents” not “zero dollars and thirty-five cents”). Likewise, do not read “and zero cents” ($4.00 is read “four dollars” and not “four dollars and zero cents”).

Read the number place value unless the question is measuring place value (refer to the large number section for details).

Application of Audio Guideline

Example 1
Four dollars and thirty-five cents

Example 2
Two dollars and fifty cents

Example 3
Five thousand three hundred ninety dollars

Angles/Triangles (∠ and ∆)

Example 1
∠RST

Example 2
∆RST

Example 3
∆R'S'T'

Audio Guideline

Read angles and shapes by leading with “angle,” “shape,” etc. and then reading letters individually.

When reading a transformed or reflected angle or shape that uses “ ' ”, describe as “prime.”

Do not reference the case of the letter unless an item includes uppercase and lowercase letters. In this instance, make reference to the uppercase letters guideline.

Application of Audio Guideline

Example 1
Angle *RST*

Example 2
Triangle *RST*

Example 3
Triangle *R* prime *S* prime *T* prime

Ratios (:)

Example
3:2

Audio Guideline

Read as "the ratio *x* to *y*."

Sometimes the ratio symbol is used for fractions. This can usually be determined by context. If this is the case, refer to the fraction guideline.

If the "the ratio of" is used in the item, read as "x to y" to avoid being redundant.

Application of Audio Guideline

Example
The ratio three to two

Equal Signs (=)

Example
2 + 3 = 5

Audio Guideline

Read as "equals."

Application of Audio Guideline

Example
Two plus three equals five.

Pi (π)

Audio Guideline

Read as "pi."

Other Greek letters

Audio Guideline

Read as the Greek letter in most cases, unless using the closest English letter is clearer.

Application of Audio Guideline

Example
sin α = 0.5 is read “sine alpha equals zero point five” but the density formula,

$$\rho = \frac{m}{V}$$

where “ρ” is the Greek letter rho, should be read “P equals fraction with …” since (a) there is no “P” in the formula, (b) the Greek letter closely resembles the English letter, and (c) use of the word “rho” is likely to be more distracting than helpful for text-to-speech users, since English readers may not know what a “rho” is. It is advisable to avoid formulas like this in item development (a “D” replaces the rho is some US textbooks), but given an item with uncommon Greek letters (other than alpha, beta, delta, theta, and perhaps a few others as may be determined on a case-by-case basis), math content specialists have found it most helpful in the past to use the closest English equivalent.

Approximately equal to (≈)

Example
π ≈ 3.14

Audio Guideline

Read as “is approximately equal to.”

Application of Audio Guideline

Example
Pi is approximately equal to three point one four.

Less than (<)

Example 1
3<5

Example 2
x<y<z

Audio Guideline

Read as “is less than.”

If there is more than one “less than” sign in a string, then read the whole relationship together. Read the last part as “is less than.”

Example 1
Three is less than five.

Example 2
X is less than y is less than *z*.

Less than or equal to (≤)

Example
$2x \leq 6$

Audio Guideline

Read as "is less than or equal to."

Application of Audio Guideline
Two x is less than or equal to six.

Greater than (>)

Example 1
$7>5$

Example 2
$x>y>z$

Audio Guideline

Read as "is greater than."
If there is more than one "greater than" sign read the whole relationship together. Start the last part as "is greater than."

Application of Audio Guideline
Example 1
Seven is greater than five.

Example 2
X is greater than y is greater than z.

Greater than or equal to (≥)

Example
$3x \geq 6$

Audio Guideline

Read as “is greater than or equal to.”

Application of Audio Guideline

Three *x* is greater than or equal to six.

Dashes (–)

Example 1
Pages 3–7

Audio Guideline

When the dash is used to reference material or as a group of conditions, use “through” for consecutive and non-consecutive numbers.

Application of Audio Guideline

Example 1
Pages three through seven

Temperatures (°F and °C)

Example 1
35°F

Example 2
25°C

Audio Guideline

Read as “degrees Fahrenheit” and “degrees Celsius.”

Application of Audio Guideline

Example 1
Thirty-five degrees Fahrenheit

Example 2
Twenty-five degrees Celsius

Parallels (RS || $\overline{XY}$)

Audio Guideline

Read as “is parallel to.”

Line segment *RS* is parallel to line segment *XY.*

Perpendiculars (⊥)

Example
RS ⊥ *XY*

Audio Guideline

Read as "is perpendicular to."

Application of Audio Guideline

Line segment *RS* is perpendicular to line segment *XY*.

Abbreviations (ft., km)

Example 1
3ft.

Example 2
What is the correct abbreviation for kilometer?
A: kl
B: K
C: km
D: klm

Audio Guideline

Present abbreviations by speaking the whole word the abbreviation represents.

If the item measures the ability to identify the meaning of the abbreviation, then read the abbreviation letter by letter.

If speaking the abbreviation violates the construct being measured, then read letter by letter.

If the item has measurements that are all uppercase or lowercase, then it is not necessary to reference the cases.

Application of Audio Guideline

Example 1
Three feet

Example 2
What is the correct abbreviation for kilometer?
A: kl
B: K
C: km
D: klm

Measurement (“ ‘ cm^2)

Example 1
6”

Example 2
12’

Example 3
4 cm^2

Example 4
5 cm^3

Audio Guideline

Present measurements by speaking the whole word the symbol represents.

Application of Audio Guideline

Example 1
Six inches

Example 2
Twelve feet

Example 3
Four square centimeters

Example 4
Five cubic centimeters

Number Signs (#)

Example
Refer to step #5.

Audio Guideline

Read as “number.”

Rule refers only to when symbol is being used to signify “number” as opposed to other non-mathematical uses of the symbol (for example, the pound key and the hash key).

Application of Audio Guideline

Example
Refer to step number five.

Empty/Unknown Boxes (□, [?])

Example 1
$4 + 2x = \square$

Example 2
$3 + y = [?]$

Audio Guideline

Refer to an empty box in a formula or equation as “blank.”
Refer to a box with a question mark in it as “question mark.”

Application of Audio Guideline

Example 1
Four plus two *x* equals blank.

Example 2
Three plus *y* equals question mark.

Not equal to (≠)

Example
$2x \neq 7$

Audio Guideline

Read as “is not equal to.”

Application of Audio Guideline

Two *x* is not equal to seven.

Arc (⌒)

Example
$\overset{\frown}{RT}$

Audio Guideline

Read as “arc.”

Application of Audio Guideline

Example
Arc RT

Infinity (∞)

Example
As $x \rightarrow \infty$, $f(x) \rightarrow -\infty$

Audio Guideline

Read as “infinity.”

Application of Audio Guideline

Example
As *x* approaches infinity, *f* of *x* approaches negative infinity.

Percent (%)

Example
35%

Audio Guideline

Read as “percent.”

Application of Audio Guideline

Thirty-five percent

Lines: Line Segment, Line, and Ray

Example 1: Line Segment
$\overline{FG}$

Example 2: Line
$\overleftrightarrow{JK}$

Example 3: Ray
$\overrightarrow{LM}$

Audio Guideline

Read as “line segment,” “line,” or “ray” when they appear above letters or numbers.

Application of Audio Guideline

Example 1
Line segment *FG*

Example 2
line *JK*

Example 3
ray *LM*

Similar to (~)

Example
$\Delta EFG \sim \Delta JKL$

Audio Guideline

Read as “is similar to.”

Application of Audio Guideline

Example
Triangle *EFG* is similar to triangle *JKL*.

Therefore (∴)

Example
$A=B$ and $B=C \therefore A=C$

Audio Guideline

Read as “therefore.”

Application of Audio Guideline

Example
A equals *B* and *B* equals *C*, therefore *A* equals *C*.

Congruent (≅)

Example
$\angle FGH \cong \angle JKL$

Audio Guideline

Read as “is congruent to.”

Application of Audio Guideline

Example
Angle *FGH* is congruent to angle *JKL*.

Factorial (!)

Example
5! = *x*

Audio Guideline

Read as “factorial.”

Application of Audio Guideline

Example
Five factorial equals *x*.

Plus or Minus (±)

Example
The margin of error is 4.5 ± .8

Audio Guideline

Read as “plus or minus.”

Application of Audio Guideline

Example
The margin of error is four point five plus or minus point eight.

Subscript (A_i)

Example
A_i represents the maximum amount of interest.

Audio Guideline

Read as “*x* subscript *y*.”

Application of Audio Guideline

A subscript *i* represents the maximum amount of interest.

Numbers

Negative/Positive Numbers

Example 1
–4

Example 2
4 – –5

Example 3
What is the distance between +4 and –3 on the number line?

Audio Guideline

Read as “negative.” Do not read the negative sign as a minus sign.

In most cases, consecutive negatives that are intended to show the negative of a negative will be represented with a set of parentheses. If this is the case, then refer to the parentheses section.

If the negative of a negative does not include parentheses, read as “negative (pause) negative.”

Two consecutive negatives should not be read as “negative negative *X*” if the operation is focused on subtraction. In this case, read as “minus negative *X*.” Note that this rule refers to numbers only. If, instead of a number, *X* is actually a variable or expression that includes variables, refer to the section entitled “Variables/Letters” below for the correct reading of expressions like –*y*.

If a positive sign precedes a number and is not part of an operation, then read as “positive.”

Application of Audio Guideline

Example 1
Negative four

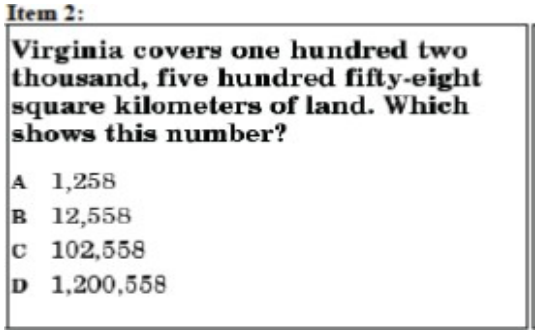
Example 2
Four minus negative five

Example 3
What is the distance between positive four and negative three on the number line?

Large Whole Numbers

Example 1
103,457

Example 2


Item 2:

Virginia covers one hundred two thousand, five hundred fifty-eight square kilometers of land. Which shows this number?

A 1,258
B 12,558
C 102,558
D 1,200,558

Audio Guidelines

For items not measuring place value, read large numbers by referencing all of the number place values.

If the item measures place value knowledge, read the number digit by digit using commas.

If reading the number as a whole number violates the construct being measured, read the number digit by digit.

Application of Audio Guideline

Example 1
One hundred three thousand, four hundred fifty-seven
Note: Use this application unless cueing occurs; then use the application in Example 2.

Example 2
A: one comma two five eight
B: one two comma five five eight
C: one zero two comma five five eight
D: one comma two zero zero comma five five eight

Fractions/Improper Fractions

Example 1
$\frac{1}{2}+\frac{3}{8}$

Example 2
$\frac{3}{14}+\frac{15}{100}+\frac{x}{2y}$

Example 3
$\frac{3x+y}{z}$

Example 4
$\frac{6}{3}$

Example 5
$\frac{3x}{5} + x^2$

Audio Guidelines

Read common fractions by presenting the numerator as the number it represents and the denominator as the ordinal number using two words for the whole presentation.

Read any fraction with numerator of_____(pause) and denominator of_____.

If the denominator is between 2 and 10 then read it is as one third, one fourth, one sixth, one sixth, one seventh, one eighth, one ninth, or one tenth.

An exception to the first guideline is $\frac{1}{2}$, which should always be read as one-half.

An exception to the first guideline is 1 in the denominator. For example, $\frac{3}{1}$ should be read as numerator of 3 (pause) and denominator of 1.

When a fraction is complex (e.g., has more than one number in the numerator/denominator, includes an arithmetic operation, or involves parentheses/exponents) denote the numerator and denominator using the language "fraction with numerator of … (pause) and denominator of …"

When an operation follows a fraction, pause between the fraction and the next operation.

Application of Audio Guidelines

Example 1
One-half plus three-eighths

Example 2
Fraction with numerator of 3 (pause) and denominator of 14 (pause) plus fraction with numerator of fifteen (pause) and denominator of one hundred (pause) minus fraction with numerator of *x* (pause) and denominator of two *y*

Example 3
Fraction with numerator of three *X* plus *Y* (pause) and denominator of *Z*

Example 4
Six-thirds

Example 5
Fraction with numerator of three *x* (pause) and denominator of 5 (pause) plus *x* squared

Mixed Numbers

Example 1
$4\frac{3}{4}$

Example 2
$5\frac{13}{28}$

Audio Guidelines

Read with “and” between the whole number and the fraction.
Use fraction audio guidelines for reading fraction portion of mixed numbers.

Application of Audio Guidelines

Example 1
Four and three fourths

Example 2
Five and (pause) fraction with numerator of thirteen (pause) and denominator of 28

Decimal Points

Example 1
40.6500

Example 2
0.100000

Example 3
0.0000000002

Example 4
0.333. . .

Example 5
3,450.0844397

Audio Guidelines

If there are up to six repeating zeroes or numbers before or after the decimal point, read them as “zero and three repeating.”

If there are more than six repeating zeroes or numbers after the decimal point (beyond millionths), say “point” and read the digits in order from left to right.

Read “repeating” where “…” represents the number of group of numbers that repeats.

Application of Audio Guidelines

Example 1
Forty point six five zero zero

Example 2
Zero point one zero zero zero zero zero

Example 3
Zero point zero zero zero (pause) zero zero zero (pause) zero zero zero two

Example 4
Zero point three repeating

Example 5
Three thousand four hundred fifty point zero eight four (pause) four three nine seven

Roman Numerals

Example 1
Find the point in quadrant II that is furthest from the origin.

Example 2
V. Three students walked to school taking different routes.

Example 3
What is the numeric value of Roman numeral VII?

Audio Guidelines

If an item uses Roman numerals but is not measuring knowledge of Roman numerals, read the Roman numeral reference and then the number.

If the item measures knowledge of Roman numeral value, read “Roman numeral” followed by the letters one at a time.

Application of Audio Guidelines

Example 1
Find the point in quadrant two that is furthest from the origin.

Example 2
Question five. Three students walked to school taking different routes.

Example 3
What is the numeric value of Roman numeral V I I?

Time

Example 1
6:30

Example 2
9 a.m.

Example 3
5:45

Example 4
5:00 p.m.

Audio Guidelines

Read the time literally without using shortcuts or reading the time in reference to a different version of time (e.g., noon, quarter of six, ten after five).

Read a.m. and p.m. without adding language about the time of day (e.g., “in the morning” or “at night.”)

Application of Audio Guidelines

Example 1
Six thirty

Example 2
Nine a m

Example 3
Five forty five

Example 4
Five o’clock p m

Date

Example 1
1976

Example 2
Feb. 5, 2003

Example 3

Population of Two Cities from 1975 to 2025

City	1975	2000	2010	2025
Tokyo	26.6 million	34.4 million	36.9 million	37.1 million
Delhi	4.4 million	15.7 million	21.9 million	28.6 million

Audio Guidelines

Read years as they would be read in plain language usage. For years after 1999, read "two thousand six" (for example) before 2010 and "twenty twelve" for years after 2009. However, when years comprise the axis of a graph or a sequence of table cells, maintain consistency in going from 2009 … 2010 … 2011 and use either convention (both are acceptable usage), except do not use the "two-thousand" style for years after 2019. For years after 2099, use the same style as for years between 1900 and 1999.

Read months as the full name even if abbreviations are presented in text.

Read days as you would when reading a date instead of reading the day as number (e.g., "second" instead of "two," "third" instead of "three," or "fourth" instead of "four").

Application of Audio Guidelines

Example 1
Nineteen seventy six

Example 2
February sixth, two thousand three

Example 3
… city … nineteen seventy five … two thousand … two thousand ten … twenty twenty five … (Refer to the section entitled "Tables" for more information.)

Ordered Pairs

Example
Point *X* is (–2, 4)

Audio Guideline

Read coordinate pairs as "ordered pair X, Y."

Application of Audio Guideline

Point X is ordered pair negative two, four.

Probability

Example
P(orange) = $\frac{1}{6}$

Audio Guideline

"P(text)" is the notation for probability. When reading a probability, do not read parentheses as "open parenthesis/close parenthesis." Read as "P of" word in parentheses "is" remaining text.

Application of Audio Guideline

Example
P of orange is one-sixth

Expressions/Equations/Operations

Multiplication

Example 1
3 x 5 = *X*

Example 2
xy + 4*x* = 10

Example 3
(3 + *x*)(*y* – 2)

Audio Guidelines

Read the multiplication symbol as "times" when it appears in a math item.

When a number, symbol, or another set of parentheses appears before a set of parentheses, read the number or symbol as is and "open parenthesis" before what is within the parentheses. When multiple sets of parentheses appear consecutively, read as "open parenthesis" and "close parenthesis."

If there are two variables or a variable and a number consecutively, do not read "times" to represent implied multiplication.

Application of Audio Guidelines

Example 1
Three times five equals *X*.

Example 2
Xy plus four x equals ten.

Example 3
Open parenthesis three plus *x*, close parenthesis, (pause) open parenthesis *y* minus two, close parenthesis.

Addition

Example
4 + 2 + 3

Audio Guideline

Read as “plus.”

Application of Audio Guideline

Four plus two plus three

Subtraction

Example
5 – 3

Audio Guideline

Read as “minus.”

Application of Audio Guideline

Five minus three

Division

Example 1

12 ÷ 4

Example 2

What is 57 ÷ 5
A: 10 R7
B: 11 R2
C: 12

Audio Guideline

Read as “divided by.”

If the item presents the remainder as “R” read as “remainder” unless the item is measuring the meaning of “R.” In this case, read it as “R.”

Application of Audio Guideline

Example 1
Twelve divided by four

Example 2
What is fifty-seven divided by five?
A: ten, remainder seven
B: eleven, remainder two
C: twelve

Parentheses

Example 1
$3(x + y) = 6$

Example 2
$2(x + 3) + \frac{(y-2)}{3} = 9$

Example 3
$(x + 4)[(x + 4) – (x – 2)]$

Audio Guideline

Read the parentheses by referring to the opening of the parentheses using the language “open parenthesis” and the closing of the parentheses using the language “close parenthesis.”

It is important to reference the close of the parentheses to be clear on when the parenthetical expression ends.

When reading an equation or expression with multiple parts and sets of parentheses, pause to help differentiate between sections.

Read brackets using the same language as parentheses in the first guideline.

Application of Audio Guideline

Example 1
Three (pause) open parenthesis *x* plus *y* close parenthesis (pause) equals six.

Example 2
Two (pause) open parenthesis *x* plus three close parenthesis (pause) plus (pause) fraction with numerator of open parenthesis *y* minus two close parenthesis (pause) and denominator of three (pause) equals nine.

Example 3
Open parenthesis *x* plus four close parenthesis, open bracket, open parenthesis, *x* plus four close parenthesis minus open parenthesis x minus two close parenthesis, close bracket.

Mathematical Exponents (x^2, x^3, 4^5)

Example 1
$y = x^2$

Example 2
$y = 4^5 + 2$

Example 3
$y = 2^{x+5} + 3$

Example 4
$16^{3/2} = 8^2$

Example 5
$3^{5.5} = (z+8)^{x/z}$

Audio Guidelines

Read the base first—the base can be either a numeral or the variable.

If the exponent has a value of 2, then read "squared." If the exponent has a value of 3, read "cubed;" otherwise, read "raised to the_ power" where_ denotes either the ordinal of the number (fourth, sixth, negative seventy-sixth, etc.) if the exponent is an integer or the expression, as specified elsewhere in these guidelines, if the exponent is anything other than an integer.

To indicate a return to the base, use a pause.

Read fraction exponents following the fractions rule.

Application of Audio Guidelines

Example 1
Y equals *x* squared.

Example 2
Y equals four raised to the sixth power (pause) plus two.

Example 3
Y equals two raised to the *x* plus five power (pause) plus three.

Example 4
Sixteen raised to the three halves power equals eight squared.

Example 5
Three raised to the five point five power equals open parenthesis *Z* plus 8 close parenthesis, raised to the fraction with numerator of *x* and denominator of 2 power.

Variables/Letters

Example 1
$x + y = 3$

Example 2
In the triangle, what is the measurement of angle *A* that is opposite side *a*?

Example 3
$N + 4$

Example 4
$-x^3$

Audio Guideline

Read lowercase variables in a math item without referring to case.

If uppercase variables are used in a math item along with lowercase variables, then specify both cases using the language “lowercase” and “uppercase.”

If an uppercase variable appears in a math item without a lowercase variable, then do not specify uppercase.

If a variable is preceded by a negative sign, read as “opposite of” the variable, rather than the “negative of” the variable.

Application of Audio Guideline

Example 1
X plus *y* equals three.

Example 2
In the triangle below, what is the measurement of angle uppercase *A* that is opposite side lowercase *a*?

Example 3
N plus four

Example 4
Opposite of *x* cubed

Logs

Example 1
$Log_{10}100 = 2$

Example 2

69 **If $\log 2 \approx 0.301$ and $\log 3 \approx 0.477$, what is the approximate value of log 72?**

A 0.051

B 0.778

C 0.861

D 1.857

Example 3
ln *x*

Audio Guidelines

Read "log" followed by the base, the word "of," and then the number or variable.

If the log is shown without an explicit base, then read as "log" and the number or variable shown. Do not interpret the implied base of 10 if it is not written.

Read "ln *x*" as "natural log of *x*."

Application of Audio Guidelines

Example 1
Log base ten of one hundred equals two.

Example 2
If log two is approximately equal to zero point three zero one and log three is approximately equal to zero point four seven seven, what is the approximate value of log seventy-two?

Example 3
Natural log of x

Radicals

Example 1
$\sqrt{7}$

Example 2
$\sqrt[4]{144} = \sqrt[x]{288}$

Example 3
$\sqrt[m+n]{x+y}$

Example 4

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Audio Guidelines

For radicals with an implied radical index of two, read as "the square root of *x*."

For radicals with a radical index of three, read as "the cube root of *x*."

For radicals with a number for a radical index other than two or three, start by reading the index as "the *X*th root of."

If the radical index is a variable, read as "the *x* root of *y*."

When multiplying numbers by radicals (e.g.,), say "*x* times the square root of *y*."

Application of Audio Guidelines

Example 1
The square root of two

Example 2
The fourth root of one hundred forty-four equals the x root of two hundred eighty-eight.

Example 3
The *m* plus n root of quantity *x* plus *y*

Example 4
X equals, fraction with numerator of, opposite of *B*, plus or minus the square root of quantity, *B* squared minus four *A C*, and denominator of two *A*.

Absolute Values

Example 1
|–16|

Example 2
|2 + 7|

Example 3
|*x*| + 1

Audio Guidelines

Read as "the absolute value of."

Pause if an absolute value is part of a larger expression or equation.

Application of Audio Guidelines

Example 1
The absolute value of negative sixteen

Example 2
The absolute value of quantity two plus seven

Example 3
The absolute value of x (pause) plus one.

Functions ($f(x)$)

Example 1
$f(x) = 5$

Example 2
$f(x + 1)$

Example 3
$f(g(x))$

Example 4
$f^{-1}(x) = -x - 2$

Audio Guidelines

For function notation in general, read the first letter shown then the word "of," followed by the variable and/or number in parentheses.

When the expression inside the parentheses is more complex or includes another function, use the same rule of reading the letter first, then the word "of," followed by the variable or expression in parentheses.

When the inverse of a function is presented, read it as "f inverse of x."

Application of Audio Guidelines

Example 1
F of *x* equals five

Example 2
F of open parenthesis *x* plus one close parenthesis

Example 3
F of g of *x*

Example 4
The inverse of *f* of *x* equals negative two-thirds *x* minus two.

For function tables where one column/row is paired with one row/column:

The table should be read as it is organized, as (*x*, *y*) pairs, according to p. 44 (If the orientation of the table lends itself to reading the table information column by column and this is a more logical manner to present the table, then do so.)

Example
This table shows a relationship between x and y:

x	*y*
3	14
7	30
9	38

"The table has two columns and three rows. The first column heading is, *x*; the second column heading is, *y*. First row, 3, 14; second row, 7, 30; third row, 9, 38."

Example
This table shows a relationship between *x* and *y*:

x	3	7	9
y	14	30	38

"The table has two rows and three columns. The first row heading is, *x*; the second row heading is, *y*. First column, 3, 14; second column, 7, 30; third column, 9, 38."

System of Equations/Inequalities

Example 1

$$\begin{cases} x + y = 4 \\ x - y = 2 \end{cases}$$

What is the solution to the system of equations?

Example 2

7 Which point lies in the solution set for the system $\begin{cases} 2y - x \geq -6 \\ 2y - 3x < -6 \end{cases}$?

A $(-4, -1)$
B $(3, 1)$
C $(0, -3)$
D $(4, 3)$

Audio Guidelines

Start by reading “system of equations” or “system of inequalities.” Then read the information in the system starting from the top to the bottom; reference the row position and insert a pause between rows.

Read equations and inequalities according to equation and inequality guidelines above.

Application of Audio Guidelines

Example 1
System of equations. Top row, x plus y equals four (pause) bottom row, *x* minus *y* equals two. What is the solution to the system of equations?

Example 2
Which point lies in the solution set for the system, top row, two *y* minus *x* is greater than or equal to negative six (pause) bottom row, two *y* minus three *x* is less than negative six.

Trigonometry

Example 1
sin15°=cos75°

Example 2
tanθ = –1

Audio Guidelines

Read the abbreviated versions of trigonometry functions in full words if doing so does not violate the construct being measured.

If the item is measuring knowledge of these abbreviations read letter by letter.

Use the Greek alphabet in reading trigonometric functions and items. The most used letter is theta (Θ)

Application of Audio Guidelines

Example 1
Sine fifteen degrees equals cosine seventy five degrees

Example 2
Tangent theta equals negative 1

Tables

Example 1

Seashell Collection

Size	Number of Seashells
Small	3
Medium	6
Large	4

Example 2

Rock Types

	Shiny	Air Holes	Flat Layers	Fossils
Metamorphic	X		X	X
Igneous	X	X		
Sedimentary			X	X

Audio Guideline

Text Only

Read the table title only. Allow for all content elements in the table to be read on demand.

Text and Graphics

Read the table title, and then state the number of rows and columns. Then read the column headings from left to right followed by reading the information in each row from left to right.

If the orientation of the table lends itself to reading table information column by column and this is a more logical manner to present the table, then do so.

Read the units of measure for each cell unless they are not specified in the table.

When reading a data table that has blank cells, skip over them if they are unnecessary to answer the question. Blank cells should be read if this information is essential to answer the item.

Remain consistent with the style of reading from table to table. Using a standardized version will help students better understand the patterns of the descriptions.

Many charts that are set up in a table format can be read in the manner described. Determine the layout of such charts before deciding the best way to read the information being presented.

Application of Audio guidelines

Example 1
The table title is Seashell Collection. The table has two columns and three rows. The first column heading is Size, the second column heading is Number of Seashells; first row, Small, three seashells; second row, Medium, six seashells; third row, Large, four seashells.

Example 2
The table title is Rock Types. The table has four columns and three rows. The first column heading is Shiny, the second column heading is Air Holes, the third column heading is Flat Layers, and the fourth column heading is Fossils; first row, Metamorphic, Shiny, Flat Layers, Fossils; second row, igneous, Shiny, Air Holes; third row, Sedimentary, Flat Layers, Fossils.

Tally Charts

Example

Name	Number of Votes
Tigers	𝍸 |
Rockets	|||
Sharks	𝍸 ||
Bobcats	||||

Audio Guideline

Text Only

Read the tally chart title only. Allow for all content elements in the chart except for the tally marks to be read on demand.

Text and Graphics

Read the tally chart title, column headings, and row headings.

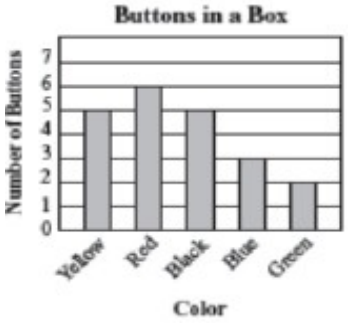
Read the number of tally marks only if it does not violate the construct being measured. If reading tally marks does violate the construct being measured, tactile representation is required to make this item accessible to blind students and some low-vision students.

Application of Text and Graphics Guidelines

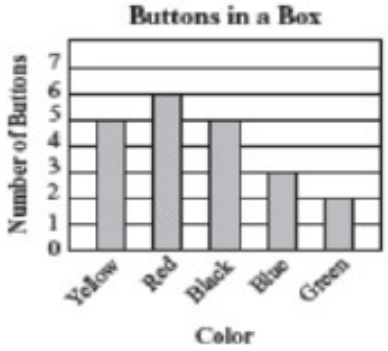
Example
The tally chart has two columns and four rows. The first column heading is Name, and the second column heading is Number of Votes; first row, Tigers, six votes; second row, Rockets, three votes; third row, Sharks, seven votes; fourth row, Bobcats, four votes.

Bar Graphs

Example 1



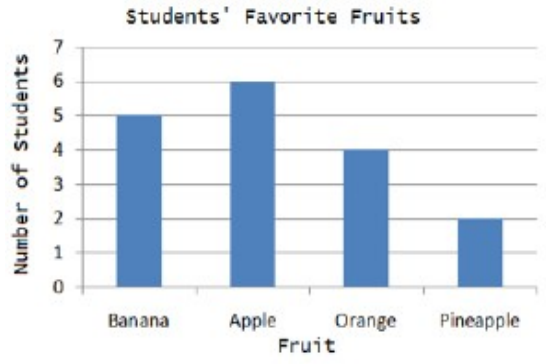
Example 2



How many red buttons are in the box?

Example 3

Kate asked the students in her class what their favorite fruit was. The results of her survey are shown in the graph below.



Audio Guideline

Text Only

Read the bar graph title. Allow for all words and numbers on the bar graph to be available to be read on demand.

Text and Graphics

Read the bar graph title first, followed by the x-axis label and the y-axis label. Do not read values on either axes until describing the bars.

Describe each bar, being careful to take into account the question, so as not to violate the construct being measured. In each description, use the units of measure for the values on the x- and y- axes if applicable.

If a bar is between two horizontal lines, then do not estimate or approximate numbers. Instead, use more general language such as “a little less than,” “a little more than,” and “midway between.”

If the item measures the student’s ability to identify the number associated with the bar, then describe the graph without noting the heights of the bars. In this case, tactile representation is required to make this item accessible to blind students and some low-vision students.

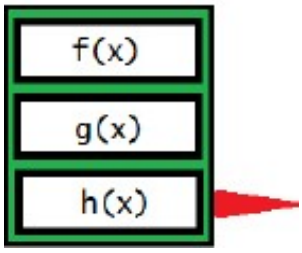
Application of Text and Graphics Guidelines

Example 1
The bar graph title is Buttons in a Box. The x-axis label is Color and the y-axis label is Number of Buttons; Yellow bar, five buttons; Red bar, six buttons; Black bar, five buttons; Blue bar, three buttons; Green bar, two buttons.

Example 2 (item specifically asks students to identify the value associated with a bar)
The bar graph title is Buttons in a Box. The x-axis label is Color and shows five colors: Yellow, Red, Black, Blue, and Green. The y-axis label is Number of Buttons.

Example 3
The bar graph title is Students’ favorite fruits. The x-axis label is Fruit, and the y-axis label is Number of students. Four bars are shown, from left to right, banana, apple, orange, pineapple.

Three functions plotted on a graph



If this graph is described with a tool like that above used to select different graphs on the same coordinate plane, it should be read as follows:

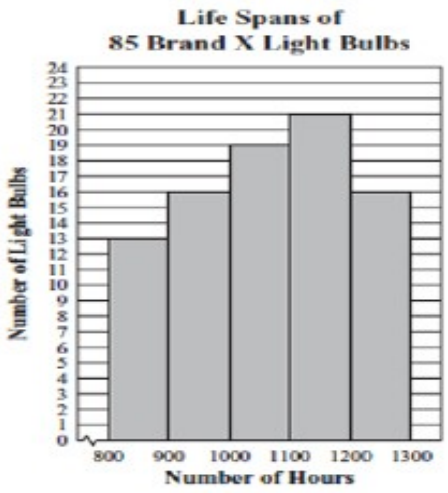
First row, F of X; second row, G of X; third row, H of X.

Note: If only two types of graph can be selected with the tool, it may be appropriate to read according to instructions beginning on page 43 for systems of equations (top row … bottom row …).

Histograms

Example 1

13 Abe tested 85 Brand X light bulbs to determine their life spans. The histogram below shows the results of his test.

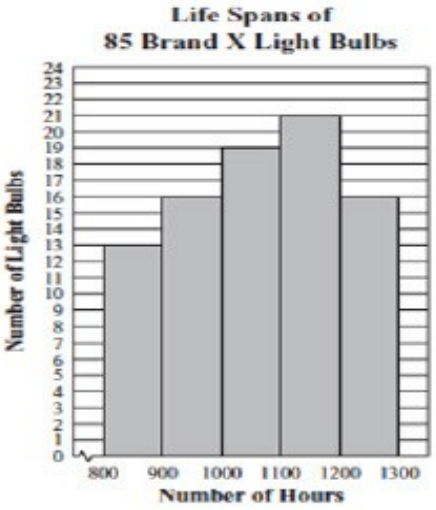


What was the total number of Brand X light bulbs that had life spans greater than or equal to 1000 hours?

A. 72
B. 56
C. 51
D. 21

Example 2

13 Abe tested 85 Brand X light bulbs to determine their life spans. The histogram below shows the results of his test.



What was the total number of Brand X light bulbs that had life spans greater than or equal to 1000 hours?

A. 72
B. 56
C. 51
D. 21

Audio Guideline

Text Only

Read the histogram title. Allow for all words and numbers on the histogram to be available to be read on demand.

Text and Graphics

Read the histogram title first, followed by the x-axis label and the y-axis label.

Describe each bar range on the x-axis, being careful to take into account the question, so as not to violate the construct being measured. In each description use the units of measure on the x- and y-axis labels if applicable.

If a bar is between two horizontal lines, then do not estimate or approximate numbers. Instead, use more general language such as “a little less than,” “a little more than,” and “midway between.”

If the item measures the student’s ability to identify the number associated with the bar, then describe the graph without noting the heights of the bars. In this case, this item is not accessible to blind and some low-vision students without tactile representation.

If there are a large number of bars (more than 10) consider associating bars together or focusing on trends or more general frequency in your description.

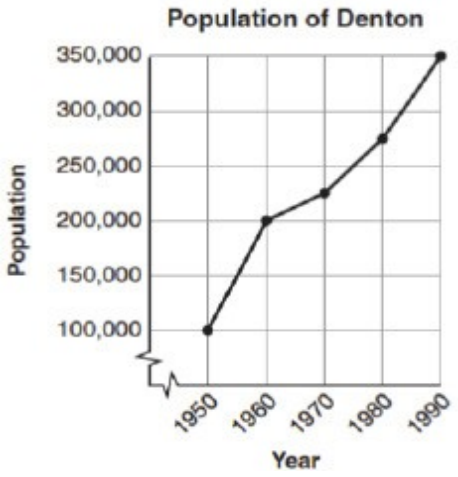
Application of Text and Graphics Guidelines

Example 1
The histogram title is Life Spans of Eighty-Five Brand X Light Bulbs. The x-axis label is Number of Hours and the y-axis label is Number of Light Bulbs; bar one, eight hundred through eight hundred ninety nine hours, thirteen light bulbs; bar two, nine hundred through nine hundred ninety nine hours, sixteen light bulbs; bar three, one thousand through one thousand ninety nine hours, nineteen light bulbs; bar four, one thousand one hundred through one thousand one hundred ninety nine hours, twenty one light bulbs; bar five, one thousand two hundred through one thousand two hundred ninety nine hours, sixteen light bulbs.

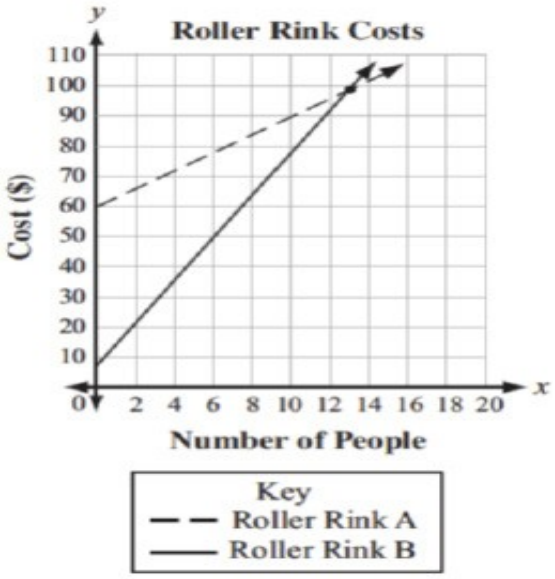
Example 2 (item specifically asks student to read information from one of the bars)
The histogram title is Life Spans of Eighty-Five Brand X Light Bulbs. The x-axis label is Number of Hours and the y-axis label is Number of Light Bulbs. Five bars show the number of light bulbs with a life span of eight hundred through eight hundred ninety nine hours, nine hundred through nine hundred ninety nine hours, one thousand through one thousand ninety nine hours, one thousand one hundred through one thousand one hundred ninety nine hours, one thousand two hundred through one thousand two hundred ninety nine hours.

Line Graphs

Example 1



Example 2



Audio Guidelines

Text Only

Read the graph title only. Allow for all words and numbers in the graph area to be available to be read on demand.

Text and Graphics

For all graphs, read the title first.

Read the Key title and then key section (refer to Key rule specifically).

Read the axis labels.

When describing the graph, be as concise as possible while providing the necessary information to understand and answer the question.

If a line or point being described falls between two marked x- or y-axis values, then do not estimate or approximate numbers Instead, use more general language such as “a little less than,” “a little more than,” and “midway between.”

It is not necessary to describe the visual attributes of the graph unless there is an explicit need, such as a key that references line types or an item referencing the attributes or if doing so would help the student is reading a tactile or a magnified version of the test.

If the description violates the construct being measured, then consider amending it to give less specific information. In this case, tactile representation is required to make this item accessible to blind students and some low-vision students.

When possible, reference the starting and ending point of the line segments or starting points of rays to provide context to the student.

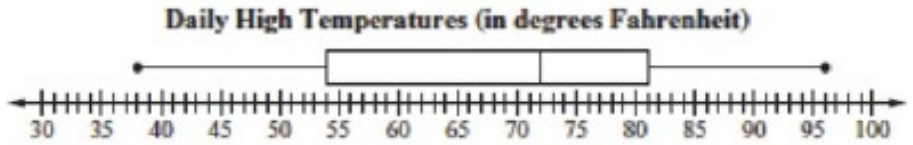
Application of Text and Graphics Guidelines

Example 1
The graph title is Population of Denton. The *x*-axis label is Year and the *y*-axis label is Population. The line starts at nineteen fifty, one hundred thousand, rises to nineteen sixty, two hundred thousand, then nineteen seventy, midway between two hundred and two hundred fifteen thousand, then nineteen eighty, midway between two hundred fifty and three hundred thousand, and ends at nineteen ninety, three hundred fifty thousand.

Example 2
The graph title is Roller Rink Costs. Key, dashed line represents Roller Rink A, solid line represents Roller Rink B. The *x*-axis is labeled Number of People. The *y*-axis is labeled Cost (in dollars). The dashed line starts at zero people, sixty dollars and moves up through midway between twelve and fourteen people, one hundred dollars and fourteen people, a little more than one hundred dollars. The solid line starts at zero people, a little less than ten dollars and moves up through between twelve and fourteen people, one hundred dollars and fourteen people, a little less than one hundred ten dollars.

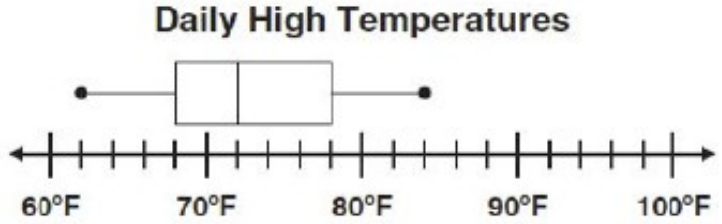
Box Plots

Example 1
The box plot shows the distribution of the daily high temperatures, in degrees Fahrenheit, in the town of Clifton during the year 2004.



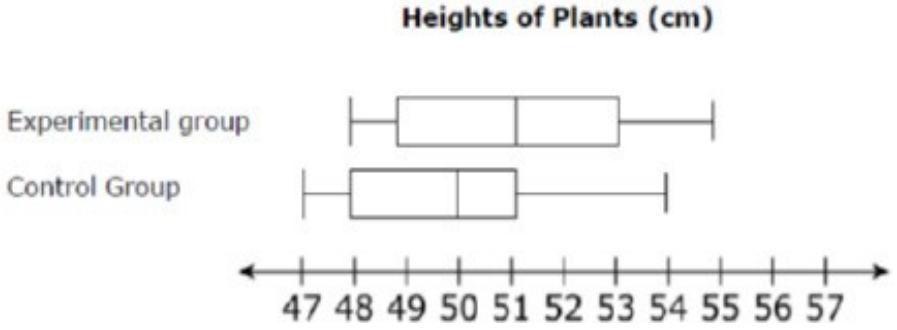
Based on the box plot, in which of the intervals of temperatures is it most likely that exactly 50% of the daily high temperatures are located?

Example 2
The box plot represents the daily high temperatures at a beach in April



What was the median daily high temperature?

Example 3



Audio Guidelines

Read the box plot title. Allow for all words and numbers on the box plot to be available to be read on demand.

Text and Graphics

Start by reading the title of the plot and reference that it is a box plot. Read the box titles or any other words on the plot if applicable.

Read the information along the bottom of the graph from left to right.

If the item measures knowledge of the box plot or if the description violates the construct being measured, then describe the box plot without using specific terminology (e.g., whiskers, quartiles, or median). In this case, tactile representation is required to make this item accessible to blind students and some low-vision students.

If a line or point being described falls between two marked values, then do not estimate or approximate number. Instead use more general language such as “a little less than,” “a little more than,” and “midway between.”

Describe the graph elements using specific box plot terminology—including whiskers, quartiles, box, and median—unless doing so violates the construct being measured.

Application of Text and Graphics Guidelines

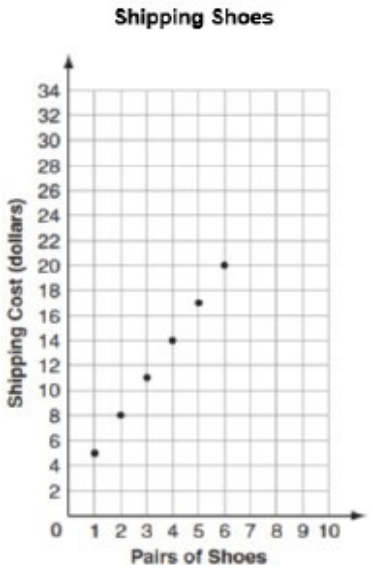
Example 1
The title of the box plot is Daily High Temperatures (in degrees Fahrenheit). The number line ranges from thirty degrees Fahrenheit to one hundred degrees Fahrenheit. The whiskers range from thirty-eight degrees to ninety-six degrees and the box ranges from fifty-four to eighty-one degrees with a median of seventy-two degrees.

Example 2
The title of the box plot is Daily High Temperatures. The number line ranges from sixty degrees Fahrenheit to one hundred degrees Fahrenheit with markers every ten degrees. The whiskers range from sixty-two degrees to eighty-four degrees and the box ranges from sixty-eight degrees to seventy-eight degrees with an **interior vertical line segment** at seventy-two degrees.

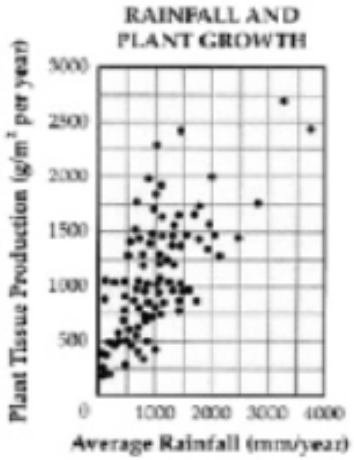
Example 3
The title of the box plot is Heights of Plants (centimeters). The number line ranges from 47 to 57 with markers every whole number. For the experimental group, the whiskers range from 48 centimeters to 55 centimeters and the box ranges from 49 centimeters to 53 centimeters with a median of 51 centimeters. For the control group, the whiskers range from 47 centimeters to 54 centimeters and the box ranges from 48 centimeters to 51 centimeters with a median of 50 centimeters.

Scatter Plots

Example 1



Example 2



Audio Guidelines

Text Only

Read the title of the scatter plot. Allow for all words and numbers on the scatter plot to be available to be read on demand.

Text and Graphics

For scatter plots, start by reading the title and x-axis and y-axis labels. Include the x- and y-axes ranges if necessary to access the item. In some cases, the rightmost extension of the x-axis and/ or topmost extension of the y-axis has no value specified. When specifying the ranges, use either the greatest number listed or the actual value at the rightmost or topmost extension of the axes, whichever is more appropriate.

For a scatter plot with fewer than ten data points, reference each data point. Include units of measure while describing data points only if deemed relevant.

If a line or point being described falls between two marked x- or y-axes values do not estimate or approximate numbers. Instead use more general language such as “a little less than,” “a little more than,” and “midway between.”

If a scatter plot has more than ten data points, then focus on the change of concentration. When possible, read at least a couple of data points (first and last preferably) to put the plot into context.

For some items with scatter plots, tactile representation is required to make the item accessible to blind students and some low-vision students.

Application of Text and Graphics Guidelines

Example 1
The graph is a scatter plot titled “Shipping Shoes.” The x-axis is labeled Pairs of Shoes and ranges from zero to ten in increments of one. The y-axis is labeled Shipping Cost (dollars) and ranges from

zero to thirty-four in increments of two. The scatter plot has points at one, midway between four and six; two, eight; three, midway between ten and eleven, four, fourteen; five, midway between sixteen and eighteen; and six, twenty.

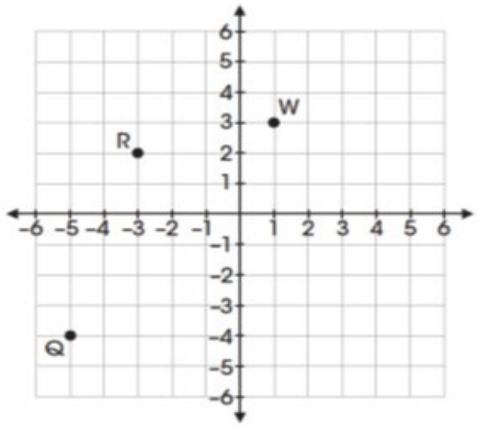
Example 2

The graph is a scatter plot titled Rainfall and Plant Growth. The x-axis is labeled Average Rainfall and ranges from zero to four thousand, in units of millimeters per year, in increments of one thousand. The y-axis is labeled Plan Tissue Production in units of grams per meter squared per year, ranging from zero to three thousand, in increments of five hundred. The graph has approximately eighty-five points scattered in a pattern beginning in the lower-left corner where Plant Tissue Production and Average Rainfall are the lowest. The pattern extends toward the upper-right corner where Plant Tissue Production and Average Rainfall are the highest. The majority of points is concentrated in the lower-left corner and diminishes in concentration as the pattern extends toward the upper-right corner.

Coordinate Planes

Example 1

23. Points *Q*, *R*, and *W* are plotted on the coordinate grid.

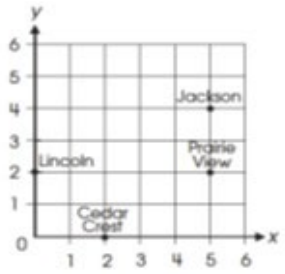


Where should point *Z* be plotted so that parallelogram QRWZ is formed?

A. (–2, –6)
B. (–1, –3)
C. (3, –2)
D. (2, –1)

Example 2

18. Mr. Yang is driving to the school located at (2, 0) on the coordinate grid.

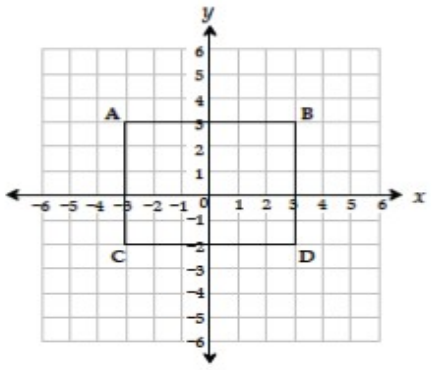


Which school is located at (2, 0)?

- A. Cedar Crest
- B. Jackson
- C. Lincoln
- D. Prairie View

Example 3

Use the diagram below to answer question 7.



7. **Which ordered pair identifies the location of vertex C?**

A	$(-3, -2)$ *	
B	$(-3, 3)$	*vertex A*
C	$(3, -2)$	*vertex D*
D	$(-2, -3)$	*vertex C reversed*

Audio Guidelines

Text Only

Start by reading the title of the coordinate plane. Allow for all words and numbers on the coordinate plane to be available to be read on demand.

Text and Graphics

Read the title of the coordinate plane first.

Read the range of each axis. In some cases, the extensions of the *x*- and/or *y*-axis have no value specified. When specifying the ranges, use either the greatest (or least for bottom and left extensions) number listed or the actual value at the furthest extension of the axes, whichever is more appropriate.

Read the points or words on the grid in a logical manner (clockwise, following the listing of a shape, etc.) referencing their location on the grid.

If a line or point being described falls between two marked *x*- or *y*-axis values, then do not estimate or approximate numbers. Instead, use more general language such as "a little less than," "a little more than," and "midway between."

If reading the location of the points violates the construct being measured, do not read the point, but reference that they are on the grid. In this case, tactile representation is required to make the item accessible to blind students and some low-vision students.

If there is a shape on the grid, then read the type of shape or name of it first, and then reference the axis points of all sides, if relevant. If referencing the axis points violates the construct being measured, then provide a description of the shape without these points.

If an empty grid is presented in an item as part of the prompt, question. Or answer, then read the title and the *x*- and *y*- axes scale.

Application of Text and Graphic Guidelines

Example 1
A coordinate plane with *x*- and *y*-axes ranging from negative six to six; point *Q*, negative five, negative four; point R, negative three, two; and point *W*, one, three.

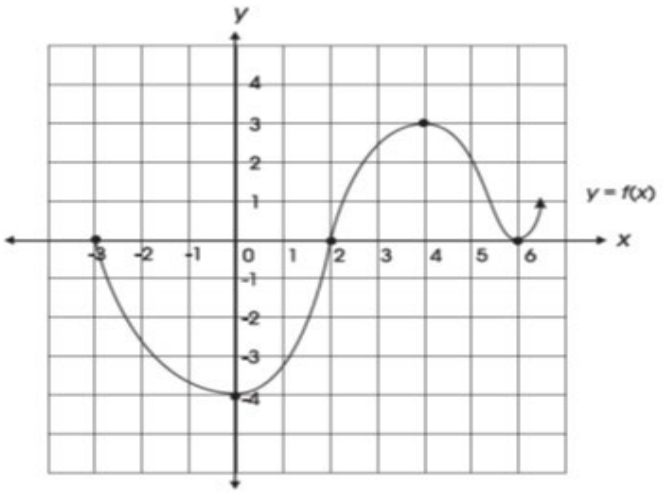
Example 2
A coordinate plane with x- and y-axes ranging from zero to six. The grid shows the location of the four schools: Jackson, Prairie View, Cedar Crest, and Lincoln.

Example 3
A coordinate plane with *x*- and *y*-axes ranging from negative six to six. Rectangle *ABCD* is shown on the grid.

Exponential/Linear Function Graphs

Example 1

31. The graph of the function $f(x)$ is shown below.

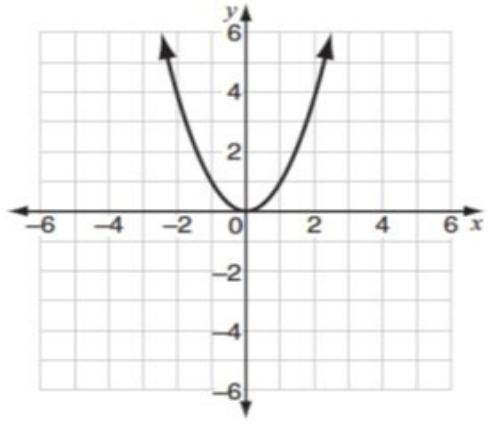


Which of the following is NOT a zero of $f(x)$?

A. –4
B. –3
C. 2
D. 6

Example 2

10 Look at this graph of $y = x^2$.



If $y = x - 2$ is graphed on the same coordinate plane, at how many points would the two graphs intersect?

A. 0
B. 1
C. 2
D. 3

Audio Guidelines

Text Only

Start by reading the title of the graph. Allow for all words and numbers on the graph to be available to be read on demand.

Text and Graphics

Read the title of the graph first.

Read the range of each axes and any words or symbols that are on the graph. In some cases, the extensions of the x- and/or y-axis have no value specified. When specifying the ranges, use either the greatest (or least for bottom and left extensions) number listed or the actual value at the furthest extension of the axes, whichever is more appropriate.

Describe the shape of the graph. Use relevant points including starting and ending points or x or y intersection points to aid the description.

If a line or point being described falls between two marked x- or y-axes values, then do not estimate or approximate numbers. Instead use more general language such as “a little less than,” “a little more than,” and “midway between.”

If reading the location of any points violates the construct being measured, then do not read these points. If describing the shape or direction of the graph violates the construct, then do not read the details of the shape of the graph. In this case, tactile representation is required to make the item accessible to blind students and some low-vision students.

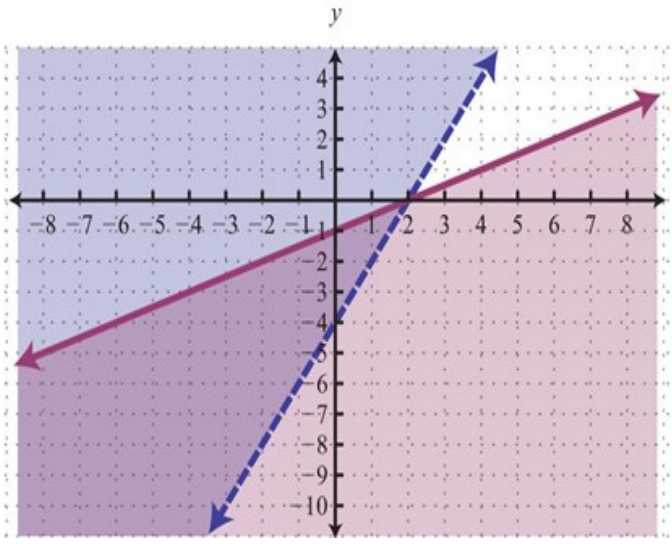
Application of Text and Graphics Guidelines

Example 1
A graph showing the function *y* equals *f* of *x*. The x-axis ranges from negative four (or three) to seven (or six), and the *y*-axis ranges from negative six (or negative four) to five (or four). The graph is in the shape of a wave. The graph starts at negative three zero, goes through zero negative four, then two zero, then four three, then six zero, and ends with an arrow pointing up at a midway between six and seven, one.

Example 2
A graph showing y equals x squared. The *x*- and *y*-axes ranges from negative six to six. The graph is a parabola that starts with an arrow at midway between negative two and negative three, six, and then the line moves down through zero zero, and ends with an arrow at midway between two and three, six.

System of inequalities

Example
Which graph represents the solution to this system of inequalities?
$y > 2x - 4$
$3x - 6y \geq 6$



Application of Audio Guidelines

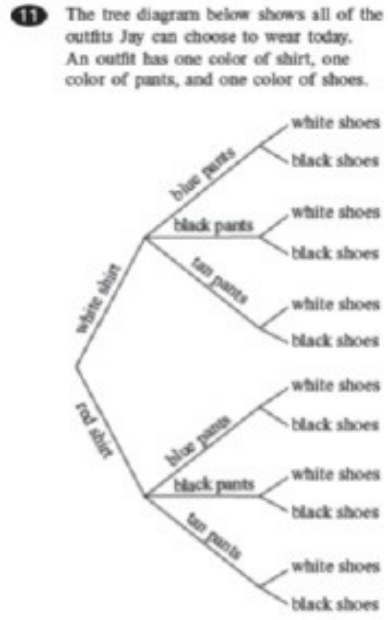
Text and Graphics

Which graph represents the solution to this system of inequalities, top row, Y is greater than 2 X minus 4; bottom row, 3 X minus 6 Y is greater than or equal to 6. A. A graph showing two lines and shaded regions. The X axis ranges from negative 9 to 9. The y axis ranges from negative 11 to 5. The purple line is solid and starts at negative 9, a little less than negative 5; rises to zero, negative 1; then 2, zero; and ends at 9, a little more than 3. The area below the solid line is shaded purple. The blue line is dashed and starts at a little less than negative 3, negative 11; rises to zero, negative 4; then 2, zero; and ends at a little more than 4, 5. The area to the left of the dashed line is shaded blue. The area in between the solid purple line and the dashed blue line is shaded light gray.

Diagrams/Figures/Keys

Tree Diagram

Example 1



Audio Guidelines

Text Only

Read the tree diagram title. Allow for all words and numbers on the tree diagram to be available to be read on demand.
Text and Graphics

Read the tree diagram title and brief description along with stating the direction of the tree diagram.

Start with the innermost parts of the tree and describe the different limbs in an order that is easy to follow.

Describe all of the elements of the tree diagram with standardized language.

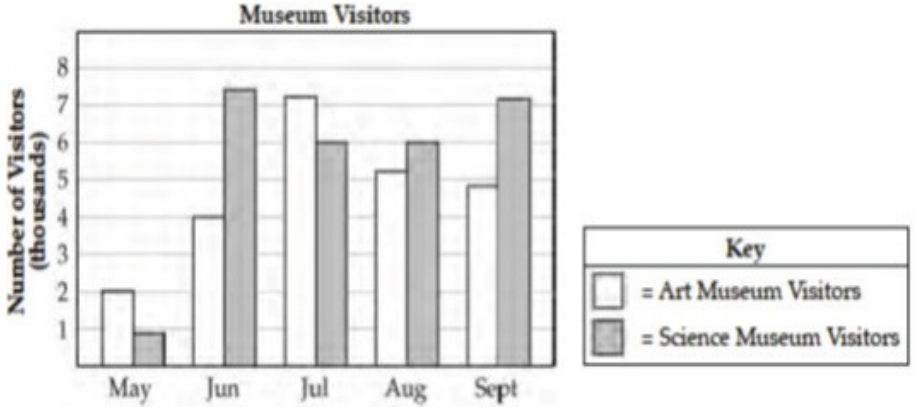
Application of Text and Graphics Guidelines

A tree diagram showing outfit combinations of shirts, pants, and shoes. The diagram displays information from left to right starting with shirts on the leftmost branches. On the top half of the

tree, white shirt branches to blue pants, black pants, and tan pants. Each of these pants branches stems to the outermost branches of white shoes and black shoes. On the bottom half of the tree, red shirt branches to blue pants, black pants, and tan pants. Each of these pants branches stems to the outermost branches of white shoes and black shoes.

Keys

Example



Audio Guidelines

Text Only

Read the word Key after reading the graph/diagram title. Allow for all words and numbers in the key to be available to be read on demand.

Text and Graphics Guidelines

Read the graph/diagram title and then the key.

Describe the key in detail, including shapes, shades, and so on. Use “represents” to associate icon with text. (e.g., –10 miles. Dashed line represents ten miles.)

Read the graph/diagram using the key symbols. (e.g., May, white bar, two; May, gray bar, a little less than one)

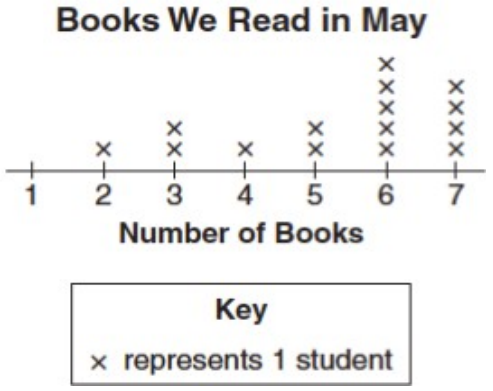
Application of Text and Graphics Guidelines

Example
The bar graph title is Museum Visitors. In the Key, the white bar represents Art Museum Visitors, while the gray bar represents Science Museum Visitors. The x-axis shows five months; the y-axis is labeled Number of Visitors (thousands); May, white bar, two; May, gray bar, a little less than one; June, white bar, four; June, gray bar, midway between seven and eight; July, white bar, a little more than seven; July, gray bar, six; August, white bar, a little more than five; August, gray bar, six; September, white bar, a little less than five; September, gray bar, a little more than seven.

Line Plots

Example

16 Look at this line plot.



Audio Guideline

Text Only

Read the line plot title. Allow for all words and numbers on the line plot and on the key to be available to be read on demand.

Text and Graphics

Read the title of the line plot, the key, and then the *x*-axis title (refer to this as the number line plot title if the term "axes" has not been taught in the grade being assessed).

Use the key symbol to describe the line plot instead of interpreting the symbol.

If there are no x's or symbols above a number, then read this as zero instead of skipping it.

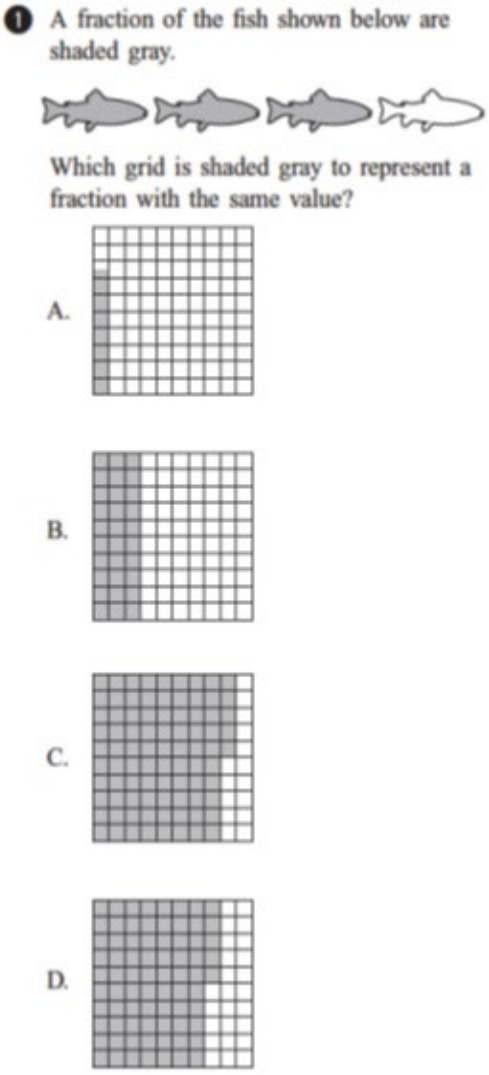
Be careful not to violate the construct being measured. Read the range of numbers on the *x*-axis without reading the data, if necessary. In this case, tactile representation is required to make the item accessible to blind students and some low-vision students.

Application of Text and Graphics Guidelines

Example
The title of the line plot is Books We Read in May. The key shows that an x represents one student. The number line title is Number of Books and ranges from one to seven in increments of one; at line plot one, zero x's are shown; at line plot two, one x is shown; at line plot three, two x's are shown; at line plot four, one x is shown; at line plot five, two x's are shown; at line plot six, five x's are shown; and at line plot seven, four x's are shown.

Shaded Figures (Grids, Bars, and Shapes)

Example



Audio Guidelines

Text Only

Read the title of the shaded figure. Allow for all words and numbers in the figure to be available to be read on demand.

Text and Graphics

Read the title if there is one, and then describe the dimensions of the figure first. If possible, read the dimensions of the figure (ten by ten) instead of just the number of boxes.

Explain how many boxes are shaded, but do not use the terminology "*x* of *y*" boxes are shaded. This creates the fraction for the student and will often violate the construct being measured.

Do not state the total number of boxes shaded when information can be provided that students should use to determine the number of boxes shaded. (e.g., seven columns of ten boxes shaded, instead of seventy boxes)

Application of Text and Graphics Guidelines

Example
A fraction of the fish shown below is shaded gray. The graphic shows four fish. Three of them are shaded gray.
Which grid below is shaded gray to represent a fraction with the same value?
A: ten by ten box grid with seven boxes shaded
B: ten by ten box grid with three columns of ten boxes shaded
C: ten by ten box grid with eight columns of ten boxes shaded and five additional boxes shaded
D: ten by ten box grid with seven columns of ten boxes shaded and five additional boxes shaded

Pictographs

Examples

Dogs at the Park

Type of Dog	Number of Dogs
Beagle	
Collie	
Poodle	
Dalmatian	

Key
 represents 1 dog

Audio Guidelines

Text Only

Read the title of the pictograph. Allow for all words and numbers in the pictograph or key to be available to be read on demand.

Text and Graphics

Start by reading the title of the pictograph and then the key.

If the pictograph is in a table format, then refer to the table guidelines.

If the pictograph is in a graph format, then refer to the graph guidelines.

Reference the picture being used in general terms without describing it in detail. Use the key to read the pictograph without interpreting it. When the pictograph, reference “picture of x,” since the scale may not be one to one.

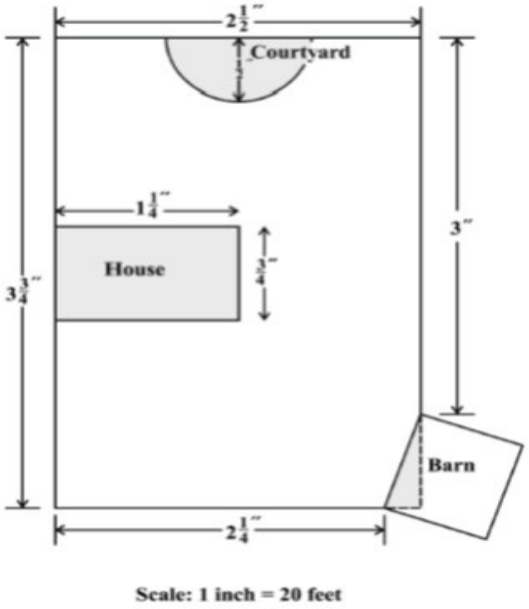
In some cases, tactile representation is required to make the item accessible to blind students and some low-vision students.

Application of Text and Graphics Guidelines

Example
The pictograph title is Dogs at the Park. The Key shows a picture of a dog represents one dog. The table has two columns and four rows; column heading one is Type of Dog; column heading two is Number of Dogs; row one, Beagle, picture of two dogs; row two, Collie, picture of three dogs; row three, Poodle, picture of one dog; row four, Dalmatian, picture of four dogs.

Figures/Illustrations

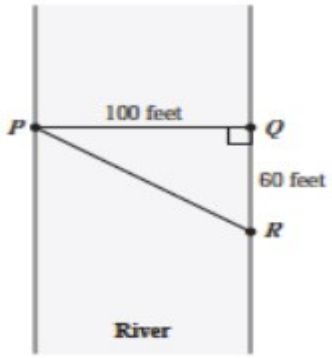
Example 1



Use the scale to find the actual dimensions, in feet, of the house. Show or explain how you found your answer.

Example 2

40 Triangle PQR in the diagram below represents Pam's trip across a river.



In the diagram, $\overline{PQ}$ represents her planned trip across the river, and $\overline{PR}$ represents her actual trip across the river.

Based on the dimensions in the diagram, which of the following is closest to the length of $\overline{PR}$?

A. 104 feet
B. 117 feet
C. 120 feet
D. 160 feet

Audio Guidelines

Text Only

Read the title of the figure/illustration or any caption that is being used in the title format. Allow for all words and numbers in the pictograph or key to be available to be read on demand.

Text and Graphics

Read the title of the figure or illustration. Include the caption in the description if it is not included in the surrounding text.

Read any scale before describing parts of the figure.

Separate the information into pieces using sentences, bullet points, or lists.

Use similar language to describe all parts of the diagram or illustration. Standardized language will help ensure comprehension.

Remember that the goal is to help the student understand the pertinent information in the diagram. Try to include descriptions of all shapes and figures, but try not to overload the student with descriptions that are overly wordy or not needed to answer the question.

Application of Text and Graphics Guideline

Example 1

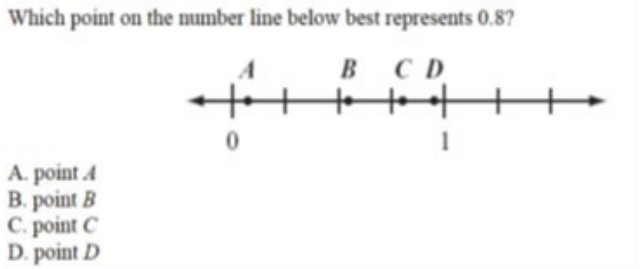
A drawing showing a rectangular plot of land is illustrated. The scale shows that one inch equals twenty feet. The left and right sides of the plot are three and three-fourths inches, and the top and bottom sides of the plot are two and a half inches. The rectangular house has side lengths of one and one-fourth inches and three-fourths of an inch. The barn is a square, mostly outside the plot, with a shaded right triangle inside the plot. The hypotenuse of the right triangle and the side of the square inside the plot are the same line segment. One corner of the triangle is at the two and one-fourth inch line at the bottom of the plot and another corner is at the three inch line on the side of the plot. The courtyard is a semicircle with a radius of one-half inch.

Example 2

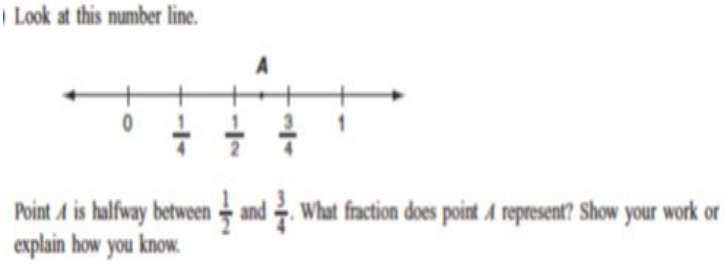
A diagram showing a rectangular section of a river is illustrated. Triangle PQR shows Pam's trip across the river with all three points of the triangle touching a side of the river. Point P is on the left side of the river, and points Q and R are on the right side of the river. Point Q is the vertex of a right angle. The distance from P to Q is one hundred feet. The distance from Q to R is sixty feet.

Number Lines

Example 1

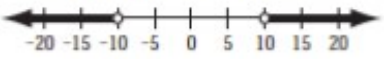


Example 2



Example 3

37 The graph below is the solution of which of the following inequalities?



A. $|x| > 10$

B. $|x| < 10$

C. $x > 10$

D. $x < -10$

Audio Guidelines

Text Only

Read the title of the number line only or any caption that is being used in the title format. Allow all letters, words, and number on the number line to be available on demand.

Text and Graphics

Start by reading the title of the number line.

Read the range on the bottom along with the increments displayed.

Read the letters or words on the number line along with their location. Be careful not to violate the construct being measured in doing so. In some cases, tactile representation is required to make the item accessible to blind students and some low-vision students.

If a line or point being described falls between two marked values, then do not estimate or approximate numbers. Instead, use more general language such as “is located a little after,” “is located a little before,” “is closer to,” and “is midway between.”

For bolded number lines, describe which parts are bolded.

Application of Text and Graphics Guidelines

Example 1
A number line is shown with points *A*, *B*, *C*, and *D* and three equally spaced tick marks between the values of zero and one. Point *A* is located between zero and the first tick mark, and is closer to zero; point B is located between the second and third tick marks, and is much closer to the second tick mark; while point *C* and point *D* are closer to the value one.

Example 2
A number line shows zero and one with three tick marks in between: one-fourth, one-half, and three-fourths. Point *A* is marked midway between one-half and three-fourths.

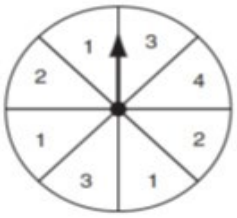
Example 3

A number line shows from negative twenty to positive twenty in increments of five. The areas from negative twenty to negative ten and positive ten to positive twenty are bolded with open circles at negative ten and positive ten. There are bolded arrows to the left of negative twenty and to the right of positive twenty.

Spinners

Example 1

10 Look at this spinner.

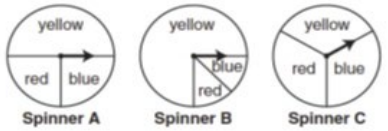


On what number is the arrow **least** likely to land?

- A. 1
- B. 2
- C. 3
- D. 4

Example 2

15 Look at these spinners.



Spinner A Spinner B Spinner C

Julie, Greg, and Lori each used a different spinner to record the results of 40 spins.

a. This table shows Julie's results.

Julie's Spinner Results

Color	Frequency
yellow	12
blue	14
red	14

Which spinner did Julie **most likely** use? Show your work or explain how you know.

b. This table shows Greg's results.

Greg's Spinner Results

Color	Frequency
yellow	30
blue	5
red	5

Which spinner did Greg **most likely** use? Show your work or explain how you know.

c. Lori used the remaining spinner. Make a table to show the **most likely** results of Lori's 40 spins. Explain your reasoning.

Audio Guidelines

Text Only

Read the title of the spinner only. Allow for all letters, words, and numbers on the spinner to be available on demand.

Text and Graphics

Read the title of the spinner and reference it as a spinner.

Read any words, symbols, or numbers in the spinner, starting at the top and moving clockwise.

If necessary, describe the sizes of each section. Be sure not to violate the construct being measured in doing so. In some cases, tactile representation is required to make the item accessible to blind students and some low-vision students.

When describing the size of sections, do not estimate or approximate a specific size if it is not labeled. Instead, use more general language such as “less than,” “more than,” and “half of.” Exceptions are for one-fourth, one-third, one-half, two-thirds, and three-fourths that are immediately apparent.

Application of Text and Graphics Guidelines

Example 1

Grades 7 and lower: A spinner is divided into eight sections of the same size. One number in each section is shown. From the top moving clockwise, the sections read three, four, two, one, three, one, two, one.

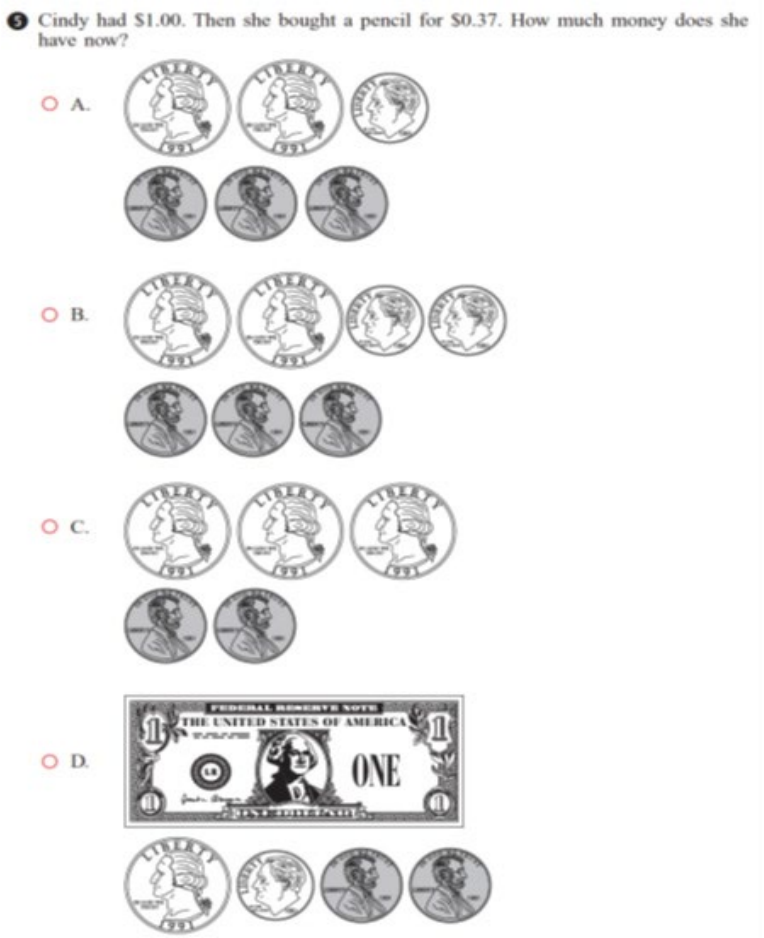
Grades 8 and higher: A spinner divided into eight congruent sections. One number in each section is shown. From the top moving clockwise, the sections read three, four, two, one, three, one, two, one.

Example 2

There are three spinners shown labeled Spinner A, Spinner B, and Spinner C. Each spinner is divided into three sections. In Spinner A, one-half of the spinner is labeled yellow, one-fourth of the spinner is labeled blue, and one-fourth of the spinner is labeled red. In Spinner B, three-fourths of the spinner is labeled yellow, and the other part is divided evenly and labeled blue and red. In Spinner C, about one-third of the spinner is labeled yellow, about one-third of the spinner is labeled red, and about one-third of the spinner is labeled blue.

Coins and Dollars

Example



Audio Guidelines

Text and Graphics

Describe the money using standard language (penny, dime, quarter, or dollar).

Be sure to read each currency symbol as a symbol and not to interpret the value. (e.g., two quarters instead of fifty cents, or three dimes instead of thirty cents)

If reading the currency symbols violates the construct being measured, tactile representation is required to make the item accessible to blind students and some low-vision students.

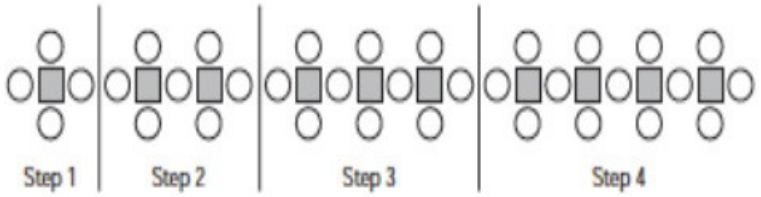
Application of Audio Guidelines

Example
A shows two quarters, one dime, and three pennies.
B shows two quarters, two dimes, and three pennies.
C shows three quarters and two pennies.
D shows one one-dollar bill, one quarter, one dime, and two pennies.

Numbered/Step Diagrams

Example

9 Don made a pattern using circles and squares. The first four steps of his pattern are shown below.



If Don continues his pattern, what is the total number of circles he will need to make Step 10?

A. 30
B. 31
C. 38
D. 40

Audio Guideline

Text Only

Read the title of the diagram only. Allow for all letters, words, and numbers on the diagram to be available to be read on demand.

Text and Graphics
Read the title of the diagram and a brief orientation of what the diagram shows.

In logical order (left to right or top to bottom), read the steps or diagram numbers along with a description of the figures in each step.

Describe the figures with enough detail to understand the item. Unless necessary, do not detail the specific characteristics of the figures being used. (e.g., color, size, location, shape, etc.)

If the description violates the construct being measured (e.g., if the question asked "How many circles are in step 1?"), then adjust the description to be vague. In this case, tactile representation is required to make the item accessible to blind students and some low-vision students.

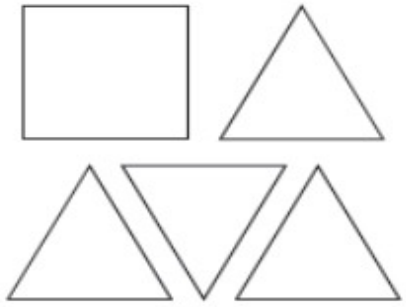
Application of Audio Guidelines

Example
A diagram shows four steps of a pattern using circles and squares. Step one shows a square and four circles, step two shows two squares and seven circles, step three shows three squares and ten circles, and step four shows four squares and thirteen circles.

Geometric Figures

Example 1

These shapes are the 5 faces of a three-dimensional figure.

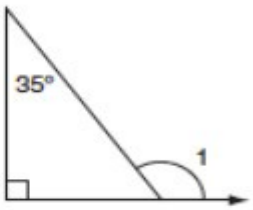


What is the three-dimensional figure?

A. cube
B. cone
C. prism
D. pyramid

Example 2

Look at this diagram.

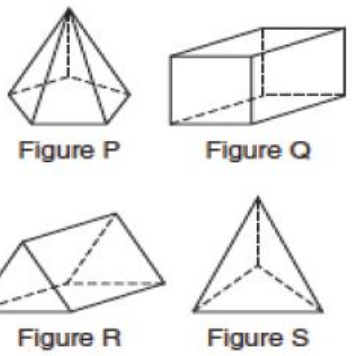


What is the measure of ∠1?

A. 55°
B. 115°
C. 125°
D. 135°

Example 3

Look at these figures.



Which two figures have the same number of faces?

A. Figure P and Figure Q
B. Figure S and Figure R
C. Figure P and Figure R
D. Figure S and Figure Q

Look at these figures.

Audio Guidelines

Text Only

Read the title of the shape(s) only. Allow for all labels of sides or angles to be available on demand.

Text and Graphics

Simple shapes (any 2D shape with eight sides or fewer): Reference simple shapes as is, unless the item is measuring identification of a shape. If the item contains a simple shape, reference it without description. If there are unique attributes to the shape, describe what type of shape it is in as few words as possible. Be sure to reference labels of s ides, angles, and so on.

3D shapes/figures: Reference the type of figure. If relevant and does not violate the construct being measured, describe the figure including the number of sides. In some cases, if a certain description would violate the construct, tactile representation is required to make the item accessible to blind students and some low-vision students.

Be sure to reference labels of sides, angles, and so on.

Refer to the coordinate plane section for reading shapes on coordinate planes.

Application of Text and Graphics Guidelines

Example 1
A square and four equally sized triangles are shown.

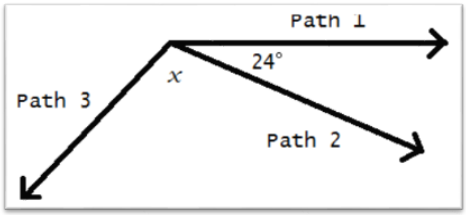
Example 2
A diagram shows a right triangle. The triangle shows a right angle in the left corner, a thirty-five degree angle at the top, with no angle reference in the bottom-right corner. Outside the bottom-right corner of the triangle there is a symbol for angle one, which arcs from the unknown angle in the triangle to touch the ray.

Example 3
Four figures are shown. Figure P is a pentagonal pyramid, Figure Q is a rectangular prism, Figure R is a triangular prism, and Figure S is a triangular pyramid.

For geometric figures with multiple lines

Diagrams with internal angles should generally be described clockwise, beginning at the 12:00 position or a logical point of origin in the diagram.

Example
Bicyclists at National Park can choose one of three bike paths from the visitors' center, as shown in this diagram.



A diagram shows three rays, each originating at the same point. The first ray, drawn horizontally to the right, is labeled Path 1. The second ray, labeled Path 2, is drawn downward and toward the right. The angle that includes Path 1 and Path 2 is labeled 24 degrees. The third ray, labeled Path 3, is drawn downward and to the left. The angle that includes Path 2 and Path 3 is labeled *x*.

References

Smarter Balanced Assessment Consortium: Mathematics Audio Guidelines. http://www.smarterbalanced.org/assessments/development/.

Appendix H: Human Signer Guidelines

Test Administration Protocol for the Human Signer Accommodation for English Language Arts (ELA) Assessments, and the Human Signer Accessibility Feature for Mathematics Assessments

In cases where a student requires a sign language accommodation on the English language arts (ELA) assessments and/or a sign language accessibility feature on the mathematics assessments, and for whom the American Sign Language (ASL) video accommodation is not appropriate, a human signer must be provided. Human signers for 2023 Math and ELA assessments must follow these procedures during testing to ensure the standardization of the signed presentation to the students.

Procedures for Human Signers Providing the Human Signer Accommodation for the ELA Assessments or the Human Signer Accessibility Feature for the Mathematics Assessments

1. Signers must be trained on test administration policies by local Test Coordinators, as indicated in the *Test Administrator Manuals* (TAM). Signers must sign the Staff Confidentiality Agreement available at https://webnew.ped.state.nm.us/wp-content/uploads/2021/07/StaffConfidentialityAgreement.pdf.
2. Signers should use signs that are conceptually accurate (except for SEE2 users), with or without simultaneous voicing, translating only the content that is printed in the test book or on the computer screen without changing, emphasizing, or adding information. Signers may not clarify (except for test directions), provide additional information, assist, or influence the student's selection of a response in any way. Signers must do their best to use the same signs if the student requests a portion repeated.
3. Signers must sign (or sign and speak when using Sim-Com [Simultaneous Communication]) in a clear and consistent manner throughout test administration, using correct production, and without inflections that may provide clues to, or mislead, a student. Signers should be provided a copy of the test and the *Test Administrator's Manual* (which includes the test administrator's directions) two school days prior to the start of testing, in order to become familiar with the words, terms, symbols, signs, and/or graphics that will be signed to the student. Review of the test materials must occur in a SECURE ENVIRONMENT.
4. Signers should emphasize only the words printed in boldface, italics, or capital letters and inform the student that the words are printed that way. No other emphasis or inflection is permitted.
5. Signers may repeat passages, test items, and response options, as requested, according to the needs of the student. Signers should not rush through the test and should ask the student if they are ready to move to the next item.
6. Signers may not attempt to solve mathematics problems, or determine the correct answer to a test item while signing, as this may result in pauses or changes in inflection which may mislead the student.
7. Signers must use facial expressions consistent with sign language delivery and must not use expressions which may be interpreted by the student as approval or disapproval of the student's answers.
8. Test Administrators must be familiar with the student's Individualized Education Plan (IEP) or 504 plan, and should know in advance which accommodations are required by the student,

and for which test (NM-ASR, NM-MSSA Math, NM-MSSA ELA, and/or NM-MSSA SLA) the student is designated to receive a human signer. Test Administrators must be aware of whether a student requires additional tools, devices, or adaptive equipment that has been approved for use during the test, such as a magnifier, closed circuit television (CCTV), abacus, brailler, slate, stylus, etc., and if use of these tools impacts the translation of the test, the signer should be made aware of this.

9. Upon review of the test, if a human signer is unsure how to sign and/or pronounce an unfamiliar word, the signer should collaborate with an ASL-fluent content expert (if available) which sign is most appropriate to use. If the signer is unable to obtain this information before the test, the signer should advise the student of the uncertainty and spell the word.

10. When using an ASL sign that can represent more than one concept or English word, the signer must adequately contextualize the word, in order to reduce ambiguity. The signer may also spell the word after signing it, if there is any doubt about which word is intended.

11. Signers must spell any words requested by the student during the test administration.

12. When test items refer to a particular line, or lines, of a passage, resign the lines before signing the question and answer choices. For example, the signer should sign, "Question X refers to the following lines…," then sign the lines to the student, followed by question X and the response options.

13. When signing selected response items, signers must be careful to give equal emphasis to each response option and to sign options before waiting for the student's response.

14. When response choices will be scribed, the signer should inform the student at the beginning of the test that if the student designates a response choice by letter only ("D", for example), the signer will ask the student if he/she would like the response to be signed again before the answer is recorded in the answer booklet or the computer-based test.

15. If the student chooses an answer before the signer has signed all the answer choices, the human signer must ask if the student wants the other response options to be signed.

16. After the signer finishes signing a test item and all response options, the signer must allow the student to pause before responding. If the pause has been lengthy, ask: "Do you want me to sign the question or any part of it again?" When signing questions again, signers must avoid emphasis on words not bolded, italicized, or capitalized.

17. Signers should refer to the ASL Glossary for technical vocabulary (signs used on the ASL video accommodation) for consistency in providing the accommodation.

Procedures for Providing the Human Signer Accommodation for ELA Assessments or the Human Signer Accessibility Feature for the Mathematics Assessments to a Small Group of Students

Human signers may sign the test to a small group of students, rather than individually, provided that each student has the human signer accommodation/accessibility feature listed in an IEP or 504 Plan. See PED policy for group size and TA to student ratios.

The following procedures must be followed:

- Check individual state policies on the maximum allowable number of students in a human signer small group.
- Students with the human signer accessibility feature for mathematics or human signer accommodation for ELA that will be grouped together must be administered the SAME TEST FORM, since test questions will differ on each form of the test. In Spring 2023, all paper forms are the same.
- Students not receiving the human accessibility feature for mathematics or human signer accommodation for ELA may not be tested in the same location as students who are receiving the human signer accessibility feature for mathematics or human signer accommodation for ELA.

Sign-System-Specific Procedures

Human signers must deliver the accommodation in the language or communication mode used by the student according to the student's IEP or 504 plan.

American Sign Language (ASL)

Human signers delivering the accommodation via ASL must use appropriate ASL features (including signs, sentence structure, non-manual markers, classifiers, etc.) while protecting the construct being measured by the assessment. Although it is necessary for a human signer to use appropriate non-manual markers to ensure proper delivery of test content in ASL, the human signer must be careful not to cue students while doing so.

English-Based Sign Systems (SEE2, CASE, Sim-Com, etc.)

Human signers delivering the accommodation via an English-based signing system must use the features of the communication mode used by the student. Human signers delivering the test in Signing Exact English (SEE2) should use the rules of that signing system (e.g. specific signed vocabulary, prefixes, suffixes, etc.). Human Signers delivering the test in other English-based signing systems (CASE, Sim-Com, etc.) should use the rules of those signing systems (conceptually accurate signs, English word order, etc.), with or without simultaneous voicing.

Mathematics Sign Language Glossary

Human signers should refer to the online Mathematics Sign Language Glossary for guidance on how to deliver mathematics symbols and terms. The guidance provided in the glossary is the same as what has been used in development of the ASL video accommodated 2023 Math and ELA assessments and provides a standardized approach for students who use sign language accommodations. The glossary provides signs that can be used for both ASL and English-Based Sign Systems.

Appendix I: The 2023 Science/Math/ELA Assessments for Students with Visual Impairments, Including Blindness

2023 Science, Math, and ELA Assessments and Students with Visual Impairment, Including Blindness

I. Purpose of this Guidance

The 2023 Science, Math, and ELA Assessments are provided online, in regular print, large print and braille. This document is for Test Coordinators, Test Administrators, test transcribers and teachers to clarify issues and potential questions for students with visual impairments, including blindness. Given the innovative approach to the 2023 Science, Math, and ELA assessments, students with visual impairments who receive instructional and assessment accommodations, and those professionals that work with them, will need to plan ahead for testing to ensure that students have all necessary tools and materials available to complete assessment tasks. All accommodations must be documented in the student's Individualized Education Program (IEP) or 504 plan.

II. Frequently Asked Questions (FAQ)

1. Who is an Eligible Test Administrator?
   In general, the following individuals may serve as a Test Administrator:

   - Individuals employed by the district as teachers
   - District and school-level administrators
   - Other certified educational professionals

   Eligible Test Administrators and proctors must attend training and follow test procedures and protocol.

2. What is included in the braille/large print versions of the tests? What additional materials do I need?

   Large print and braille versions of the tests are used by students who have this presentation format identified in their IEPs or 504 plans for instruction and assessment. Charts in Section III of the Test Administrator Manual identify the materials packaged with each large print and braille test and additional needed materials. Additional materials needed must be documented in the student's IEP or 504 plan, except for the following items:

   - Test Administrator Manual
   - No. 2 pencils with erasers
   - Blank scratch paper
     - Blank scratch paper may include: abacus, slate, stylus, Braille Math Window or Brannan Cubarithm.
   - Highlighter
   - Graph paper
   - Calculator
     - Use of a grade-level appropriate calculator is available to all students during designated portions of the mathematics assessment.

- Students who have calculators identified as a needed accommodation in an IEP or 504 plan may use the calculator on all portions of the mathematics assessment.

3. What special issues exist regarding the use of optical or electronic magnification of the test?

   Electronic magnification systems enlarge print materials in black/white or color combinations. Magnification for viewing text and graphics can be increased up to 800% with option for changing font colors, background colors, using a line marker, etc. They come in a variety of models – desktop or handheld, near or distance, stand alone or connected to a computer. Electronic magnification systems provide students with access to all printed materials, and the size of the print can be customized for the task. Students who require magnification by using an electronic magnification system can use a regular paper-based test book.

   - If the electronic magnification system used by the student has the ability to capture images, these images must be deleted at the end of the test session.
   - Graphics enlarged on an electronic magnification system may be problematic for some students with low vision. When an image is magnified, the student may not be able to see the whole graphic at once. If the student has difficulty with graphics, a large print test should be ordered. Large print is the regular print book enlarged to 150% which is equivalent to 18 point font size.

4. What special issues should be considered regarding students with a visual impairment, including blindness who may take the online test?

   For any student taking the online test, it will be delivered using iTester.

   Screen readers

   A screen reader is a software application, separate from text-to-speech embedded in iTester, which conveys web content through audio. Screen readers are appropriate for students who are experienced with using the software, including those who are blind or have a visual impairment. Students who take the 2023 Science, Math, and ELA assessments online using a screen reader must be able to independently navigate the online testing environment. Professionals who work with students who are screen reader users are encouraged to work with students during instructional activities to ensure that they have independent computer-access skills. The skills used to navigate the 2023 Science, Math, and ELA assessments are the same needed to access a variety of internet resources, including the ability to navigate by regions and headings and the ability to use keyboard shortcuts and lists, such as link lists. See a more comprehensive list of prerequisite skills in Section IV of this document.

   As with all students taking a 2023 Science, Math, and ELA assessment, students with a visual impairment, including blindness are encouraged to use the practice tests which include screen reader, large print or access to Braille Ready Files (.brf) to download a braille practice test. Practice tests are currently posted in the following locations on the New Mexico Help and Support site:
   NM-MSSA: https://newmexico.onlinehelp.cognia.org/practice-tests-nm-mssa/
   NM-ASR: https://newmexico.onlinehelp.cognia.org/practice-tests-nm-asr/

   For more information about prerequisite skills, refer to the Technology Skills Checklist below.

Refreshable Braille Display

Students who use a screen reader can also access the English language arts (ELA), Spanish language arts (SLA), and Mathematics assessments using a refreshable braille display. Students who choose to take advantage of refreshable braille during the assessment should be comfortable and independent with using a refreshable braille display in instructional activities prior to using one in an assessment environment. As stated above, students and professionals are encouraged to use the Practice Tests in order to become familiar and comfortable with the Computer Based Assessments.

For more information about prerequisite skills, refer to the Technology Skills Checklist below.

Screen enlargement

The online 2023 Science, Math, and ELA assessments come with a built-in screen zoom/magnifier that can be used by all students at any time during the assessment period. The screen zoom enlarges the entire screen by increments of 150, 200, and 300%.

Note that some graphical information may become "pixelated" at very high magnification. Students and teachers should explore the Practice Items in order to determine the efficacy of using the kiosk-based screen zoom/magnification tool in a testing environment. Note that screen zoom/magnification is not available in the browser-based practice test, therefore students needing to practice with this tool should access the practice test using the kiosk.

For students who will use screen enlargement software with a Human Reader, refer to the *Accessibility Features and Accommodations Manual*, Appendix A: Test Administrator Protocol for the Human Reader Accommodation for English Language Arts (ELA) Assessments, and the Human Reader Accessibility Feature for Mathematics Assessments.

For more information about prerequisite skills, refer to the Technology Skills Checklist below.

Color Contrast/Reverse Contrast

The iTester system provides a built-in method for changing the color contrast settings and is available to all students. Currently, there are twelve color contrast options students can choose from and the option to reverse the color contrast.

Braillers and Braille Note-Takers

Students who are accustomed to using a brailler, slate and stylus or a braille note- taker to produce work during instructional activities will be able to do so with the online test. In these cases, the student will produce their answers and transcribe them into iTester or have them transcribed into the iTester.

5. Who can transcribe the tests?

Only an Eligible Test Administrator who is a certified Teacher of Students with Visual Impairment, including Blindness or someone working under the direct supervision of an Eligible Test Administrator who is a certified Teacher of Students with Visual Impairment, including Blindness may transcribe the student's responses into the test booklet, answer document or online form of the 2023 Science, Math, and ELA assessments.

Answers written on braille paper must be transcribed onto the standard-size paper form of the 2023 Science, Math, and ELA assessment. If responses are written on an electronic braille note-taker, they should be printed and transcribed into a standard-size paper test booklet or answer document. The file in the electronic braille note-taker must be deleted following successful transcription of the student's responses. **Note:** A student response can be embossed for their reviews, after which copies must be securely shredded after transcription.

III. Testing Materials

Science			
Materials/Language	**Large Print**	**Braille**	**Online**
Included with the Test English	• Large Print Test Booklet – English version • Large Print test administrator special instructions • Standard size test booklet - English • Standard size answer document for transcription • Grade 11 Periodic Table Large Print – English version	• Braille Test Booklet – English version • Braille test administrator special instructions • Standard size test booklet – English • Standard size answer document for transcription • Grade 11 Periodic Table Braille – English version	• Embedded grade 11 periodic table – English
Spanish	• Large Print Test Booklet – Spanish version • Large Print test administrator special instructions • Standard size test booklet - Spanish • Standard size answer document for transcription – Spanish • Grade 11 Periodic Table Large Print – Spanish version • Spanish Glossary of Science Terms for grades 5, 8, 11 Large Print version	• Braille Test Booklet – Spanish version • Braille test administrator special instructions • Standard size test booklet - Spanish • Standard size answer document for transcription – Spanish • Grade 11 Periodic Table Braille – Spanish version • Spanish Glossary of Science Terms for grades 5, 8, 11	• Embedded grade 11 periodic table – Spanish • Embedded Spanish glossary of science terms for grades 5, 8, and 11
Additional Materials Needed English and Spanish	• Test Administrator Manual • No. 2 pencils with erasers • Other materials included in student's IEP or 504 plan, such as Large Print writing devices, etc. • Scratch paper	• Test Administrator Manual • No. 2 pencils with erasers • Other materials included in student's IEP or 504 plan, such as Braille writing devices, etc. • Scratch paper	• Test Administrator Manual • No. 2 pencils with erasers • Other materials included in student's IEP or 504 plan, such as Large Print and Braille writing devices, etc. • Scratch paper • Student's preferred access technology

English Language Arts (ELA)			
Materials	**Large Print**	**Braille**	**Online**
Included with the Test	• Large Print Test Booklet • Standard Print Test Booklet or Answer Document for transcription	• Braille test booklet or answer document with embedded tactile graphics (certain forms) • Standard Print Test Booklet or Answer Document for transcription	• Tactile graphics – must order a Braille kit.
Additional Materials Needed	• Test Administrator Manual • No. 2 pencils with erasers • Blank scratch paper • Highlighter • Other materials included in the student's IEP or 504 plan	• Test Administrator Manual • No. 2 pencils with erasers • Other materials included in student's IEP or 504 plan, such as braille writing devices	• Test Administrator Manual • No. 2 pencils with erasers • Other materials included in student's IEP or 504 plan, such as braille writing devices • Student's preferred access technology

Mathematics			
Materials	**Large Print English or Large Print Spanish**	**Braille**	**Online**
Included with the Test	• Instructions for Large Print • Administration, including Test Administrator Scripts • Large Print Test Booklet • Standard Print Test Booklet or Answer Document for transcription	• Instructions for Braille • Administration, including Test Administrator Scripts • Braille test booklet or answer document with embedded tactile graphics • Standard Print Test Booklet or answer document for transcription	• Tactile graphics – must order a Braille kit.
Additional Materials Needed	• Test Administrator Manual • No. 2 pencils with erasers • Blank scratch paper • Highlighter • Regular classroom compass • Grade-level appropriate calculator – four-function or scientific • Other materials included in the student's IEP or 504 plan	• Test Administrator Manual • No. 2 pencils with erasers • Braille writing devices, such as a Perkins Brailler or an electronic braille note-taker • Grade-level appropriate tactile compass • Grade-level appropriate calculator – four-function or scientific • Braille materials that can be used as scratch paper • Cranmer Abacus • Braille Math Window • Brannan Cubarithm	• Test Administrator Manual • No. 2 pencils with erasers • Blank scratch paper or braille materials that can be used as scratch paper • Cranmer Abacus • Braille Math Window • Brannan Cubarithm • Grade-level appropriate tactile compass • Grade-level appropriate calculator – four-function or scientific • Other materials included in the student's IEP or 504 plan • Student's preferred access technology

IV. Technology Skills Checklist

Accessibility of testing materials for all students is an important part of the 2023 Science, Math, and ELA assessments. For a student with visual impairment, including blindness to take the online test, he or she will need to have a minimum level of skills with computer technology and the assistive technology he or she uses to access instructional materials. The following is a list of skills a student should be using regularly during instructional activities and be proficient with on the day of testing in order to independently access the 2023 Science, Math, and ELA assessments online. Students should, at a minimum, be able to complete these tasks independently and should be given multiple opportunities to practice using the Practice Tests and the Sample Items available on the New Mexico Help and Support Site at https://newmexico.onlinehelp.cognia.org/.

Screen Reader

- Use arrow keys to navigate
- Navigate by headings
- Access and use the Headings List
- Access and use the Links List
- Activate Links using keyboard commands
- Activate Buttons
- Adjust voice settings
- Select text using keyboard commands
- Copy text to clipboard
- Paste text from clipboard
- Access edit fields
- Use check boxes
- Use radio buttons
- Enter and exit forms mode
- Navigate, locate and read text on a webpage
- Navigate and understand a table

Refreshable Braille Display

- Complete all of the functions listed under Screen Reader
- Use corresponding commands to run a screen reader with a supported refreshable braille display

Screen Magnification

- Adjust color and contrast settings
- Adjust magnification settings

NEW MEXICO
Public Education Department

NM-ASR

NEW MEXICO ASSESSMENT
OF SCIENCE READINESS

NM-MSSA

NEW MEXICO MEASURES
OF STUDENT SUCCESS AND
ACHIEVEMENT

APPENDIX F
WRITING PROMPT ITEM-WRITING WORKSHOP PARTICIPANT PROFILES

Table F-1. Writing Prompt Item-Writing Workshop Participant Profiles

First Name	Last Name	Current Position	District Name
Grades 3–5			
Amaya	Alevesque Arau	4th grade teacher	Santa Fe Public Schools
MONICA	ALMANZA	Lead Teacher	Artesia Public Schools
Christell	Begay	Instruction/Social Emotional Learning Administrator	Central Consolidated School District #22
Lisa	Burnham	K-5 Instructional Coach	Aztec Municipal School District
Julius	Catapang	Elementary Teacher	Grants Cibola County Schools
Jacqueline	Griego	3rd and 4th Grade Teacher	Los Alamos Public Schools
Rodney King	Lisondra	Art education Teacher	Grants Cibola County Schools
Melissa	Nakai	5th Grade Teacher	Central Consolidated Schools #22
Alba	Pages-Fortuny	4th Class Teacher	SFPS
Cynthia	Perez	Instructional Coach	Gadsden ISD
Tara	Sterneker	Reading Specialist	Clovis Municipal Schools
Sadie	Wheeler	5th science/social studies teacher	Pecos Cyber Academy
Lucinda	Valencia	5th Grade Teacher	Cobre Consolidated Schools
Patricia	Pinnere	Director of Curriculum & Instruction	Albuquerque Collegiate Charter School
Carrie	Rowe	District Data Coach and Testing Coordinator	Alamogordo
Christina	Meyer Dzurec	Structured Literacy Student Focused Coach	Espanola
Sarah	Gleason	k-5 Instructional Coach	North Valley Academy
Cintia	Llinas Mahugo	4th grade bilingual teacher	Santa Fe Public Schools
Breanna	Leseberg		Socorro Consolidated Schools
Grades 6–8			
Ariel	Becenti	6th grade teacher	Central consolidated schools
Mark	Forman	Administrator	Socorro Consolidated Schools
Cristina	Garcia Sanchez	6th grade teacher	Santa Fe Public Schools
Andrea	Irangan	Reading Teacher	Cuba Independent Schools District
Janelle	Maestas	ELA Teacher	Los Alamos Public Schools
Darian	Muniz	Teacher	Public Academy for Performing Arts
Kathy	Owensby	6th grade teacher	Des Moines Municipal Schools
Aaron	Romero	Instructional Coach	Moriarty Edgewood School District
Kristina	Saiz	Instructional Coach/District Testing Coordinator	Santa Rosa Consolidated Schools
Jaynelle	Trujillo	7th & 8th grade ELA	Questa Independent School District
Sharon	West	Literacy Director	Santa Rosa Consolidated Schools
Melissa	Waldrop	6th Grade ELA Teacher	Pecos Cyber Academy
Laura	Kirkpatrick	Literacy Coach	Alamogordo Public Schools
Jennifer	Duran	Instructional Coach	Cobre Consolidated School District
Gwyn	Del Toro	7th grade ELA	Clovis Municipal Schools
Daniel	Busse	MS Instructional Coach	North Valley Academy
Rosalie	Capillo	ELA	Gallup Mckinley County School
Brandi	Lindsay	ELA Teacher	Truth or Consequences
Laryssa	Thomas	ELA	Truth or Consequences

APPENDIX G
COMMITTEE MEMBERSHIP

Table G-1. New Mexico Participants in the Cognia 2023 Bias Review Meetings by Content Area and Grade

Content Area	First Name	Last
ELA & Mathematics	Mary Faith	Silva
	Martha	Soliz
Science	Anthony	Chipre
	David	Martinez
PED Observer (ELA)	Eugene	Deaton
PED Observer (Mathematics)	Nicholas	Salazar
PED Observer (Science)	Shafiq	Chaudhary

Table G-2. New Mexico Participants in the Cognia 2023 Item Content Review Meetings by Content Area and Grade

Content Area	Grade	Name
ELA	3–4	Sonny Sapien
		Liana Croley
	5–6	Kelli Furney
		Tamara Lopez
	7–9	Christine Eisenmann
Ped Observer	All	Eugene Deaton
Mathematics	3–4	Juliana Burrola
		Antonio Gonzalez
	5–6	Lauren Ingham
		Cresta Hooser
	7–8	Cris Anthony Rabino
		Robert Gallegos
	9	Shyra Rabino
		Dustie Gonzalez
Ped Observer	All	Nicholas Salazar
Science	5	Nicole Hahn
		Deb Novak
	8	Tamara Wilburn
		Jennifer Neakrase
	11	Anaastacia Cadena
Ped Observer	All	Shafiq Chaudhary

APPENDIX H
SCORER QUALIFICATION RATES

Tables H-1 and H-2 summarize the qualification rates for the 2023 operational assessment for NM-MSSA Mathematics and NM-ASR Science respectively. Rates of success during qualification varied. Multiple factors determine the success of a scorer during qualification. These include familiarity with the assessment, grade level, and variation of item types. Please note that not all scorers who failed Qual 1 attempted Qual 2 and that the team scoring the Spanish-language queues qualified in English before scoring the Spanish responses and are included below.

Table H-1. Qualification Summary for NM-MSSA Mathematics

Grade 3	532101 Qual 1	532101 Qual 2	Scorers Qualified 532101	619276 Qual 1	619276 Qual 2	Scorers Qualified 619276	619288 Qual 1	619288 Qual 2	Scorers Qualified 619288	785068 Qual 1	785068 Qual 2	Scorers Qualified 785068
Total Passed	25	1	26	25	N/A	25	24	1	25	26	N/A	26
Total Failed	1	0	0	0	N/A	0	1	0	0	0	N/A	0
Grade 4	411965 Qual 1	411965 Qual 2	Scorers Qualified 411965	540658 Qual 1	540658 Qual 2	Scorers Qualified 540658	630481 Qual 1	630481 Qual 2	Scorers Qualified 630481	630485 Qual 1	630485 Qual 2	Scorers Qualified 630485
Total Passed	26	1	27	27	N/A	27	26	1	27	25	2	27
Total Failed	1	0	0	0	N/A	0	1	0	0	2	0	0
Grade 5	412281 Qual 1	412281 Qual 2	Scorers Qualified 412281	695227 Qual 1	695227 Qual 2	Scorers Qualified 695227	695233 Qual 1	695233 Qual 2	Scorers Qualified 695233	783563 Qual 1	783563 Qual 2	Scorers Qualified 785563
Total Passed	23	1	24	24	N/A	24	15	5	20	20	N/A	20
Total Failed	1	0	0	0	N/A	0	7	2	2	0	N/A	0
Grade 6	412531 Qual 1	412531 Qual 2	Scorers Qualified 412531	532604 Qual 1	532604 Qual 2	Scorers Qualified 532604	539624 Qual 1	539624 Qual 2	Scorers Qualified 539624	540196 Qual 1	540196 Qual 2	Scorers Qualified 540196
Total Passed	22	N/A	22	22	N/A	22	18	5	23	23	N/A	23
Total Failed	0	N/A	0	0	N/A	0	6	1	1	0	N/A	0
Grade 7	406233 Qual 1	406233 Qual 2	Scorers Qualified 406233	412656 Qual 1	412656 Qual 2	Scorers Qualified 412656	607222 Qual 1	607222 Qual 2	Scorers Qualified 607222	691595 Qual 1	691595 Qual 2	Scorers Qualified 691595
Total Passed	23	N/A	23	24	0	24	21	3	24	23	1	24
Total Failed	0	N/A	0	1	1	1	3	0	0	2	1	1
Grade 8	551249 Qual 1	551249 Qual 2	Scorers Qualified 551249	615320 Qual 1	615320 Qual 2	Scorers Qualified 615320	615411 Qual 1	615411 Qual 2	Scorers Qualified 615411	615422 Qual 1	615422 Qual 2	Scorers Qualified 615422
Total Passed	20	2	22	33	12	45	19	N/A	19	19	N/A	19
Total Failed	2	0	0	17	5	5	0	N/A	0	0	N/A	0

Table H-2. Qualification Summary for NM-ASR Science

Grade 5	661177 Qual 1	661177 Qual 2	Scorers Qualified 661177	666120 Qual 1	666120 Qual 2	Scorers Qualified 666120	697164 Qual 1	697164 Qual 2	Scorers Qualified 697164
Total Passed	19	4	23	29	3	32	15	5	20
Total Failed	4	0	0	3	0	0	5	0	0
Grade 8	663576 Qual 1	663576 Qual 2	Scorers Qualified 663576	697245 Qual 1	697245 Qual 2	Scorers Qualified 697245	717529 Qual 1	717529 Qual 2	Scorers Qualified 717529
Total Passed	33	5	38	21	2	23	27	5	32
Total Failed	9	4	4	6	4	4	7	2	2
Grade 11	666236 Qual 1	666236 Qual 2	Scorers Qualified 666236	710876 Qual 1	710876 Qual 2	Scorers Qualified 710876	735374 Qual 1	735374 Qual 2	Scorers Qualified 735374
Total Passed	15	2	17	18	1	19	18	1	19
Total Failed	3	1	1	1	0	0	3	2	2

Table H-3. Qualification Summary for NM-MSSA Writing

Grade 3	WRCC001 Qual 1	Scorers Qualified WRCC001	WRCC002 Qual 1	Scorers Qualified WRCC002	WRCC003 Qual 1	Scorers Qualified WRCC003
Total Passed	7	7	5	5	6	6
Total Failed	2	2	1	1	2	2
Grade 4	WRCC001 Qual 1	Scorers Qualified WRCC001	WRCC002 Qual 1	Scorers Qualified WRCC002	WRCC003 Qual 1	Scorers Qualified
Total Passed	14	14	10	10	17	17
Total Failed	0	0	0	0	0	0
Grade 5	WRCC001 Qual 1	Scorers Qualified WRCC001	WRCC002 Qual 1	Scorers Qualified WRCC002	WRCC003 Qual 1	Scorers Qualified
Total Passed	12	12	11	11	14	14
14	3	3	1	1	2	2
Grade 6	WRCC001 Qual 1	Scorers Qualified WRCC001	WRCC002 Qual 1	Scorers Qualified WRCC002	WRCC003 Qual 1	Scorers Qualified
Total Passed	17	17	8	8	9	9
Total Failed	2	2	1	1	1	1
Grade 7	WRCC001 Qual 1	Scorers Qualified WRCC001	WRCC002 Qual 1	Scorers Qualified WRCC002	WRCC003 Qual 1	Scorers Qualified
Total Passed	12	12	12	12	10	10
Total Failed	0	0	0	0	0	0
Grade 8	WRCC001 Qual 1	Scorers Qualified WRCC001	WRCC002 Qual 1	Scorers Qualified WRCC002	WRCC003 Qual 1	Scorers Qualified
Total Passed	8	8	13	13	16	16
Total Failed	5	5	1	1	0	0

APPENDIX I
CLASSICAL ITEM STATISTICS

Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.

Table I-1. Classical Item Statistics for the Operational Items on NM-MSSA ELA Grade 3*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
507621	MC	1	21,037	0.60	0.40
507623	MC	1	21,037	0.36	0.25
507628	MC	1	21,037	0.78	0.51
507633	MC	1	21,037	0.46	0.42
507637	MC	1	21,037	0.37	0.35
535773	MC	1	21,037	0.43	0.39
535779	MC	1	21,037	0.51	0.39
535783	MC	1	21,037	0.61	0.47
535785	MC	1	21,037	0.61	0.55
535787	MC	1	21,037	0.55	0.52
552233	MC	1	21,037	0.67	0.57
552235	MC	1	21,037	0.64	0.56
552251	MC	1	21,037	0.62	0.49
552255	MC	1	21,037	0.63	0.61
568986	MC	1	21,037	0.60	0.58
714498	MC	1	21,037	0.46	0.31
714500	MC	1	21,037	0.51	0.51
714507	MC	1	21,037	0.45	0.48
714509	MC	1	21,037	0.57	0.53
714511	MC	1	21,037	0.64	0.60
714518	MC	1	21,037	0.53	0.43
758004	MC	1	21,037	0.37	0.34
758006	MC	1	21,037	0.43	0.40
758012	MC	1	21,037	0.54	0.54
758018	MC	1	21,037	0.41	0.25
760410	MC	1	21,037	0.43	0.40
760412	MC	1	21,037	0.50	0.44
507631	EBSR	2	21,037	0.46	0.54
535797	EBSR	2	21,037	0.48	0.63
552223	EBSR	2	21,037	0.44	0.61
714494	EBSR	2	21,037	0.55	0.66
758008	EBSR	2	21,037	0.33	0.55
NM100834A	WP	3	6,802	0.19	0.55
NM100834B	WP	3	6,802	0.19	0.56
NM100978A	WP	3	6,780	0.13	0.58
NM100978B	WP	3	6,780	0.12	0.59
NM102996A	WP	3	7,455	0.11	0.55
NM102996B	WP	3	7,455	0.10	0.56

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-2. Classical Item Statistics for the Operational Items on NM-MSSA ELA Grade 4*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
475858	MC	1	21,494	0.70	0.54
475882	MC	1	21,494	0.30	0.15
475895	MC	1	21,494	0.46	0.43
475901	MC	1	21,494	0.39	0.37
486740	MC	1	21,494	0.44	0.28
507745	MC	1	21,494	0.35	0.37
507749	MC	1	21,494	0.68	0.43
507753	MC	1	21,494	0.69	0.47
507755	MC	1	21,494	0.64	0.51
507757	MC	1	21,494	0.74	0.60
543905	MC	1	21,494	0.57	0.53
543909	MC	1	21,494	0.55	0.36
543913	MC	1	21,494	0.53	0.38
543915	MC	1	21,494	0.60	0.33
543919	MC	1	21,494	0.67	0.53
552931	MC	1	21,494	0.55	0.53
552933	MC	1	21,494	0.67	0.40
552940	MC	1	21,494	0.68	0.60
552946	MC	1	21,494	0.62	0.56
552948	MC	1	21,494	0.58	0.56
691525	MC	1	21,494	0.74	0.55
691533	MC	1	21,494	0.57	0.52
691535	MC	1	21,494	0.35	0.27
691541	MC	1	21,494	0.44	0.41
691547	MC	1	21,494	0.54	0.42
787293	MC	1	21,494	0.46	0.34
691529	MS	1	21,494	0.31	0.42
475897	EBSR	2	21,494	0.53	0.67
507759	EBSR	2	21,494	0.44	0.53
543911	EBSR	2	21,494	0.49	0.47
552927	EBSR	2	21,494	0.45	0.67
691523	EBSR	2	21,494	0.28	0.35
NM100890A	WP	3	6,911	0.11	0.50
NM100890B	WP	3	6,911	0.11	0.50
NM100945A	WP	3	6,903	0.13	0.60
NM100945B	WP	3	6,903	0.13	0.61
NM103012A	WP	3	7,680	0.15	0.58
NM103012B	WP	3	7,680	0.14	0.59

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-3. Classical Item Statistics for the Operational Items on NM-MSSA ELA Grade 5*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
506900	MC	1	22,038	0.56	0.30
506916	MC	1	22,038	0.49	0.32
506919	MC	1	22,038	0.35	0.32
506936	MC	1	22,038	0.59	0.46
507226	MC	1	22,038	0.54	0.46
536393	MC	1	22,038	0.70	0.47
536395	MC	1	22,038	0.42	0.37
536397	MC	1	22,038	0.23	0.25
536405	MC	1	22,038	0.46	0.47
536411	MC	1	22,038	0.77	0.48
633769	MC	1	22,038	0.56	0.42
633778	MC	1	22,038	0.51	0.51
633783	MC	1	22,038	0.73	0.53
633789	MC	1	22,038	0.39	0.35
633791	MC	1	22,038	0.35	0.25
633795	MC	1	22,038	0.61	0.48
691843	MC	1	22,038	0.42	0.41
691849	MC	1	22,038	0.45	0.44
691851	MC	1	22,038	0.58	0.44
692491	MC	1	22,038	0.40	0.34
692493	MC	1	22,038	0.61	0.58
692497	MC	1	22,038	0.40	0.29
692499	MC	1	22,038	0.68	0.61
692505	MC	1	22,038	0.41	0.36
784114	MC	1	22,038	0.40	0.29
784116	MC	1	22,038	0.50	0.44
691847	MS	1	22,038	0.26	0.46
531590	EBSR	2	22,038	0.44	0.49
536391	EBSR	2	22,038	0.51	0.47
633799	EBSR	2	22,038	0.37	0.48
691837	EBSR	2	22,038	0.38	0.52
691938	EBSR	2	22,038	0.58	0.64
NM101321A	WP	3	7,049	0.17	0.64
NM101321B	WP	3	7,049	0.16	0.64
NM101356A	WP	3	7,907	0.29	0.67
NM101356B	WP	3	7,907	0.28	0.67
NM101363A	WP	3	7,082	0.21	0.68
NM101363B	WP	3	7,082	0.21	0.69

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-4. Classical Item Statistics for the Operational Items on NM-MSSA ELA Grade 6*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
505553	MC	1	22,208	0.76	0.34
505555	MC	1	22,208	0.86	0.43
505557	MC	1	22,208	0.67	0.43
505561	MC	1	22,208	0.80	0.41
505563	MC	1	22,208	0.37	0.31
553112	MC	1	22,208	0.44	0.31
553116	MC	1	22,208	0.63	0.53
553120	MC	1	22,208	0.47	0.34
553126	MC	1	22,208	0.40	0.25
553128	MC	1	22,208	0.37	0.27
553130	MC	1	22,208	0.47	0.39
602894	MC	1	22,208	0.63	0.40
602904	MC	1	22,208	0.69	0.54
602906	MC	1	22,208	0.84	0.50
602908	MC	1	22,208	0.59	0.41
602913	MC	1	22,208	0.79	0.29
758991	MC	1	22,208	0.42	0.42
758993	MC	1	22,208	0.42	0.41
758995	MC	1	22,208	0.61	0.37
759007	MC	1	22,208	0.54	0.44
759033	MC	1	22,208	0.41	0.30
759035	MC	1	22,208	0.34	0.41
759041	MC	1	22,208	0.41	0.51
759045	MC	1	22,208	0.55	0.44
759047	MC	1	22,208	0.48	0.39
759011	MS	1	22,208	0.25	0.43
759037	MS	1	22,208	0.19	0.39
505559	EBSR	2	22,208	0.50	0.52
553108	EBSR	2	22,208	0.36	0.44
602892	EBSR	2	22,208	0.39	0.50
759003	EBSR	2	22,208	0.40	0.53
759027	EBSR	2	22,208	0.48	0.56
NM101282A	WP	3	7,127	0.22	0.62
NM101282B	WP	3	7,127	0.22	0.62
NM101296A	WP	3	7,107	0.22	0.62
NM101296B	WP	3	7,107	0.22	0.62
NM101667A	WP	3	7,974	0.24	0.70
NM101667B	WP	3	7,974	0.24	0.70

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-5. Classical Item Statistics for the Operational Items on NM-MSSA ELA Grade 7*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
506279	MC	1	22,572	0.44	0.48
506282	MC	1	22,572	0.72	0.52
506285	MC	1	22,572	0.86	0.48
506287	MC	1	22,572	0.43	0.30
506302	MC	1	22,572	0.32	0.21
537120	MC	1	22,572	0.54	0.26
537122	MC	1	22,572	0.50	0.30
537124	MC	1	22,572	0.85	0.45
537134	MC	1	22,572	0.62	0.36
537138	MC	1	22,572	0.77	0.49
560013	MC	1	22,572	0.57	0.43
560015	MC	1	22,572	0.46	0.42
560017	MC	1	22,572	0.58	0.44
560023	MC	1	22,572	0.52	0.32
560027	MC	1	22,572	0.43	0.37
560029	MC	1	22,572	0.49	0.21
635295	MC	1	22,572	0.34	0.26
635299	MC	1	22,572	0.51	0.45
635303	MC	1	22,572	0.53	0.35
635307	MC	1	22,572	0.60	0.44
635313	MC	1	22,572	0.29	0.10
743356	MC	1	22,572	0.62	0.53
743360	MC	1	22,572	0.70	0.40
743368	MC	1	22,572	0.63	0.42
743372	MC	1	22,572	0.33	0.38
635309	MS	1	22,572	0.20	0.37
743366	MS	1	22,572	0.37	0.14
506297	EBSR	2	22,572	0.28	0.39
537130	EBSR	2	22,572	0.70	0.58
560009	EBSR	2	22,572	0.45	0.45
635291	EBSR	2	22,572	0.23	0.34
743350	EBSR	2	22,572	0.35	0.48
NM101540A	WP	3	7,298	0.22	0.65
NM101540B	WP	3	7,298	0.22	0.65
NM101709A	WP	3	7,973	0.27	0.65
NM101709B	WP	3	7,973	0.27	0.66
NM103314A	WP	3	7,301	0.30	0.64
NM103314B	WP	3	7,301	0.31	0.64

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-6. Classical Item Statistics for the Operational Items on NM-MSSA ELA Grade 8*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
402075	MC	1	23,836	0.42	0.24
402077	MC	1	23,836	0.34	0.28
402111	MC	1	23,836	0.80	0.34
402116	MC	1	23,836	0.74	0.45
402118	MC	1	23,836	0.60	0.46
546059	MC	1	23,836	0.58	0.48
546065	MC	1	23,836	0.49	0.32
546080	MC	1	23,836	0.56	0.48
546082	MC	1	23,836	0.59	0.55
546084	MC	1	23,836	0.79	0.44
560416	MC	1	23,836	0.61	0.40
560420	MC	1	23,836	0.58	0.39
560428	MC	1	23,836	0.59	0.35
560433	MC	1	23,836	0.62	0.31
560440	MC	1	23,836	0.49	0.34
560442	MC	1	23,836	0.76	0.50
641557	MC	1	23,836	0.68	0.44
641559	MC	1	23,836	0.50	0.52
641563	MC	1	23,836	0.65	0.60
641565	MC	1	23,836	0.56	0.52
641579	MC	1	23,836	0.29	0.26
NM110388	MC	1	23,836	0.54	0.24
NM110390	MC	1	23,836	0.63	0.48
NM110394	MC	1	23,836	0.62	0.59
NM110396	MC	1	23,836	0.32	0.28
NM110398	MC	1	23,836	0.33	0.37
NM110400	MC	1	23,836	0.31	0.33
402079	EBSR	2	23,836	0.42	0.47
546073	EBSR	2	23,836	0.32	0.48
560404	EBSR	2	23,836	0.47	0.45
641567	EBSR	2	23,836	0.43	0.52
NM110392	EBSR	2	23,836	0.29	0.35
NM101422A	WP	3	7,709	0.23	0.68
NM101422B	WP	3	7,709	0.23	0.69
NM101473A	WP	3	7,722	0.17	0.65
NM101473B	WP	3	7,722	0.18	0.66
NM101480A	WP	3	8,405	0.26	0.66
NM101480B	WP	3	8,405	0.28	0.67

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-7. Classical Item Statistics for the Operational Items on NM-MSSA Mathematics Grade 3*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
400604	MC	1	21,044	0.54	0.45
408129	MC	1	21,044	0.33	0.42
408165	MC	1	21,044	0.38	0.41
411119	MC	1	21,044	0.54	0.56
411764	MC	1	21,044	0.43	0.37
413036	MC	1	21,044	0.31	0.41
417040	MC	1	21,044	0.73	0.48
462672	MC	1	21,044	0.53	0.51
464204	MC	1	21,044	0.53	0.52
464268	MC	1	21,044	0.53	0.43
539890	MC	1	21,044	0.38	0.24
539903	MC	1	21,044	0.45	0.48
539940	MC	1	21,044	0.84	0.38
557246	MC	1	21,044	0.40	0.37
619075	MC	1	21,044	0.26	0.10
619084	MC	1	20,859	0.56	0.38
619098	MC	1	21,044	0.61	0.28
619106	MC	1	21,044	0.34	0.16
619137	MC	1	21,044	0.62	0.40
619174	MC	1	21,044	0.63	0.53
619192	MC	1	21,044	0.68	0.41
619211	MC	1	21,044	0.28	0.37
619217	MC	1	21,044	0.40	0.40
619227	MC	1	21,044	0.55	0.49
619235	MC	1	21,044	0.34	0.21
619242	MC	1	21,044	0.49	0.49
690977	MC	1	21,044	0.40	0.15
691034	MC	1	21,044	0.55	0.52
691055	MC	1	21,044	0.77	0.47
619276B	CR	1	21,044	0.13	0.50
619288B	CR	1	21,044	0.29	0.48
532101B	CR	2	21,044	0.18	0.62
619276A	CR	2	21,044	0.37	0.67
619288A	CR	2	21,044	0.14	0.55
785068B	CR	2	21,044	0.12	0.61
532101A	CR	4	21,044	0.17	0.67
785068A	CR	4	21,044	0.28	0.78
464225	MS	1	21,044	0.13	0.29
541508	MS	1	21,044	0.47	0.50
691038	MS	1	21,044	0.26	0.55
691053	MS	1	21,044	0.30	0.60

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-8. Classical Item Statistics for the Operational Items on NM-MSSA Mathematics Grade 4*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
407489	MC	1	21,541	0.63	0.46
407491	MC	1	21,541	0.48	0.44
462873	MC	1	21,541	0.27	0.17
462916	MC	1	21,541	0.29	0.43
465876	MC	1	21,541	0.30	0.18
540283	MC	1	21,541	0.31	0.40
540312	MC	1	21,541	0.25	0.39
540324	MC	1	21,541	0.55	0.40
540589	MC	1	21,541	0.37	0.42
541517	MC	1	21,541	0.42	0.36
560945	MC	1	21,541	0.50	0.39
629029	MC	1	21,541	0.56	0.35
629038	MC	1	21,541	0.29	0.16
629048	MC	1	21,541	0.26	0.14
629066	MC	1	21,541	0.33	0.47
629094	MC	1	21,541	0.43	0.22
629111	MC	1	21,541	0.48	0.48
629123	MC	1	21,541	0.27	0.12
691166	MC	1	21,541	0.42	0.32
691215	MC	1	21,541	0.40	0.27
740654	MC	1	21,541	0.34	0.33
740658	MC	1	21,541	0.52	0.41
740662	MC	1	21,541	0.38	0.53
740664	MC	1	21,541	0.52	0.38
740668	MC	1	21,541	0.58	0.39
740694	MC	1	21,541	0.45	0.37
124856A	MC	1	21,541	0.42	0.43
126020A	MC	1	21,541	0.21	0.55
127388A	MC	1	21,541	0.30	0.41
127466A	MC	1	21,541	0.25	0.25
540658B	CR	1	21,541	0.12	0.48
630485B	CR	1	21,541	0.26	0.60
411965B	CR	2	21,541	0.09	0.58
540658A	CR	2	21,541	0.33	0.66
630481B	CR	2	21,541	0.30	0.64
630485A	CR	2	21,541	0.34	0.66
411965A	CR	4	21,541	0.18	0.74
630481A	CR	4	21,541	0.26	0.75
540599	MS	1	21,541	0.18	0.59
629036	MS	1	21,541	0.57	0.55
629073	MS	1	21,541	0.13	0.55

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-9. Classical Item Statistics for the Operational Items on NM-MSSA Mathematics Grade 5*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
400300	MC	1	22,041	0.35	0.04
400488	MC	1	22,041	0.39	0.37
400650	MC	1	22,041	0.41	0.23
400711	MC	1	22,041	0.70	0.46
405943	MC	1	22,041	0.43	0.48
408463	MC	1	22,041	0.23	0.13
411149	MC	1	22,041	0.42	0.53
411304	MC	1	22,041	0.67	0.54
413871	MC	1	22,041	0.34	0.31
464071	MC	1	22,041	0.73	0.43
464399	MC	1	22,041	0.43	0.54
465792	MC	1	22,041	0.42	0.32
532490	MC	1	22,041	0.44	0.48
539162	MC	1	22,041	0.57	0.45
539164	MC	1	22,041	0.32	0.51
539188	MC	1	22,041	0.51	0.12
540635	MC	1	22,041	0.39	0.43
540637	MC	1	22,041	0.72	0.43
540710	MC	1	22,041	0.43	0.58
558693	MC	1	22,041	0.54	0.46
607336	MC	1	22,041	0.49	0.49
607394	MC	1	22,041	0.53	0.53
607495	MC	1	22,041	0.29	0.35
607514	MC	1	22,041	0.54	0.45
607538	MC	1	22,041	0.49	0.38
692838	MC	1	22,041	0.39	0.35
692846	MC	1	22,041	0.39	0.27
693074	MC	1	22,041	0.48	0.30
695196	MC	1	22,041	0.69	0.42
695213	MC	1	22,041	0.42	0.36
124077A	MC	1	22,041	0.61	0.37
125951A	MC	1	22,041	0.58	0.44
695227B	CR	1	22,041	0.08	0.46
783563B	CR	1	22,041	0.07	0.37
412281B	CR	2	22,041	0.31	0.74
695227A	CR	2	22,041	0.17	0.65
695233B	CR	2	22,041	0.14	0.65
783563A	CR	2	22,041	0.40	0.67
412281A	CR	4	22,041	0.39	0.77
695233A	CR	4	22,041	0.19	0.76
781233	MS	1	22,041	0.40	0.55

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-10. Classical Item Statistics for the Operational Items on NM-MSSA Mathematics Grade 6*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
406099	MC	1	22,227	0.37	0.26
411834	MC	1	22,227	0.57	0.37
412144	MC	1	22,227	0.41	0.38
412393	MC	1	22,227	0.21	0.31
412439	MC	1	22,227	0.27	0.17
412462	MC	1	22,227	0.50	0.39
464785	MC	1	22,227	0.23	0.57
539618	MC	1	22,227	0.27	0.36
539622	MC	1	22,227	0.35	0.26
539643	MC	1	22,227	0.26	0.33
539649	MC	1	22,227	0.58	0.43
539809	MC	1	22,227	0.59	0.41
540132	MC	1	22,227	0.44	0.36
540725	MC	1	22,227	0.32	0.31
558413	MC	1	22,227	0.28	0.23
607665	MC	1	22,227	0.55	0.35
607688	MC	1	22,227	0.38	0.30
607721	MC	1	22,227	0.62	0.42
607725	MC	1	22,227	0.59	0.46
607751	MC	1	22,227	0.44	0.52
607773	MC	1	22,227	0.23	0.30
607775	MC	1	22,227	0.75	0.36
607782	MC	1	22,227	0.31	0.15
695581	MC	1	22,227	0.62	0.32
695587	MC	1	22,227	0.37	0.23
695595	MC	1	22,227	0.34	0.28
695600	MC	1	22,227	0.26	0.33
695606	MC	1	22,227	0.34	0.30
700976	MC	1	22,227	0.38	0.19
706548	MC	1	22,227	0.66	0.47
124799A	MC	1	22,227	0.44	0.56
127179A	MC	1	22,227	0.50	0.39
127738A	MC	1	22,227	0.47	0.42
532604B	CR	1	22,227	0.03	0.37
540196B	CR	1	22,227	0.12	0.48
412531B	CR	2	22,227	0.08	0.63
532604A	CR	2	22,227	0.10	0.41
539624B	CR	2	22,227	0.18	0.65
540196A	CR	2	22,227	0.30	0.62
412531A	CR	4	22,227	0.14	0.75
539624A	CR	4	22,227	0.18	0.75
607785	MS	1	22,227	0.17	0.21

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-11. Classical Item Statistics for the Operational Items on NM-MSSA Mathematics Grade 7*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
412224	MC	1	22,543	0.40	0.42
412251	MC	1	22,543	0.63	0.42
412395	MC	1	22,543	0.27	0.41
412425	MC	1	22,543	0.50	0.48
412486	MC	1	22,543	0.40	0.35
412493	MC	1	22,543	0.37	0.52
467208	MC	1	22,543	0.43	0.54
467772	MC	1	22,543	0.33	0.09
467826	MC	1	22,543	0.23	0.21
539394	MC	1	22,543	0.52	0.27
539407	MC	1	22,543	0.33	0.43
539444	MC	1	22,543	0.33	0.34
539450	MC	1	22,543	0.45	0.44
540128	MC	1	22,543	0.61	0.44
557976	MC	1	22,543	0.30	0.14
607105	MC	1	22,543	0.32	0.29
607107	MC	1	22,543	0.39	0.27
607119	MC	1	22,543	0.57	0.49
607131	MC	1	22,543	0.35	0.36
607135	MC	1	22,543	0.56	0.12
607155	MC	1	22,543	0.40	0.32
607157	MC	1	22,543	0.18	0.16
607161	MC	1	22,543	0.45	0.57
607163	MC	1	22,543	0.51	0.39
607203	MC	1	22,543	0.36	0.23
607205	MC	1	22,543	0.54	0.35
607213	MC	1	22,543	0.56	0.44
607215	MC	1	22,543	0.35	0.43
687771	MC	1	22,543	0.18	0.55
687773	MC	1	22,543	0.15	0.49
687793	MC	1	22,543	0.40	0.19
687860	MC	1	22,543	0.18	0.11
690152	MC	1	22,543	0.68	0.42
691316	MC	1	22,543	0.35	0.24
406243B	CR	1	22,543	0.17	0.63
412656B	CR	1	22,543	0.31	0.53
406243A	CR	2	22,543	0.28	0.70
412656A	CR	2	22,543	0.21	0.69
607222B	CR	2	22,543	0.10	0.63
691595B	CR	2	22,543	0.10	0.64
607222A	CR	4	22,543	0.18	0.73
691595A	CR	4	22,543	0.17	0.72

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-12. Classical Item Statistics for the Operational Items on NM-MSSA Mathematics Grade 8*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
400766	MC	1	23,843	0.30	0.16
400771	MC	1	23,843	0.44	0.29
408524	MC	1	23,843	0.31	0.34
413137	MC	1	23,843	0.32	0.08
413229	MC	1	23,843	0.72	0.40
414834	MC	1	23,843	0.23	0.24
414849	MC	1	23,843	0.34	0.17
414957	MC	1	23,843	0.41	0.20
465465	MC	1	23,843	0.31	0.30
468359	MC	1	23,843	0.36	0.44
483010	MC	1	23,843	0.53	0.29
483452	MC	1	23,843	0.35	0.19
540844	MC	1	23,843	0.30	0.16
540892	MC	1	23,843	0.69	0.40
540918	MC	1	23,843	0.49	0.32
540955	MC	1	23,843	0.53	0.41
541134	MC	1	23,843	0.19	0.34
614704	MC	1	23,843	0.42	0.12
614780	MC	1	23,843	0.29	0.38
614837	MC	1	23,843	0.26	0.39
614943	MC	1	23,843	0.21	0.26
615097	MC	1	23,843	0.34	0.28
615300	MC	1	23,843	0.43	0.17
631977	MC	1	23,843	0.28	0.41
696121	MC	1	23,843	0.53	0.28
696123	MC	1	23,843	0.47	0.42
696125	MC	1	23,843	0.56	0.35
696131	MC	1	23,843	0.82	0.34
696139	MC	1	23,843	0.25	0.27
696141	MC	1	23,843	0.28	0.36
697997	MC	1	23,843	0.41	0.39
126335A	MC	1	23,843	0.20	0.28
126398A	MC	1	23,843	0.15	0.12
127164A	MC	1	23,843	0.19	0.40
551249B	CR	1	23,843	0.06	0.50
615422B	CR	1	23,843	0.23	0.50
551249A	CR	2	23,843	0.08	0.58
615320B	CR	2	23,843	0.19	0.64
615411B	CR	2	23,843	0.08	0.65
615422A	CR	2	23,843	0.32	0.62
615320A	CR	4	23,843	0.06	0.65
615411A	CR	4	23,843	0.10	0.64

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-13. Classical Item Statistics for the Operational Items on NM-ASR Science Grade 5*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
632724	MC	1	22,001	0.40	0.15
637807	MC	1	22,001	0.71	0.44
637951	MC	1	22,001	0.47	0.41
638656	MC	1	22,001	0.24	0.18
638658	MC	1	22,001	0.40	0.40
706138	MC	1	22,001	0.37	0.28
706149	MC	1	22,001	0.46	0.40
706792	MC	1	22,001	0.42	0.30
744445	MC	1	22,001	0.43	0.35
784744	MC	1	2,640	0.59	0.54
848745	MC	1	2,640	0.59	0.54
661177	CR	4	22,001	0.29	0.68
666120	CR	4	22,001	0.16	0.50
697164	CR	4	22,001	0.27	0.66
634127	MCMS	2	7,373	0.35	0.47
636189	MCMS	2	22,001	0.20	0.45
743165	MCMS	2	7,373	0.38	0.57
744451	MS	1	22,001	0.31	0.36
633906	MSMC	2	14,628	0.23	0.44
756457	MSMS	2	14,628	0.19	0.51
626442	PMC	2	14,628	0.67	0.52
631577	PMC	2	7,373	0.53	0.60
632426	PMC	2	7,373	0.33	0.30
632454	PMC	2	14,628	0.35	0.44
635879	PMC	2	14,628	0.53	0.41
635886	PMC	2	7,373	0.20	0.22
636211	PMC	2	22,001	0.49	0.48
638354	PMC	2	22,001	0.47	0.61
638558	PMC	2	22,001	0.61	0.64
638639	PMC	2	22,001	0.51	0.34
639474	PMC	2	7,373	0.49	0.63
706119	PMC	2	22,001	0.53	0.53
706135	PMC	2	22,001	0.41	0.52
706765	PMC	2	22,001	0.53	0.59
706847	PMC	2	22,001	0.37	0.54
737916	PMC	2	7,373	0.60	0.49
762758	PMC	2	22,001	0.37	0.46
784722	PMC	2	2,640	0.40	0.51
784830	PMC	2	2,640	0.62	0.56
784847	PMC	2	2,640	0.47	0.49
785041	PMC	2	2,640	0.45	0.48
848747	PMC	2	2,640	0.65	0.57
632837	TEI	1	19,361	0.72	0.45
706801	TEI	1	19,361	0.42	0.59
626478	TEI	2	11,988	0.39	0.44
638526	TEI	2	19,361	0.53	0.55
639510	TEI	2	11,988	0.37	0.58
697027	TEI	2	11,988	0.56	0.52
697044	TEI	2	7,373	0.56	0.39
744455	TEI	2	19,361	0.71	0.53

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-14. Classical Item Statistics for the Operational Items on NM-ASR Science Grade 8*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
636837	MC	1	23,808	0.39	0.45
636843	MC	1	23,808	0.41	0.33
641873	MC	1	23,808	0.35	0.33
641894	MC	1	23,808	0.27	0.27
709292	MC	1	23,808	0.32	0.27
709609	MC	1	23,808	0.29	0.36
758938	MC	1	23,808	0.48	0.31
847907	MC	1	3,048	0.43	0.31
847909	MC	1	3,048	0.58	0.40
847911	MC	1	3,048	0.39	0.42
847927	MC	1	3,048	0.52	0.41
663576	CR	4	23,808	0.14	0.68
697245	CR	4	23,808	0.16	0.62
717529	CR	4	23,808	0.22	0.63
643622	MCMS	2	8,609	0.29	0.49
709306	MS	1	23,808	0.16	0.26
758919	MSMC	2	23,808	0.31	0.55
642855	MSMS	2	8,609	0.54	0.63
631360	PMC	2	8,609	0.53	0.40
636830	PMC	2	23,808	0.55	0.47
636852	PMC	2	23,808	0.40	0.44
637562	PMC	2	15,199	0.50	0.54
637622	PMC	2	15,199	0.53	0.53
637635	PMC	2	8,609	0.54	0.33
640163	PMC	2	15,199	0.39	0.33
640740	PMC	2	15,199	0.51	0.47
641845	PMC	2	23,808	0.24	0.36
642091	PMC	2	23,808	0.30	0.32
696483	PMC	2	15,199	0.38	0.37
707172	PMC	2	15,199	0.39	0.37
709309	PMC	2	23,808	0.42	0.52
709622	PMC	2	23,808	0.55	0.57
712986	PMC	2	15,199	0.45	0.44
713386	PMC	2	23,808	0.43	0.51
713388	PMC	2	23,808	0.42	0.32
716040	PMC	2	15,199	0.55	0.49
731140	PMC	2	8,609	0.49	0.51
786048	PMC	2	3,048	0.25	0.15
847905	PMC	2	3,048	0.30	0.45
847934	PMC	2	3,048	0.45	0.48
709617	TEI	1	20,760	0.38	0.41
713686	TEI	1	20,760	0.47	0.42
713695	TEI	1	20,760	0.24	0.45
758880	TEI	1	20,760	0.49	0.45
641866	TEI	2	8,609	0.74	0.47
709294	TEI	2	20,760	0.18	0.30
709604	TEI	2	20,760	0.20	0.24
741297	TEI	2	8,609	0.43	0.58
752254	TEI	2	8,609	0.45	0.50
763243	TEI	2	20,760	0.45	0.42

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

Table I-15. Classical Item Statistics for the Operational Items on NM-ASR Science Grade 11*

PsyItemNumber	Item Type	Max Points	Number of Students	P-Value	Item-Total Correlation
633246	MC	1	21,426	0.61	0.45
633315	MC	1	21,426	0.46	0.42
637994	MC	1	21,426	0.38	0.24
639319	MC	1	21,426	0.51	0.48
642634	MC	1	21,426	0.37	0.27
643598	MC	1	21,426	0.58	0.40
705787	MC	1	21,426	0.25	0.25
705815	MC	1	21,426	0.42	0.23
753774	MC	1	21,426	0.52	0.43
753780	MC	1	21,426	0.23	0.22
782469	MC	1	2,913	0.34	0.27
782471	MC	1	2,913	0.42	0.46
666236	CR	4	21,426	0.27	0.67
710876	CR	4	21,426	0.28	0.52
735374	CR	4	21,426	0.21	0.68
636258	MSMC	2	8,401	0.49	0.61
640447	MSMC	2	13,025	0.29	0.36
642454	MSMC	2	21,426	0.32	0.62
696546	MSMC	2	13,025	0.19	0.27
NM100318	MSMC	2	8,401	0.29	0.44
626027	PMC	2	13,025	0.46	0.51
628033	PMC	2	13,025	0.54	0.58
632730	PMC	2	21,426	0.36	0.47
633116	PMC	2	13,025	0.39	0.49
633266	PMC	2	21,426	0.36	0.41
637608	PMC	2	8,401	0.34	0.25
637610	PMC	2	13,025	0.26	0.37
639346	PMC	2	21,426	0.31	0.48
640641	PMC	2	13,025	0.32	0.45
642245	PMC	2	8,401	0.55	0.43
642377	PMC	2	8,401	0.33	0.35
705738	PMC	2	21,426	0.21	0.18
705807	PMC	2	21,426	0.34	0.36
706583	PMC	2	21,426	0.31	0.37
709638	PMC	2	8,401	0.40	0.48
733448	PMC	2	8,401	0.42	0.33
762012	PMC	2	21,426	0.47	0.50
762916	PMC	2	21,426	0.42	0.39
782467	PMC	2	2,913	0.48	0.54
782473	PMC	2	2,913	0.42	0.54
782475	PMC	2	2,913	0.34	0.50
847612	PMC	2	2,913	0.54	0.37
847619	PMC	2	2,913	0.34	0.40
847621	PMC	2	2,913	0.37	0.33
706534	TEI	1	18,513	0.20	0.31
706670	TEI	1	18,513	0.39	0.48
627081	TEI	2	10,112	0.32	0.47
639344	TEI	2	18,513	0.53	0.43
642533	TEI	2	10,112	0.45	0.52
706468	TEI	2	18,513	0.47	0.54
710663	TEI	2	8,401	0.45	0.59
738031	TEI	2	8,401	0.24	0.42
748109	TEI	2	10,112	0.45	0.46
755860	TEI	2	8,401	0.49	0.39
762922	TEI	2	18,513	0.31	0.28

**Calculations based on those students attempting 5 or more items on the English forms of the given NM-MSSA & ASR assessments. For 1-point items, the item-total correlation is the point-biserial. For 2 or more-point items, the item-total correlation is the point-polyserial.*

APPENDIX J
ITEM RESPONSE THEORY PARAMETERS

Table J-1. IRT Parameters for Operational Items on the NM-MSSA Grade 3 ELA Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*
507621	0.59662	-0.89722	0.12975				
507623	0.73753	1.27242	0.17872				
507628	1.66847	-0.82964	0.21426				
507633	1.24690	0.01544	0.16517				
507637	0.60388	1.34538	0.16665				
535773	1.41886	0.72505	0.29601				
535779	0.71815	0.19134	0.21394				
535783	0.82051	-0.44170	0.19459				
535785	1.43373	-0.18431	0.36269				
535787	1.23853	0.36233	0.24507				
552233	1.41813	-0.32446	0.22621				
552235	0.84310	-0.71083	0.08758				
552251	0.88685	-0.80997	0.07513				
552255	1.63805	-0.19938	0.18059				
568986	0.98347	-0.44496	0.09311				
714498	0.87562	1.43668	0.23305				
714500	1.30710	0.69176	0.18440				
714507	1.08715	1.04048	0.20110				
714509	1.93190	0.67388	0.26451				
714511	1.64514	-0.09584	0.13009				
714518	0.82548	0.82943	0.15961				
758004	1.31310	1.30327	0.25239				
758006	0.88711	0.73811	0.18185				
758012	1.71941	0.10455	0.20268				
758018	0.65231	1.43835	0.27774				
760410	1.26621	1.00766	0.26137				
760412	0.90409	0.80876	0.25801				
507631	0.67472	0.13392	0.00000	0.42202	-0.42202	0.00000	
535797	0.88420	-0.05546	0.00000	0.52528	-0.52528	0.00000	
552223	0.99411	-0.10920	0.00000	0.21364	-0.21364	0.00000	
714494	0.92106	0.05782	0.00000	0.20887	-0.20887	0.00000	
758008	0.57373	1.12262	0.00000	0.44527	-0.44527	0.00000	
NM100834A	0.79552	2.02617	0.00000	2.11285	-0.11627	-1.99657	0.00000
NM100834B	0.80983	1.91209	0.00000	2.00221	-0.21551	-1.78671	0.00000
NM100978A	1.02204	2.07487	0.00000	1.40828	-0.07022	-1.33806	0.00000
NM100978B	1.02606	2.11111	0.00000	1.44685	-0.16851	-1.27834	0.00000
NM102996A	0.95757	1.93417	0.00000	1.03916	0.01425	-1.05341	0.00000
NM102996B	0.98201	2.00871	0.00000	1.12970	0.00960	-1.13931	0.00000

Table J-2. IRT Parameters for Operational Items on the NM-MSSA Grade 4 ELA Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*
475858	1.12416	-0.28249	0.21306				
475882	0.62797	1.51250	0.23678				
475895	1.18575	0.50358	0.22008				
475901	1.24171	0.98465	0.20492				
486740	1.28335	0.92055	0.31354				
507745	0.91785	0.73178	0.14130				
507749	0.82058	-0.16948	0.28941				
507753	0.87331	-1.01212	0.23717				
507755	1.02566	-0.56391	0.16685				
507757	1.64203	-0.98566	0.14969				
543905	1.18483	0.13849	0.16182				
543909	0.47018	0.33815	0.21662				
543913	0.84959	0.57456	0.23694				
543915	0.57838	0.14166	0.17779				
543919	1.40433	-0.57500	0.20840				
552931	1.23378	0.09629	0.18836				
552933	0.58004	-0.88966	0.00000				
552940	1.47121	-0.52494	0.15187				
552946	1.32044	-0.04411	0.20669				
552948	0.95154	-0.14121	0.07554				
691525	1.42021	-0.60525	0.28291				
691529	0.52364	0.98049	0.00000				
691533	1.38577	0.07870	0.25613				
691535	1.22181	1.21944	0.24454				
691541	1.00021	0.74036	0.22546				
691547	0.63049	0.36799	0.19770				
787293	0.53474	0.57995	0.16318				
475897	0.86816	-0.07860	0.00000	0.54268	-0.54268	0.00000	
507759	0.57871	-0.00378	0.00000	0.44622	-0.44622	0.00000	
543911	0.45302	-0.42587	0.00000	0.43861	-0.43861	0.00000	
552927	1.02753	0.02368	0.00000	0.20947	-0.20947	0.00000	
691523	0.28001	2.16721	0.00000	0.75754	-0.75754	0.00000	
NM100890A	0.83342	2.34862	0.00000	1.54149	0.16753	-1.70902	0.00000
NM100890B	0.84080	2.34202	0.00000	1.54153	0.12634	-1.66787	0.00000
NM100945A	1.34686	1.64780	0.00000	1.04862	-0.02481	-1.02381	0.00000
NM100945B	1.38202	1.58987	0.00000	1.00079	-0.00251	-0.99828	0.00000
NM103012A	1.03631	1.66601	0.00000	1.17051	0.03829	-1.20880	0.00000
NM103012B	1.06002	1.75469	0.00000	1.26986	0.04342	-1.31327	0.00000

Table J-3. IRT Parameters for Operational Items on the NM-MSSA Grade 5 ELA Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*
506900	0.47433	0.26598	0.20858				
506916	0.87345	-0.24385	0.26340				
506919	0.58349	0.70116	0.13087				
506936	0.66453	-0.33102	0.20284				
507226	1.07853	-0.02329	0.23305				
536393	0.88678	-1.03216	0.23012				
536395	0.76348	0.63977	0.24257				
536397	0.60988	1.66400	0.10245				
536405	0.88372	0.19567	0.22762				
536411	1.06741	-0.75188	0.24245				
633769	0.44227	-0.13209	0.16843				
633778	1.02914	0.18566	0.20086				
633783	0.98319	-0.82140	0.14677				
633789	0.42072	0.20308	0.00000				
633791	0.45290	1.60281	0.12640				
633795	0.87255	-0.27687	0.15873				
691843	1.59527	0.94708	0.26254				
691847	0.84889	0.85161	0.00000				
691849	1.08404	0.53046	0.20116				
691851	1.21572	-0.00027	0.25611				
692491	1.62824	0.76556	0.23316				
692493	1.20867	-0.40318	0.10522				
692497	1.03152	0.86720	0.24770				
692499	1.39731	-0.60130	0.11281				
692505	0.61256	0.55515	0.08356				
784114	0.92925	1.33177	0.27519				
784116	0.90500	0.30199	0.18551				
531590	0.62469	0.18086	0.00000	0.46560	-0.46560	0.00000	
536391	0.50983	-0.38859	0.00000	0.12324	-0.12324	0.00000	
633799	0.70292	0.24328	0.00000	0.31771	-0.31771	0.00000	
691837	0.62351	0.50946	0.00000	0.67207	-0.67207	0.00000	
691938	1.06701	-0.63269	0.00000	0.19199	-0.19199	0.00000	
NM101321A	1.14698	1.35604	0.00000	0.66106	0.13331	-0.79436	0.00000
NM101321B	1.15299	1.40862	0.00000	0.71687	0.13501	-0.85187	0.00000
NM101356A	1.23886	0.65262	0.00000	1.24164	-0.12702	-1.11462	0.00000
NM101356B	1.25424	0.67416	0.00000	1.26290	-0.10851	-1.15439	0.00000
NM101363A	1.69460	0.87729	0.00000	1.03072	-0.11486	-0.91586	0.00000
NM101363B	1.73743	0.92028	0.00000	1.07562	-0.10784	-0.96778	0.00000

Table J-4. IRT Parameters for Operational Items on the NM-MSSA Grade 6 ELA Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*
505553	0.80069	-0.57118	0.29343				
505555	1.21410	-1.25717	0.14752				
505557	0.64474	-0.92450	0.09776				
505561	0.85377	-1.31549	0.04844				
505563	0.90232	0.95626	0.23038				
553112	0.45827	1.39147	0.21136				
553116	1.22937	-0.45519	0.16151				
553120	0.33573	-0.07131	0.00000				
553126	1.16113	1.68768	0.28426				
553128	0.45520	0.67331	0.12076				
553130	0.52134	0.38755	0.08609				
602894	0.48036	-0.50471	0.18248				
602904	1.44206	-0.46466	0.21754				
602906	1.45723	-0.91787	0.22978				
602908	1.03537	0.46262	0.33437				
602913	0.40234	-1.25677	0.24155				
758991	0.80155	0.69244	0.12177				
758993	0.58830	0.70288	0.12748				
758995	0.94825	0.40343	0.34082				
759007	0.84626	0.67542	0.21179				
759011	0.75696	1.21179	0.00000				
759033	0.95833	0.97393	0.21591				
759035	1.27208	0.78167	0.11036				
759037	0.60935	1.94724	0.00000				
759041	1.63917	0.31093	0.15391				
759045	1.10645	-0.00438	0.23567				
759047	0.56734	0.18019	0.13037				
505559	0.76162	0.09858	0.00000	0.57699	-0.57699	0.00000	
553108	0.58795	0.11939	0.00000	0.09877	-0.09877	0.00000	
602892	0.47391	0.41131	0.00000	0.69716	-0.69716	0.00000	
759003	0.61469	0.53541	0.00000	0.62184	-0.62184	0.00000	
759027	0.69244	-0.04609	0.00000	0.25023	-0.25023	0.00000	
NM101282A	1.48652	1.17694	0.00000	1.11738	-0.09188	-1.02549	0.00000
NM101282B	1.49402	1.19526	0.00000	1.13619	-0.12367	-1.01251	0.00000
NM101296A	1.20632	1.49700	0.00000	1.56478	-0.18007	-1.38471	0.00000
NM101296B	1.20761	1.49760	0.00000	1.56539	-0.18270	-1.38269	0.00000
NM101667A	2.04088	0.79974	0.00000	0.86302	-0.06189	-0.80113	0.00000
NM101667B	2.06434	0.82150	0.00000	0.88567	-0.05137	-0.83430	0.00000

Table J-5. IRT Parameters for Operational Items on the NM-MSSA Grade 7 ELA Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*
506279	1.29897	0.51877	0.19382				
506282	0.93887	-0.87456	0.20531				
506285	1.33434	-1.41766	0.19775				
506287	0.72342	0.79350	0.18455				
506302	0.70224	1.67742	0.26593				
537120	0.30675	0.82106	0.25063				
537122	0.75057	0.64777	0.31496				
537124	1.10057	-1.23281	0.26754				
537134	0.47934	-0.30239	0.16437				
537138	1.07306	-0.77403	0.24967				
560013	0.75633	-0.24308	0.18500				
560015	0.58664	0.37248	0.12796				
560017	0.91955	0.24008	0.16944				
560023	0.47470	0.89175	0.22888				
560027	0.76670	1.06651	0.22937				
560029	1.41870	1.56512	0.31621				
635295	0.70035	1.39277	0.19029				
635299	0.63033	0.42786	0.13619				
635303	0.59690	0.22129	0.28428				
635307	0.52642	-0.40116	0.14943				
635309	0.73054	1.48901	0.00000				
635313	0.87372	1.60193	0.26356				
743356	1.60080	0.25107	0.27941				
743360	1.20676	-0.01650	0.33741				
743366	0.22615	1.92058	0.00000				
743368	0.97584	0.06077	0.29714				
743372	1.93828	1.04578	0.18024				
506297	0.39712	1.56584	0.00000	1.47581	-1.47581	0.00000	
537130	0.81771	-0.89258	0.00000	0.24935	-0.24935	0.00000	
560009	0.70802	-0.19666	0.00000	0.21305	-0.21305	0.00000	
635291	0.56153	1.08608	0.00000	0.57852	-0.57852	0.00000	
743350	0.41865	1.45396	0.00000	0.66198	-0.66198	0.00000	
NM101540A	1.18699	1.12493	0.00000	1.06561	-0.09538	-0.97023	0.00000
NM101540B	1.20499	1.10851	0.00000	1.04976	-0.07332	-0.97644	0.00000
NM101709A	1.11152	1.13949	0.00000	2.01475	-0.33735	-1.67740	0.00000
NM101709B	1.12278	1.09047	0.00000	1.96374	-0.35409	-1.60965	0.00000
NM103314A	1.00606	0.92643	0.00000	1.77713	-0.28617	-1.49096	0.00000
NM103314B	1.01266	0.91737	0.00000	1.76595	-0.27671	-1.48924	0.00000

Table J-6. IRT Parameters for Operational Items on the NM-MSSA Grade 8 ELA Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*
402075	0.33370	-0.23090	0.10559				
402077	0.92175	1.34470	0.17955				
402111	0.55580	-1.72053	0.00000				
402116	0.81859	-0.97802	0.05162				
402118	0.79469	-0.37100	0.07989				
546059	1.18705	0.11134	0.22303				
546065	0.63638	0.01474	0.17495				
546080	0.86398	-0.03888	0.22303				
546082	1.19646	-0.13890	0.10753				
546084	0.94486	-1.17153	0.04988				
560416	0.96182	-0.10844	0.19668				
560420	0.52814	0.15437	0.15426				
560428	0.69013	-0.28790	0.21886				
560433	0.53230	0.62934	0.31666				
560440	0.67047	0.56184	0.18048				
560442	0.97505	-0.73333	0.23961				
641557	0.62043	-0.73981	0.24383				
641559	1.25164	0.26083	0.29332				
641563	1.10268	-0.46287	0.24669				
641565	0.77761	-0.46151	0.13879				
641579	0.49110	2.14605	0.20164				
743491	1.15326	1.33652	0.18674				
743493	0.47234	0.98429	0.27289				
743499	0.86978	1.42474	0.22154				
743504	1.12135	0.83264	0.14057				
743508	1.53164	-0.29804	0.18344				
743510	0.84018	-0.29724	0.12955				
402079	0.67022	0.06441	0.00000	0.45776	-0.45776	0.00000	
546073	0.62949	0.68853	0.00000	0.41057	-0.41057	0.00000	
560404	0.55740	0.29893	0.00000	1.02532	-1.02532	0.00000	
641567	0.62076	0.11442	0.00000	0.18266	-0.18266	0.00000	
NM110392	0.32584	1.67887	0.00000	1.20072	-1.20072	0.00000	
NM101422A	1.70784	0.82778	0.00000	0.80016	-0.02024	-0.77992	0.00000
NM101422B	1.81564	0.75544	0.00000	0.73281	0.03920	-0.77201	0.00000
NM101473A	1.74361	1.12116	0.00000	0.81266	-0.01302	-0.79964	0.00000
NM101473B	1.77962	1.04474	0.00000	0.73973	0.05423	-0.79397	0.00000
NM101480A	1.26246	0.91574	0.00000	1.41994	-0.19580	-1.22414	0.00000
NM101480B	1.27605	0.76806	0.00000	1.27172	-0.13170	-1.14001	0.00000

Table J-7. IRT Parameters for Operational Items on the NM-MSSA Grade 3 Mathematics Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
400604	0.77295	0.15502	0.23344					
408129	1.21198	0.83531	0.16994					
408165	1.41468	0.83346	0.23907					
411119	0.80831	-0.46046	0.08694					
411764	0.84376	0.95669	0.17507					
413036	1.33719	0.87708	0.22145					
417040	1.00279	-1.07412	0.25513					
462672	1.03955	-0.47844	0.10207					
464204	1.15216	-0.15448	0.16485					
464225	0.41206	2.34714	0.00000					
464268	1.06338	0.21647	0.29154					
539890	1.24785	1.14369	0.26562					
539903	0.86501	0.19996	0.21398					
539940	1.01875	-1.38545	0.20424					
541508	0.60994	-0.48337	0.00000					
557246	0.91650	0.76682	0.16902					
619075	1.20172	2.13958	0.19995					
619084	1.22141	0.33374	0.43351					
619098	0.32644	-1.33036	0.00000					
619106	1.23039	1.84550	0.24583					
619137	1.38028	0.39715	0.49680					
619174	1.09432	-0.51282	0.29902					
619192	0.83943	-0.30900	0.27783					
619211	0.78871	1.06932	0.13361					
619217	1.04069	0.73171	0.19273					
619227	0.87729	-0.26352	0.16102					
619235	1.58617	1.45311	0.33531					
619242	0.88938	-0.32627	0.11219					
619276B	0.76484	1.15964	0.00000					
619288B	0.56389	0.91261	0.00000					
690977	0.61729	2.72891	0.37077					
691034	0.96900	-0.13351	0.13911					
691038	0.77719	0.75605	0.00000					
691053	0.99135	0.21861	0.00000					
691055	0.98323	-1.39226	0.19873					
532101B	0.97658	0.80465	0.00000	0.65622	-0.65622	0.00000		
619276A	0.83773	0.04209	0.00000	0.80540	-0.80540	0.00000		
619288A	1.02123	1.38061	0.00000	0.64187	-0.64187	0.00000		
785068B	0.99159	1.70403	0.00000	0.78479	-0.78479	0.00000		
532101A	0.94395	0.81004	0.00000	1.04052	0.29028	-0.31012	-1.02069	0.00000
785068A	1.15516	0.66554	0.00000	1.09662	0.54845	-0.30956	-1.33551	0.00000

Table J-8. IRT Parameters for Operational Items on the NM-MSSA Grade 4 Mathematics Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
124856A	0.89753	0.78451	0.16573					
126020A	1.13609	0.96475	0.11599					
127388A	0.85444	1.34870	0.14000					
127466A	1.38525	1.62226	0.14704					
407489	1.19982	-0.23301	0.30782					
407491	0.78345	0.30589	0.10085					
462873	1.51153	1.67529	0.19465					
462916	1.40369	0.87278	0.17963					
465876	0.28676	3.46525	0.20000					
540283	0.96554	1.58183	0.12785					
540312	0.84926	1.38737	0.15211					
540324	0.69893	0.09077	0.24021					
540589	1.21716	1.05363	0.21918					
540599	1.08582	1.11936	0.00000					
540658B	1.01721	1.38976	0.00000					
541517	1.04609	0.42218	0.23908					
560945	0.64262	0.88310	0.23248					
629029	0.70001	-0.00773	0.24316					
629036	1.10810	-0.46271	0.00000					
629038	1.10668	1.64849	0.27785					
629048	0.94359	1.86535	0.21223					
629066	1.14351	0.51385	0.12449					
629073	1.12537	1.26735	0.00000					
629094	1.44477	1.58738	0.27326					
629111	0.90093	-0.01836	0.19078					
629123	0.70683	2.39831	0.23462					
630485B	1.11940	0.28442	0.00000					
691166	0.62177	1.04193	0.24947					
691215	0.74516	1.47860	0.24029					
740654	1.01858	1.59055	0.24524					
740658	0.73049	-0.00018	0.20694					
740662	1.77799	0.50527	0.14843					
740664	1.07531	0.54011	0.31192					
740668	0.65612	0.27984	0.26545					
740694	1.00522	0.87741	0.26306					
411965B	0.99083	1.73972	0.00000	0.44719	-0.44719	0.00000		
540658A	1.03935	0.44286	0.00000	0.70565	-0.70565	0.00000		
630481B	1.05789	0.49360	0.00000	1.12425	-1.12425	0.00000		
630485A	0.98666	0.22857	0.00000	0.51454	-0.51454	0.00000		
411965A	1.14814	0.98013	0.00000	0.94080	0.29182	-0.33263	-0.89998	0.00000
630481A	1.18092	0.85932	0.00000	1.53665	0.55614	-0.46754	-1.62524	0.00000

Table J-9. IRT Parameters for Operational Items on the NM-MSSA Grade 5 Mathematics Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
400300	0.28337	4.66411	0.20000					
400488	1.37744	0.80542	0.17295					
400650	1.57056	1.57787	0.35576					
400711	0.95677	-1.03244	0.10624					
405943	1.27412	0.24073	0.18722					
408463	1.36853	1.80166	0.17193					
411149	1.29981	0.41710	0.22185					
411304	1.33004	-0.69030	0.14475					
413871	1.37431	1.28749	0.15912					
464071	0.96732	-0.72449	0.24649					
464399	1.58832	0.47412	0.22059					
465792	0.99272	1.17370	0.23560					
532490	1.16540	0.35123	0.21691					
539162	0.88489	0.01670	0.28319					
539164	1.04984	0.66932	0.10488					
539188	1.30813	1.66606	0.43712					
540635	0.61325	0.48727	0.18081					
540637	0.72155	-1.21573	0.14256					
540710	1.37836	0.60508	0.13071					
558693	1.29237	0.17457	0.28249					
607336	0.91007	-0.10288	0.13444					
607394	1.03575	-0.27889	0.20112					
607495	1.04036	1.73310	0.13850					
607514	1.14076	0.47687	0.32937					
607538	0.84185	0.59737	0.25981					
692838	0.83196	0.99948	0.30377					
692846	1.99810	1.05048	0.28326					
693074	0.81005	0.95968	0.28107					
695196	0.73500	-0.48910	0.30119					
695213	0.89512	1.08185	0.25114					
695227B	1.38200	1.56734	0.00000					
781233	0.86728	-0.04617	0.00000					
783563B	0.81482	2.34457	0.00000					
412281B	1.40278	0.44428	0.00000	0.47448	-0.47448	0.00000		
695227A	1.29753	1.08666	0.00000	0.45125	-0.45125	0.00000		
695233B	1.29915	1.36279	0.00000	0.44267	-0.44267	0.00000		
783563A	1.04430	0.24016	0.00000	0.08217	-0.08217	0.00000		
412281A	1.42376	0.26765	0.00000	0.64372	0.24327	-0.10096	-0.78602	0.00000
695233A	1.48930	0.97475	0.00000	1.00807	0.41799	-0.51294	-0.91312	0.00000

Table J-10. IRT Parameters for Operational Items on the NM-MSSA Grade 6 Mathematics Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
124799A	1.14516	-0.68946	0.04400					
127179A	1.45318	0.15922	0.42448					
127738A	1.19485	0.31310	0.28178					
406099	0.52026	0.99136	0.18075					
411834	0.77947	0.04636	0.24506					
412144	0.65403	0.32307	0.11413					
412393	1.24813	1.31766	0.14814					
412439	0.90972	2.06124	0.18001					
412462	0.65202	0.01646	0.18622					
464785	1.94245	0.64906	0.07749					
532604B	1.05394	2.22743	0.00000					
539618	1.23135	1.01758	0.20155					
539622	1.01929	0.89872	0.22317					
539643	2.20986	1.19269	0.11250					
539649	1.00178	-0.02580	0.32345					
539809	0.95984	-0.73097	0.25413					
540132	1.06453	0.72927	0.27286					
540196B	0.94027	1.15074	0.00000					
540725	1.63269	1.19692	0.26648					
558413	0.90881	1.57586	0.22176					
607665	0.59237	-0.08441	0.27064					
607688	1.12000	0.85156	0.23408					
607721	1.02259	-0.79680	0.22891					
607725	1.07710	-0.56256	0.25166					
607751	0.92025	-0.15864	0.09901					
607773	1.65294	1.15107	0.16715					
607775	0.82188	-1.15297	0.18581					
607782	0.54619	2.74407	0.30060					
607785	0.27672	3.33411	0.00000					
695581	0.48594	-0.79437	0.16764					
695587	1.50665	1.24694	0.34156					
695595	1.91497	1.17275	0.27260					
695600	1.86502	0.95409	0.15643					
695606	0.83052	1.27841	0.21680					
697210	0.51785	1.18999	0.00000					
700748	0.91704	0.78678	0.00000					
700976	1.71862	1.23067	0.35300					
706548	1.08987	-1.08522	0.28826					
412531B	1.50942	1.25546	0.00000	0.39149	-0.39149	0.00000		
532604A	0.59696	2.44813	0.00000	0.87122	-0.87122	0.00000		
539624B	0.97344	0.42306	0.00000	0.45449	-0.45449	0.00000		
540196A	0.88624	-0.03070	0.00000	0.29134	-0.29134	0.00000		
412531A	1.29727	0.86804	0.00000	0.92807	0.30992	-0.26832	-0.96967	0.00000
539624A	1.16235	0.65110	0.00000	1.44101	0.53879	-0.73568	-1.24413	0.00000

Table J-11. IRT Parameters for Operational Items on the NM-MSSA Grade 7 Mathematics Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
406243B	1.44001	0.33896	0.00000					
412224	1.05991	0.53346	0.28665					
412251	1.24080	-0.32214	0.29490					
412395	1.76031	0.95898	0.11343					
412425	0.83278	-0.61524	0.16100					
412486	1.64388	0.62485	0.29237					
412493	1.21473	0.16538	0.12359					
412656B	1.13688	0.24199	0.00000					
467208	1.55212	0.15405	0.21132					
467772	1.80057	1.84966	0.27623					
467826	1.49738	1.58260	0.21348					
539394	1.42700	0.78538	0.50375					
539407	1.03167	0.63599	0.21435					
539444	1.02201	0.78447	0.20284					
539450	1.05989	0.07403	0.17885					
540128	1.35964	-0.37361	0.27729					
557976	1.24811	1.68957	0.27463					
607105	1.15474	0.83277	0.24010					
607107	1.70733	1.10485	0.31527					
607119	1.16481	-0.65674	0.20009					
607131	1.18381	0.34359	0.16778					
607135	1.88866	1.59095	0.45928					
607155	0.47000	0.56725	0.17244					
607157	1.88323	1.48935	0.19259					
607161	1.81844	-0.15942	0.17320					
607163	0.78018	-0.24218	0.18012					
607203	1.00940	1.30031	0.33310					
607205	0.74931	-0.38622	0.22352					
607213	0.91739	-0.75556	0.23443					
607215	1.30308	0.41722	0.17607					
687771	1.94213	0.78528	0.07512					
687773	2.22449	1.09876	0.07428					
687793	1.74220	1.22061	0.37656					
687860	1.86211	1.75155	0.13491					
690152	0.71495	-1.37908	0.21295					
691314	1.02454	1.81841	0.00000					
691316	2.18396	1.23250	0.31402					
781526	0.85328	-0.81581	0.23748					
406243A	1.38149	-0.18515	0.00000	0.29503	-0.29503	0.00000		
412656A	1.38877	0.53852	0.00000	0.50552	-0.50552	0.00000		
607222B	1.19621	1.30752	0.00000	0.65330	-0.65330	0.00000		
691595B	1.44795	1.05068	0.00000	0.60912	-0.60912	0.00000		
607222A	1.31059	0.47099	0.00000	0.54172	0.34284	-0.34748	-0.53708	0.00000
691595A	0.97581	0.90454	0.00000	1.35388	0.04005	-0.33716	-1.05677	0.00000

Table J-12. IRT Parameters for Operational Items on the NM-MSSA Grade 8 Mathematics Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
126335A	1.05521	1.48006	0.13964					
126398A	1.56383	1.49274	0.11396					
127164A	1.65517	1.15960	0.15706					
400766	0.60541	1.90436	0.19883					
400771	0.40996	0.31564	0.12831					
408524	1.05436	0.71358	0.21198					
413137	1.47901	2.00441	0.22304					
413229	0.74671	-1.07382	0.00000					
414834	1.20803	1.49891	0.16951					
414849	0.83680	3.20628	0.34539					
414957	0.51464	1.10948	0.28025					
465465	0.98915	0.91087	0.21744					
468006	0.97057	1.99540	0.00000					
468359	0.65119	1.15921	0.12668					
483010	0.50784	0.31069	0.28332					
483452	0.63025	0.95902	0.21998					
540844	1.77215	1.68057	0.22785					
540892	1.60826	-0.31803	0.37446					
540918	0.86816	0.80448	0.34703					
540955	0.76456	-0.37982	0.14052					
541134	0.80523	0.97638	0.10604					
551249B	1.48575	1.61675	0.00000					
614704	0.52375	2.39446	0.37395					
614780	0.53858	0.82498	0.11990					
614837	1.42308	0.73366	0.19535					
614856	1.01402	0.38035	0.00000					
614943	1.25567	2.00586	0.19989					
615097	0.94712	1.47257	0.18109					
615111	0.81070	0.63950	0.00000					
615300	1.49443	1.29621	0.41767					
615422B	0.61012	1.02292	0.00000					
631977	1.00831	0.80760	0.14531					
696121	1.78063	0.89723	0.43953					
696123	0.70766	0.20381	0.23102					
696125	0.66231	-0.49325	0.17599					
696131	1.03094	-1.53953	0.19752					
696139	1.06742	1.48475	0.19982					
696141	1.10342	1.22581	0.19349					
697997	0.67292	0.20946	0.16389					
551249A	1.31409	1.37297	0.00000	0.27972	-0.27972	0.00000		
615320B	1.08486	0.64931	0.00000	0.49877	-0.49877	0.00000		
615411B	1.36956	1.20927	0.00000	0.47595	-0.47595	0.00000		
615422A	0.91391	0.26326	0.00000	0.96919	-0.96919	0.00000		
615320A	1.43491	1.45480	0.00000	0.73678	0.32981	-0.36981	-0.69678	0.00000
615411A	1.02978	1.48921	0.00000	1.02099	0.19414	-0.21723	-0.99790	0.00000

Table J-13. IRT Parameters for Operational Items on the ASR Grade 5 Science Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
637807	0.62288	-1.35588	0.00000					
637951	0.94836	0.77058	0.22658					
638656	1.01219	2.34572	0.19770					
638658	0.63882	0.79630	0.09667					
706138	0.51892	1.59056	0.17098					
706149	0.41471	-0.19523	0.00000					
706792	0.83975	1.37537	0.26479					
706801	0.99637	0.32987	0.00000					
744445	0.57361	0.93911	0.15286					
744451	0.43881	1.26637	0.00000					
784744	1.38303	0.15901	0.22321					
848745	1.14502	-0.08301	0.11585					
626442	0.73417	-0.85181	0.00000	0.68121	-0.68121	0.00000		
626478	0.49019	0.78409	0.00000	1.42087	-1.42087	0.00000		
631577	0.81699	-0.01798	0.00000	0.80992	-0.80992	0.00000		
632426	0.41017	1.47637	0.00000	1.59690	-1.59690	0.00000		
632454	0.47781	1.06192	0.00000	1.36879	-1.36879	0.00000		
633906	0.52965	2.17526	0.00000	1.83215	-1.83215	0.00000		
634127	0.60273	1.28859	0.00000	1.79429	-1.79429	0.00000		
635879	0.55434	-0.23204	0.00000	0.39888	-0.39888	0.00000		
635886	0.24564	4.63068	0.00000	3.04072	-3.04072	0.00000		
636189	0.58582	2.44211	0.00000	1.73361	-1.73361	0.00000		
636211	0.55585	0.06320	0.00000	0.94133	-0.94133	0.00000		
638354	0.81846	0.10290	0.00000	0.53019	-0.53019	0.00000		
638526	0.69458	-0.13430	0.00000	0.73171	-0.73171	0.00000		
638558	1.03441	-0.45257	0.00000	0.60170	-0.60170	0.00000		
638639	0.32562	-0.07699	0.00000	1.32820	-1.32820	0.00000		
639474	0.89836	0.15411	0.00000	0.79037	-0.79037	0.00000		
639510	0.74656	0.67090	0.00000	0.97387	-0.97387	0.00000		
697027	0.64913	-0.42395	0.00000	0.96675	-0.96675	0.00000		
697044	0.46279	-0.29765	0.00000	1.78577	-1.78577	0.00000		
706119	0.61514	-0.17125	0.00000	1.03488	-1.03488	0.00000		
706135	0.63263	0.51028	0.00000	1.01185	-1.01185	0.00000		
706765	0.79008	-0.08108	0.00000	0.71768	-0.71768	0.00000		
706847	0.66793	0.72810	0.00000	0.98715	-0.98715	0.00000		
737916	0.59711	-0.43392	0.00000	1.07544	-1.07544	0.00000		
743165	0.77342	0.79238	0.00000	1.07838	-1.07838	0.00000		
744455	0.77711	-1.01047	0.00000	0.80538	-0.80538	0.00000		
756457	0.69889	1.85712	0.00000	1.19726	-1.19726	0.00000		
762758	0.51585	0.88948	0.00000	1.26419	-1.26419	0.00000		
784722	0.61696	0.67545	0.00000	1.25742	-1.25742	0.00000		
784830	0.72861	-0.49765	0.00000	0.80524	-0.80524	0.00000		
784847	0.55221	0.24596	0.00000	1.13963	-1.13963	0.00000		
785041	0.55916	0.40336	0.00000	1.28645	-1.28645	0.00000		
848747	0.77990	-0.63973	0.00000	0.78018	-0.78018	0.00000		
661177	1.01161	1.53694	0.00000	2.34981	0.98451	-0.83217	-2.50215	0.00000
666120	0.62360	3.29063	0.00000	3.12968	1.13869	-0.96869	-3.29969	0.00000
697164	0.93395	1.37909	0.00000	1.89904	0.64813	-0.51044	-2.03674	0.00000

Table J-14. IRT Parameters for Operational Items on the ASR Grade 8 Science Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
636837	1.24502	0.88903	0.20065					
636843	0.82189	1.19908	0.23931					
641873	0.85801	1.37787	0.21168					
641894	0.28539	1.46677	0.00000					
709292	0.73517	1.87278	0.21466					
709306	0.37800	2.77656	0.00000					
709609	0.70080	1.40146	0.10364					
709617	0.58397	0.54659	0.00000					
713686	0.59822	0.11082	0.00000					
713695	0.73355	1.15720	0.00000					
758880	0.67003	0.04384	0.00000					
758938	1.02121	1.12092	0.34469					
847907	1.14936	1.34244	0.29435					
847909	1.05179	0.52412	0.35618					
847911	0.84247	0.98580	0.15255					
847927	1.11545	0.72687	0.32027					
631360	0.45462	-0.08329	0.00000	1.44254	-1.44254	0.00000		
636830	0.58204	-0.28410	0.00000	1.11864	-1.11864	0.00000		
636852	0.48058	0.62908	0.00000	1.14652	-1.14652	0.00000		
637562	0.74441	-0.05665	0.00000	0.85441	-0.85441	0.00000		
637622	0.70843	-0.17387	0.00000	0.83869	-0.83869	0.00000		
637635	0.34298	-0.32363	0.00000	2.12381	-2.12381	0.00000		
640163	0.32268	1.00587	0.00000	1.79723	-1.79723	0.00000		
640740	0.58809	-0.12792	0.00000	0.84869	-0.84869	0.00000		
641845	0.40293	2.25513	0.00000	1.71096	-1.71096	0.00000		
641866	0.72230	-1.48305	0.00000	1.47667	-1.47667	0.00000		
642091	0.32482	2.04713	0.00000	1.92193	-1.92193	0.00000		
642855	1.08455	-0.07076	0.00000	0.33712	-0.33712	0.00000		
643622	0.60187	1.39915	0.00000	1.22298	-1.22298	0.00000		
696483	0.40621	1.00668	0.00000	1.74304	-1.74304	0.00000		
707172	0.51382	0.97170	0.00000	2.13148	-2.13148	0.00000		
709294	0.40649	3.21875	0.00000	2.05849	-2.05849	0.00000		
709309	0.64171	0.41131	0.00000	1.05488	-1.05488	0.00000		
709604	0.33183	3.36256	0.00000	2.11492	-2.11492	0.00000		
709622	0.81753	-0.21239	0.00000	0.79440	-0.79440	0.00000		
712986	0.49914	0.28444	0.00000	1.40211	-1.40211	0.00000		
713386	0.67185	0.37076	0.00000	0.99105	-0.99105	0.00000		
713388	0.31410	0.88684	0.00000	2.23302	-2.23302	0.00000		
716040	0.64656	-0.31981	0.00000	1.23043	-1.23043	0.00000		
731140	0.66375	0.09718	0.00000	0.79210	-0.79210	0.00000		
741297	0.81002	0.40976	0.00000	0.92535	-0.92535	0.00000		
752254	0.64346	0.34647	0.00000	1.08352	-1.08352	0.00000		
758919	0.79424	1.07421	0.00000	1.24373	-1.24373	0.00000		
763243	0.47989	0.36736	0.00000	1.45840	-1.45840	0.00000		
786048	0.15757	0.04204	0.00000	3.17044	-3.17044	0.00000		
847905	0.55021	1.31665	0.00000	1.18868	-1.18868	0.00000		
847934	0.57383	0.37724	0.00000	1.15132	-1.15132	0.00000		
663576	1.17534	2.16833	0.00000	1.95642	0.64121	-0.58147	-2.01615	0.00000
697245	0.95052	2.31301	0.00000	2.19963	0.67501	-0.74059	-2.13405	0.00000
717529	0.94598	1.73954	0.00000	2.08847	0.61388	-0.64822	-2.05413	0.00000

Table J-15. IRT Parameters for Operational Items on the ASR Grade 11 Science Assessment

Item ID	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
633246	0.96478	-0.16949	0.23050					
633315	1.07057	0.72270	0.26507					
637994	0.70552	1.87717	0.27908					
639319	0.88077	0.03580	0.11812					
642634	0.58688	1.64355	0.22960					
643598	0.80955	0.10797	0.26409					
705787	0.51262	2.20020	0.11792					
705815	0.56550	1.85560	0.30390					
706534	0.40906	2.10731	0.00000					
706670	0.68128	0.33302	0.00000					
753774	0.71758	0.03096	0.13042					
753780	1.02408	2.19229	0.17602					
782469	1.03077	1.76116	0.26074					
782471	1.05919	0.65608	0.18642					
626027	0.66291	0.03377	0.00000	0.98072	-0.98072	0.00000		
627081	0.63793	0.93026	0.00000	1.13172	-1.13172	0.00000		
628033	0.86686	-0.30662	0.00000	0.76366	-0.76366	0.00000		
632730	0.63675	0.74758	0.00000	1.19477	-1.19477	0.00000		
633116	0.59157	0.53132	0.00000	1.17633	-1.17633	0.00000		
633266	0.44124	0.91661	0.00000	1.37493	-1.37493	0.00000		
636258	0.89574	-0.05648	0.00000	0.73651	-0.73651	0.00000		
637608	0.28352	1.79230	0.00000	2.31758	-2.31758	0.00000		
637610	0.39515	2.07339	0.00000	1.98026	-1.98026	0.00000		
639344	0.48912	-0.33432	0.00000	1.39477	-1.39477	0.00000		
639346	0.57702	1.04480	0.00000	1.11586	-1.11586	0.00000		
640447	0.41864	1.87538	0.00000	2.15293	-2.15293	0.00000		
640641	0.53125	1.06443	0.00000	1.30728	-1.30728	0.00000		
642245	0.51275	-0.44047	0.00000	1.13713	-1.13713	0.00000		
642377	0.36397	1.57542	0.00000	1.94544	-1.94544	0.00000		
642454	0.89617	0.64585	0.00000	0.83600	-0.83600	0.00000		
642533	0.65514	0.10687	0.00000	1.06042	-1.06042	0.00000		
696546	0.34190	3.65973	0.00000	2.69937	-2.69937	0.00000		
705738	0.22567	4.72835	0.00000	3.29800	-3.29800	0.00000		
705807	0.44018	1.57377	0.00000	2.40979	-2.40979	0.00000		
706468	0.69722	0.00100	0.00000	1.04956	-1.04956	0.00000		
706583	0.40086	1.51632	0.00000	1.74171	-1.74171	0.00000		
709638	0.56553	0.55154	0.00000	1.31949	-1.31949	0.00000		
710663	0.82237	0.09484	0.00000	0.67613	-0.67613	0.00000		
733448	0.33083	0.76398	0.00000	2.09123	-2.09123	0.00000		
738031	0.48744	1.90326	0.00000	1.53826	-1.53826	0.00000		
748109	0.55603	0.19015	0.00000	1.35926	-1.35926	0.00000		
755860	0.44223	-0.02481	0.00000	1.62641	-1.62641	0.00000		
762012	0.61362	0.01422	0.00000	1.06998	-1.06998	0.00000		
762916	0.42610	0.53328	0.00000	1.52480	-1.52480	0.00000		
762922	0.29663	2.33457	0.00000	2.79378	-2.79378	0.00000		
782467	0.73148	-0.01423	0.00000	0.91948	-0.91948	0.00000		
782473	0.72142	0.27466	0.00000	1.00053	-1.00053	0.00000		
782475	0.59645	0.82111	0.00000	1.03452	-1.03452	0.00000		
847612	0.41227	-0.37789	0.00000	1.65668	-1.65668	0.00000		
847619	0.44668	1.17249	0.00000	1.47266	-1.47266	0.00000		
847621	0.40791	1.19059	0.00000	2.22766	-2.22766	0.00000		
NM100318	0.51000	1.34947	0.00000	1.40473	-1.40473	0.00000		
666236	0.99709	1.45232	0.00000	2.18447	1.06249	-0.50034	-2.74663	0.00000
710876	0.52588	1.28889	0.00000	1.79352	0.53457	-0.60506	-1.72301	0.00000
735374	1.03598	1.60185	0.00000	1.98687	0.45377	-0.22145	-2.21919	0.00000

APPENDIX K
TEST CHARACTERISTIC CURVES
AND
CONDITIONAL STANDARD ERRORS OF MEASUREMENT

Figure K-1. Test Characteristic Curve for ELA—Grade 3

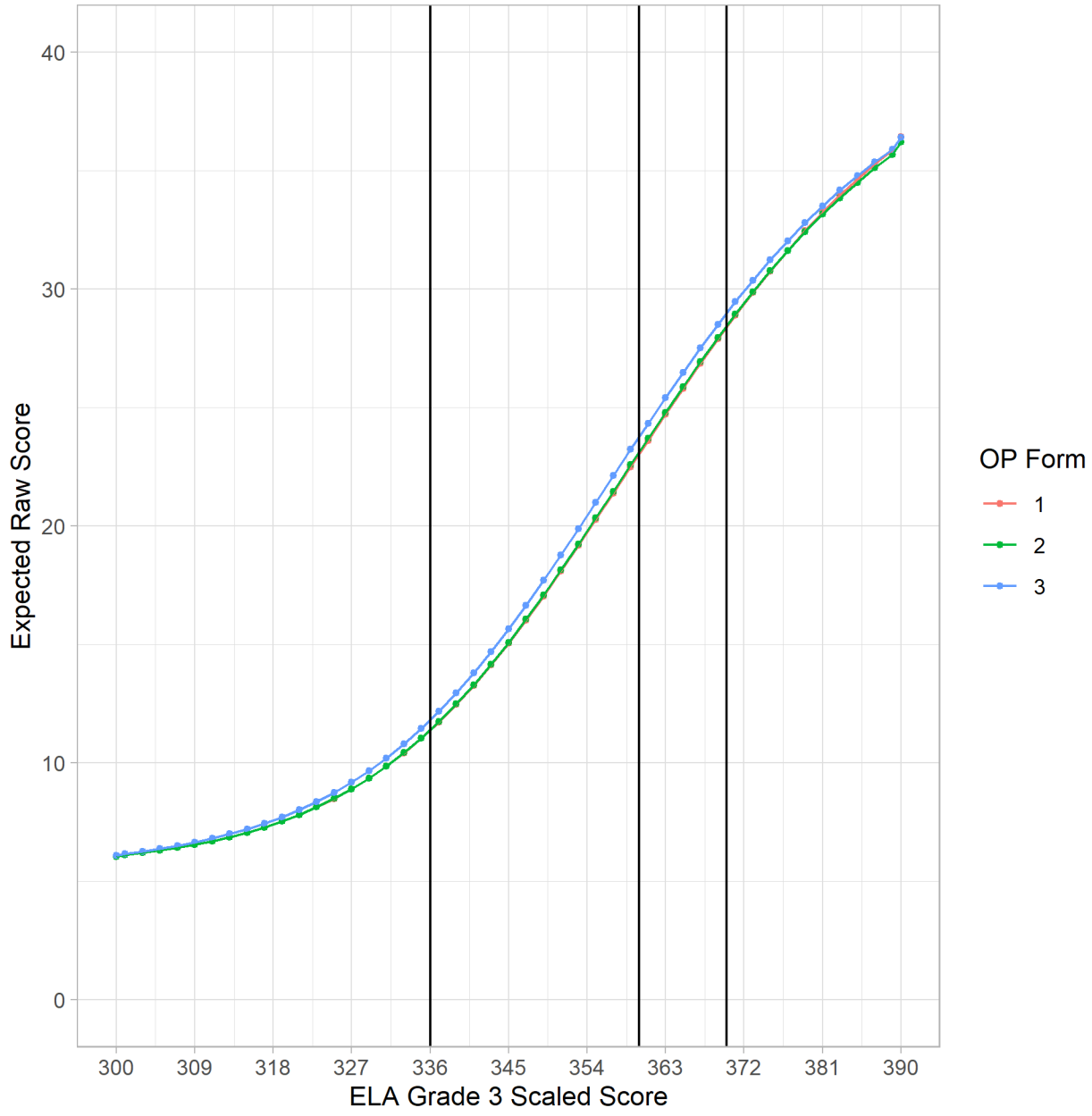


Figure K-2. Test Characteristic Curve for ELA—Grade 4

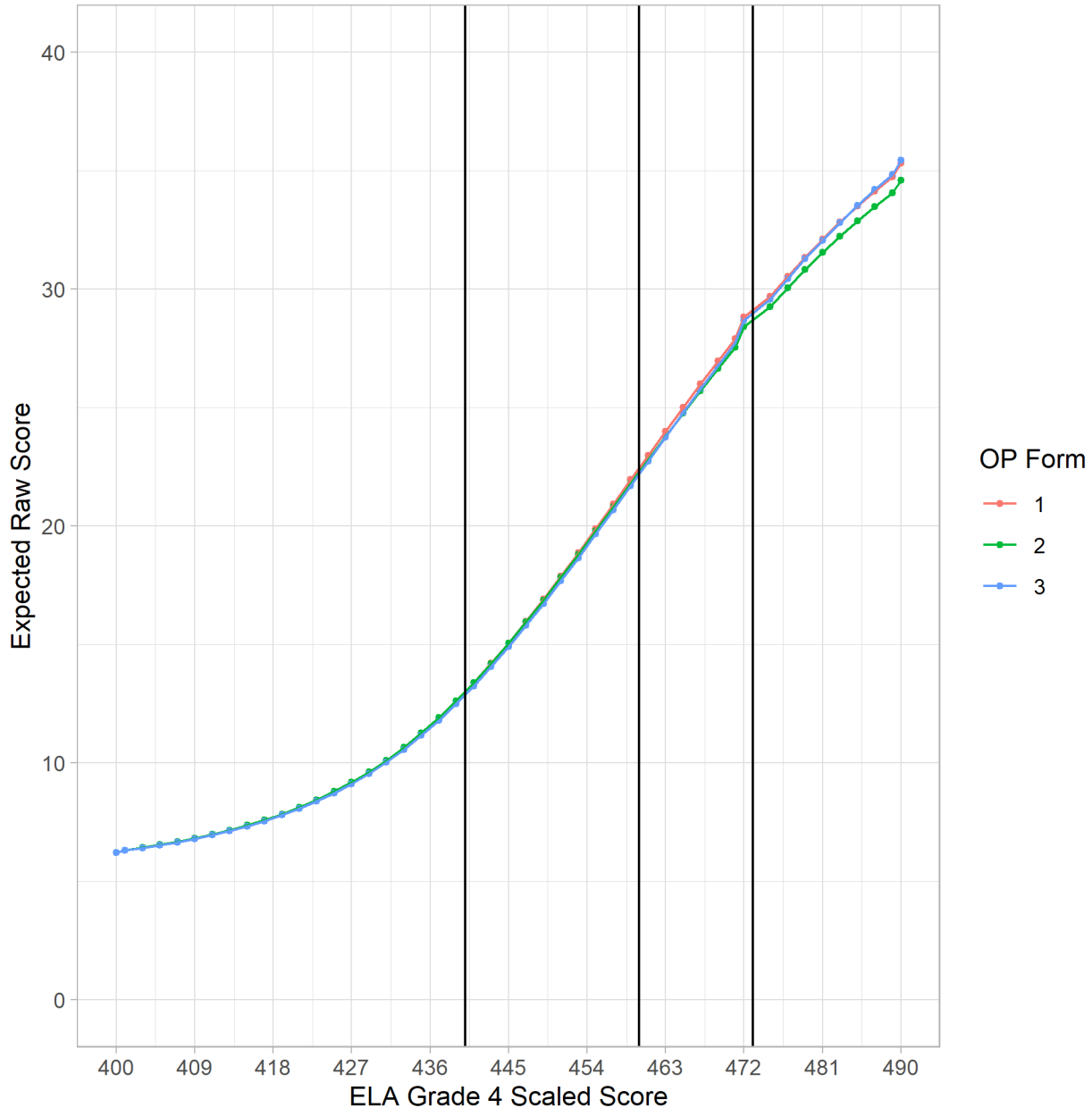


Figure K-3. Test Characteristic Curve for ELA—Grade 5

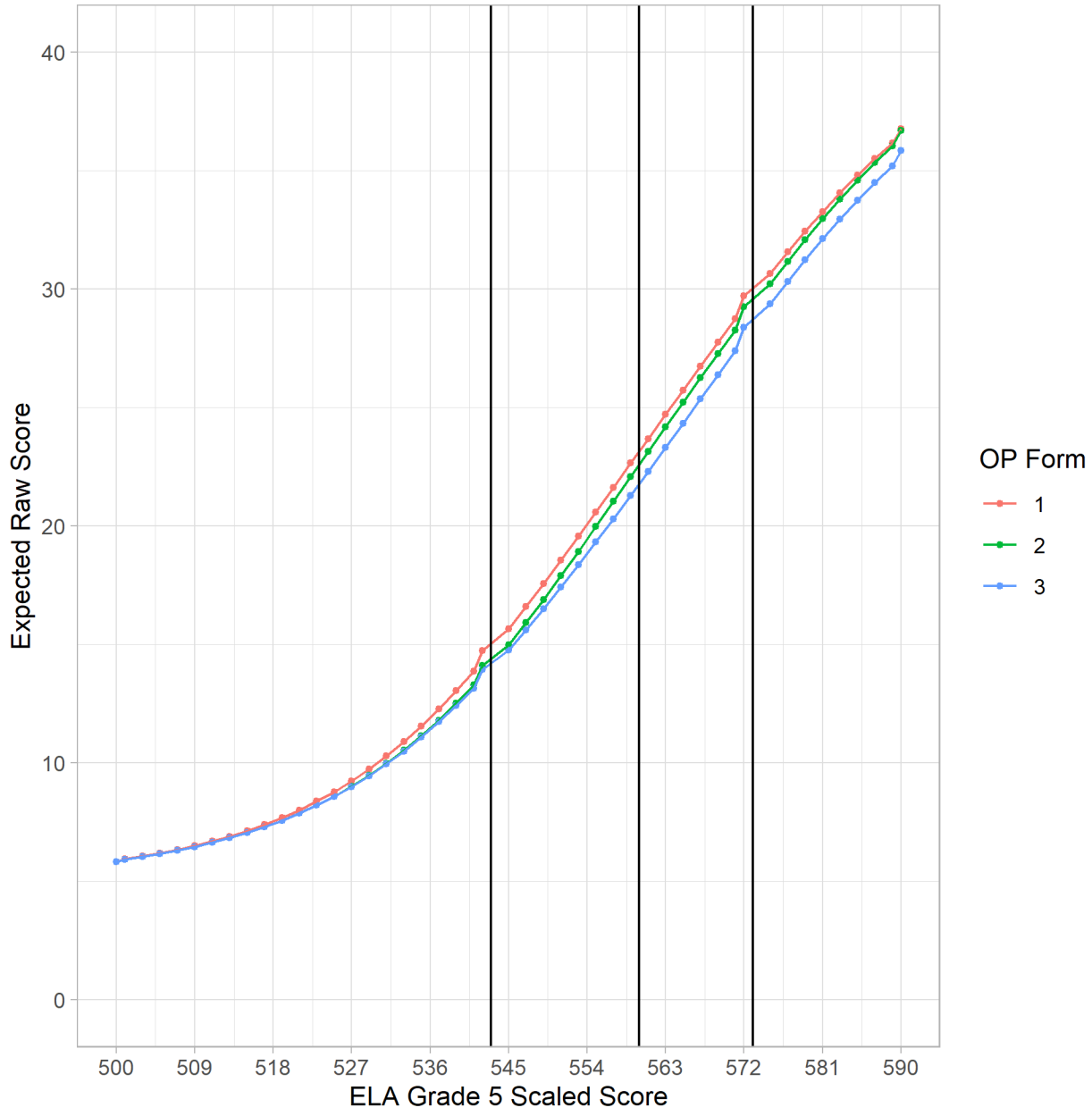


Figure K-4. Test Characteristic Curve for ELA—Grade 6

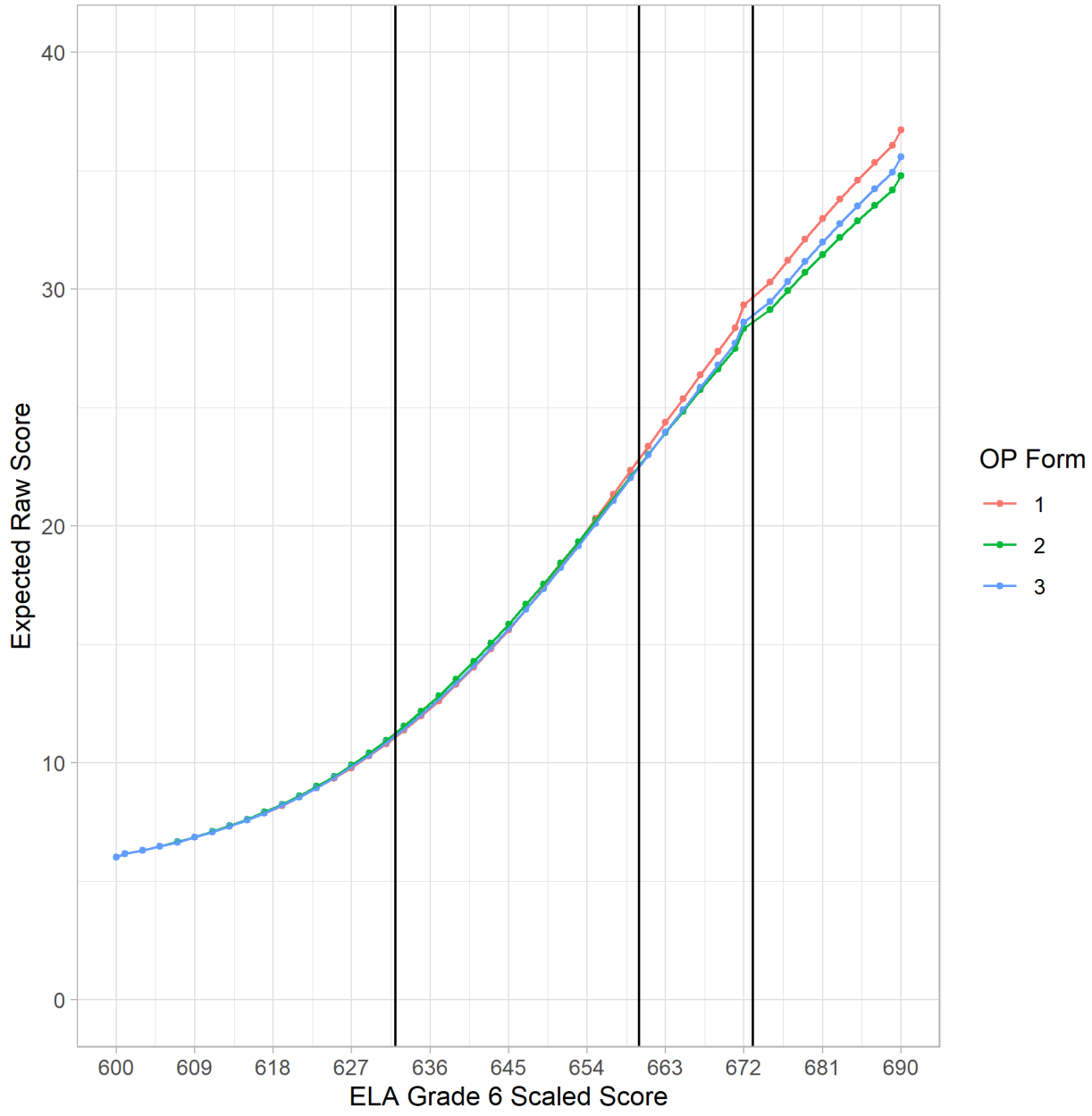


Figure K-5. Test Characteristic Curve for ELA—Grade 7

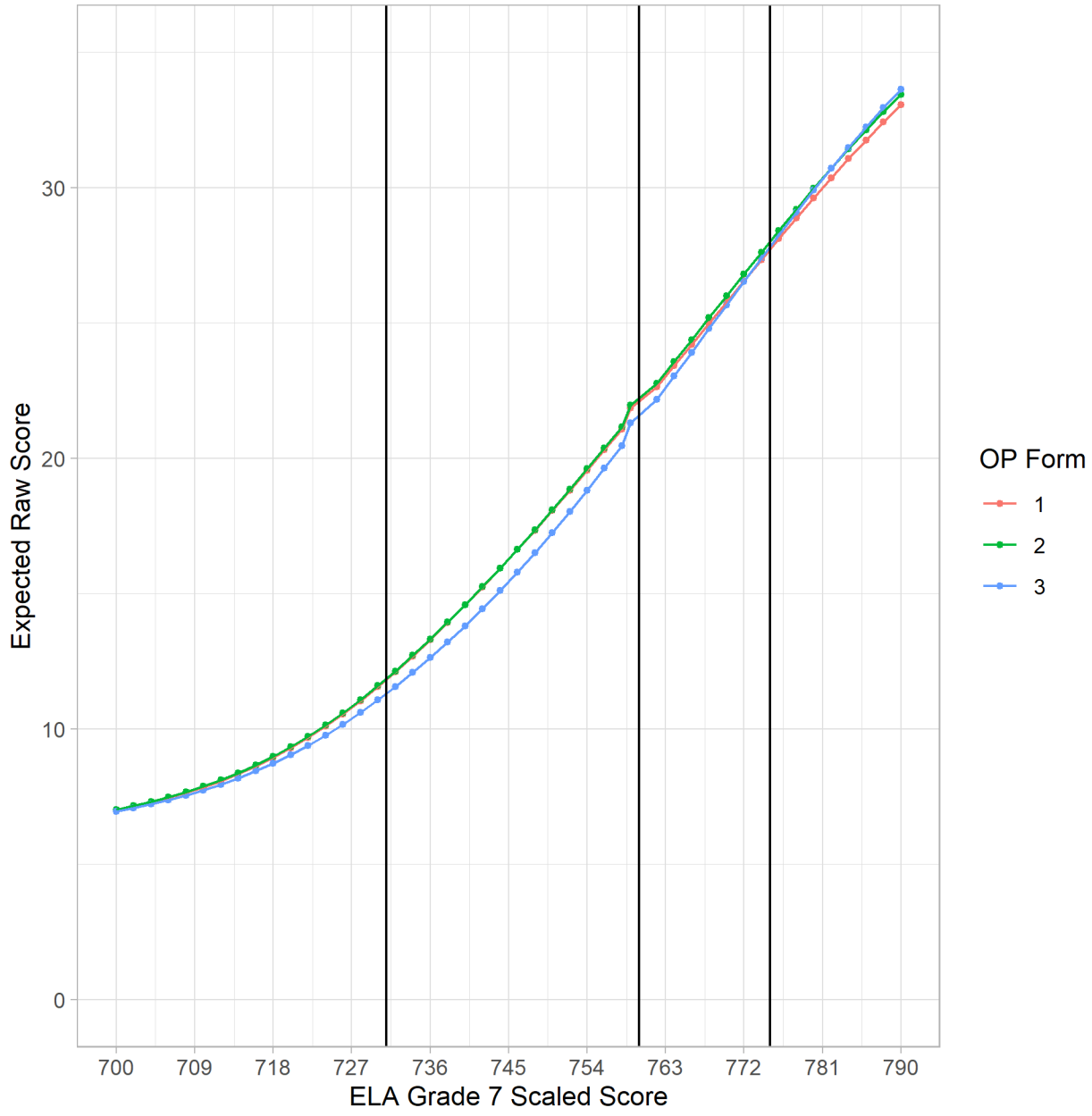


Figure K-6. Test Characteristic Curve for ELA—Grade 8

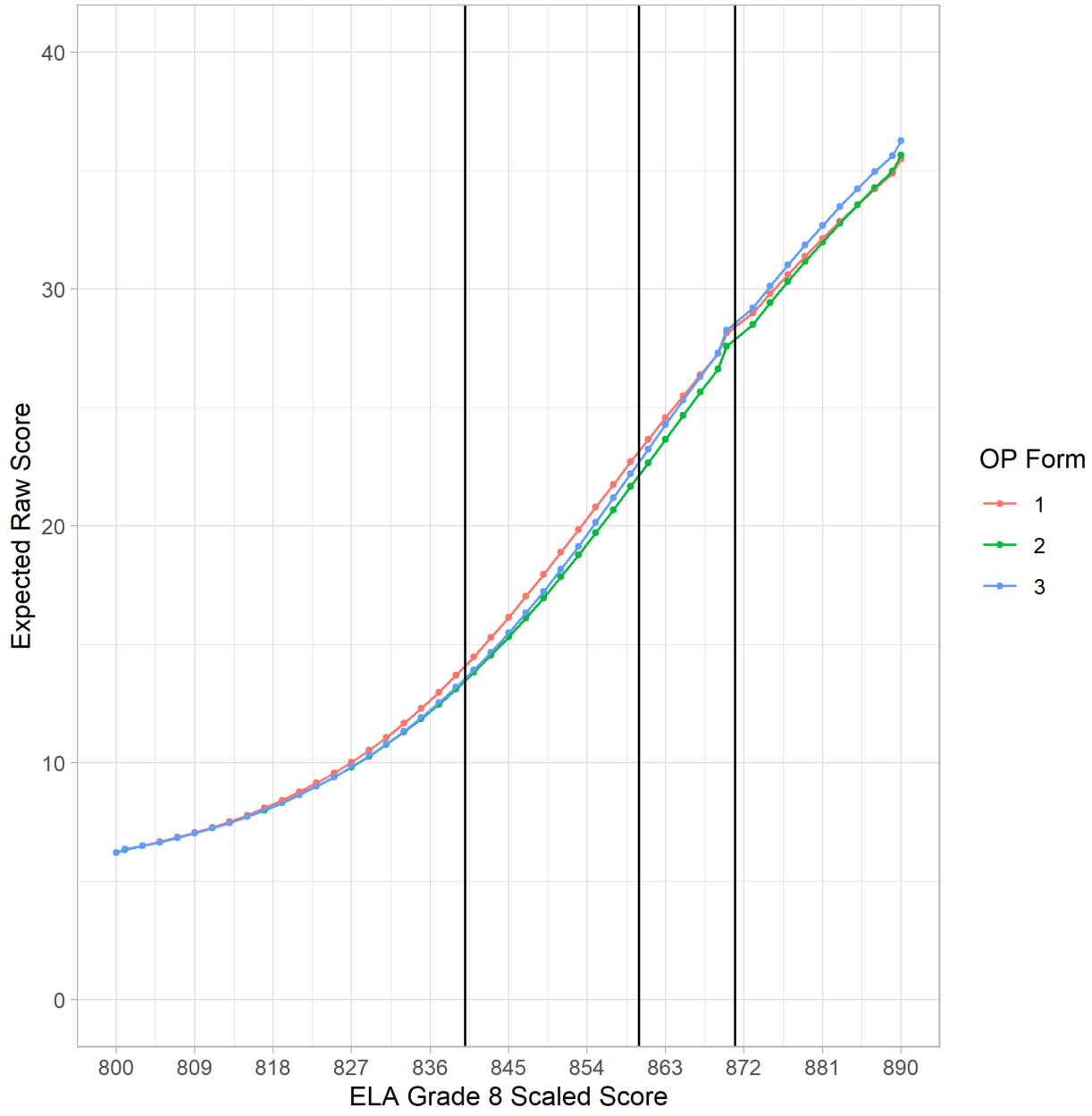


Figure K-7. Test Characteristic Curve for Mathematics—Grade 3

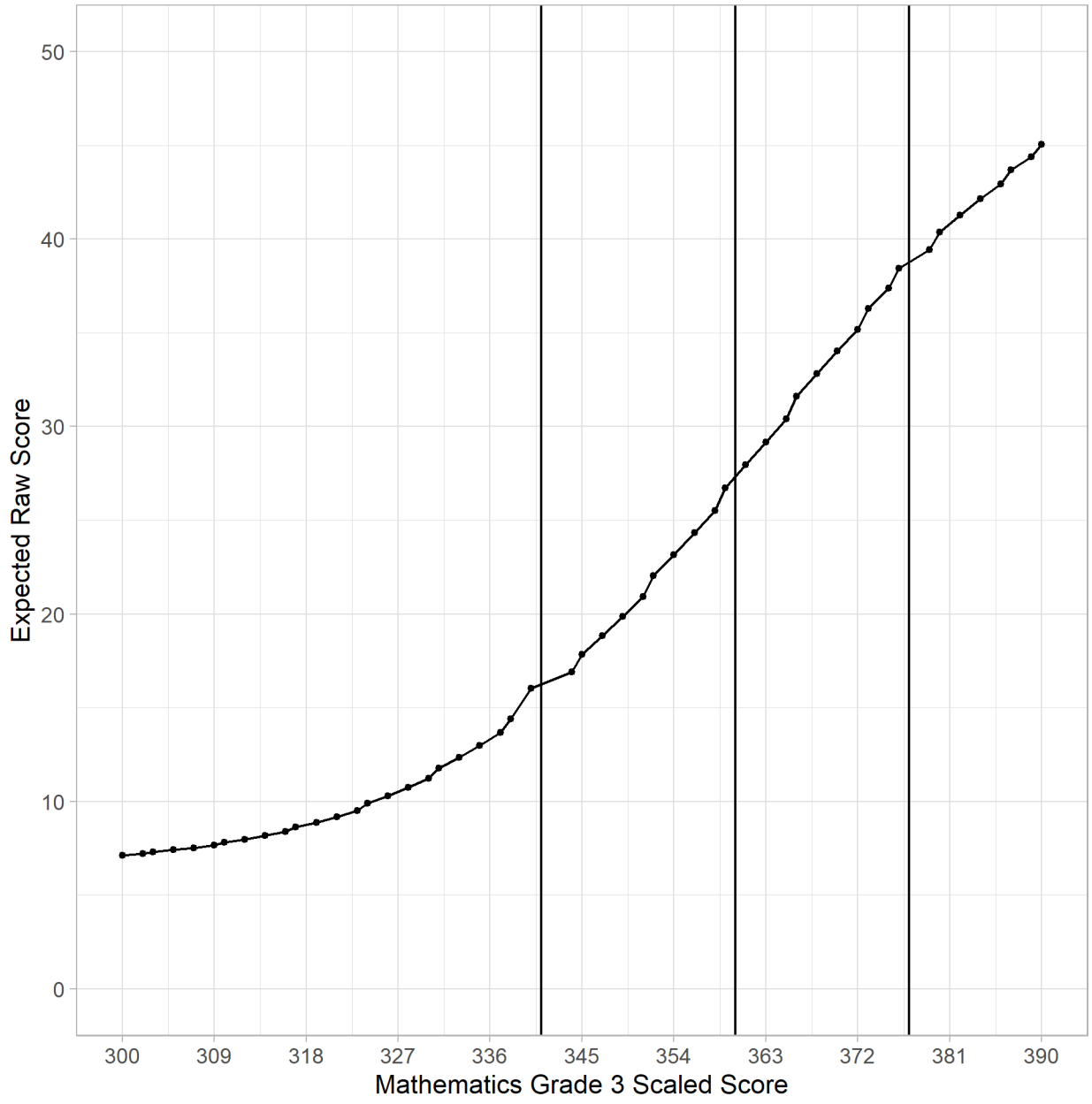


Figure K-8. Test Characteristic Curve for Mathematics—Grade 4

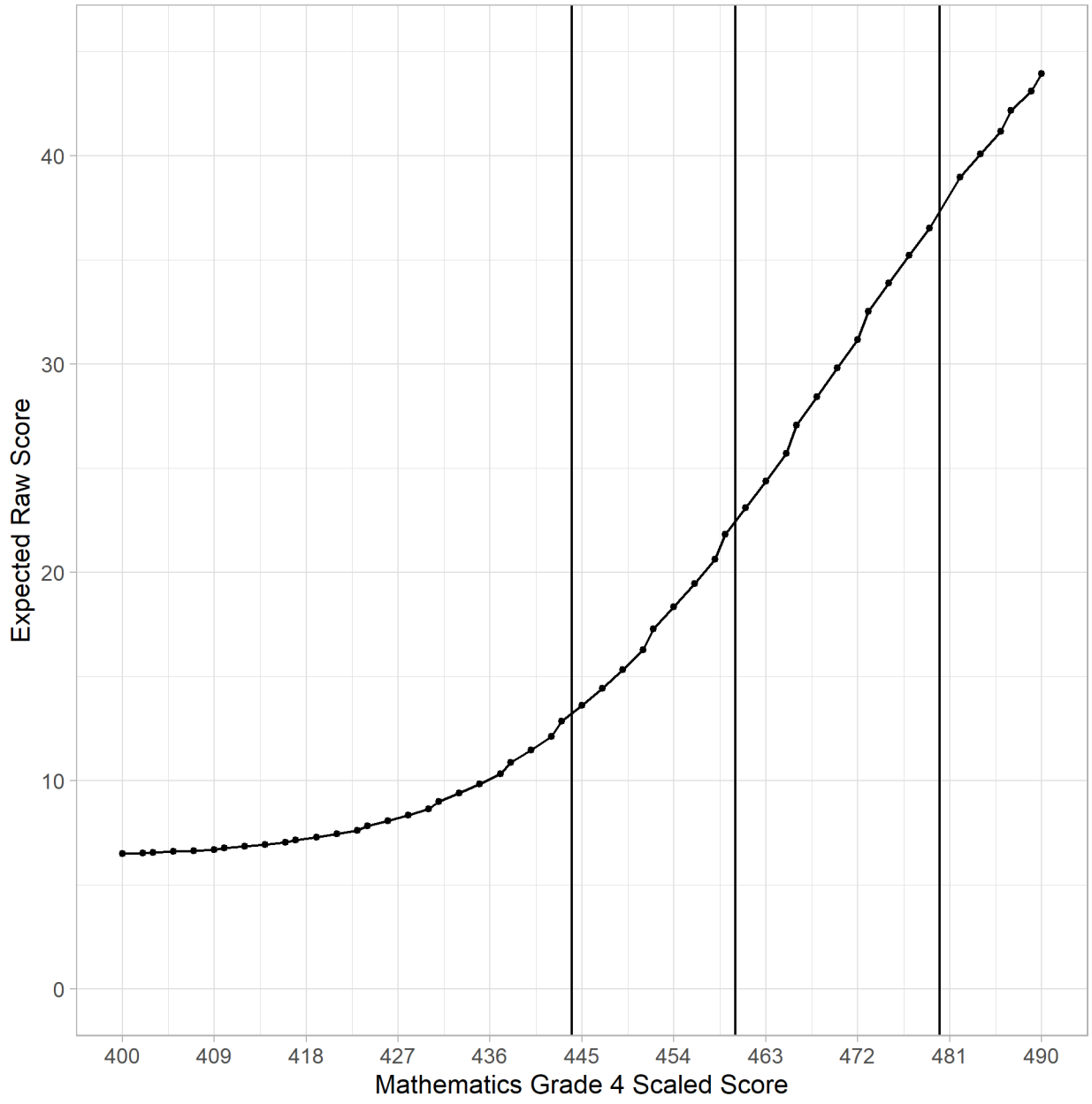


Figure K-9. Test Characteristic Curve for Mathematics—Grade 5

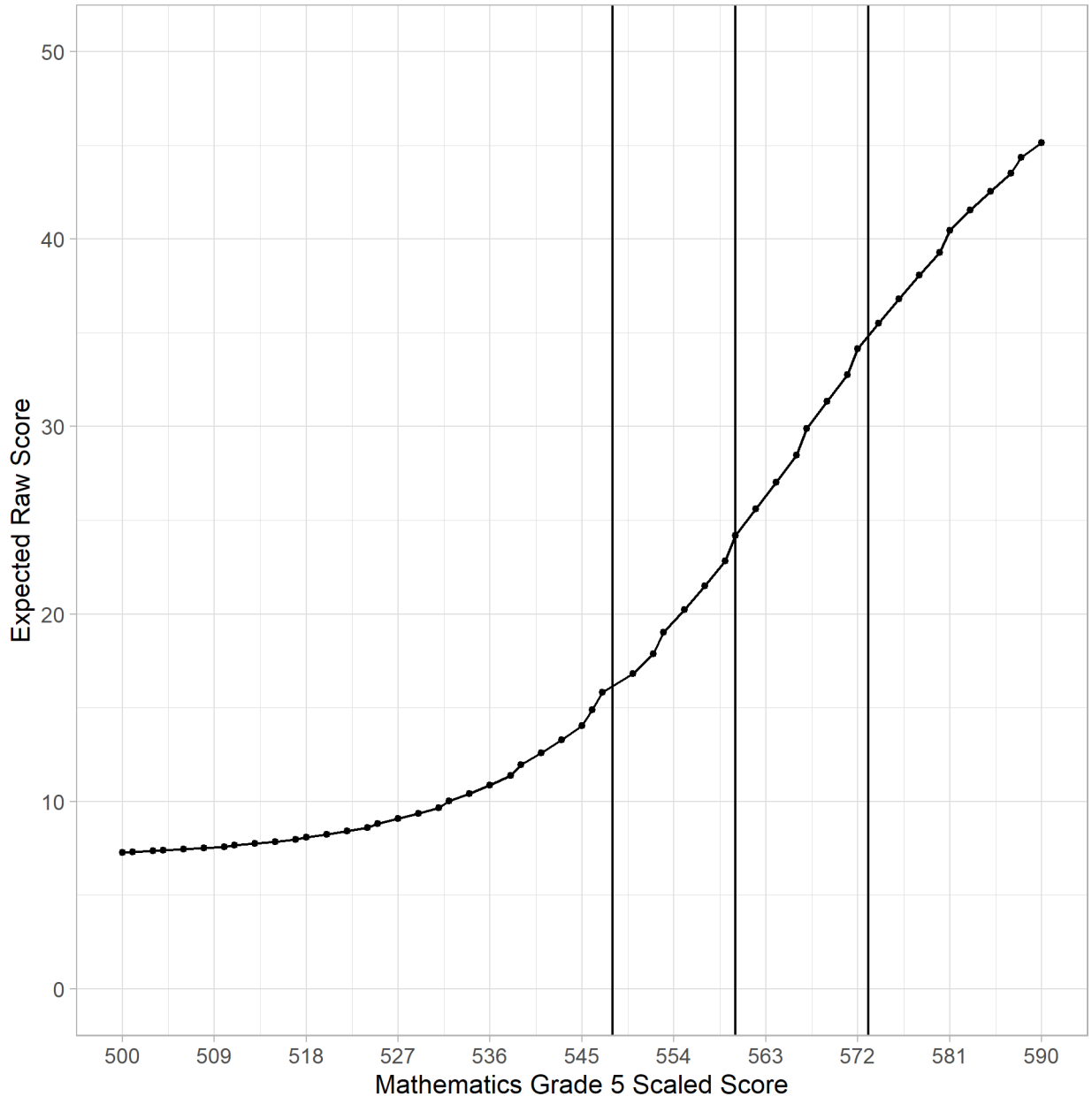


Figure K-10. Test Characteristic Curve for Mathematics—Grade 6

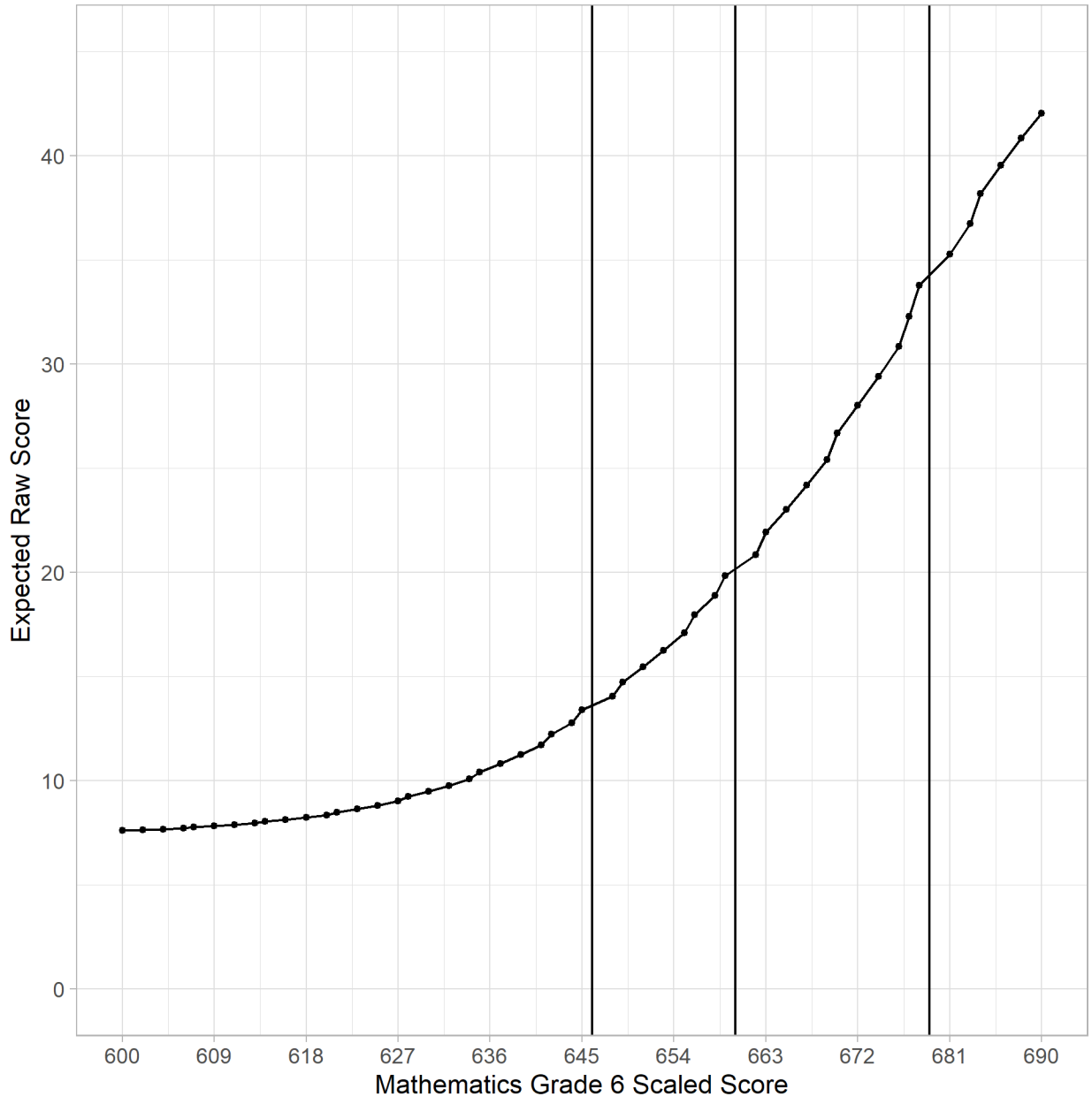


Figure K-11. Test Characteristic Curve for Mathematics—Grade 7

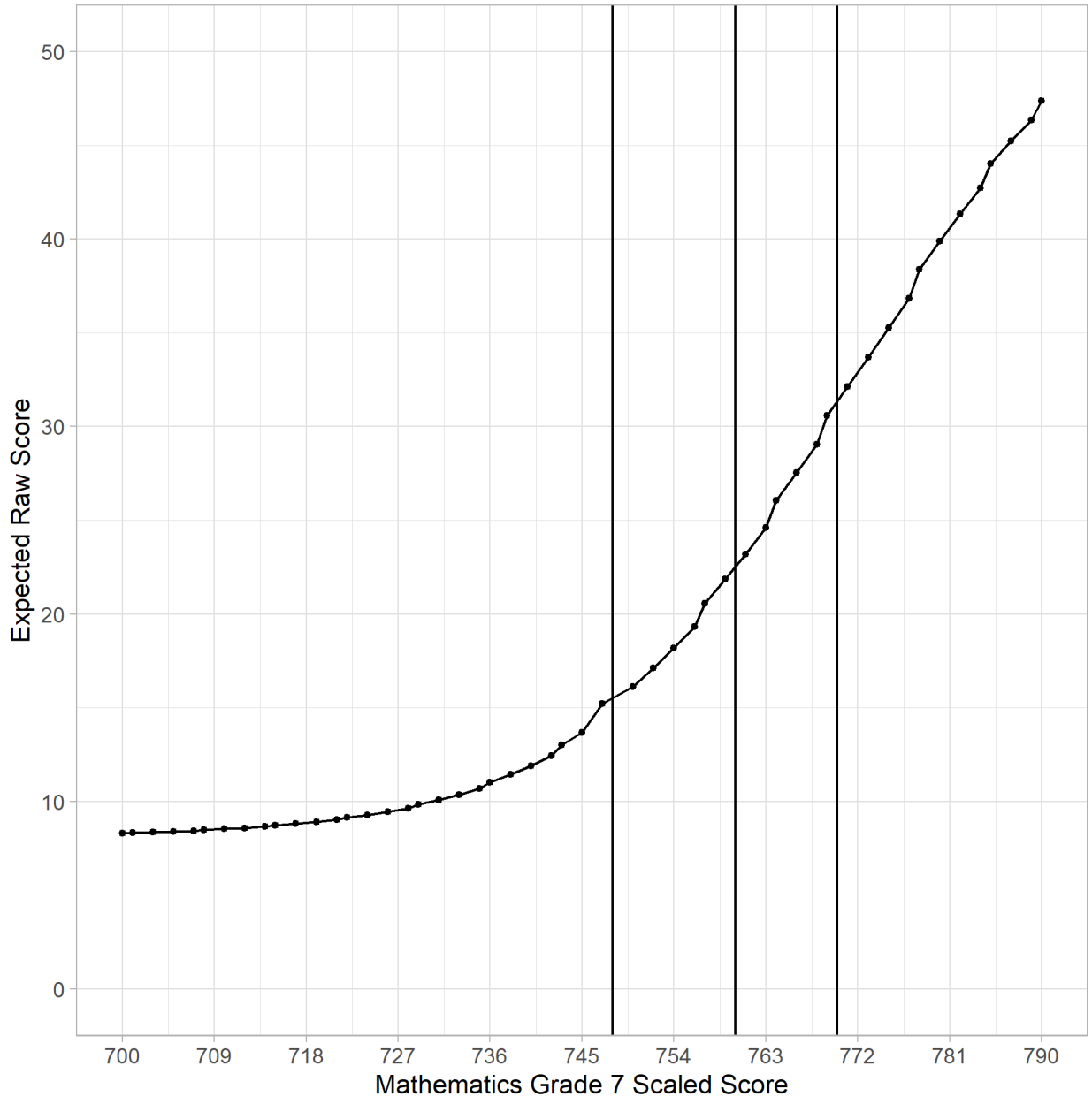


Figure K-12. Test Characteristic Curve for Mathematics—Grade 8

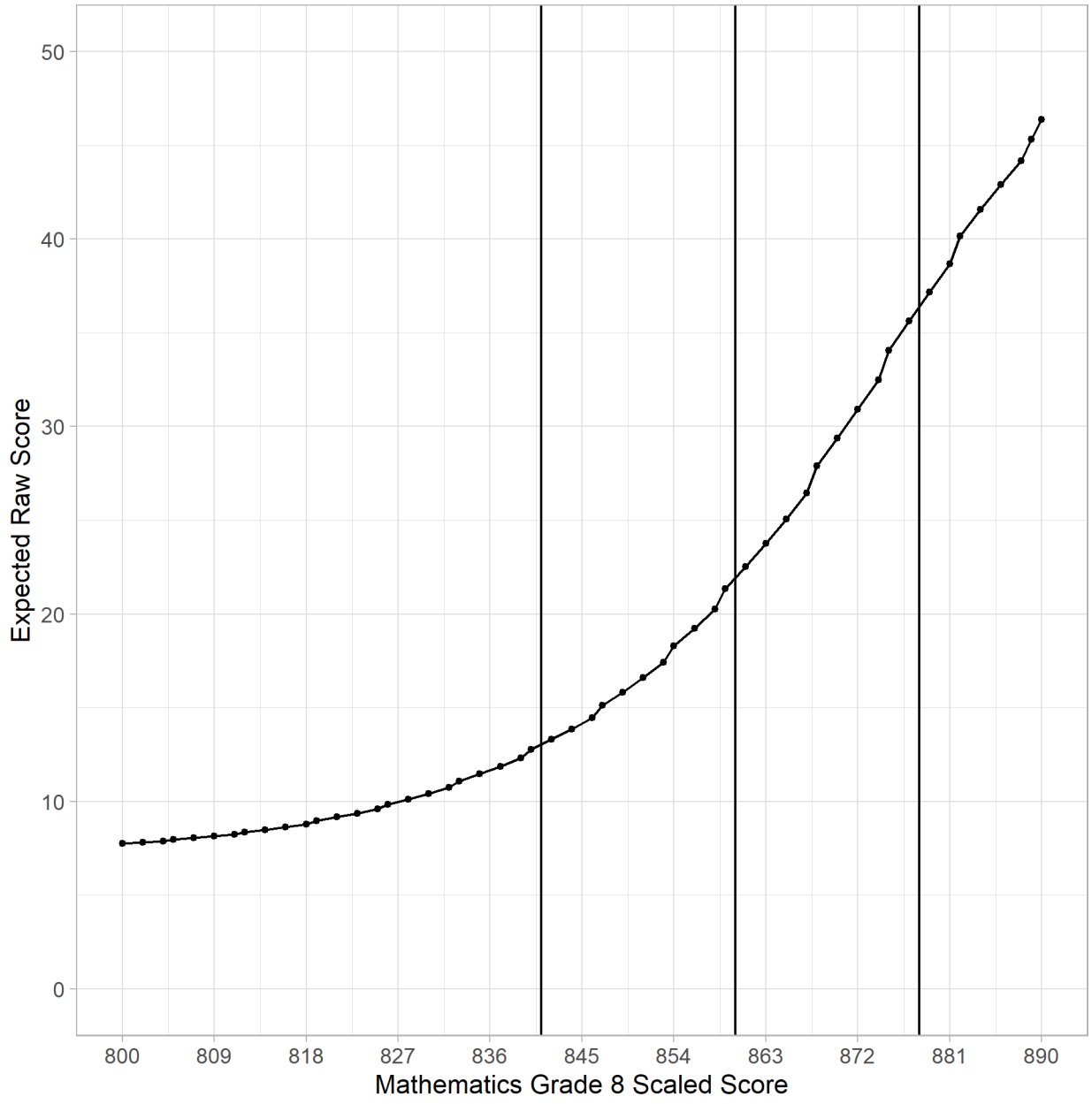


Figure K-13. Test Characteristic Curve for Science—Grade 5

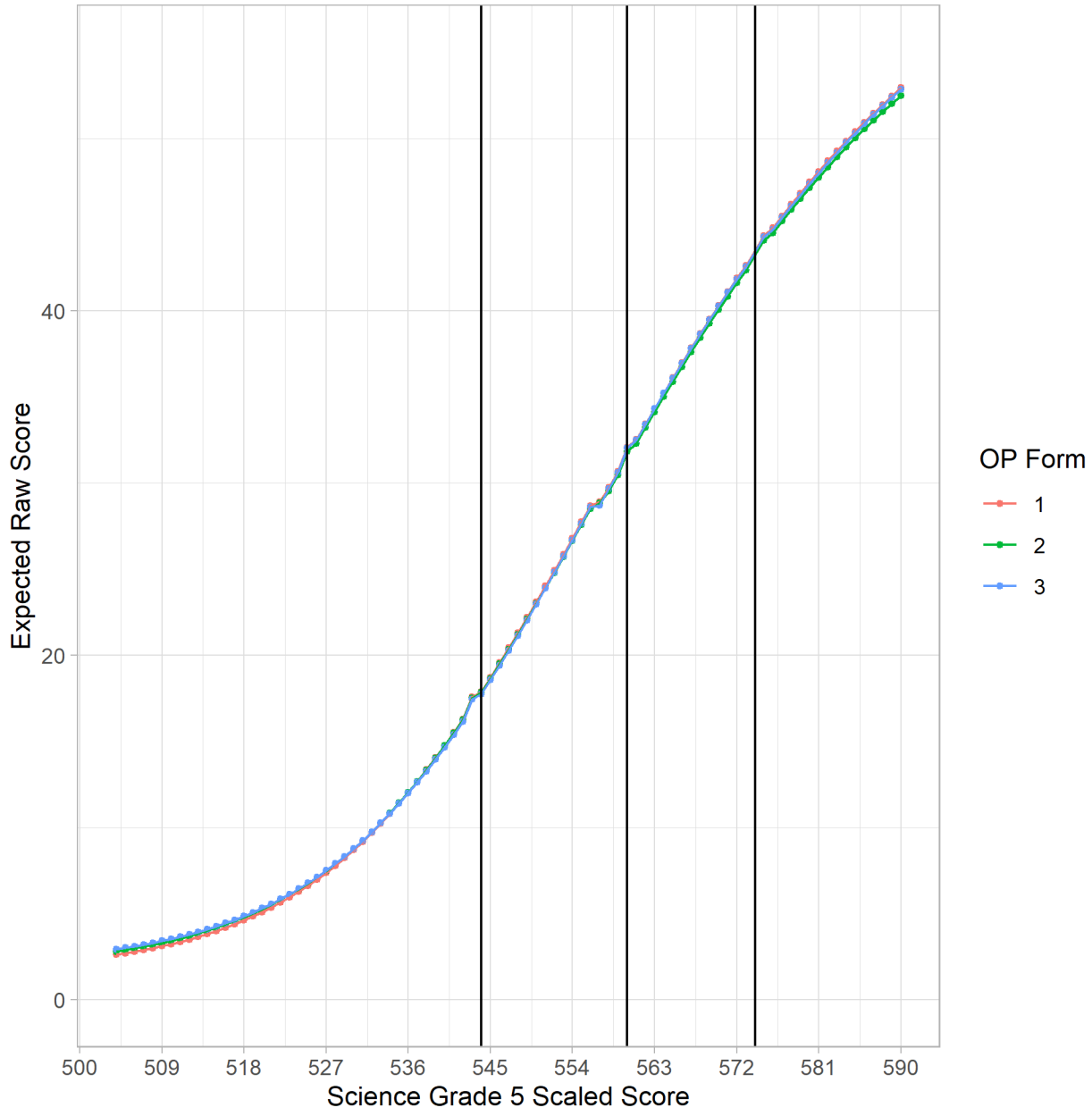


Figure K-14. Test Characteristic Curve for Science—Grade 8

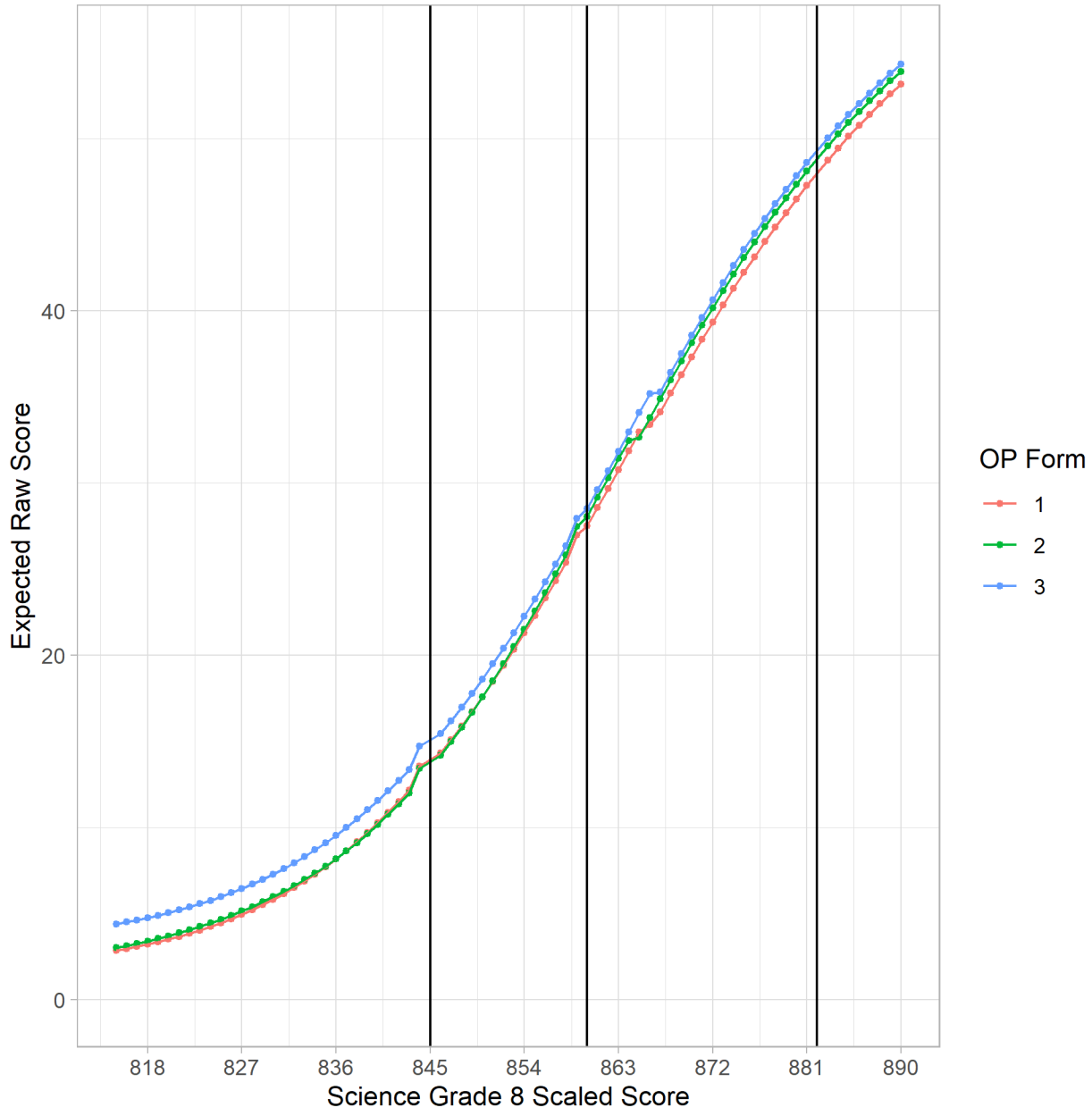


Figure K-15. Test Characteristic Curve for Science—Grade 11

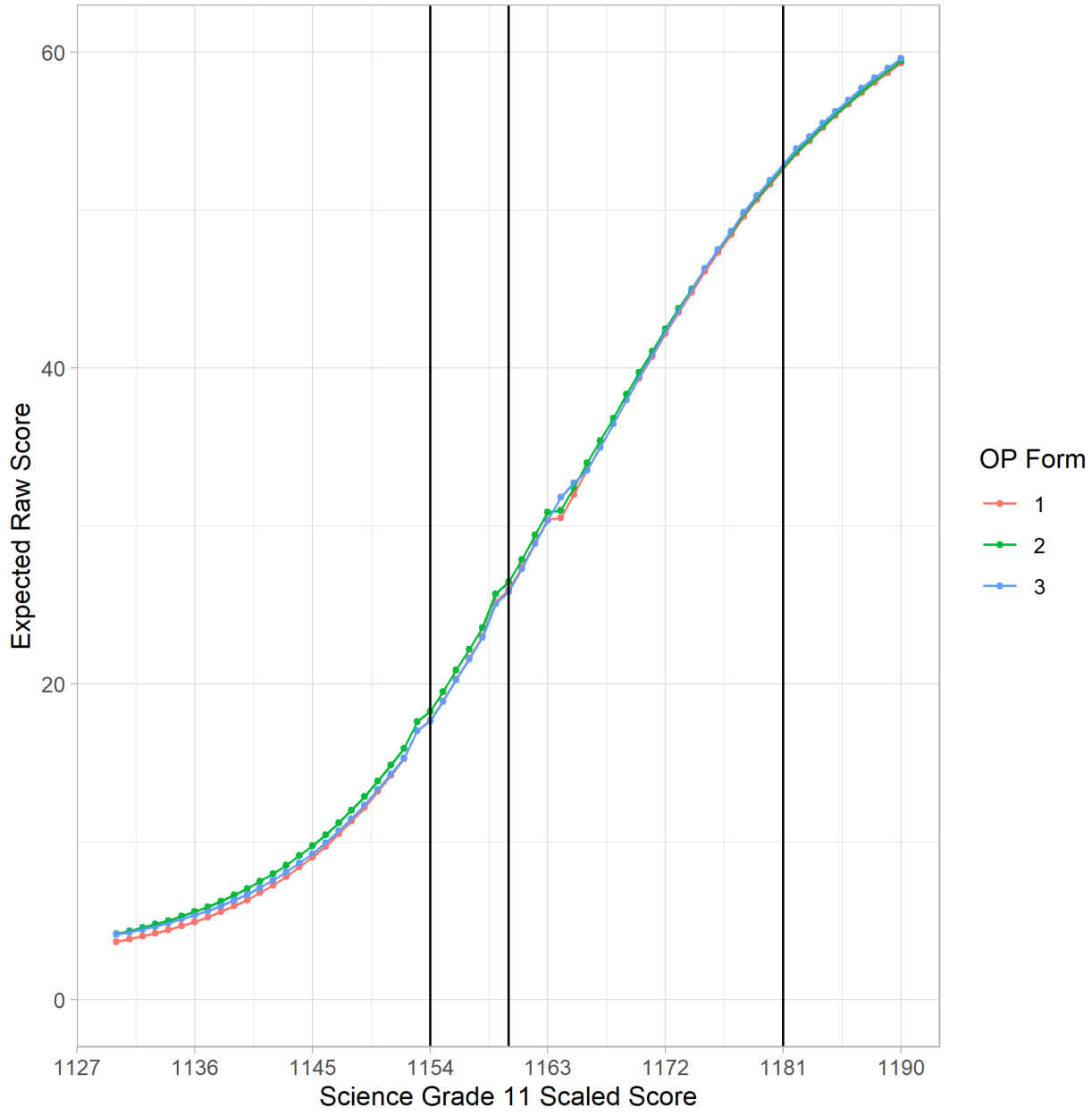


Figure K-16. Conditional Standard Errors of Measurement for ELA—Grade 3

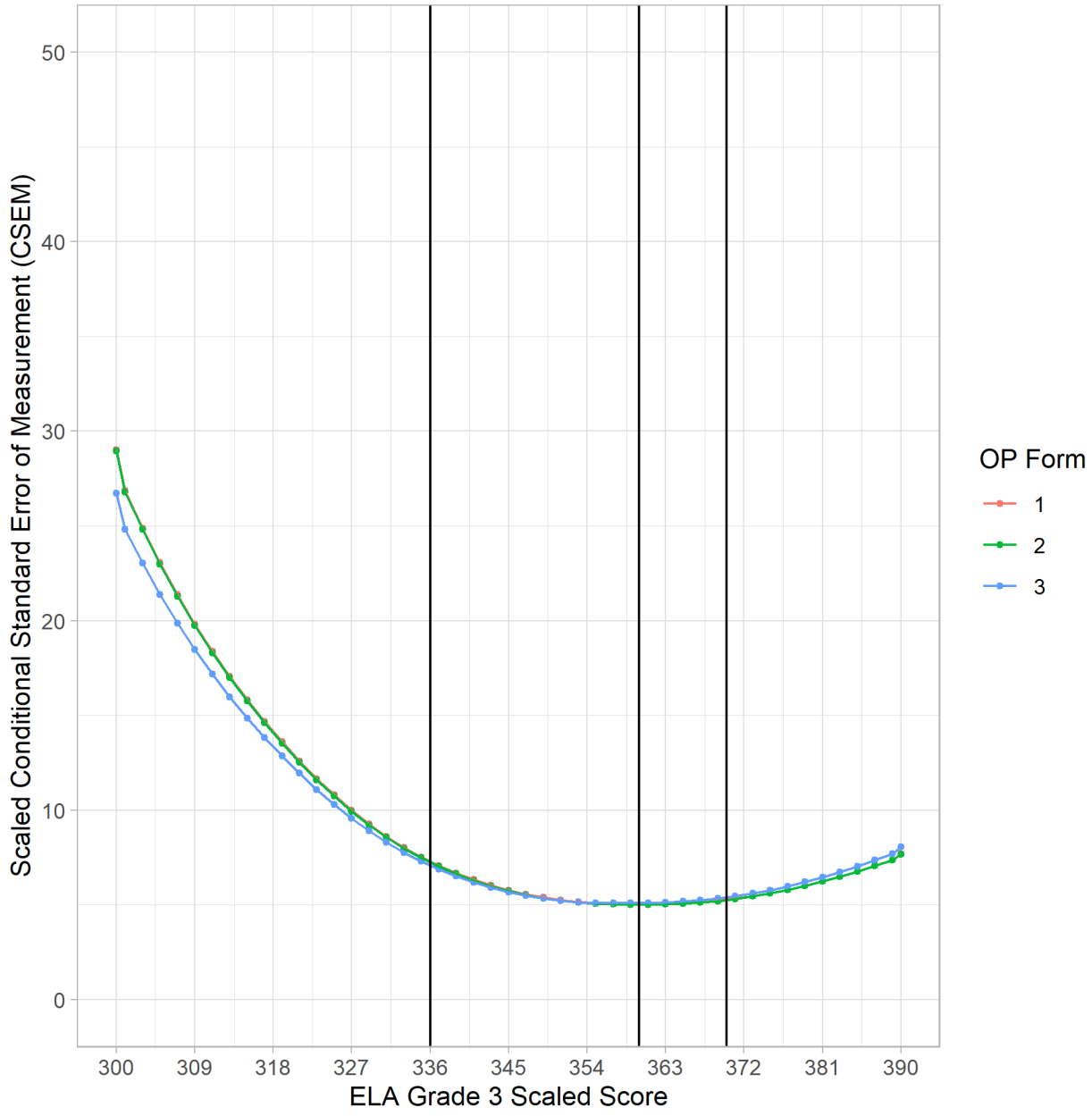


Figure K-17. Conditional Standard Errors of Measurement for ELA—Grade 4

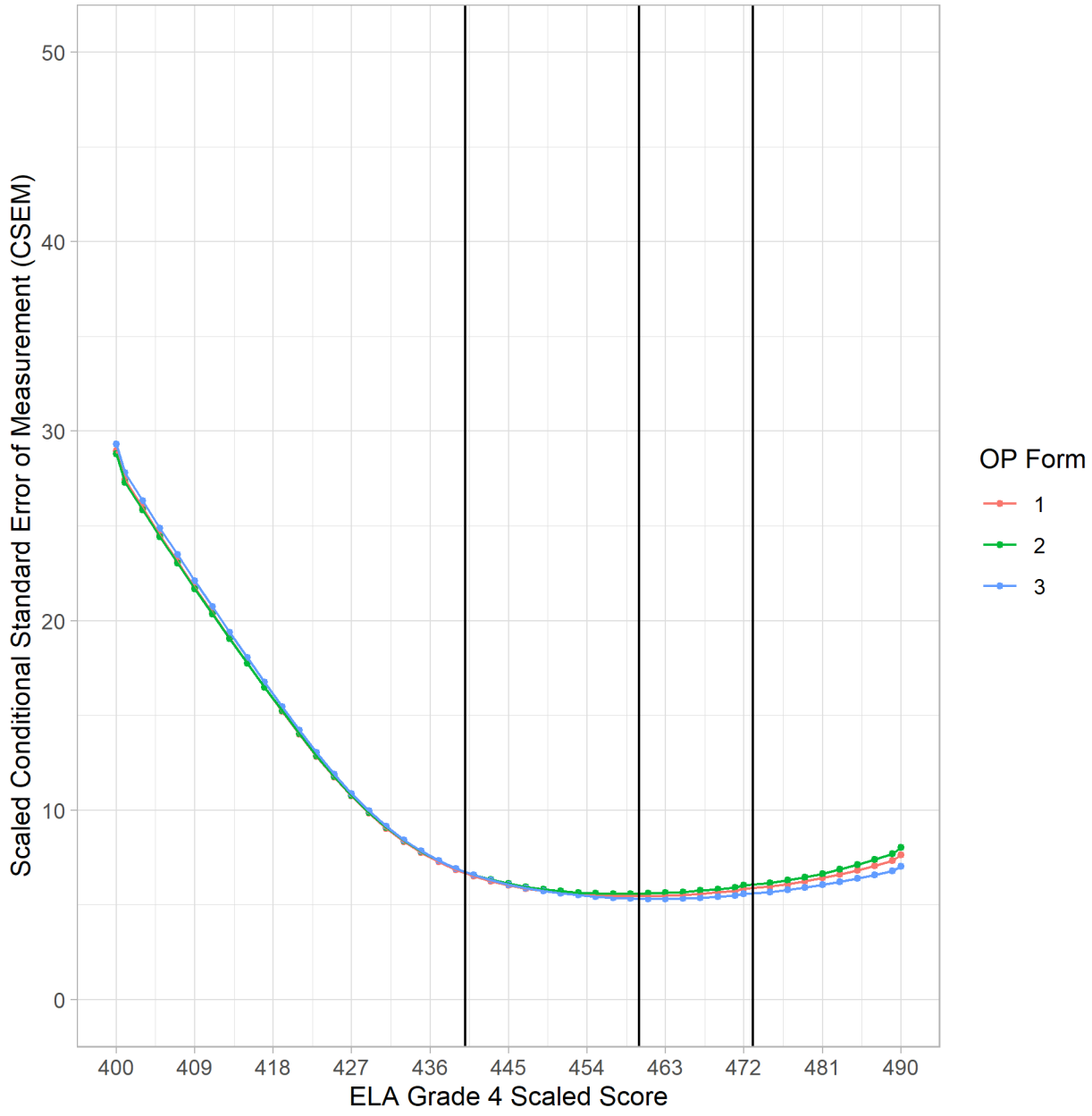


Figure K-18. Conditional Standard Errors of Measurement for ELA—Grade 5

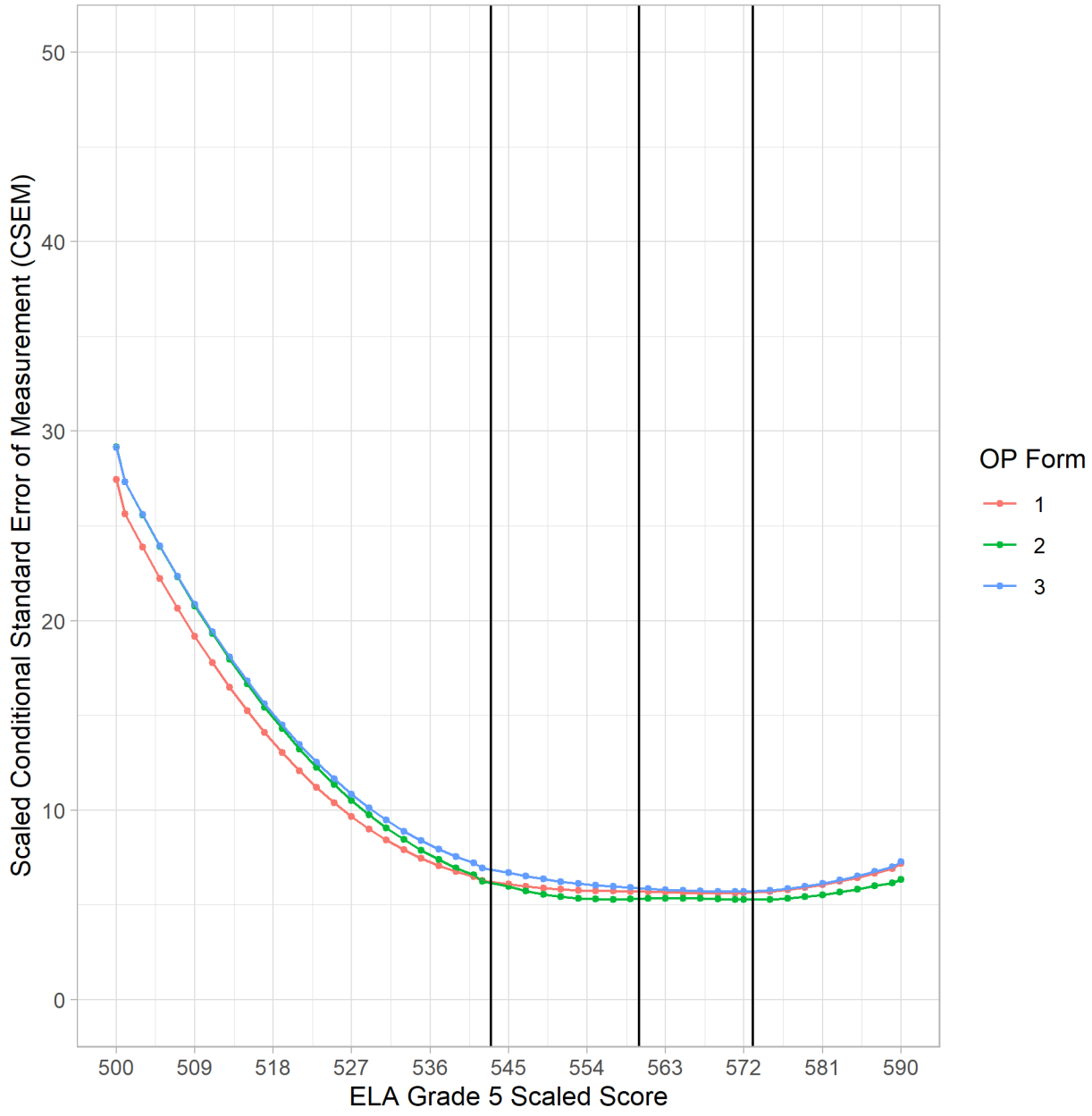


Figure K-19. Conditional Standard Errors of Measurement for ELA—Grade 6

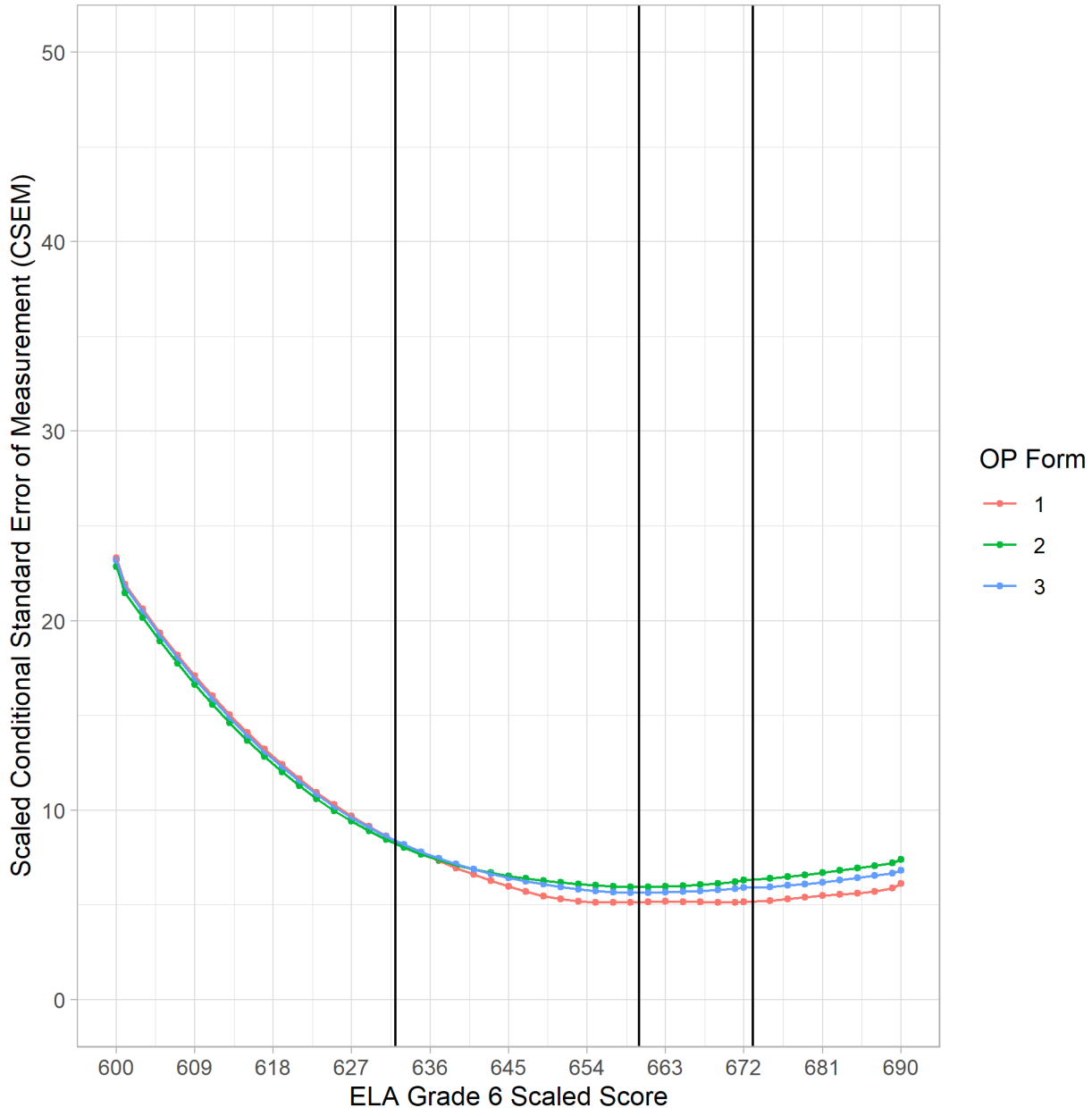


Figure K-20. Conditional Standard Errors of Measurement for ELA—Grade 7

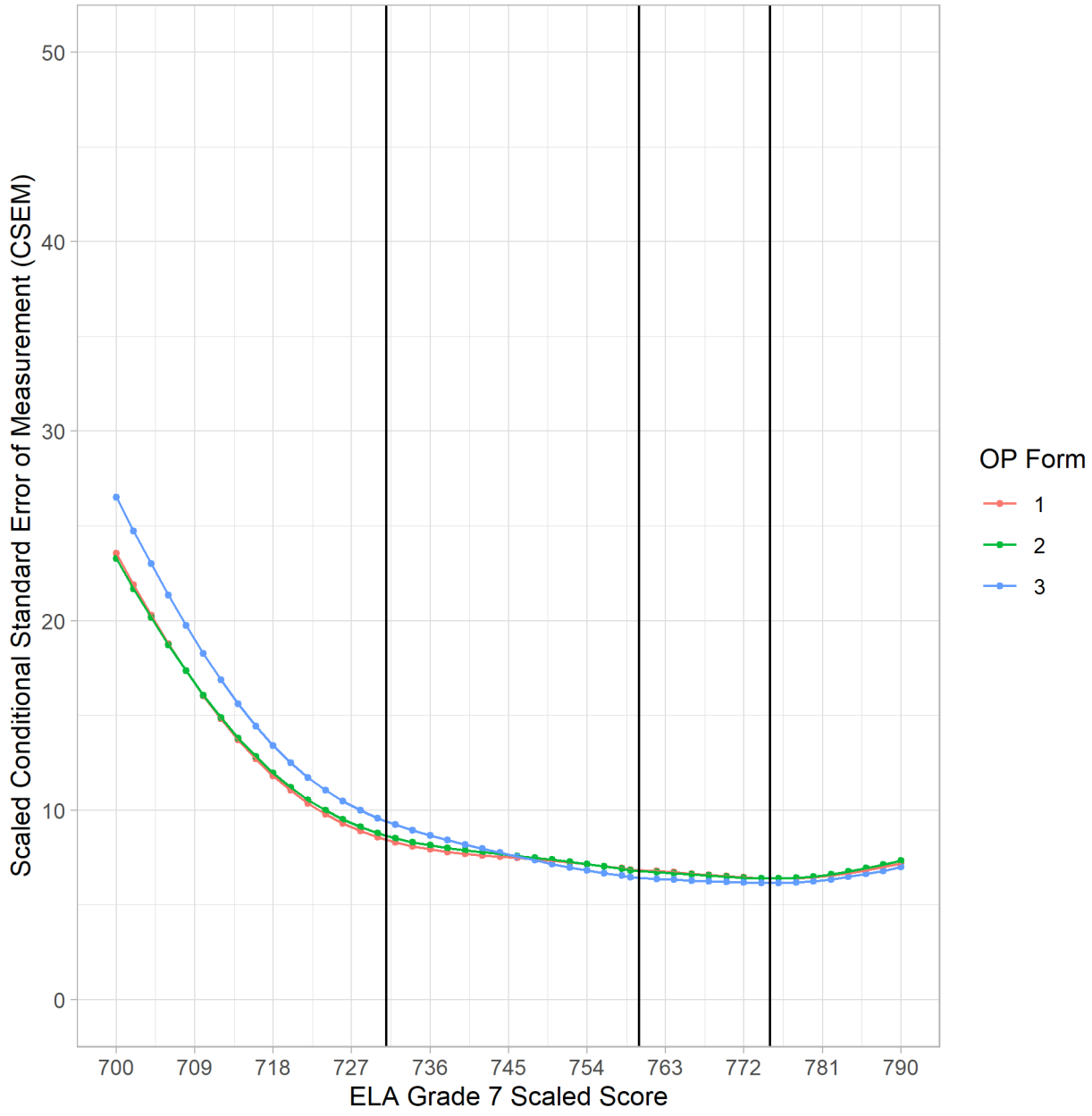


Figure K-21. Conditional Standard Errors of Measurement for ELA—Grade 8

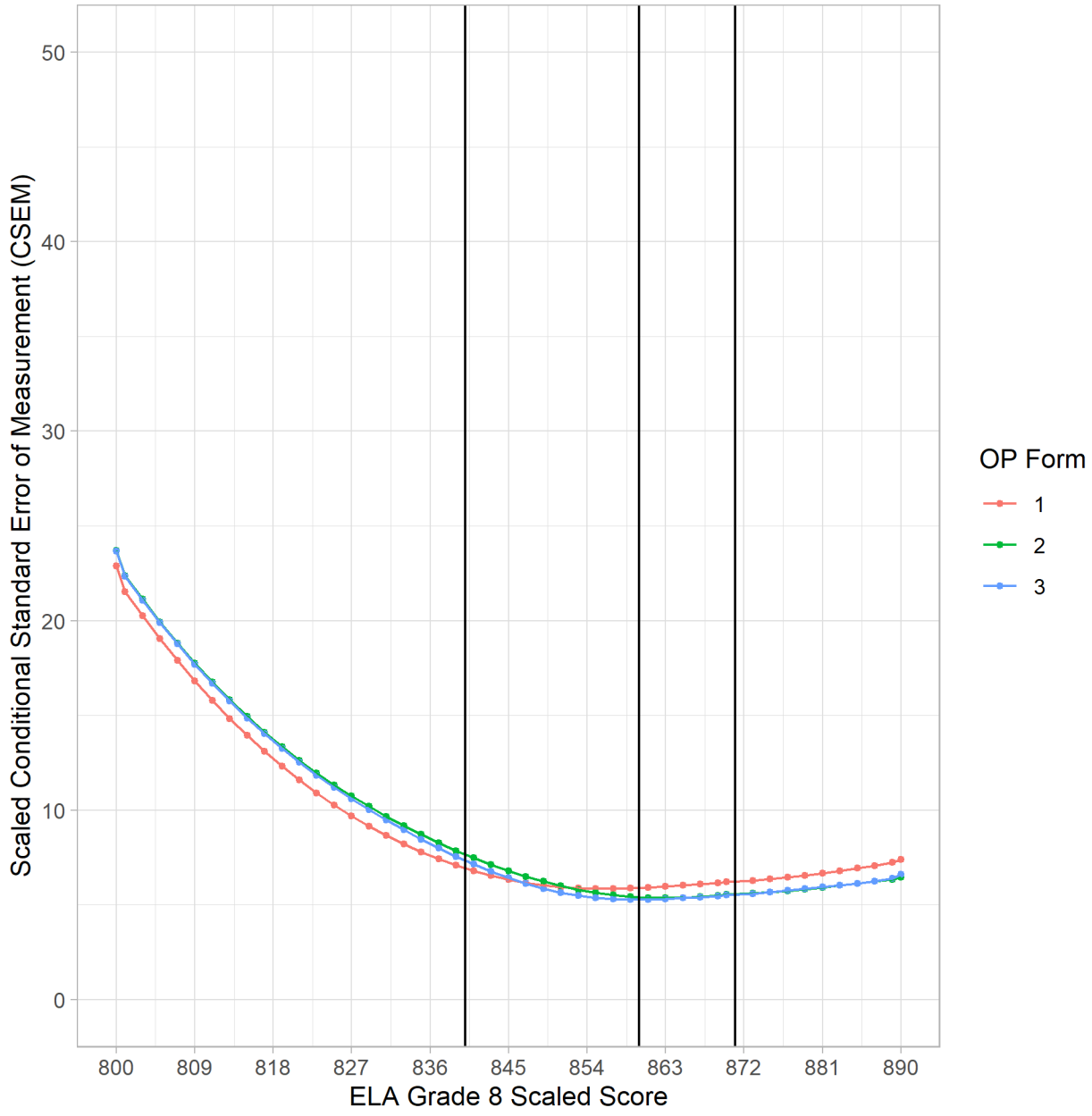


Figure K-22. Conditional Standard Errors of Measurement for Mathematics—Grade 3

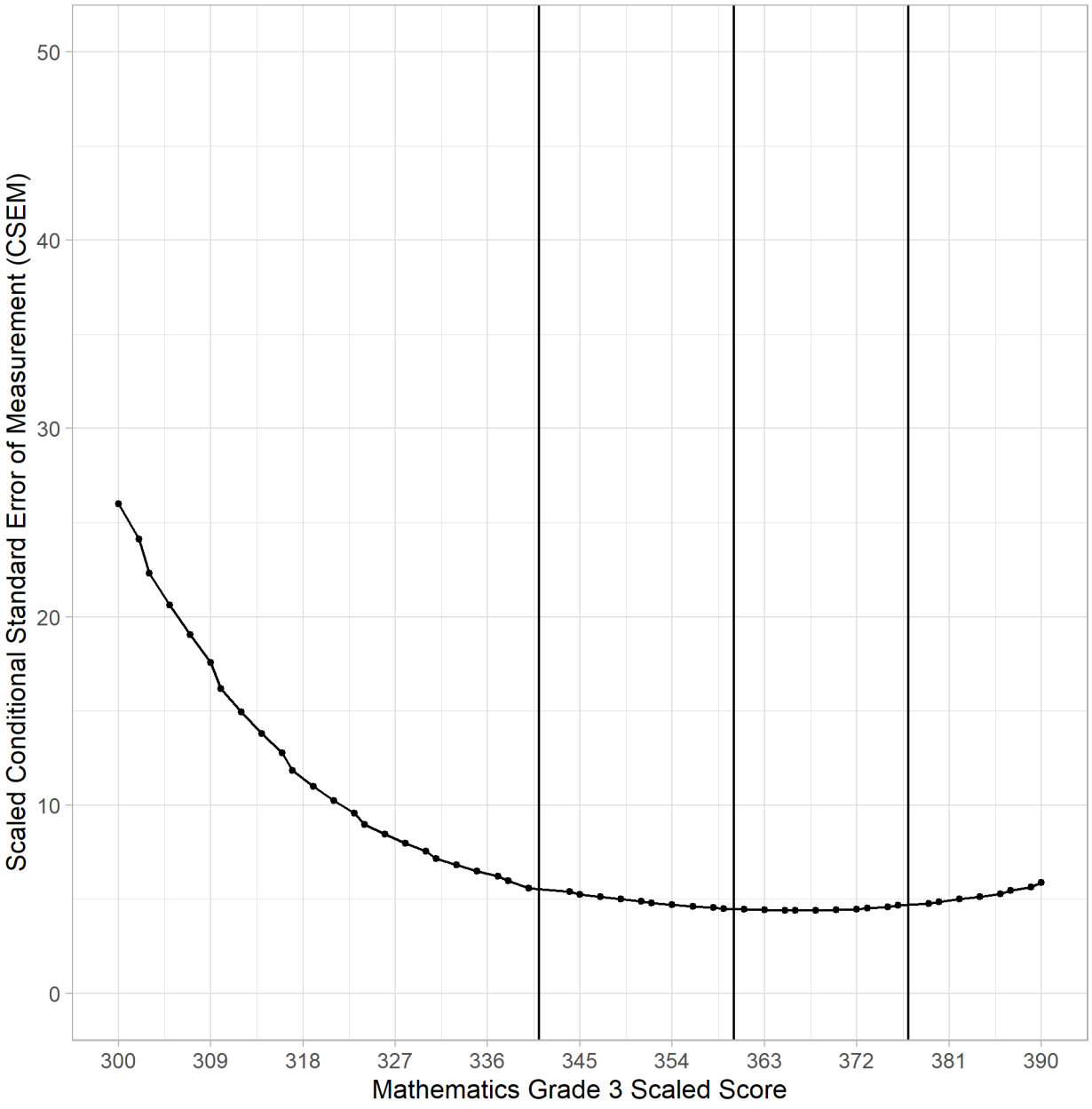


Figure K-23. Conditional Standard Errors of Measurement for Mathematics—Grade 4

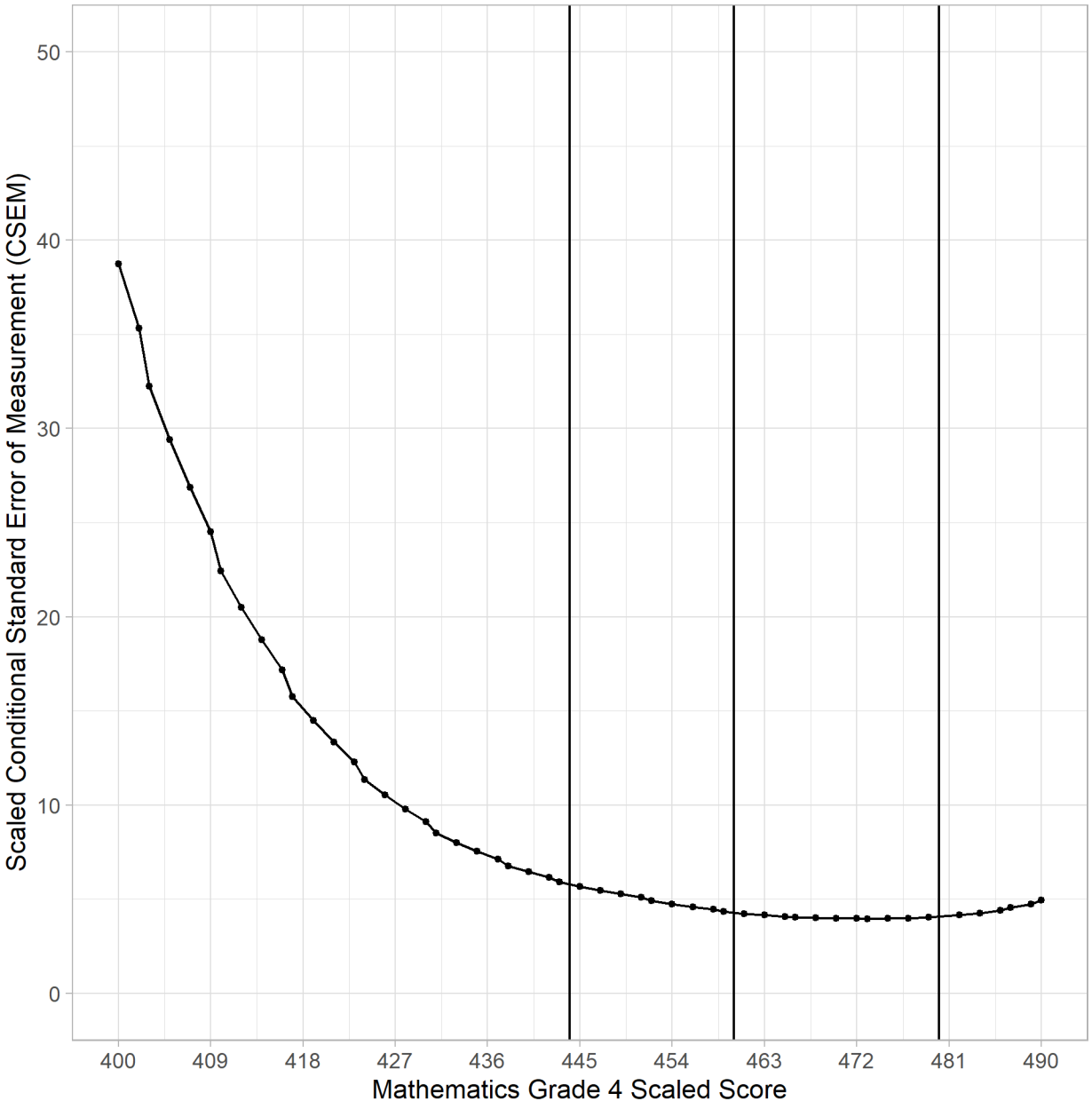


Figure K-24. Conditional Standard Errors of Measurement for Mathematics—Grade 5

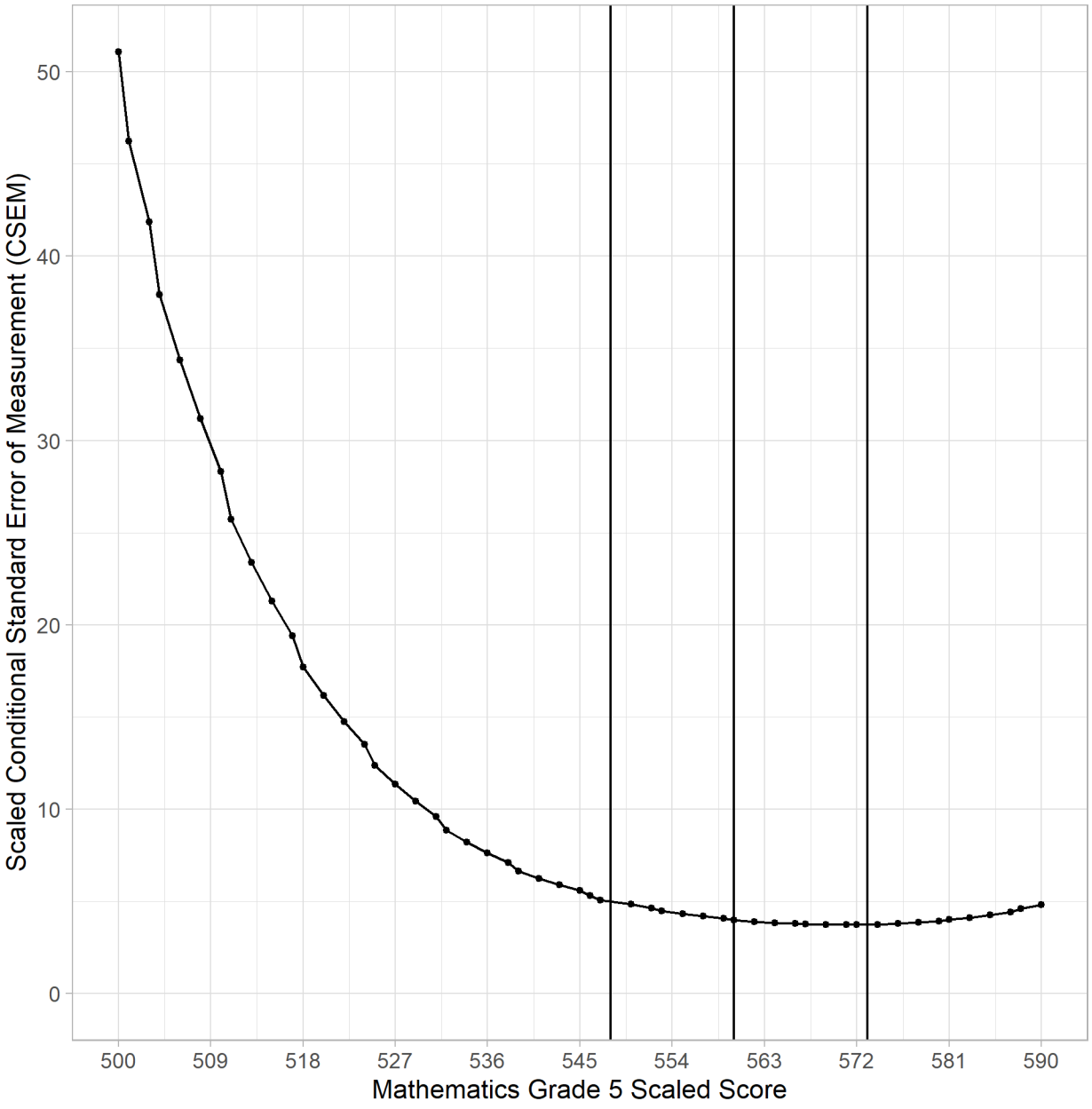


Figure K-25. Conditional Standard Errors of Measurement for Mathematics—Grade 6

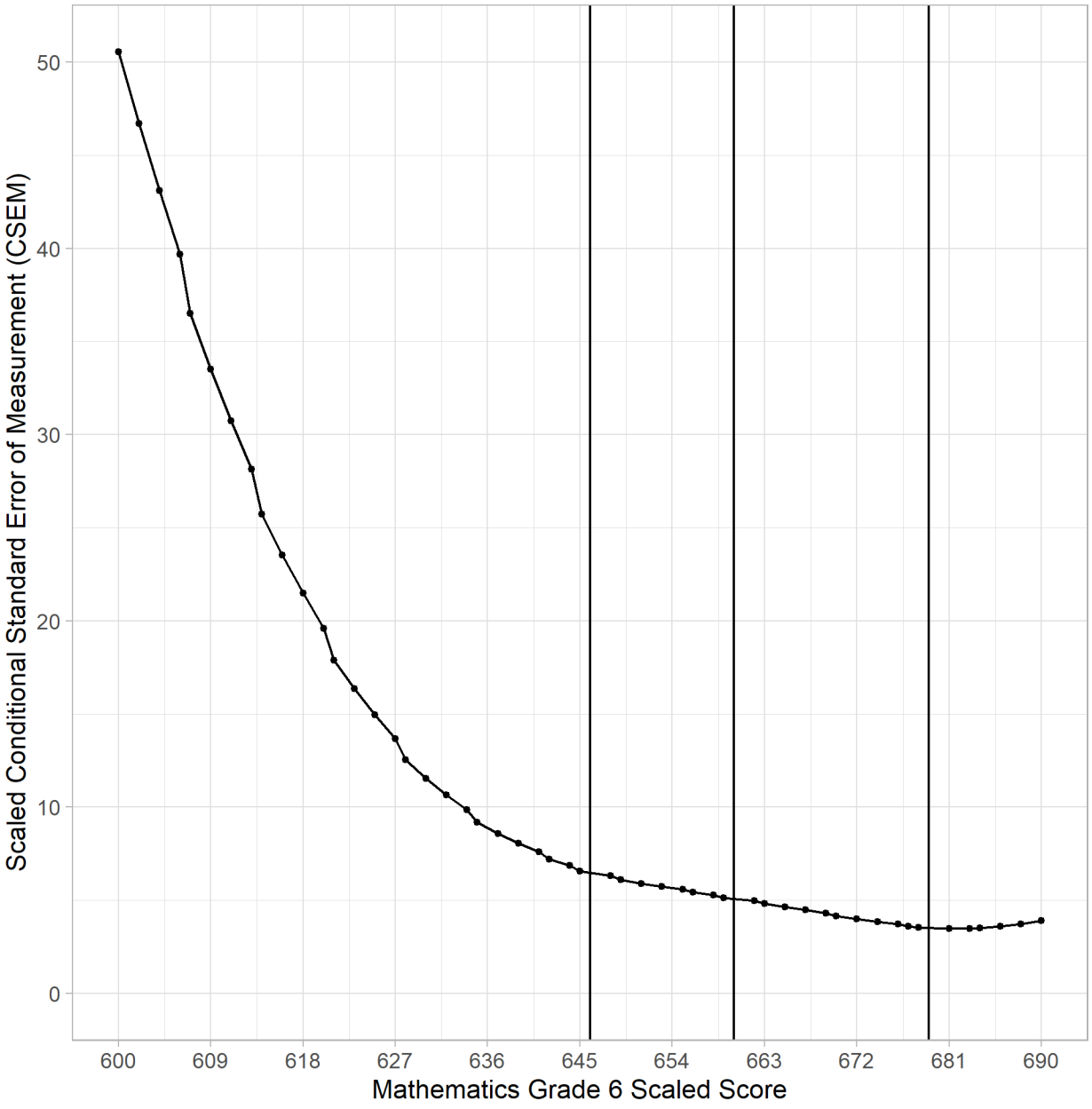


Figure K-26. Conditional Standard Errors of Measurement for Mathematics—Grade 7

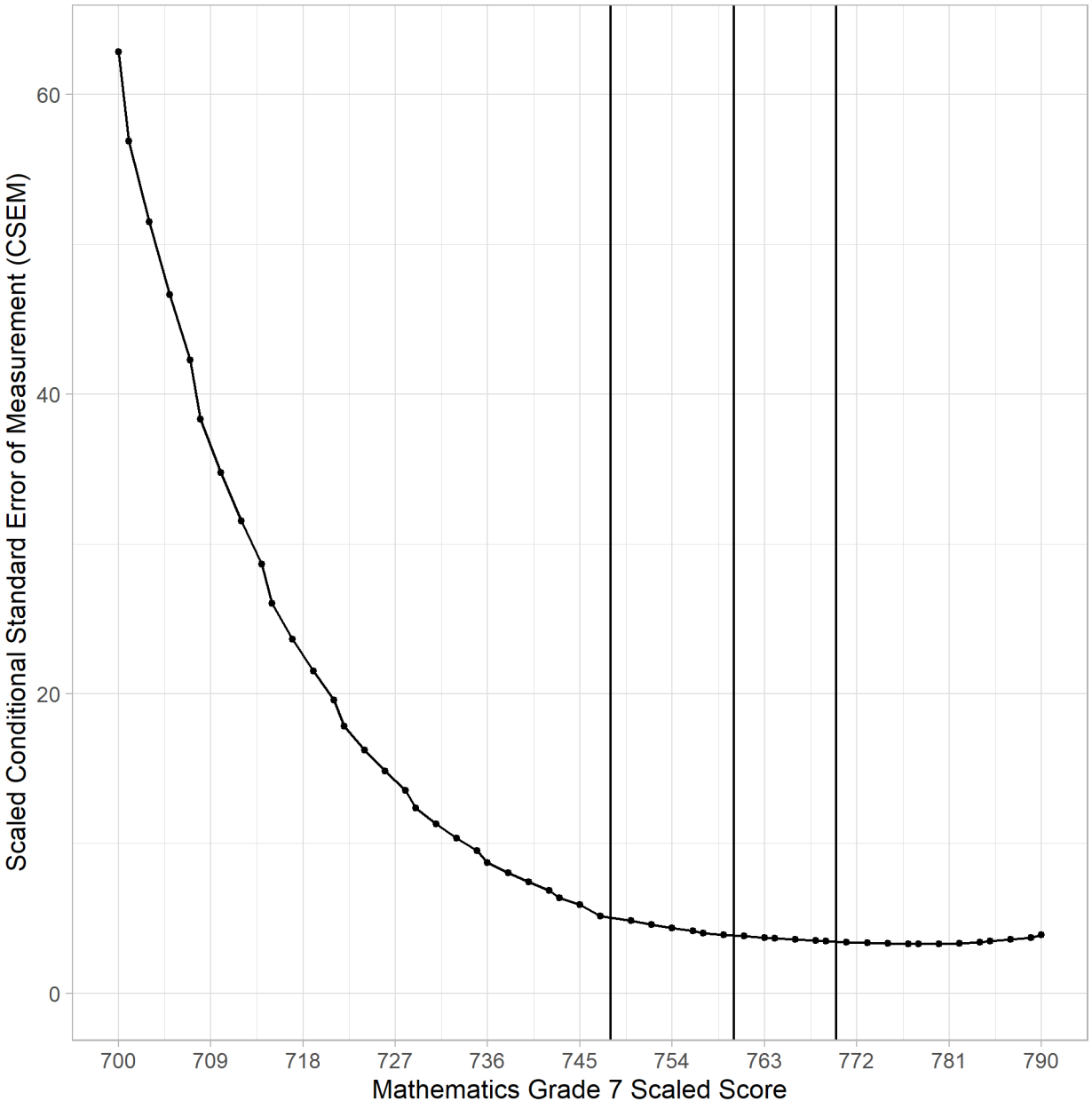


Figure K-27. Conditional Standard Errors of Measurement for Mathematics—Grade 8

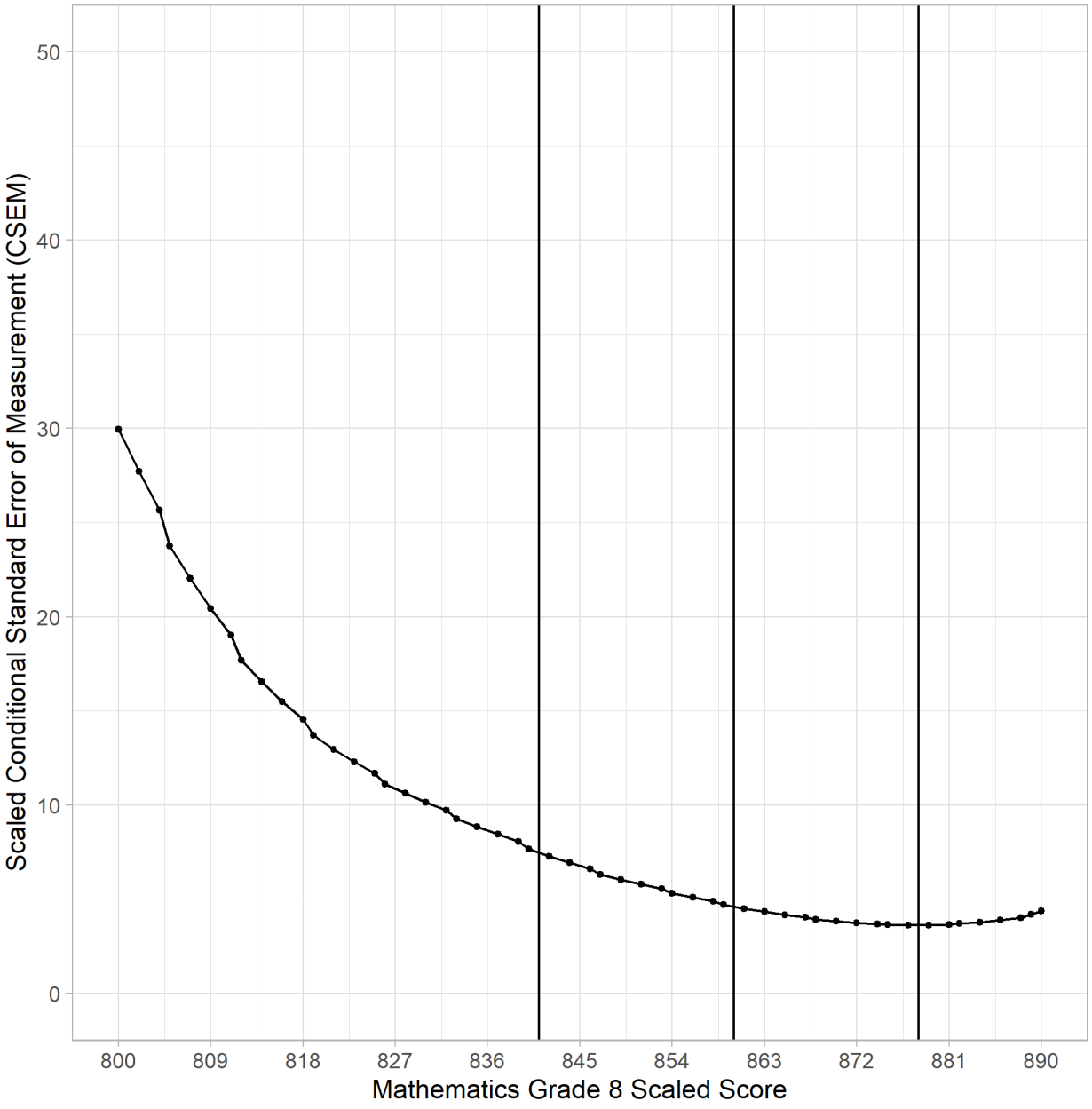


Figure K-28. Conditional Standard Errors of Measurement for Science—Grade 5

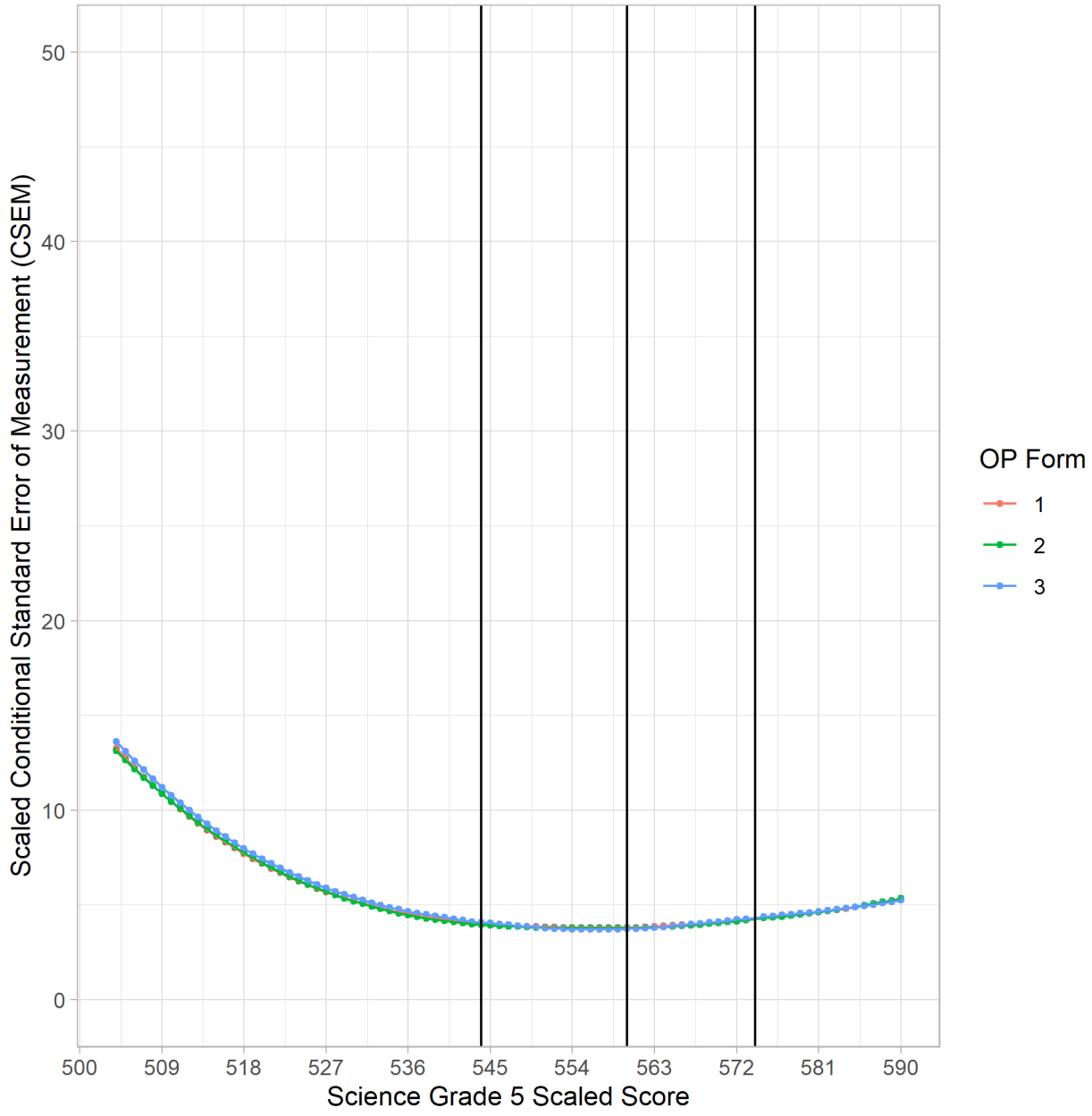


Figure K-29. Conditional Standard Errors of Measurement for Science—Grade 8

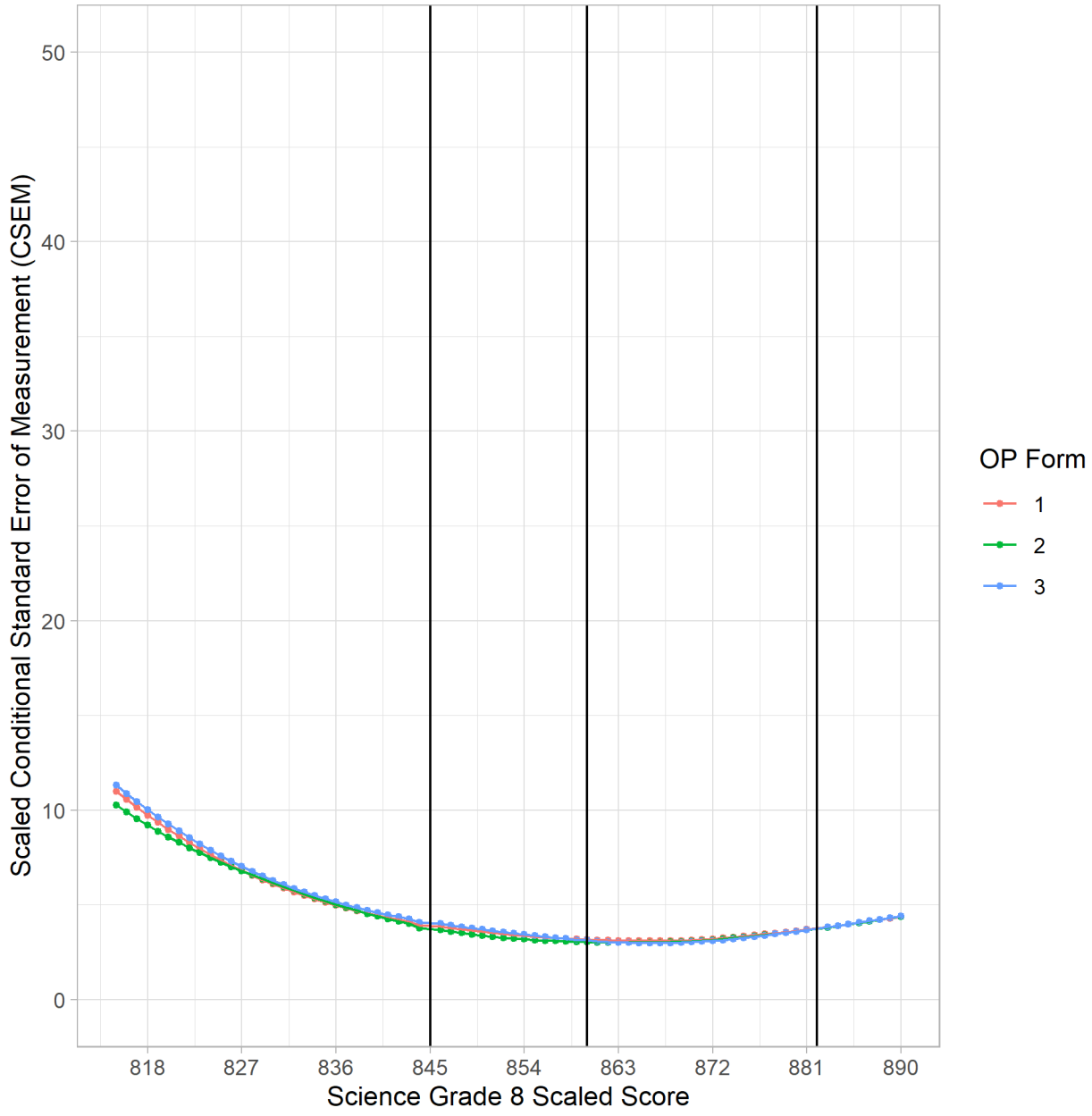
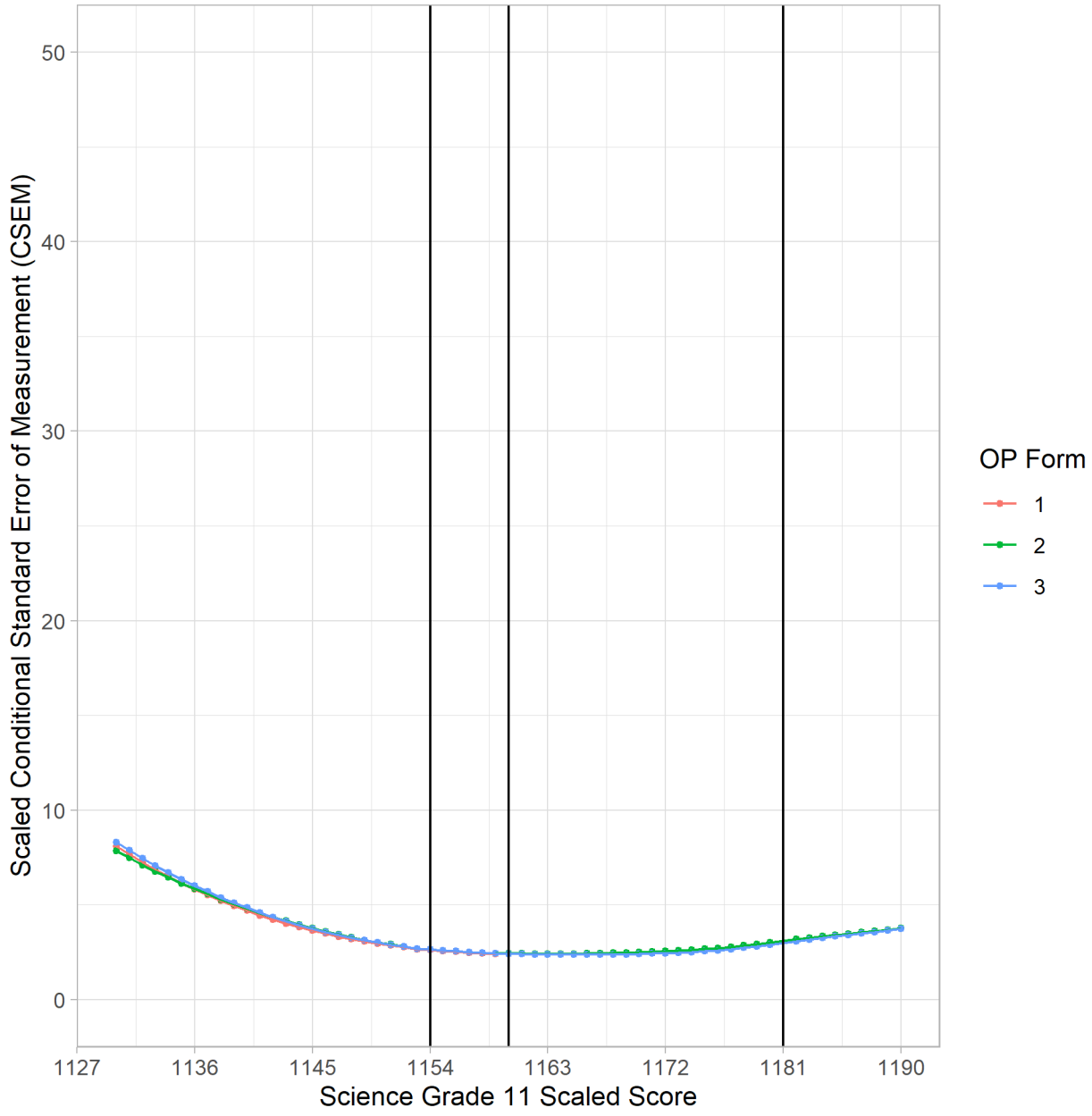


Figure K-30. Conditional Standard Errors of Measurement for Science—Grade 11



APPENDIX L
RAW TO SCALED SCORE LOOKUP TABLES

Table L-1. Raw to Scaled Score Look-up Table—ELA Grade 3 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	3.86412	77.28239
1	-3.79036	300	1	3.31799	66.35974
2	-3.58072	300	1	2.84166	56.83325
3	-3.37108	300	1	2.42779	48.55584
4	-3.16144	300	1	2.06987	41.39737
5	-2.95180	300	1	1.76209	35.24171
6	-2.74216	300	1	1.49913	29.98266
7	-1.93020	314	1	0.80935	16.18696
8	-1.53860	322	1	0.60102	12.02049
9	-1.27501	327	1	0.49034	9.80675
10	-1.07276	331	1	0.42240	8.44802
11	-0.90578	335	1	0.37747	7.54931
12	-0.76135	338	2	0.34570	6.91401
13	-0.63243	340	2	0.32198	6.43951
14	-0.51470	343	2	0.30364	6.07274
15	-0.40528	345	2	0.28929	5.78587
16	-0.30209	347	2	0.27815	5.56299
17	-0.20356	349	2	0.26967	5.39344
18	-0.10846	351	2	0.26344	5.26871
19	-0.01578	353	2	0.25906	5.18120
20	0.07530	354	2	0.25618	5.12369
21	0.16543	356	2	0.25447	5.08933
22	0.25516	358	2	0.25360	5.07201
23	0.34497	359	2	0.25335	5.06709
24	0.43528	362	3	0.25362	5.07230
25	0.52652	363	3	0.25441	5.08828
26	0.61914	365	3	0.25592	5.11846
27	0.71371	367	3	0.25841	5.16818
28	0.81089	369	3	0.26217	5.24336
29	0.91148	371	4	0.26747	5.34944
30	1.01641	373	4	0.27455	5.49093
31	1.12680	375	4	0.28359	5.67179
32	1.24400	378	4	0.29482	5.89648
33	1.36962	380	4	0.30858	6.17152
34	1.50575	383	4	0.32535	6.50705
35	1.65507	386	4	0.34591	6.91811
36	1.82121	389	4	0.37127	7.42548
37	2.00913	389	4	0.40279	8.05590
38	2.22585	389	4	0.44203	8.84060
39	2.48170	389	4	0.49057	9.81147
40	2.79330	389	4	0.55067	11.01339
41	3.19427	389	4	0.63404	12.68076
42	3.79553	389	4	0.82046	16.40918
43	4.00000	390	4	0.91485	18.29710

Table L-2. Raw to Scaled Score Look-up Table—ELA Grade 3 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	3.47194	69.43885
1	-3.80145	300	1	3.01330	60.26603
2	-3.60290	300	1	2.61048	52.20962
3	-3.40434	300	1	2.25759	45.15176
4	-3.20579	300	1	1.94943	38.98857
5	-3.00724	300	1	1.68142	33.62838
6	-2.80869	300	1	1.44945	28.98907
7	-1.99611	313	1	0.79657	15.93143
8	-1.60273	321	1	0.59837	11.96745
9	-1.33766	326	1	0.49193	9.83863
10	-1.13448	330	1	0.42516	8.50327
11	-0.96705	334	1	0.38035	7.60704
12	-0.82250	336	2	0.34854	6.97090
13	-0.69367	339	2	0.32479	6.49584
14	-0.57617	341	2	0.30638	6.12760
15	-0.46707	344	2	0.29185	5.83695
16	-0.36429	346	2	0.28039	5.60774
17	-0.26627	348	2	0.27150	5.43009
18	-0.17175	349	2	0.26484	5.29683
19	-0.07970	351	2	0.26009	5.20183
20	0.01073	353	2	0.25696	5.13927
21	0.10029	355	2	0.25517	5.10333
22	0.18957	357	2	0.25441	5.08826
23	0.27913	359	2	0.25444	5.08884
24	0.36946	360	3	0.25506	5.10124
25	0.46101	362	3	0.25620	5.12402
26	0.55431	364	3	0.25794	5.15887
27	0.64992	366	3	0.26053	5.21052
28	0.74857	368	3	0.26429	5.28583
29	0.85114	369	3	0.26962	5.39243
30	0.95874	372	4	0.27688	5.53755
31	1.07273	374	4	0.28639	5.72785
32	1.19480	377	4	0.29852	5.97042
33	1.32711	379	4	0.31375	6.27494
34	1.47248	382	4	0.33284	6.65686
35	1.63480	386	4	0.35706	7.14113
36	1.81956	389	4	0.38831	7.76622
37	2.03483	389	4	0.42944	8.58876
38	2.29274	389	4	0.48448	9.68955
39	2.61217	389	4	0.55871	11.17426
40	3.02277	389	4	0.65585	13.11691
41	3.57010	389	4	0.77104	15.42080
42	4.00000	390	4	0.85924	17.18488
43	4.00000	390	4	0.85924	17.18488

Table L-3. Raw to Scaled Score Look-up Table—ELA Grade 3 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	3.86200	77.23998
1	-3.79028	300	1	3.31498	66.29959
2	-3.58056	300	1	2.83800	56.76001
3	-3.37083	300	1	2.42368	48.47368
4	-3.16111	300	1	2.06549	41.30987
5	-2.95139	300	1	1.75759	35.15185
6	-2.74167	300	1	1.49463	29.89266
7	-1.93037	314	1	0.80593	16.11854
8	-1.53927	322	1	0.59820	11.96398
9	-1.27610	327	1	0.48795	9.75902
10	-1.07425	331	1	0.42028	8.40558
11	-0.90765	335	1	0.37549	7.50973
12	-0.76361	338	2	0.34380	6.87603
13	-0.63509	340	2	0.32012	6.40243
14	-0.51777	343	2	0.30180	6.03605
15	-0.40877	345	2	0.28745	5.74905
16	-0.30603	347	2	0.27627	5.52550
17	-0.20796	349	2	0.26774	5.35474
18	-0.11333	351	2	0.26142	5.22839
19	-0.02113	353	2	0.25695	5.13903
20	0.06947	354	2	0.25399	5.07972
21	0.15913	356	2	0.25220	5.04392
22	0.24845	358	2	0.25129	5.02581
23	0.33791	359	2	0.25105	5.02096
24	0.42798	361	3	0.25135	5.02710
25	0.51912	363	3	0.25224	5.04478
26	0.61186	365	3	0.25387	5.07734
27	0.70682	367	3	0.25651	5.13017
28	0.80473	369	3	0.26048	5.20955
29	0.90653	371	4	0.26608	5.32154
30	1.01331	373	4	0.27357	5.47145
31	1.12640	375	4	0.28320	5.66407
32	1.24741	378	4	0.29523	5.90461
33	1.37834	380	4	0.31001	6.20012
34	1.52171	383	4	0.32805	6.56094
35	1.68075	387	4	0.35008	7.00154
36	1.85969	389	4	0.37704	7.54088
37	2.06412	389	4	0.41015	8.20293
38	2.30159	389	4	0.45090	9.01803
39	2.58280	389	4	0.50070	10.01406
40	2.92355	389	4	0.55884	11.17690
41	3.35168	389	4	0.62835	12.56710
42	3.96185	389	4	0.78180	15.63598
43	4.00000	390	4	0.79596	15.91917

Table L-4. Raw to Scaled Score Look-up Table—ELA Grade 4 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	400	1	2.46862	49.37241
1	-3.86031	400	1	2.31628	46.32563
2	-3.72062	400	1	2.17144	43.42871
3	-3.58093	400	1	2.03363	40.67253
4	-3.44124	400	1	1.90240	38.04809
5	-3.30155	400	1	1.77732	35.54647
6	-3.16187	400	1	1.65795	33.15891
7	-2.27859	412	1	1.00576	20.11520
8	-1.83229	420	1	0.71963	14.39255
9	-1.53789	426	1	0.55535	11.10706
10	-1.31715	431	1	0.45816	9.16319
11	-1.13792	434	1	0.39848	7.96956
12	-0.98435	437	1	0.35997	7.19948
13	-0.84773	439	1	0.33395	6.67903
14	-0.72283	443	2	0.31569	6.31387
15	-0.60633	445	2	0.30254	6.05071
16	-0.49599	447	2	0.29290	5.85797
17	-0.39019	449	2	0.28580	5.71602
18	-0.28773	451	2	0.28061	5.61226
19	-0.18769	453	2	0.27692	5.53845
20	-0.08929	455	2	0.27446	5.48928
21	0.00812	457	2	0.27307	5.46139
22	0.10514	459	2	0.27263	5.45266
23	0.20233	461	3	0.27309	5.46176
24	0.30024	463	3	0.27439	5.48781
25	0.39940	465	3	0.27651	5.53026
26	0.50040	467	3	0.27945	5.58894
27	0.60386	469	3	0.28321	5.66425
28	0.71045	471	3	0.28788	5.75757
29	0.82098	474	4	0.29358	5.87165
30	0.93640	476	4	0.30056	6.01117
31	1.05793	478	4	0.30917	6.18332
32	1.18711	481	4	0.31992	6.39834
33	1.32594	484	4	0.33352	6.67037
34	1.47715	487	4	0.35092	7.01841
35	1.64441	489	4	0.37337	7.46745
36	1.83276	489	4	0.40241	8.04814
37	2.04918	489	4	0.43949	8.78989
38	2.30298	489	4	0.48482	9.69640
39	2.60621	489	4	0.53612	10.72241
40	2.97737	489	4	0.59656	11.93116
41	3.46741	489	4	0.70932	14.18634
42	4.00000	490	4	0.93266	18.65328
43	4.00000	490	4	0.93266	18.65328

Table L-5. Raw to Scaled Score Look-up Table—ELA Grade 4 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	400	1	2.45219	49.04375
1	-3.86141	400	1	2.30172	46.03439
2	-3.72283	400	1	2.15872	43.17447
3	-3.58424	400	1	2.02276	40.45510
4	-3.44566	400	1	1.89338	37.86751
5	-3.30707	400	1	1.77015	35.40296
6	-3.16849	400	1	1.65264	33.05281
7	-2.28549	411	1	1.00821	20.16428
8	-1.83826	420	1	0.72458	14.49158
9	-1.54280	426	1	0.56015	11.20307
10	-1.32107	431	1	0.46230	9.24590
11	-1.14091	434	1	0.40213	8.04261
12	-0.98642	437	1	0.36339	7.26778
13	-0.84882	439	1	0.33732	6.74644
14	-0.72286	443	2	0.31916	6.38321
15	-0.60518	445	2	0.30621	6.12412
16	-0.49350	447	2	0.29686	5.93712
17	-0.38617	449	2	0.29011	5.80229
18	-0.28198	451	2	0.28534	5.70678
19	-0.17995	453	2	0.28211	5.64220
20	-0.07930	456	2	0.28015	5.60309
21	0.02066	458	2	0.27930	5.58599
22	0.12057	459	2	0.27944	5.58870
23	0.22102	462	3	0.28049	5.60981
24	0.32258	464	3	0.28242	5.64835
25	0.42587	466	3	0.28519	5.70377
26	0.53149	468	3	0.28881	5.77616
27	0.64016	470	3	0.29333	5.86661
28	0.75266	472	3	0.29890	5.97794
29	0.86997	474	4	0.30577	6.11542
30	0.99332	477	4	0.31438	6.28757
31	1.12433	480	4	0.32535	6.50709
32	1.26514	482	4	0.33959	6.79180
33	1.41871	485	4	0.35829	7.16587
34	1.58909	489	4	0.38306	7.66117
35	1.78188	489	4	0.41593	8.31852
36	2.00488	489	4	0.45940	9.18790
37	2.26896	489	4	0.51634	10.32684
38	2.58949	489	4	0.58945	11.78905
39	2.98738	489	4	0.67692	13.53836
40	3.48578	489	4	0.76089	15.21788
41	4.00000	490	4	0.82603	16.52070
42	4.00000	490	4	0.82603	16.52070
43	4.00000	490	4	0.82603	16.52070

Table L-6. Raw to Scaled Score Look-up Table—ELA Grade 4 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	400	1	2.48280	49.65609
1	-3.85960	400	1	2.33041	46.60819
2	-3.71920	400	1	2.18556	43.71120
3	-3.57880	400	1	2.04778	40.95561
4	-3.43840	400	1	1.91660	38.33193
5	-3.29799	400	1	1.79154	35.83073
6	-3.15759	400	1	1.67213	33.44264
7	-2.27077	412	1	1.01682	20.33631
8	-1.82246	421	1	0.72493	14.49859
9	-1.52673	427	1	0.55728	11.14567
10	-1.30489	431	1	0.45880	9.17599
11	-1.12459	435	1	0.39870	7.97392
12	-0.96991	438	1	0.36004	7.20088
13	-0.83212	440	2	0.33392	6.67834
14	-0.70600	443	2	0.31551	6.31020
15	-0.58826	445	2	0.30212	6.04233
16	-0.47666	448	2	0.29213	5.84262
17	-0.36965	450	2	0.28455	5.69109
18	-0.26608	452	2	0.27874	5.57489
19	-0.16504	454	2	0.27429	5.48572
20	-0.06585	456	2	0.27092	5.41838
21	0.03212	458	2	0.26849	5.36981
22	0.12940	459	2	0.26692	5.33841
23	0.22650	462	3	0.26618	5.32357
24	0.32394	464	3	0.26627	5.32536
25	0.42223	466	3	0.26722	5.34435
26	0.52195	468	3	0.26907	5.38144
27	0.62372	470	3	0.27189	5.43782
28	0.72825	472	3	0.27574	5.51480
29	0.83632	474	4	0.28069	5.61380
30	0.94888	476	4	0.28682	5.73637
31	1.06699	478	4	0.29423	5.88450
32	1.19193	481	4	0.30308	6.06151
33	1.32524	484	4	0.31368	6.27364
34	1.46890	486	4	0.32662	6.53237
35	1.62558	489	4	0.34280	6.85591
36	1.79894	489	4	0.36328	7.26555
37	1.99398	489	4	0.38846	7.76915
38	2.21705	489	4	0.41688	8.33756
39	2.47660	489	4	0.44789	8.95771
40	2.79063	489	4	0.49542	9.90832
41	3.22151	489	4	0.62102	12.42048
42	4.00000	490	4	1.11702	22.34043
43	4.00000	490	4	1.11702	22.34043

Table L-7. Raw to Scaled Score Look-up Table—ELA Grade 5 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	500	1	2.67395	53.47891
1	-3.79056	500	1	2.38007	47.60141
2	-3.58112	500	1	2.10966	42.19328
3	-3.37167	500	1	1.86086	37.21714
4	-3.16223	500	1	1.63218	32.64362
5	-2.95279	500	1	1.42264	28.45283
6	-2.74335	502	1	1.23174	24.63484
7	-2.15160	514	1	0.79335	15.86703
8	-1.80104	521	1	0.60443	12.08863
9	-1.54917	526	1	0.50001	10.00021
10	-1.34993	530	1	0.43479	8.69582
11	-1.18271	533	1	0.39108	7.82169
12	-1.03658	536	1	0.36052	7.21031
13	-0.90499	539	1	0.33859	6.77179
14	-0.78376	541	1	0.32267	6.45341
15	-0.67000	543	2	0.31109	6.22185
16	-0.56167	546	2	0.30274	6.05475
17	-0.45725	548	2	0.29681	5.93614
18	-0.35563	550	2	0.29270	5.85396
19	-0.25594	552	2	0.28993	5.79861
20	-0.15753	554	2	0.28810	5.76209
21	-0.05993	556	2	0.28688	5.73759
22	0.03725	558	2	0.28597	5.71931
23	0.13427	559	2	0.28513	5.70266
24	0.23136	561	3	0.28423	5.68460
25	0.32872	563	3	0.28321	5.66416
26	0.42657	565	3	0.28214	5.64284
27	0.52520	567	3	0.28124	5.62477
28	0.62498	569	3	0.28082	5.61630
29	0.72645	571	3	0.28127	5.62537
30	0.83031	572	3	0.28303	5.66062
31	0.93746	576	4	0.28653	5.73063
32	1.04906	578	4	0.29216	5.84326
33	1.16654	580	4	0.30026	6.00528
34	1.29165	583	4	0.31111	6.22212
35	1.42659	585	4	0.32494	6.49880
36	1.57417	588	4	0.34217	6.84333
37	1.73837	589	4	0.36380	7.27592
38	1.92542	589	4	0.39240	7.84795
39	2.14648	589	4	0.43397	8.67947
40	2.42432	589	4	0.50248	10.04951
41	2.81425	589	4	0.63530	12.70595
42	3.50392	589	4	0.98016	19.60328
43	4.00000	590	4	1.31516	26.30327

Table L-8. Raw to Scaled Score Look-up Table—ELA Grade 5 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	500	1	2.73589	54.71776
1	-3.78757	500	1	2.44426	48.88530
2	-3.57513	500	1	2.17640	43.52796
3	-3.36270	500	1	1.92993	38.59855
4	-3.15027	500	1	1.70284	34.05681
5	-2.93784	500	1	1.49361	29.87213
6	-2.72540	502	1	1.30130	26.02598
7	-2.12078	514	1	0.84598	16.91952
8	-1.75987	522	1	0.64144	12.82889
9	-1.49950	527	1	0.52520	10.50398
10	-1.29328	531	1	0.45082	9.01641
11	-1.12048	534	1	0.39964	7.99274
12	-0.97009	537	1	0.36266	7.25315
13	-0.83559	540	1	0.33506	6.70114
14	-0.71272	542	1	0.31407	6.28133
15	-0.59858	545	2	0.29802	5.96035
16	-0.49101	547	2	0.28587	5.71736
17	-0.38834	549	2	0.27694	5.53871
18	-0.28922	551	2	0.27073	5.41463
19	-0.19252	553	2	0.26686	5.33718
20	-0.09730	555	2	0.26493	5.29866
21	-0.00276	557	2	0.26453	5.29061
22	0.09172	559	2	0.26516	5.30312
23	0.18660	561	3	0.26625	5.32497
24	0.28218	562	3	0.26724	5.34479
25	0.37861	564	3	0.26767	5.35331
26	0.47594	566	3	0.26731	5.34611
27	0.57421	568	3	0.26627	5.32543
28	0.67352	570	3	0.26499	5.29990
29	0.77414	572	3	0.26412	5.28237
30	0.87659	574	4	0.26435	5.28695
31	0.98166	576	4	0.26632	5.32631
32	1.09045	579	4	0.27049	5.40972
33	1.20437	581	4	0.27705	5.54092
34	1.32505	583	4	0.28580	5.71610
35	1.45433	586	4	0.29621	5.92419
36	1.59435	589	4	0.30783	6.15654
37	1.74804	589	4	0.32159	6.43176
38	1.92066	589	4	0.34150	6.83007
39	2.12332	589	4	0.37672	7.53438
40	2.38113	589	4	0.44674	8.93480
41	2.75745	589	4	0.60059	12.01181
42	3.45859	589	4	1.00040	20.00806
43	4.00000	590	4	1.37818	27.56352

Table L-9. Raw to Scaled Score Look-up Table—ELA Grade 5 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	500	1	2.73018	54.60353
1	-3.78790	500	1	2.43910	48.78202
2	-3.57580	500	1	2.17196	43.43914
3	-3.36369	500	1	1.92646	38.52924
4	-3.15159	500	1	1.70067	34.01350
5	-2.93949	500	1	1.49311	29.86225
6	-2.72739	502	1	1.30286	26.05721
7	-2.12282	514	1	0.85449	17.08973
8	-1.76101	522	1	0.65453	13.09053
9	-1.49894	527	1	0.54176	10.83529
10	-1.29023	531	1	0.47049	9.40989
11	-1.11412	535	1	0.42235	8.44704
12	-0.95955	538	1	0.38843	7.76869
13	-0.81994	540	1	0.36387	7.27740
14	-0.69107	542	1	0.34575	6.91498
15	-0.57009	545	2	0.33221	6.64423
16	-0.45500	548	2	0.32200	6.43995
17	-0.34436	550	2	0.31421	6.28430
18	-0.23711	552	2	0.30821	6.16420
19	-0.13249	554	2	0.30349	6.06976
20	-0.02992	556	2	0.29967	5.99337
21	0.07104	558	2	0.29646	5.92916
22	0.17076	560	3	0.29365	5.87294
23	0.26952	562	3	0.29111	5.82223
24	0.36763	564	3	0.28883	5.77655
25	0.46541	566	3	0.28688	5.73752
26	0.56323	568	3	0.28544	5.70881
27	0.66156	570	3	0.28479	5.69574
28	0.76097	572	3	0.28523	5.70470
29	0.86217	574	4	0.28713	5.74253
30	0.96602	576	4	0.29081	5.81612
31	1.07360	578	4	0.29660	5.93204
32	1.18615	581	4	0.30482	6.09643
33	1.30521	583	4	0.31574	6.31472
34	1.43264	585	4	0.32957	6.59142
35	1.57074	588	4	0.34650	6.92992
36	1.72241	589	4	0.36666	7.33322
37	1.89157	589	4	0.39041	7.80824
38	2.08395	589	4	0.41900	8.38004
39	2.30920	589	4	0.45642	9.12830
40	2.58663	589	4	0.51386	10.27726
41	2.96413	589	4	0.62432	12.48636
42	3.61177	589	4	0.92762	18.55239
43	4.00000	590	4	1.18516	23.70327

Table L-10. Raw to Scaled Score Look-up Table—ELA Grade 6 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	600	1	2.33683	46.73660
1	-3.80157	600	1	2.09547	41.90932
2	-3.60313	600	1	1.87377	37.47536
3	-3.40470	600	1	1.67137	33.42739
4	-3.20627	600	1	1.48768	29.75361
5	-3.00784	600	1	1.32175	26.43494
6	-2.80940	600	1	1.17225	23.44502
7	-2.23035	610	1	0.81674	16.33484
8	-1.85864	618	1	0.64398	12.87950
9	-1.58115	623	1	0.54082	10.81633
10	-1.35630	628	1	0.47210	9.44200
11	-1.16460	631	1	0.42293	8.45851
12	-0.99559	635	2	0.38560	7.71200
13	-0.84322	638	2	0.35570	7.11393
14	-0.70382	641	2	0.33074	6.61481
15	-0.57493	643	2	0.30960	6.19205
16	-0.45476	646	2	0.29199	5.83987
17	-0.34171	648	2	0.27799	5.55976
18	-0.23427	650	2	0.26766	5.35317
19	-0.13097	652	2	0.26089	5.21772
20	-0.03042	654	2	0.25729	5.14589
21	0.06858	656	2	0.25624	5.12489
22	0.16698	658	2	0.25686	5.13712
23	0.26542	659	2	0.25810	5.16204
24	0.36421	662	3	0.25902	5.18040
25	0.46341	664	3	0.25903	5.18056
26	0.56298	666	3	0.25816	5.16318
27	0.66296	668	3	0.25704	5.14080
28	0.76364	670	3	0.25660	5.13190
29	0.86558	672	3	0.25765	5.15294
30	0.96964	674	4	0.26057	5.21145
31	1.07680	676	4	0.26508	5.30151
32	1.18798	678	4	0.27018	5.40363
33	1.30394	681	4	0.27475	5.49493
34	1.42540	683	4	0.27858	5.57157
35	1.55359	686	4	0.28344	5.66873
36	1.69124	689	4	0.29303	5.86057
37	1.84385	689	4	0.31259	6.25172
38	2.02150	689	4	0.34936	6.98724
39	2.24209	689	4	0.41443	8.28857
40	2.53811	689	4	0.52527	10.50539
41	2.97606	689	4	0.71408	14.28160
42	3.75528	689	4	1.10795	22.15896
43	4.00000	690	4	1.25473	25.09463

Table L-11. Raw to Scaled Score Look-up Table—ELA Grade 6 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	600	1	2.33422	46.68432
1	-3.80173	600	1	2.09259	41.85178
2	-3.60347	600	1	1.87056	37.41112
3	-3.40520	600	1	1.66776	33.35515
4	-3.20694	600	1	1.48361	29.67226
5	-3.00867	600	1	1.31719	26.34371
6	-2.81041	600	1	1.16718	23.34368
7	-2.23267	610	1	0.81099	16.21974
8	-1.86243	617	1	0.63872	12.77440
9	-1.58651	623	1	0.53682	10.73639
10	-1.36320	627	1	0.47006	9.40124
11	-1.17286	631	1	0.42366	8.47320
12	-1.00481	635	2	0.39008	7.80162
13	-0.85268	638	2	0.36493	7.29869
14	-0.71247	640	2	0.34550	6.90994
15	-0.58149	643	2	0.33006	6.60112
16	-0.45781	646	2	0.31755	6.35098
17	-0.34000	648	2	0.30734	6.14676
18	-0.22686	650	2	0.29906	5.98126
19	-0.11740	652	2	0.29254	5.85071
20	-0.01068	654	2	0.28766	5.75320
21	0.09413	657	2	0.28437	5.68749
22	0.19785	659	2	0.28260	5.65210
23	0.30122	661	3	0.28223	5.64466
24	0.40490	663	3	0.28308	5.66154
25	0.50951	665	3	0.28489	5.69774
26	0.61554	667	3	0.28737	5.74731
27	0.72341	669	3	0.29022	5.80445
28	0.83346	671	3	0.29326	5.86521
29	0.94603	672	3	0.29645	5.92909
30	1.06152	676	4	0.30000	5.99997
31	1.18047	678	4	0.30427	6.08534
32	1.30370	681	4	0.30971	6.19411
33	1.43229	683	4	0.31666	6.33318
34	1.56767	686	4	0.32520	6.50400
35	1.71158	689	4	0.33509	6.70174
36	1.86619	689	4	0.34611	6.92221
37	2.03455	689	4	0.35910	7.18194
38	2.22188	689	4	0.37748	7.54966
39	2.43874	689	4	0.40938	8.18762
40	2.70875	689	4	0.47243	9.44865
41	3.09240	689	4	0.61291	12.25815
42	3.79852	689	4	1.00747	20.14946
43	4.00000	690	4	1.14559	22.91171

Table L-12. Raw to Scaled Score Look-up Table—ELA Grade 6 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	600	1	2.32041	46.40819
1	-3.80249	600	1	2.07886	41.57716
2	-3.60498	600	1	1.85688	37.13763
3	-3.40747	600	1	1.65414	33.08284
4	-3.20996	600	1	1.47008	29.40157
5	-3.01245	600	1	1.30378	26.07563
6	-2.81494	600	1	1.15400	23.08001
7	-2.24041	610	1	0.79987	15.99731
8	-1.87310	617	1	0.62985	12.59694
9	-1.59990	623	1	0.52993	10.59857
10	-1.37917	627	1	0.46493	9.29869
11	-1.19121	631	1	0.42021	8.40416
12	-1.02527	634	2	0.38837	7.76738
13	-0.87488	637	2	0.36516	7.30315
14	-0.73587	640	2	0.34791	6.95820
15	-0.60541	643	2	0.33488	6.69756
16	-0.48148	645	2	0.32488	6.49755
17	-0.36258	647	2	0.31708	6.34167
18	-0.24753	650	2	0.31093	6.21850
19	-0.13540	652	2	0.30604	6.12084
20	-0.02538	654	2	0.30225	6.04503
21	0.08324	656	2	0.29952	5.99037
22	0.19115	659	2	0.29790	5.95802
23	0.29902	661	3	0.29749	5.94978
24	0.40751	663	3	0.29834	5.96687
25	0.51731	665	3	0.30046	6.00910
26	0.62906	667	3	0.30373	6.07461
27	0.74340	670	3	0.30800	6.16008
28	0.86093	672	3	0.31307	6.26148
29	0.98226	674	4	0.31875	6.37508
30	1.10800	677	4	0.32493	6.49862
31	1.23882	679	4	0.33163	6.63267
32	1.37557	682	4	0.33910	6.78202
33	1.51939	685	4	0.34784	6.95677
34	1.67188	688	4	0.35861	7.17230
35	1.83537	689	4	0.37236	7.44714
36	2.01310	689	4	0.38986	7.79712
37	2.20942	689	4	0.41128	8.22567
38	2.42994	689	4	0.43599	8.71972
39	2.68250	689	4	0.46449	9.28982
40	2.98190	689	4	0.50558	10.11155
41	3.36913	689	4	0.59263	11.85253
42	4.00000	690	4	0.86737	17.34735
43	4.00000	690	4	0.86737	17.34735

Table L-13. Raw to Scaled Score Look-up Table—ELA Grade 7 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	700	1	2.51124	50.22479
1	-3.82903	700	1	2.28043	45.60857
2	-3.65805	700	1	2.06415	41.28310
3	-3.48708	700	1	1.86194	37.23884
4	-3.31611	700	1	1.67317	33.46336
5	-3.14514	700	1	1.49711	29.94230
6	-2.97416	700	1	1.33309	26.66172
7	-2.80319	700	1	1.18061	23.61217
8	-2.23337	711	1	0.76083	15.21670
9	-1.88377	718	1	0.58427	11.68535
10	-1.62305	723	1	0.49567	9.91337
11	-1.40804	727	1	0.44658	8.93152
12	-1.21973	730	1	0.41773	8.35457
13	-1.04832	735	2	0.40022	8.00447
14	-0.88821	738	2	0.38936	7.78716
15	-0.73601	741	2	0.38229	7.64572
16	-0.58958	744	2	0.37705	7.54097
17	-0.44758	747	2	0.37223	7.44460
18	-0.30917	749	2	0.36694	7.33876
19	-0.17376	752	2	0.36089	7.21775
20	-0.04089	755	2	0.35437	7.08748
21	0.08994	757	2	0.34802	6.96043
22	0.21922	760	3	0.34238	6.84767
23	0.34748	763	3	0.33762	6.75243
24	0.47516	765	3	0.33344	6.66872
25	0.60261	768	3	0.32934	6.58684
26	0.73009	770	3	0.32518	6.50353
27	0.85796	773	3	0.32149	6.42972
28	0.98680	774	3	0.31942	6.38834
29	1.11759	778	4	0.32018	6.40353
30	1.25169	781	4	0.32451	6.49026
31	1.39070	783	4	0.33256	6.65125
32	1.53643	786	4	0.34407	6.88133
33	1.69091	789	4	0.35876	7.17511
34	1.85649	789	4	0.37656	7.53128
35	2.03603	789	4	0.39747	7.94947
36	2.23301	789	4	0.42124	8.42471
37	2.45191	789	4	0.44784	8.95690
38	2.69957	789	4	0.47974	9.59490
39	2.98932	789	4	0.52611	10.52217
40	3.35242	789	4	0.61113	12.22269
41	3.87482	789	4	0.80463	16.09259
42	4.00000	790	4	0.86408	17.28159
43	4.00000	790	4	0.86408	17.28159

Table L-14. Raw to Scaled Score Look-up Table—ELA Grade 7 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	700	1	2.45579	49.11587
1	-3.83120	700	1	2.23247	44.64939
2	-3.66240	700	1	2.02397	40.47939
3	-3.49360	700	1	1.82973	36.59461
4	-3.32479	700	1	1.64903	32.98056
5	-3.15599	700	1	1.48101	29.62028
6	-2.98719	700	1	1.32481	26.49630
7	-2.81839	700	1	1.17971	23.59417
8	-2.24759	711	1	0.77203	15.44061
9	-1.89587	718	1	0.59682	11.93647
10	-1.63289	723	1	0.50790	10.15806
11	-1.41566	727	1	0.45843	9.16855
12	-1.22530	730	1	0.42918	8.58367
13	-1.05214	735	2	0.41107	8.22147
14	-0.89072	738	2	0.39924	7.98477
15	-0.73777	741	2	0.39078	7.81562
16	-0.59127	744	2	0.38381	7.67621
17	-0.44991	747	2	0.37709	7.54172
18	-0.31282	749	2	0.36994	7.39881
19	-0.17938	752	2	0.36228	7.24569
20	-0.04900	755	2	0.35451	7.09023
21	0.07885	757	2	0.34724	6.94479
22	0.20478	759	2	0.34098	6.81965
23	0.32934	762	3	0.33590	6.71794
24	0.45308	765	3	0.33174	6.63473
25	0.57644	767	3	0.32805	6.56095
26	0.69980	770	3	0.32457	6.49141
27	0.82357	772	3	0.32160	6.43199
28	0.94837	774	3	0.32002	6.40045
29	1.07512	777	4	0.32096	6.41917
30	1.20506	780	4	0.32527	6.50543
31	1.33975	782	4	0.33328	6.66552
32	1.48091	785	4	0.34477	6.89536
33	1.63043	788	4	0.35934	7.18687
34	1.79046	789	4	0.37681	7.53622
35	1.96371	789	4	0.39745	7.94905
36	2.15390	789	4	0.42215	8.44291
37	2.36652	789	4	0.45276	9.05516
38	2.61062	789	4	0.49347	9.86941
39	2.90296	789	4	0.55390	11.07809
40	3.27902	789	4	0.65724	13.14474
41	3.82864	789	4	0.86910	17.38193
42	4.00000	790	4	0.95007	19.00135
43	4.00000	790	4	0.95007	19.00135

Table L-15. Raw to Scaled Score Look-up Table—ELA Grade 7 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	700	1	2.65559	53.11177
1	-3.82379	700	1	2.42436	48.48719
2	-3.64758	700	1	2.20723	44.14459
3	-3.47137	700	1	2.00331	40.06614
4	-3.29516	700	1	1.81146	36.22915
5	-3.11895	700	1	1.63037	32.60747
6	-2.94274	700	1	1.45879	29.17586
7	-2.76653	700	1	1.29592	25.91849
8	-2.17854	712	1	0.82956	16.59117
9	-1.81458	719	1	0.63157	12.63135
10	-1.54084	725	1	0.53482	10.69630
11	-1.31377	729	1	0.48147	9.62934
12	-1.11486	733	2	0.44826	8.96511
13	-0.93508	737	2	0.42470	8.49407
14	-0.76958	740	2	0.40586	8.11719
15	-0.61540	743	2	0.38946	7.78925
16	-0.47055	746	2	0.37463	7.49259
17	-0.33343	749	2	0.36117	7.22334
18	-0.20268	752	2	0.34920	6.98406
19	-0.07706	754	2	0.33897	6.77932
20	0.04457	757	2	0.33064	6.61279
21	0.16327	759	2	0.32426	6.48513
22	0.27994	761	3	0.31964	6.39271
23	0.39539	764	3	0.31637	6.32734
24	0.51022	766	3	0.31389	6.27786
25	0.62490	768	3	0.31171	6.23413
26	0.73980	770	3	0.30962	6.19232
27	0.85533	773	3	0.30793	6.15852
28	0.97200	774	3	0.30737	6.14731
29	1.09060	777	4	0.30878	6.17564
30	1.21214	780	4	0.31280	6.25597
31	1.33786	782	4	0.31963	6.39255
32	1.46917	785	4	0.32911	6.58211
33	1.60768	788	4	0.34094	6.81887
34	1.75532	789	4	0.35509	7.10179
35	1.91466	789	4	0.37211	7.44219
36	2.08955	789	4	0.39359	7.87175
37	2.28615	789	4	0.42275	8.45501
38	2.51512	789	4	0.46580	9.31592
39	2.79642	789	4	0.53506	10.70112
40	3.17136	789	4	0.65762	13.15248
41	3.73970	789	4	0.90358	18.07161
42	4.00000	790	4	1.03637	20.72737
43	4.00000	790	4	1.03637	20.72737

Table L-16. Raw to Scaled Score Look-up Table—ELA Grade 8 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	800	1	2.08010	41.60199
1	-3.84657	800	1	1.92775	38.55510
2	-3.69315	800	1	1.78288	35.65755
3	-3.53972	800	1	1.64504	32.90072
4	-3.38629	800	1	1.51400	30.28000
5	-3.23286	800	1	1.38974	27.79474
6	-3.07944	800	1	1.27237	25.44744
7	-2.42667	809	1	0.85501	17.10011
8	-2.02683	817	1	0.66593	13.31857
9	-1.73596	822	1	0.55721	11.14430
10	-1.50525	827	1	0.48609	9.72181
11	-1.31231	831	1	0.43579	8.71579
12	-1.14499	834	1	0.39847	7.96939
13	-0.99589	837	1	0.37000	7.39997
14	-0.86015	839	1	0.34799	6.95986
15	-0.73432	842	2	0.33097	6.61942
16	-0.61586	845	2	0.31794	6.35889
17	-0.50283	847	2	0.30821	6.16424
18	-0.39367	849	2	0.30124	6.02488
19	-0.28710	851	2	0.29662	5.93230
20	-0.18208	853	2	0.29396	5.87924
21	-0.07771	856	2	0.29296	5.85914
22	0.02674	858	2	0.29330	5.86592
23	0.13190	859	2	0.29469	5.89389
24	0.23829	862	3	0.29689	5.93785
25	0.34635	864	3	0.29967	5.99337
26	0.45643	866	3	0.30286	6.05714
27	0.56888	868	3	0.30637	6.12744
28	0.68404	870	3	0.31023	6.20450
29	0.80228	873	4	0.31453	6.29060
30	0.92413	876	4	0.31949	6.38982
31	1.05026	878	4	0.32537	6.50736
32	1.18157	881	4	0.33243	6.64857
33	1.31926	884	4	0.34091	6.81830
34	1.46489	886	4	0.35106	7.02113
35	1.62057	889	4	0.36321	7.26414
36	1.78924	889	4	0.37821	7.56411
37	1.97546	889	4	0.39807	7.96144
38	2.18706	889	4	0.42728	8.54551
39	2.43901	889	4	0.47541	9.50815
40	2.76362	889	4	0.56481	11.29619
41	3.24396	889	4	0.75898	15.17951
42	4.00000	890	4	1.19471	23.89413
43	4.00000	890	4	1.19471	23.89413

Table L-17. Raw to Scaled Score Look-up Table—ELA Grade 8 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	800	1	2.10235	42.04702
1	-3.84536	800	1	1.95130	39.02610
2	-3.69072	800	1	1.80795	36.15894
3	-3.53609	800	1	1.67179	33.43580
4	-3.38145	800	1	1.54255	30.85094
5	-3.22681	800	1	1.42013	28.40266
6	-3.07217	800	1	1.30463	26.09266
7	-2.41084	809	1	0.89319	17.86383
8	-2.00143	817	1	0.70628	14.12561
9	-1.70032	823	1	0.59798	11.95963
10	-1.45897	828	1	0.52560	10.51194
11	-1.25531	832	1	0.47236	9.44727
12	-1.07751	836	1	0.43068	8.61359
13	-0.91850	839	1	0.39685	7.93701
14	-0.77367	842	2	0.36896	7.37926
15	-0.63981	844	2	0.34590	6.91793
16	-0.51458	847	2	0.32686	6.53724
17	-0.39619	849	2	0.31123	6.22454
18	-0.28324	851	2	0.29848	5.96953
19	-0.17460	854	2	0.28822	5.76444
20	-0.06934	856	2	0.28021	5.60415
21	0.03339	858	2	0.27429	5.48584
22	0.13440	859	2	0.27040	5.40803
23	0.23444	862	3	0.26846	5.36928
24	0.33430	864	3	0.26833	5.36670
25	0.43469	866	3	0.26974	5.39484
26	0.53629	868	3	0.27227	5.44539
27	0.63967	870	3	0.27542	5.50846
28	0.74532	872	4	0.27879	5.57584
29	0.85367	874	4	0.28225	5.64492
30	0.96523	876	4	0.28602	5.72039
31	1.08069	879	4	0.29057	5.81145
32	1.20096	881	4	0.29626	5.92513
33	1.32716	884	4	0.30297	6.05944
34	1.46058	886	4	0.31014	6.20282
35	1.60265	889	4	0.31739	6.34780
36	1.75550	889	4	0.32597	6.51943
37	1.92324	889	4	0.34009	6.80170
38	2.11467	889	4	0.36781	7.35615
39	2.34841	889	4	0.42395	8.47908
40	2.66535	889	4	0.54001	10.80011
41	3.16385	889	4	0.78972	15.79433
42	4.00000	890	4	1.28984	25.79685
43	4.00000	890	4	1.28984	25.79685

Table L-18. Raw to Scaled Score Look-up Table—ELA Grade 8 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	800	1	2.10206	42.04112
1	-3.84538	800	1	1.95095	39.01904
2	-3.69077	800	1	1.80752	36.15033
3	-3.53615	800	1	1.67126	33.42517
4	-3.38153	800	1	1.54189	30.83776
5	-3.22691	800	1	1.41932	28.38634
6	-3.07230	800	1	1.30363	26.07252
7	-2.41127	809	1	0.89092	17.81849
8	-2.00245	817	1	0.70256	14.05123
9	-1.70225	823	1	0.59257	11.85139
10	-1.46220	828	1	0.51829	10.36590
11	-1.26032	832	1	0.46309	9.26186
12	-1.08487	835	1	0.41951	8.39025
13	-0.92882	839	1	0.38400	7.67996
14	-0.78761	841	2	0.35475	7.09508
15	-0.65801	844	2	0.33075	6.61508
16	-0.53763	846	2	0.31130	6.22597
17	-0.42456	849	2	0.29584	5.91686
18	-0.31721	851	2	0.28393	5.67865
19	-0.21422	853	2	0.27517	5.50349
20	-0.11440	855	2	0.26921	5.38424
21	-0.01668	857	2	0.26569	5.31385
22	0.07989	859	2	0.26423	5.28462
23	0.17611	861	3	0.26439	5.28778
24	0.27268	863	3	0.26569	5.31375
25	0.37016	865	3	0.26766	5.35328
26	0.46901	867	3	0.26998	5.39961
27	0.56961	869	3	0.27252	5.45049
28	0.67239	870	3	0.27542	5.50840
29	0.77783	873	4	0.27892	5.57838
30	0.88655	875	4	0.28319	5.66390
31	0.99924	877	4	0.28816	5.76328
32	1.11669	879	4	0.29348	5.86954
33	1.23974	882	4	0.29884	5.97680
34	1.36951	885	4	0.30465	6.09291
35	1.50785	887	4	0.31258	6.25164
36	1.65822	889	4	0.32590	6.51800
37	1.82698	889	4	0.34963	6.99254
38	2.02568	889	4	0.39162	7.83246
39	2.27559	889	4	0.46548	9.30957
40	2.61815	889	4	0.59682	11.93635
41	3.14507	889	4	0.84079	16.81571
42	4.00000	890	4	1.31058	26.21156
43	4.00000	890	4	1.31058	26.21156

Table L-19. Raw to Scaled Score Look-up Table—Mathematics Grade 3 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	2.73878	47.92865
1	-3.88661	300	1	2.58572	45.25005
2	-3.77322	300	1	2.43497	42.61203
3	-3.65983	300	1	2.28660	40.01551
4	-3.54644	300	1	2.14075	37.46306
5	-3.43305	300	1	1.99768	34.95940
6	-3.31966	300	1	1.85781	32.51161
7	-3.20627	300	1	1.72167	30.12929
8	-2.28699	312	1	0.84476	14.78325
9	-1.86103	320	1	0.61086	10.69010
10	-1.57167	325	1	0.50376	8.81583
11	-1.34698	329	1	0.44162	7.72843
12	-1.16026	332	1	0.40059	7.01038
13	-0.99861	335	1	0.37131	6.49800
14	-0.85473	337	1	0.34937	6.11395
15	-0.72406	340	1	0.33236	5.81628
16	-0.60352	340	1	0.31884	5.57975
17	-0.49095	344	2	0.30788	5.38790
18	-0.38474	346	2	0.29882	5.22939
19	-0.28368	347	2	0.29121	5.09610
20	-0.18685	349	2	0.28470	4.98217
21	-0.09349	351	2	0.27905	4.88336
22	-0.00300	352	2	0.27409	4.79664
23	0.08512	354	2	0.26971	4.71997
24	0.17131	355	2	0.26583	4.65205
25	0.25594	357	2	0.26241	4.59218
26	0.33933	358	2	0.25943	4.54010
27	0.42182	359	2	0.25691	4.49587
28	0.50368	361	3	0.25484	4.45977
29	0.58522	363	3	0.25327	4.43222
30	0.66673	364	3	0.25221	4.41369
31	0.74850	365	3	0.25169	4.40464
32	0.83086	367	3	0.25174	4.40549
33	0.91413	368	3	0.25238	4.41660
34	0.99869	370	3	0.25362	4.43827
35	1.08491	371	3	0.25548	4.47082
36	1.17325	373	3	0.25798	4.51472
37	1.26419	374	3	0.26119	4.57075
38	1.35832	376	3	0.26516	4.64022
39	1.45632	376	3	0.27001	4.72519
40	1.55902	380	4	0.27592	4.82864
41	1.66744	382	4	0.28313	4.95472
42	1.78292	384	4	0.29196	5.10922
43	1.90725	386	4	0.30289	5.30064
44	2.04290	388	4	0.31671	5.54249
45	2.19357	389	4	0.33474	5.85791
46	2.36515	389	4	0.35937	6.28896
47	2.56791	389	4	0.39531	6.91797
48	2.82207	389	4	0.45298	7.92723
49	3.17590	389	4	0.56127	9.82219
50	3.79836	389	4	0.84810	14.84176
51	4.00000	390	4	0.97136	16.99874

Table L-20. Raw to Scaled Score Look-up Table—Mathematics Grade 3 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	2.73878	47.92867
1	-3.81131	300	1	2.48536	43.49378
2	-3.62263	300	1	2.23846	39.17310
3	-3.43394	300	1	1.99880	34.97899
4	-3.24526	300	1	1.76803	30.94045
5	-3.05657	300	1	1.54885	27.10488
6	-2.86789	302	1	1.34478	23.53361
7	-2.67920	305	1	1.15943	20.29008
8	-2.07482	316	1	0.71535	12.51857
9	-1.72255	322	1	0.55533	9.71832
10	-1.46539	327	1	0.47257	8.27000
11	-1.25855	330	1	0.42138	7.37411
12	-1.08301	333	1	0.38631	6.76039
13	-0.92883	336	1	0.36074	6.31296
14	-0.79012	339	1	0.34134	5.97352
15	-0.66305	340	1	0.32622	5.70884
16	-0.54499	343	2	0.31418	5.49809
17	-0.43405	345	2	0.30441	5.32724
18	-0.32882	347	2	0.29636	5.18636
19	-0.22822	348	2	0.28961	5.06818
20	-0.13142	350	2	0.28385	4.96734
21	-0.03775	352	2	0.27885	4.87989
22	0.05334	353	2	0.27446	4.80303
23	0.14230	355	2	0.27056	4.73486
24	0.22953	356	2	0.26710	4.67427
25	0.31537	358	2	0.26404	4.62074
26	0.40013	359	2	0.26139	4.57425
27	0.48413	361	3	0.25915	4.53510
28	0.56764	362	3	0.25736	4.50382
29	0.65098	364	3	0.25606	4.48105
30	0.73443	365	3	0.25528	4.46745
31	0.81833	367	3	0.25506	4.46359
32	0.90300	368	3	0.25543	4.46998
33	0.98883	370	3	0.25640	4.48703
34	1.07622	371	3	0.25801	4.51510
35	1.16561	373	3	0.26027	4.55472
36	1.25750	374	3	0.26324	4.60666
37	1.35249	376	3	0.26699	4.67227
38	1.45126	376	3	0.27163	4.75359
39	1.55464	380	4	0.27735	4.85364
40	1.66369	381	4	0.28438	4.97658
41	1.77973	383	4	0.29304	5.12821
42	1.90455	386	4	0.30383	5.31701
43	2.04064	388	4	0.31751	5.55649
44	2.19170	389	4	0.33541	5.86976
45	2.36362	389	4	0.35994	6.29889
46	2.56670	389	4	0.39579	6.92627
47	2.82113	389	4	0.45338	7.93418
48	3.17522	389	4	0.56161	9.82821
49	3.79796	389	4	0.84843	14.84754
50	4.00000	390	4	0.97191	17.00838

Table L-21. Raw to Scaled Score Look-up Table—Mathematics Grade 4 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	400	1	5.55944	80.00000
1	-3.73211	400	1	4.34682	76.06932
2	-3.46422	400	1	3.39626	59.43447
3	-3.19634	400	1	2.65279	46.42389
4	-2.92845	401	1	2.07283	36.27448
5	-2.66056	406	1	1.62187	28.38268
6	-2.39267	410	1	1.27271	22.27237
7	-2.12479	415	1	1.00394	17.56886
8	-1.52538	426	1	0.61322	10.73131
9	-1.19886	431	1	0.48694	8.52140
10	-0.96504	435	1	0.42250	7.39367
11	-0.77795	439	1	0.38283	6.69950
12	-0.61900	441	1	0.35554	6.22193
13	-0.47892	443	1	0.33520	5.86596
14	-0.35244	446	2	0.31904	5.58320
15	-0.23629	448	2	0.30554	5.34691
16	-0.12829	450	2	0.29383	5.14204
17	-0.02691	452	2	0.28345	4.96044
18	0.06899	453	2	0.27418	4.79810
19	0.16032	455	2	0.26591	4.65336
20	0.24780	457	2	0.25861	4.52562
21	0.33206	458	2	0.25226	4.41454
22	0.41362	459	2	0.24683	4.31955
23	0.49294	461	3	0.24227	4.23971
24	0.57044	462	3	0.23850	4.17370
25	0.64647	464	3	0.23542	4.11993
26	0.72134	465	3	0.23296	4.07672
27	0.79533	466	3	0.23100	4.04245
28	0.86869	467	3	0.22947	4.01568
29	0.94164	469	3	0.22830	3.99527
30	1.01440	470	3	0.22745	3.98042
31	1.08719	471	3	0.22690	3.97071
32	1.16022	473	3	0.22664	3.96613
33	1.23374	474	3	0.22669	3.96709
34	1.30801	475	3	0.22711	3.97437
35	1.38333	476	3	0.22796	3.98922
36	1.46006	478	3	0.22933	4.01330
37	1.53864	479	3	0.23136	4.04875
38	1.61958	481	4	0.23419	4.09827
39	1.70354	482	4	0.23801	4.16519
40	1.79133	484	4	0.24307	4.25368
41	1.88400	485	4	0.24966	4.36904
42	1.98294	487	4	0.25819	4.51826
43	2.08999	489	4	0.26920	4.71100
44	2.20776	489	4	0.28350	4.96132
45	2.34004	489	4	0.30237	5.29142
46	2.49276	489	4	0.32799	5.73974
47	2.67611	489	4	0.36465	6.38142
48	2.91001	489	4	0.42223	7.38907
49	3.24327	489	4	0.52992	9.27352
50	3.86151	489	4	0.83443	14.60244
51	4.00000	490	4	0.92526	16.19197

Table L-22. Raw to Scaled Score Look-up Table—Mathematics Grade 5 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	500	1	6.52341	80.00000
1	-3.77155	500	1	5.18696	80.00000
2	-3.54309	500	1	4.12079	72.11388
3	-3.31464	500	1	3.27457	57.30496
4	-3.08618	501	1	2.60569	45.59957
5	-2.85773	505	1	2.07860	36.37542
6	-2.62927	509	1	1.66394	29.11903
7	-2.40082	513	1	1.33775	23.41055
8	-2.17236	517	1	1.08062	18.91082
9	-1.62817	527	1	0.66411	11.62201
10	-1.30719	532	1	0.50928	8.91241
11	-1.07573	536	1	0.42792	7.48863
12	-0.89251	539	1	0.37789	6.61314
13	-0.73929	542	1	0.34400	6.02004
14	-0.60641	544	1	0.31943	5.59004
15	-0.48812	547	1	0.30070	5.26225
16	-0.38072	548	2	0.28587	5.00272
17	-0.28169	550	2	0.27379	4.79130
18	-0.18921	552	2	0.26373	4.61533
19	-0.10194	553	2	0.25523	4.46655
20	-0.01885	555	2	0.24797	4.33939
21	0.06088	556	2	0.24171	4.22997
22	0.13790	558	2	0.23632	4.13552
23	0.21274	559	2	0.23166	4.05405
24	0.28585	560	3	0.22766	3.98404
25	0.35763	561	3	0.22425	3.92429
26	0.42839	563	3	0.22136	3.87375
27	0.49845	564	3	0.21894	3.83150
28	0.56806	565	3	0.21695	3.79671
29	0.63748	566	3	0.21535	3.76866
30	0.70694	567	3	0.21411	3.74685
31	0.77666	569	3	0.21320	3.73106
32	0.84686	570	3	0.21265	3.72143
33	0.91780	571	3	0.21248	3.71844
34	0.98970	572	3	0.21274	3.72294
35	1.06285	574	4	0.21348	3.73592
36	1.13757	575	4	0.21477	3.75846
37	1.21421	576	4	0.21666	3.79161
38	1.29317	578	4	0.21922	3.83641
39	1.37496	579	4	0.22252	3.89414
40	1.46019	581	4	0.22667	3.96671
41	1.54965	582	4	0.23184	4.05728
42	1.64440	584	4	0.23835	4.17117
43	1.74596	586	4	0.24669	4.31712
44	1.85655	588	4	0.25769	4.50949
45	1.97966	589	4	0.27272	4.77252
46	2.12103	589	4	0.29428	5.14983
47	2.29114	589	4	0.32734	5.72852
48	2.51176	589	4	0.38362	6.71338
49	2.84035	589	4	0.49960	8.74301
50	3.52940	589	4	0.89131	15.59791
51	4.00000	590	4	1.29574	22.67549

Table L-23. Raw to Scaled Score Look-up Table—Mathematics Grade 6 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	600	1	4.18583	73.25201
1	-3.84329	600	1	3.74186	65.48253
2	-3.68658	600	1	3.33236	58.31629
3	-3.52987	600	1	2.95620	51.73350
4	-3.37316	603	1	2.61217	45.71305
5	-3.21645	605	1	2.29905	40.23336
6	-3.05974	608	1	2.01560	35.27301
7	-2.90303	611	1	1.76062	30.81093
8	-2.74632	614	1	1.53293	26.82621
9	-2.00417	627	1	0.78478	13.73366
10	-1.62163	633	1	0.57308	10.02899
11	-1.35387	638	1	0.47529	8.31750
12	-1.14202	642	1	0.42007	7.35130
13	-0.96296	645	1	0.38517	6.74052
14	-0.80538	648	2	0.36131	6.32293
15	-0.66301	650	2	0.34389	6.01800
16	-0.53205	652	2	0.33034	5.78088
17	-0.41010	654	2	0.31912	5.58457
18	-0.29552	656	2	0.30926	5.41212
19	-0.18718	658	2	0.30016	5.25276
20	-0.08425	660	3	0.29142	5.09988
21	0.01387	662	3	0.28285	4.94985
22	0.10767	664	3	0.27436	4.80124
23	0.19752	665	3	0.26595	4.65420
24	0.28376	667	3	0.25771	4.50994
25	0.36668	668	3	0.24972	4.37014
26	0.44659	669	3	0.24209	4.23651
27	0.52379	671	3	0.23488	4.11048
28	0.59857	672	3	0.22818	3.99311
29	0.67126	673	3	0.22201	3.88516
30	0.74216	675	3	0.21642	3.78731
31	0.81161	676	3	0.21145	3.70034
32	0.87992	677	3	0.20716	3.62528
33	0.94744	678	3	0.20362	3.56338
34	1.01455	679	4	0.20092	3.51616
35	1.08164	681	4	0.19916	3.48527
36	1.14916	682	4	0.19843	3.47246
37	1.21760	683	4	0.19883	3.47952
38	1.28752	684	4	0.20048	3.50832
39	1.35956	685	4	0.20347	3.56079
40	1.43443	687	4	0.20794	3.63901
41	1.51299	688	4	0.21402	3.74538
42	1.59624	689	4	0.22187	3.88280
43	1.68537	689	4	0.23172	4.05516
44	1.78190	689	4	0.24389	4.26803
45	1.88778	689	4	0.25885	4.52994
46	2.00566	689	4	0.27740	4.85459
47	2.13938	689	4	0.30085	5.26480
48	2.29474	689	4	0.33145	5.80031
49	2.48127	689	4	0.37340	6.53447
50	2.71605	689	4	0.43516	7.61535
51	3.03372	689	4	0.53613	9.38233
52	3.51825	689	4	0.72949	12.76611
53	4.00000	690	4	0.96454	16.87952
54	4.00000	690	4	0.96454	16.87952

Table L-24. Raw to Scaled Score Look-up Table—Mathematics Grade 7 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	700	1	6.55322	80.00000
1	-3.80127	700	1	5.36837	80.00000
2	-3.60254	700	1	4.39788	76.96295
3	-3.40380	700	1	3.60486	63.08507
4	-3.20507	703	1	2.95796	51.76431
5	-3.00634	707	1	2.43075	42.53808
6	-2.80761	710	1	2.00112	35.01966
7	-2.60887	714	1	1.65080	28.88907
8	-2.41014	717	1	1.36483	23.88452
9	-2.21141	721	1	1.13110	19.79416
10	-1.63301	731	1	0.66603	11.65554
11	-1.30865	736	1	0.50317	8.80546
12	-1.08019	740	1	0.41754	7.30688
13	-0.90187	743	1	0.36391	6.36849
14	-0.75408	746	1	0.32695	5.72161
15	-0.62663	747	1	0.29997	5.24956
16	-0.51356	750	2	0.27960	4.89302
17	-0.41106	752	2	0.26389	4.61803
18	-0.31656	754	2	0.25160	4.40309
19	-0.22823	755	2	0.24191	4.23350
20	-0.14474	757	2	0.23420	4.09857
21	-0.06509	758	2	0.22801	3.99022
22	0.01147	759	2	0.22298	3.90212
23	0.08553	761	3	0.21882	3.82928
24	0.15754	762	3	0.21530	3.76774
25	0.22785	763	3	0.21225	3.71435
26	0.29676	764	3	0.20952	3.66666
27	0.36449	766	3	0.20702	3.62286
28	0.43122	767	3	0.20466	3.58163
29	0.49710	768	3	0.20241	3.54217
30	0.56228	769	3	0.20023	3.50407
31	0.62687	769	3	0.19813	3.46731
32	0.69102	771	4	0.19612	3.43212
33	0.75483	772	4	0.19423	3.39900
34	0.81847	774	4	0.19250	3.36870
35	0.88209	775	4	0.19098	3.34213
36	0.94588	776	4	0.18974	3.32045
37	1.01005	777	4	0.18886	3.30497
38	1.07488	778	4	0.18840	3.29707
39	1.14066	779	4	0.18847	3.29820
40	1.20775	780	4	0.18913	3.30971
41	1.27656	782	4	0.19045	3.33286
42	1.34758	783	4	0.19251	3.36887
43	1.42139	784	4	0.19538	3.41917
44	1.49868	785	4	0.19919	3.48582
45	1.58035	787	4	0.20413	3.57232
46	1.66758	788	4	0.21056	3.68476
47	1.76205	789	4	0.21907	3.83369
48	1.86629	789	4	0.23069	4.03715
49	1.98432	789	4	0.24724	4.32671
50	2.12311	789	4	0.27207	4.76123
51	2.29610	789	4	0.31227	5.46470
52	2.53425	789	4	0.38636	6.76129
53	2.93931	789	4	0.56993	9.97384
54	4.00000	790	4	1.56018	27.30309

Table L-25. Raw to Scaled Score Look-up Table—Mathematics Grade 8 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	800	1	3.57676	62.59329
1	-3.83165	800	1	3.17486	55.56005
2	-3.66331	800	1	2.81287	49.22526
3	-3.49496	800	1	2.48694	43.52146
4	-3.32661	800	1	2.19392	38.39366
5	-3.15827	800	1	1.93139	33.79935
6	-2.98992	800	1	1.69755	29.70714
7	-2.82158	803	1	1.49107	26.09364
8	-2.65323	806	1	1.31078	22.93872
9	-1.88454	820	1	0.77589	13.57810
10	-1.44458	827	1	0.61947	10.84076
11	-1.12353	833	1	0.53585	9.37742
12	-0.86882	837	1	0.47564	8.32365
13	-0.65763	840	1	0.42839	7.49688
14	-0.47694	844	2	0.39174	6.85537
15	-0.31856	847	2	0.36359	6.36276
16	-0.17718	850	2	0.34152	5.97661
17	-0.04927	852	2	0.32349	5.66104
18	0.06762	854	2	0.30812	5.39207
19	0.17532	856	2	0.29463	5.15600
20	0.27527	857	2	0.28261	4.94564
21	0.36867	859	2	0.27184	4.75719
22	0.45652	861	3	0.26219	4.58836
23	0.53966	862	3	0.25358	4.43757
24	0.61883	864	3	0.24591	4.30348
25	0.69465	865	3	0.23914	4.18489
26	0.76767	866	3	0.23318	4.08069
27	0.83835	867	3	0.22799	3.98977
28	0.90709	869	3	0.22349	3.91108
29	0.97427	870	3	0.21963	3.84359
30	1.04020	871	3	0.21636	3.78636
31	1.10516	872	3	0.21363	3.73857
32	1.16943	873	3	0.21140	3.69953
33	1.23325	874	3	0.20964	3.66874
34	1.29686	875	3	0.20834	3.64593
35	1.36051	876	3	0.20749	3.63106
36	1.42444	877	3	0.20711	3.62435
37	1.48892	879	4	0.20722	3.62633
38	1.55422	880	4	0.20787	3.63779
39	1.62069	881	4	0.20914	3.65987
40	1.68868	882	4	0.21109	3.69403
41	1.75865	883	4	0.21384	3.74213
42	1.83112	885	4	0.21752	3.80656
43	1.90674	886	4	0.22231	3.89034
44	1.98633	887	4	0.22843	3.99748
45	2.07094	889	4	0.23619	4.13337
46	2.16195	889	4	0.24604	4.30566
47	2.26126	889	4	0.25860	4.52555
48	2.37157	889	4	0.27487	4.81019
49	2.49690	889	4	0.29641	5.18721
50	2.64363	889	4	0.32593	5.70379
51	2.82264	889	4	0.36842	6.44732
52	3.05477	889	4	0.43430	7.60029
53	3.38756	889	4	0.55029	9.63011
54	3.97490	889	4	0.82208	14.38643

Table L-26. Raw to Scaled Score Look-up Table—Science Grade 5 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.32534	500	1	1.29035	16.12941
1	-4.08470	503	1	1.14454	14.30679
2	-3.84407	506	1	1.01520	12.68999
3	-3.60343	509	1	0.90099	11.26240
4	-3.04662	515	1	0.68794	8.59930
5	-2.67624	520	1	0.58032	7.25397
6	-2.39479	524	1	0.51390	6.42374
7	-2.16525	527	1	0.46836	5.85450
8	-1.96961	529	1	0.43508	5.43856
9	-1.79774	531	1	0.40975	5.12192
10	-1.64337	533	1	0.38994	4.87425
11	-1.50235	535	1	0.37415	4.67693
12	-1.37176	536	1	0.36142	4.51776
13	-1.24949	538	1	0.35106	4.38823
14	-1.13394	539	1	0.34257	4.28214
15	-1.02390	541	1	0.33558	4.19479
16	-0.91841	542	1	0.32980	4.12249
17	-0.81670	543	1	0.32499	4.06233
18	-0.71817	545	2	0.32095	4.01193
19	-0.62231	546	2	0.31755	3.96943
20	-0.52869	547	2	0.31467	3.93332
21	-0.43696	548	2	0.31220	3.90244
22	-0.34683	549	2	0.31008	3.87594
23	-0.25802	550	2	0.30826	3.85324
24	-0.17030	551	2	0.30672	3.83398
25	-0.08345	553	2	0.30544	3.81800
26	0.00273	554	2	0.30442	3.80530
27	0.08844	555	2	0.30368	3.79600
28	0.17387	556	2	0.30323	3.79033
29	0.25920	557	2	0.30308	3.78853
30	0.34463	558	2	0.30327	3.79091
31	0.43035	559	2	0.30382	3.79776
32	0.51656	560	3	0.30475	3.80934
33	0.60347	561	3	0.30607	3.82590
34	0.69128	562	3	0.30781	3.84759
35	0.78022	563	3	0.30996	3.87452
36	0.87050	564	3	0.31253	3.90669
37	0.96236	566	3	0.31552	3.94400
38	1.05602	567	3	0.31890	3.98626
39	1.15175	568	3	0.32265	4.03315
40	1.24977	569	3	0.32674	4.08426
41	1.35034	570	3	0.33113	4.13913
42	1.45373	572	3	0.33579	4.19732
43	1.56020	573	3	0.34068	4.25849
44	1.67006	573	3	0.34581	4.32258
45	1.78362	576	4	0.35119	4.38987
46	1.90126	577	4	0.35690	4.46124
47	2.02341	579	4	0.36306	4.53820
48	2.15062	580	4	0.36984	4.62298
49	2.28354	582	4	0.37747	4.71840
50	2.42301	584	4	0.38621	4.82767
51	2.57005	586	4	0.39631	4.95391
52	2.72588	588	4	0.40796	5.09956
53	2.89201	589	4	0.42126	5.26581
54	2.95466	590	4	0.42646	5.33071
55	2.95466	590	4	0.42646	5.33071

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
56	2.95466	590	4	0.42646	5.33071
57	2.95466	590	4	0.42646	5.33071
58	2.95466	590	4	0.42646	5.33071
59	2.95466	590	4	0.42646	5.33071
60	2.95466	590	4	0.42646	5.33071
61	2.95466	590	4	0.42646	5.33071
62	2.95466	590	4	0.42646	5.33071
63	2.95466	590	4	0.42646	5.33071
64	2.95466	590	4	0.42646	5.33071

Table L-27. Raw to Scaled Score Look-up Table—Science Grade 5 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.32534	500	1	1.26420	15.80254
1	-4.13329	502	1	1.15544	14.44306
2	-3.94124	504	1	1.05549	13.19363
3	-3.74919	507	1	0.96386	12.04826
4	-3.13684	514	1	0.72211	9.02638
5	-2.73597	519	1	0.60058	7.50720
6	-2.43538	523	1	0.52595	6.57440
7	-2.19298	526	1	0.47507	5.93832
8	-1.98834	529	1	0.43812	5.47645
9	-1.81002	531	1	0.41019	5.12739
10	-1.65095	533	1	0.38853	4.85659
11	-1.50645	535	1	0.37143	4.64286
12	-1.37326	536	1	0.35778	4.47224
13	-1.24899	538	1	0.34680	4.33504
14	-1.13190	539	1	0.33794	4.22419
15	-1.02062	541	1	0.33075	4.13436
16	-0.91411	542	1	0.32491	4.06140
17	-0.81151	543	1	0.32016	4.00199
18	-0.71218	545	2	0.31628	3.95345
19	-0.61554	546	2	0.31309	3.91363
20	-0.52117	547	2	0.31046	3.88075
21	-0.42868	548	2	0.30827	3.85342
22	-0.33777	549	2	0.30644	3.83056
23	-0.24815	550	2	0.30491	3.81138
24	-0.15959	552	2	0.30363	3.79536
25	-0.07187	553	2	0.30258	3.78219
26	0.01520	554	2	0.30174	3.77179
27	0.10181	555	2	0.30114	3.76421
28	0.18815	556	2	0.30077	3.75965
29	0.27440	557	2	0.30067	3.75838
30	0.36074	558	2	0.30086	3.76073
31	0.44737	559	2	0.30137	3.76707
32	0.53448	560	3	0.30222	3.77774
33	0.62227	561	3	0.30345	3.79308
34	0.71096	562	3	0.30507	3.81337
35	0.80078	564	3	0.30711	3.83884
36	0.89194	565	3	0.30957	3.86963
37	0.98470	566	3	0.31247	3.90582
38	1.07931	567	3	0.31579	3.94736
39	1.17603	568	3	0.31953	3.99413
40	1.27514	570	3	0.32367	4.04592
41	1.37693	571	3	0.32820	4.10249
42	1.48170	572	3	0.33309	4.16359
43	1.58978	573	3	0.33833	4.22914

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
44	1.70151	575	4	0.34394	4.29926
45	1.81730	576	4	0.34996	4.37456
46	1.93761	578	4	0.35650	4.45619
47	2.06298	579	4	0.36369	4.54609
48	2.19409	581	4	0.37176	4.64696
49	2.33179	583	4	0.38098	4.76223
50	2.47711	585	4	0.39166	4.89575
51	2.63137	586	4	0.40410	5.05119
52	2.79613	589	4	0.41850	5.23119
53	2.95466	590	4	0.43317	5.41463
54	2.95466	590	4	0.43317	5.41463
55	2.95466	590	4	0.43317	5.41463
56	2.95466	590	4	0.43317	5.41463
57	2.95466	590	4	0.43317	5.41463
58	2.95466	590	4	0.43317	5.41463
59	2.95466	590	4	0.43317	5.41463
60	2.95466	590	4	0.43317	5.41463
61	2.95466	590	4	0.43317	5.41463
62	2.95466	590	4	0.43317	5.41463
63	2.95466	590	4	0.43317	5.41463
64	2.95466	590	4	0.43317	5.41463

Table L-28. Raw to Scaled Score Look-up Table—Science Grade 5 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.32534	500	1	1.32239	16.52986
1	-4.16977	501	1	1.22576	15.32202
2	-4.01419	503	1	1.13605	14.20065
3	-3.85862	505	1	1.05290	13.16119
4	-3.17305	514	1	0.75599	9.44987
5	-2.75123	519	1	0.62177	7.77211
6	-2.44226	523	1	0.54308	6.78850
7	-2.19567	526	1	0.49080	6.13499
8	-1.98849	529	1	0.45343	5.66792
9	-1.80833	531	1	0.42544	5.31802
10	-1.64773	533	1	0.40378	5.04728
11	-1.50188	535	1	0.38662	4.83279
12	-1.36747	536	1	0.37277	4.65969
13	-1.24213	538	1	0.36142	4.51770
14	-1.12415	540	1	0.35195	4.39941
15	-1.01219	541	1	0.34394	4.29929
16	-0.90525	542	1	0.33705	4.21307
17	-0.80251	543	1	0.33100	4.13753
18	-0.70334	545	2	0.32562	4.07021
19	-0.60720	546	2	0.32075	4.00940
20	-0.51365	547	2	0.31632	3.95397
21	-0.42232	548	2	0.31227	3.90332
22	-0.33285	549	2	0.30859	3.85733
23	-0.24496	551	2	0.30530	3.81620
24	-0.15835	552	2	0.30243	3.78035
25	-0.07275	553	2	0.30003	3.75033
26	0.01209	554	2	0.29814	3.72672
27	0.09643	555	2	0.29680	3.71000
28	0.18051	556	2	0.29604	3.70056
29	0.26459	557	2	0.29589	3.69861
30	0.34889	558	2	0.29634	3.70422

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
31	0.43367	559	2	0.29738	3.71727
32	0.51914	560	3	0.29900	3.73752
33	0.60553	561	3	0.30117	3.76462
34	0.69308	562	3	0.30385	3.79813
35	0.78199	563	3	0.30700	3.83755
36	0.87248	564	3	0.31059	3.88235
37	0.96478	566	3	0.31456	3.93194
38	1.05908	567	3	0.31886	3.98570
39	1.15561	568	3	0.32344	4.04298
40	1.25459	569	3	0.32825	4.10309
41	1.35621	571	3	0.33323	4.16538
42	1.46072	572	3	0.33834	4.22927
43	1.56834	573	3	0.34355	4.29436
44	1.67933	573	3	0.34885	4.36060
45	1.79396	576	4	0.35427	4.42840
46	1.91258	577	4	0.35990	4.49879
47	2.03557	579	4	0.36588	4.57352
48	2.16346	581	4	0.37241	4.65508
49	2.29687	582	4	0.37972	4.74649
50	2.43661	584	4	0.38809	4.85109
51	2.58366	586	4	0.39777	4.97206
52	2.73925	588	4	0.40895	5.11184
53	2.90483	589	4	0.42173	5.27163
54	2.95466	590	4	0.42572	5.32145
55	2.95466	590	4	0.42572	5.32145
56	2.95466	590	4	0.42572	5.32145
57	2.95466	590	4	0.42572	5.32145
58	2.95466	590	4	0.42572	5.32145
59	2.95466	590	4	0.42572	5.32145
60	2.95466	590	4	0.42572	5.32145
61	2.95466	590	4	0.42572	5.32145
62	2.95466	590	4	0.42572	5.32145
63	2.95466	590	4	0.42572	5.32145
64	2.95466	590	4	0.42572	5.32145

Table L-29. Raw to Scaled Score Look-up Table—Science Grade 8 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-5.56012	800	1	2.07892	20.78924
1	-5.34708	802	1	1.91345	19.13454
2	-5.13404	804	1	1.75961	17.59608
3	-3.83625	817	1	1.04165	10.41645
4	-3.18655	823	1	0.80116	8.01156
5	-2.74746	828	1	0.67527	6.75269
6	-2.41220	831	1	0.59661	5.96612
7	-2.13831	834	1	0.54246	5.42457
8	-1.90474	836	1	0.50280	5.02802
9	-1.69952	838	1	0.47249	4.72485
10	-1.51525	840	1	0.44855	4.48547
11	-1.34704	842	1	0.42916	4.29157
12	-1.19151	843	1	0.41311	4.13115
13	-1.04623	844	1	0.39961	3.99607
14	-0.90938	846	2	0.38806	3.88063
15	-0.77959	847	2	0.37808	3.78077
16	-0.65578	849	2	0.36935	3.69352
17	-0.53708	850	2	0.36167	3.61672
18	-0.42279	851	2	0.35488	3.54875

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
19	-0.31231	852	2	0.34884	3.48837
20	-0.20515	853	2	0.34346	3.43455
21	-0.10086	854	2	0.33865	3.38647
22	0.00093	855	2	0.33434	3.34340
23	0.10059	856	2	0.33047	3.30466
24	0.19840	857	2	0.32697	3.26969
25	0.29464	858	2	0.32380	3.23801
26	0.38958	859	2	0.32093	3.20929
27	0.48345	859	2	0.31834	3.18337
28	0.57647	861	3	0.31602	3.16024
29	0.66885	862	3	0.31401	3.14009
30	0.76081	863	3	0.31232	3.12325
31	0.85255	864	3	0.31101	3.11013
32	0.94429	865	3	0.31012	3.10122
33	1.03626	865	3	0.30970	3.09701
34	1.12869	866	3	0.30980	3.09800
35	1.22183	867	3	0.31046	3.10464
36	1.31593	868	3	0.31174	3.11735
37	1.41128	869	3	0.31365	3.13649
38	1.50816	870	3	0.31623	3.16232
39	1.60690	871	3	0.31950	3.19503
40	1.70783	872	3	0.32346	3.23461
41	1.81128	873	3	0.32809	3.28089
42	1.91762	874	3	0.33335	3.33351
43	2.02722	875	3	0.33919	3.39192
44	2.14046	877	3	0.34556	3.45558
45	2.25775	878	3	0.35241	3.52410
46	2.37953	879	3	0.35976	3.59757
47	2.50629	880	3	0.36768	3.67680
48	2.63860	881	3	0.37634	3.76342
49	2.77716	883	4	0.38596	3.85961
50	2.92281	884	4	0.39674	3.96738
51	3.07655	886	4	0.40873	4.08731
52	3.23948	887	4	0.42173	4.21728
53	3.41276	889	4	0.43521	4.35205
54	3.53988	890	4	0.44452	4.44525
55	3.53988	890	4	0.44452	4.44525
56	3.53988	890	4	0.44452	4.44525
57	3.53988	890	4	0.44452	4.44525
58	3.53988	890	4	0.44452	4.44525
59	3.53988	890	4	0.44452	4.44525
60	3.53988	890	4	0.44452	4.44525
61	3.53988	890	4	0.44452	4.44525
62	3.53988	890	4	0.44452	4.44525
63	3.53988	890	4	0.44452	4.44525
64	3.53988	890	4	0.44452	4.44525

Table L-30. Raw to Scaled Score Look-up Table—Science Grade 8 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-5.56012	800	1	1.92916	19.29158
1	-5.42928	801	1	1.83121	18.31212
2	-5.29845	802	1	1.73778	17.37780
3	-3.97572	815	1	1.02957	10.29570
4	-3.29608	822	1	0.80823	8.08235
5	-2.82408	827	1	0.69122	6.91220
6	-2.45895	831	1	0.61289	6.12890
7	-2.16102	833	1	0.55457	5.54570
8	-1.90956	836	1	0.50916	5.09158
9	-1.69198	838	1	0.47302	4.73018
10	-1.49991	840	1	0.44385	4.43854
11	-1.32755	842	1	0.42007	4.20067
12	-1.17071	843	1	0.40050	4.00500
13	-1.02629	844	1	0.38429	3.84292
14	-0.89195	846	2	0.37079	3.70792
15	-0.76586	847	2	0.35950	3.59495
16	-0.64659	849	2	0.35000	3.50004
17	-0.53299	850	2	0.34200	3.41999
18	-0.42413	851	2	0.33522	3.35222
19	-0.31924	852	2	0.32946	3.29464
20	-0.21768	853	2	0.32456	3.24558
21	-0.11891	854	2	0.32036	3.20363
22	-0.02248	855	2	0.31677	3.16769
23	0.07202	856	2	0.31368	3.13682
24	0.16493	857	2	0.31103	3.11028
25	0.25655	858	2	0.30874	3.08743
26	0.34714	859	2	0.30678	3.06779
27	0.43694	859	2	0.30510	3.05100
28	0.52616	860	3	0.30368	3.03684
29	0.61502	861	3	0.30252	3.02523
30	0.70369	862	3	0.30162	3.01623
31	0.79235	863	3	0.30100	3.01004
32	0.88119	864	3	0.30069	3.00695
33	0.97039	865	3	0.30073	3.00733
34	1.06015	866	3	0.30116	3.01161
35	1.15067	867	3	0.30202	3.02025
36	1.24217	868	3	0.30337	3.03370
37	1.33488	868	3	0.30524	3.05243
38	1.42907	869	3	0.30769	3.07686
39	1.52502	870	3	0.31074	3.10737
40	1.62303	871	3	0.31442	3.14423
41	1.72342	872	3	0.31876	3.18758
42	1.82657	873	3	0.32374	3.23739
43	1.93283	874	3	0.32935	3.29345
44	2.04261	876	3	0.33554	3.35543
45	2.15634	877	3	0.34230	3.42298
46	2.27446	878	3	0.34959	3.49594
47	2.39751	879	3	0.35747	3.57467
48	2.52605	880	3	0.36603	3.66027
49	2.66079	881	3	0.37547	3.75471
50	2.80260	883	4	0.38605	3.86054
51	2.95251	885	4	0.39800	3.98003
52	3.11176	886	4	0.41138	4.11376
53	3.28177	888	4	0.42591	4.25907
54	3.46403	889	4	0.44098	4.40979
55	3.53988	890	4	0.44694	4.46942

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
56	3.53988	890	4	0.44694	4.46942
57	3.53988	890	4	0.44694	4.46942
58	3.53988	890	4	0.44694	4.46942
59	3.53988	890	4	0.44694	4.46942
60	3.53988	890	4	0.44694	4.46942
61	3.53988	890	4	0.44694	4.46942
62	3.53988	890	4	0.44694	4.46942
63	3.53988	890	4	0.44694	4.46942
64	3.53988	890	4	0.44694	4.46942

Table L-31. Raw to Scaled Score Look-up Table—Science Grade 8 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-5.56012	800	1	2.10570	21.05695
1	-5.26709	802	1	1.88839	18.88387
2	-4.97407	805	1	1.68875	16.88754
3	-4.68104	808	1	1.50647	15.06469
4	-4.38801	811	1	1.34108	13.41083
5	-3.49981	820	1	0.93684	9.36841
6	-2.95747	826	1	0.75525	7.55253
7	-2.56496	829	1	0.65097	6.50967
8	-2.25462	833	1	0.58300	5.83002
9	-1.99565	835	1	0.53516	5.35158
10	-1.77157	837	1	0.49967	4.99668
11	-1.57263	839	1	0.47230	4.72300
12	-1.39260	841	1	0.45052	4.50525
13	-1.22730	843	1	0.43273	4.32730
14	-1.07381	844	1	0.41783	4.17833
15	-0.93001	846	2	0.40509	4.05086
16	-0.79431	847	2	0.39395	3.93955
17	-0.66551	848	2	0.38406	3.84060
18	-0.54264	850	2	0.37513	3.75126
19	-0.42493	851	2	0.36696	3.66959
20	-0.31175	852	2	0.35942	3.59420
21	-0.20254	853	2	0.35241	3.52413
22	-0.09682	854	2	0.34587	3.45870
23	0.00582	855	2	0.33975	3.39748
24	0.10578	856	2	0.33402	3.34019
25	0.20340	857	2	0.32867	3.28666
26	0.29899	858	2	0.32369	3.23687
27	0.39287	859	2	0.31909	3.19091
28	0.48531	859	2	0.31489	3.14894
29	0.57658	861	3	0.31112	3.11124
30	0.66693	862	3	0.30781	3.07813
31	0.75664	863	3	0.30500	3.04997
32	0.84593	864	3	0.30271	3.02715
33	0.93509	864	3	0.30100	3.01003
34	1.02435	865	3	0.29990	2.99902
35	1.11400	866	3	0.29945	2.99446
36	1.20430	867	3	0.29967	2.99674
37	1.29556	868	3	0.30062	3.00622
38	1.38808	869	3	0.30232	3.02323
39	1.48219	870	3	0.30481	3.04808
40	1.57822	871	3	0.30810	3.08102
41	1.67656	872	3	0.31222	3.12217
42	1.77759	873	3	0.31715	3.17145
43	1.88170	874	3	0.32286	3.22861

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
44	1.98933	875	3	0.32931	3.29315
45	2.10090	876	3	0.33644	3.36441
46	2.21686	877	3	0.34418	3.44176
47	2.33769	878	3	0.35248	3.52485
48	2.46393	880	3	0.36139	3.61395
49	2.59619	881	3	0.37103	3.71026
50	2.73520	882	4	0.38158	3.81584
51	2.88190	884	4	0.39331	3.93308
52	3.03739	885	4	0.40635	4.06350
53	3.20300	887	4	0.42061	4.20609
54	3.38016	889	4	0.43562	4.35622
55	3.53988	890	4	0.44839	4.48395
56	3.53988	890	4	0.44839	4.48395
57	3.53988	890	4	0.44839	4.48395
58	3.53988	890	4	0.44839	4.48395
59	3.53988	890	4	0.44839	4.48395
60	3.53988	890	4	0.44839	4.48395
61	3.53988	890	4	0.44839	4.48395
62	3.53988	890	4	0.44839	4.48395
63	3.53988	890	4	0.44839	4.48395
64	3.53988	890	4	0.44839	4.48395

Table L-32. Raw to Scaled Score Look-up Table—Science Grade 11 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-8.02951	1100	1	5.31211	39.84084
1	-6.91348	1108	1	3.55037	26.62779
2	-5.79744	1116	1	2.31968	17.39761
3	-4.68141	1125	1	1.48517	11.13877
4	-3.64207	1132	1	0.96877	7.26579
5	-3.07169	1137	1	0.76619	5.74644
6	-2.67515	1140	1	0.65276	4.89569
7	-2.36915	1142	1	0.57881	4.34106
8	-2.11840	1144	1	0.52633	3.94744
9	-1.90468	1145	1	0.48703	3.65270
10	-1.71735	1147	1	0.45651	3.42382
11	-1.54966	1148	1	0.43220	3.24146
12	-1.39704	1149	1	0.41246	3.09347
13	-1.25629	1150	1	0.39623	2.97175
14	-1.12504	1151	1	0.38276	2.87071
15	-1.00151	1152	1	0.37151	2.78631
16	-0.88432	1153	1	0.36207	2.71554
17	-0.77236	1153	1	0.35414	2.65608
18	-0.66477	1155	2	0.34749	2.60614
19	-0.56082	1156	2	0.34190	2.56424
20	-0.45990	1156	2	0.33723	2.52920
21	-0.36153	1157	2	0.33333	2.50000
22	-0.26528	1158	2	0.33011	2.47580
23	-0.17078	1158	2	0.32745	2.45590
24	-0.07772	1159	2	0.32529	2.43969
25	0.01415	1159	2	0.32356	2.42668
26	0.10509	1161	3	0.32219	2.41646
27	0.19528	1161	3	0.32116	2.40870
28	0.28494	1162	3	0.32042	2.40316
29	0.37420	1163	3	0.31995	2.39965
30	0.46324	1163	3	0.31974	2.39803
31	0.55219	1164	3	0.31976	2.39821

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
32	0.64117	1165	3	0.32001	2.40011
33	0.73033	1165	3	0.32049	2.40367
34	0.81977	1166	3	0.32118	2.40884
35	0.90963	1167	3	0.32207	2.41555
36	1.00001	1167	3	0.32317	2.42374
37	1.09105	1168	3	0.32445	2.43337
38	1.18288	1169	3	0.32592	2.44441
39	1.27564	1169	3	0.32758	2.45688
40	1.36949	1170	3	0.32945	2.47088
41	1.46460	1171	3	0.33155	2.48664
42	1.56119	1171	3	0.33393	2.50449
43	1.65948	1172	3	0.33666	2.52495
44	1.75977	1173	3	0.33983	2.54869
45	1.86238	1174	3	0.34354	2.57655
46	1.96771	1174	3	0.34794	2.60951
47	2.07620	1175	3	0.35315	2.64865
48	2.18840	1176	3	0.35934	2.69508
49	2.30494	1177	3	0.36665	2.74984
50	2.42653	1178	3	0.37517	2.81377
51	2.55398	1179	3	0.38497	2.88730
52	2.68819	1180	3	0.39603	2.97022
53	2.83012	1180	3	0.40819	3.06142
54	2.98079	1182	4	0.42117	3.15875
55	3.14122	1183	4	0.43458	3.25934
56	3.31247	1185	4	0.44810	3.36075
57	3.49569	1186	4	0.46177	3.46327
58	3.69243	1187	4	0.47639	3.57289
59	3.90510	1189	4	0.49390	3.70422
60	4.10383	1190	4	0.51380	3.85351
61	4.10383	1190	4	0.51380	3.85351
62	4.10383	1190	4	0.51380	3.85351
63	4.10383	1190	4	0.51380	3.85351
64	4.10383	1190	4	0.51380	3.85351
65	4.10383	1190	4	0.51380	3.85351
66	4.10383	1190	4	0.51380	3.85351
67	4.10383	1190	4	0.51380	3.85351
68	4.10383	1190	4	0.51380	3.85351

Table L-33. Raw to Scaled Score Look-up Table—Science Grade 11 Operational Form 2

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-8.02951	1100	1	4.40007	33.00052
1	-7.08498	1107	1	3.21920	24.14397
2	-6.14046	1114	1	2.32964	17.47227
3	-5.19594	1121	1	1.66965	12.52234
4	-4.02164	1130	1	1.09278	8.19588
5	-3.37354	1134	1	0.86236	6.46772
6	-2.92309	1138	1	0.73079	5.48095
7	-2.57685	1140	1	0.64320	4.82397
8	-2.29498	1143	1	0.57977	4.34824
9	-2.05671	1144	1	0.53145	3.98584
10	-1.84980	1146	1	0.49343	3.70070
11	-1.66635	1147	1	0.46287	3.47156
12	-1.50099	1148	1	0.43798	3.28483
13	-1.34987	1150	1	0.41750	3.13122
14	-1.21015	1151	1	0.40055	3.00410

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
15	-1.07964	1152	1	0.38647	2.89850
16	-0.95668	1153	1	0.37475	2.81060
17	-0.83992	1153	1	0.36499	2.73739
18	-0.72829	1154	2	0.35686	2.67646
19	-0.62093	1155	2	0.35011	2.62582
20	-0.51710	1156	2	0.34451	2.58384
21	-0.41621	1157	2	0.33988	2.54913
22	-0.31778	1157	2	0.33607	2.52056
23	-0.22136	1158	2	0.33295	2.49714
24	-0.12661	1159	2	0.33041	2.47807
25	-0.03321	1159	2	0.32836	2.46266
26	0.05909	1160	3	0.32672	2.45039
27	0.15054	1161	3	0.32545	2.44084
28	0.24133	1162	3	0.32449	2.43370
29	0.33166	1162	3	0.32384	2.42876
30	0.42170	1163	3	0.32345	2.42590
31	0.51159	1164	3	0.32334	2.42505
32	0.60149	1164	3	0.32349	2.42617
33	0.69154	1165	3	0.32390	2.42928
34	0.78189	1166	3	0.32458	2.43435
35	0.87266	1166	3	0.32552	2.44139
36	0.96399	1167	3	0.32672	2.45037
37	1.05604	1168	3	0.32816	2.46123
38	1.14893	1168	3	0.32986	2.47394
39	1.24283	1169	3	0.33180	2.48848
40	1.33790	1170	3	0.33398	2.50486
41	1.43433	1170	3	0.33643	2.52321
42	1.53232	1171	3	0.33917	2.54378
43	1.63210	1172	3	0.34226	2.56697
44	1.73396	1173	3	0.34579	2.59339
45	1.83821	1174	3	0.34984	2.62382
46	1.94522	1174	3	0.35456	2.65921
47	2.05544	1175	3	0.36008	2.70063
48	2.16937	1176	3	0.36656	2.74918
49	2.28762	1177	3	0.37412	2.80587
50	2.41085	1178	3	0.38286	2.87146
51	2.53980	1179	3	0.39283	2.94619
52	2.67531	1180	3	0.40394	3.02956
53	2.81821	1180	3	0.41600	3.12003
54	2.96938	1182	4	0.42866	3.21498
55	3.12968	1183	4	0.44148	3.31113
56	3.29996	1184	4	0.45412	3.40588
57	3.48118	1186	4	0.46661	3.49958
58	3.67465	1187	4	0.47978	3.59835
59	3.88251	1189	4	0.49551	3.71633
60	4.10383	1190	4	0.51646	3.87348
61	4.10383	1190	4	0.51646	3.87348
62	4.10383	1190	4	0.51646	3.87348
63	4.10383	1190	4	0.51646	3.87348
64	4.10383	1190	4	0.51646	3.87348
65	4.10383	1190	4	0.51646	3.87348
66	4.10383	1190	4	0.51646	3.87348
67	4.10383	1190	4	0.51646	3.87348
68	4.10383	1190	4	0.51646	3.87348

Table L-34. Raw to Scaled Score Look-up Table—Science Grade 11 Operational Form 3

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-8.02951	1100	1	5.21388	39.10409
1	-7.25571	1105	1	3.95656	29.67424
2	-6.48192	1111	1	2.98278	22.37088
3	-5.70813	1117	1	2.23393	16.75448
4	-4.01005	1130	1	1.15594	8.66954
5	-3.27995	1135	1	0.86169	6.46268
6	-2.81009	1139	1	0.71221	5.34157
7	-2.46237	1141	1	0.61944	4.64577
8	-2.18512	1143	1	0.55570	4.16772
9	-1.95339	1145	1	0.50913	3.81847
10	-1.75323	1147	1	0.47369	3.55267
11	-1.57605	1148	1	0.44593	3.34447
12	-1.41623	1149	1	0.42373	3.17794
13	-1.26987	1150	1	0.40569	3.04268
14	-1.13417	1151	1	0.39087	2.93154
15	-1.00704	1152	1	0.37859	2.83943
16	-0.88690	1153	1	0.36835	2.76259
17	-0.77251	1153	1	0.35976	2.69817
18	-0.66289	1155	2	0.35252	2.64394
19	-0.55723	1156	2	0.34641	2.59811
20	-0.45491	1156	2	0.34123	2.55925
21	-0.35538	1157	2	0.33682	2.52617
22	-0.25819	1158	2	0.33305	2.49790
23	-0.16297	1158	2	0.32982	2.47366
24	-0.06940	1159	2	0.32704	2.45282
25	0.02279	1159	2	0.32465	2.43490
26	0.11384	1161	3	0.32261	2.41955
27	0.20398	1161	3	0.32087	2.40653
28	0.29338	1162	3	0.31943	2.39570
29	0.38222	1163	3	0.31826	2.38698
30	0.47065	1163	3	0.31738	2.38035
31	0.55884	1164	3	0.31677	2.37577
32	0.64692	1165	3	0.31643	2.37325
33	0.73502	1165	3	0.31637	2.37275
34	0.82329	1166	3	0.31656	2.37422
35	0.91185	1167	3	0.31701	2.37756
36	1.00083	1167	3	0.31769	2.38267
37	1.09036	1168	3	0.31859	2.38945
38	1.18059	1169	3	0.31971	2.39782
39	1.27166	1169	3	0.32104	2.40778
40	1.36373	1170	3	0.32259	2.41939
41	1.45698	1171	3	0.32438	2.43288
42	1.55162	1171	3	0.32648	2.44862
43	1.64788	1172	3	0.32895	2.46715
44	1.74605	1173	3	0.33189	2.48918
45	1.84645	1174	3	0.33542	2.51562
46	1.94948	1174	3	0.33966	2.54746
47	2.05560	1175	3	0.34478	2.58584
48	2.16534	1176	3	0.35092	2.63188
49	2.27932	1177	3	0.35823	2.68669
50	2.39825	1178	3	0.36682	2.75118
51	2.52295	1179	3	0.37679	2.82589
52	2.65429	1180	3	0.38810	2.91076
53	2.79323	1180	3	0.40064	3.00481
54	2.94078	1182	4	0.41413	3.10596
55	3.09794	1183	4	0.42815	3.21111

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
56	3.26572	1184	4	0.44228	3.31712
57	3.44517	1186	4	0.45637	3.42279
58	3.63764	1187	4	0.47092	3.53191
59	3.84516	1189	4	0.48751	3.65634
60	4.07129	1189	4	0.50910	3.81823
61	4.10383	1190	4	0.51265	3.84491
62	4.10383	1190	4	0.51265	3.84491
63	4.10383	1190	4	0.51265	3.84491
64	4.10383	1190	4	0.51265	3.84491
65	4.10383	1190	4	0.51265	3.84491
66	4.10383	1190	4	0.51265	3.84491
67	4.10383	1190	4	0.51265	3.84491
68	4.10383	1190	4	0.51265	3.84491

Table L-35. Raw to Scaled Score Look-up Table—NM-MSSA SLA Grade 3 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	3.86412	77.28239
1	-3.79036	300	1	3.31799	66.35974
2	-3.58072	300	1	2.84166	56.83325
3	-3.37108	300	1	2.42779	48.55584
4	-3.16144	300	1	2.06987	41.39737
5	-2.95180	300	1	1.76209	35.24171
6	-2.74216	300	1	1.49913	29.98266
7	-1.93020	314	1	0.80935	16.18696
8	-1.53860	322	1	0.60102	12.02049
9	-1.27501	327	1	0.49034	9.80675
10	-1.07276	331	1	0.42240	8.44802
11	-0.90578	335	1	0.37747	7.54931
12	-0.76135	338	2	0.34570	6.91401
13	-0.63243	340	2	0.32198	6.43951
14	-0.51470	343	2	0.30364	6.07274
15	-0.40528	345	2	0.28929	5.78587
16	-0.30209	347	2	0.27815	5.56299
17	-0.20356	349	2	0.26967	5.39344
18	-0.10846	351	2	0.26344	5.26871
19	-0.01578	353	2	0.25906	5.18120
20	0.07530	354	2	0.25618	5.12369
21	0.16543	356	2	0.25447	5.08933
22	0.25516	358	2	0.25360	5.07201
23	0.34497	359	2	0.25335	5.06709
24	0.43528	362	3	0.25362	5.07230
25	0.52652	363	3	0.25441	5.08828
26	0.61914	365	3	0.25592	5.11846
27	0.71371	367	3	0.25841	5.16818
28	0.81089	369	3	0.26217	5.24336
29	0.91148	371	4	0.26747	5.34944
30	1.01641	373	4	0.27455	5.49093
31	1.12680	375	4	0.28359	5.67179
32	1.24400	378	4	0.29482	5.89648
33	1.36962	380	4	0.30858	6.17152
34	1.50575	383	4	0.32535	6.50705
35	1.65507	386	4	0.34591	6.91811
36	1.82121	389	4	0.37127	7.42548
37	2.00913	389	4	0.40279	8.05590
38	2.22585	389	4	0.44203	8.84060

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
39	2.48170	389	4	0.49057	9.81147
40	2.79330	389	4	0.55067	11.01339
41	3.19427	389	4	0.63404	12.68076
42	3.79553	389	4	0.82046	16.40918
43	4.00000	390	4	0.91485	18.29710

Table L-36. Raw to Scaled Score Look-up Table—NM-MSSA SLA Grade 4 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	400	1	2.46862	49.37241
1	-3.86031	400	1	2.31628	46.32563
2	-3.72062	400	1	2.17144	43.42871
3	-3.58093	400	1	2.03363	40.67253
4	-3.44124	400	1	1.90240	38.04809
5	-3.30155	400	1	1.77732	35.54647
6	-3.16187	400	1	1.65795	33.15891
7	-2.27859	412	1	1.00576	20.11520
8	-1.83229	420	1	0.71963	14.39255
9	-1.53789	426	1	0.55535	11.10706
10	-1.31715	431	1	0.45816	9.16319
11	-1.13792	434	1	0.39848	7.96956
12	-0.98435	437	1	0.35997	7.19948
13	-0.84773	439	1	0.33395	6.67903
14	-0.72283	443	2	0.31569	6.31387
15	-0.60633	445	2	0.30254	6.05071
16	-0.49599	447	2	0.29290	5.85797
17	-0.39019	449	2	0.28580	5.71602
18	-0.28773	451	2	0.28061	5.61226
19	-0.18769	453	2	0.27692	5.53845
20	-0.08929	455	2	0.27446	5.48928
21	0.00812	457	2	0.27307	5.46139
22	0.10514	459	2	0.27263	5.45266
23	0.20233	461	3	0.27309	5.46176
24	0.30024	463	3	0.27439	5.48781
25	0.39940	465	3	0.27651	5.53026
26	0.50040	467	3	0.27945	5.58894
27	0.60386	469	3	0.28321	5.66425
28	0.71045	471	3	0.28788	5.75757
29	0.82098	474	4	0.29358	5.87165
30	0.93640	476	4	0.30056	6.01117
31	1.05793	478	4	0.30917	6.18332
32	1.18711	481	4	0.31992	6.39834
33	1.32594	484	4	0.33352	6.67037
34	1.47715	487	4	0.35092	7.01841
35	1.64441	489	4	0.37337	7.46745
36	1.83276	489	4	0.40241	8.04814
37	2.04918	489	4	0.43949	8.78989
38	2.30298	489	4	0.48482	9.69640
39	2.60621	489	4	0.53612	10.72241
40	2.97737	489	4	0.59656	11.93116
41	3.46741	489	4	0.70932	14.18634
42	4.00000	490	4	0.93266	18.65328
43	4.00000	490	4	0.93266	18.65328

Table L-37. Raw to Scaled Score Look-up Table—NM-MSSA SLA Grade 5 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	500	1	2.67395	53.47891
1	-3.79056	500	1	2.38007	47.60141
2	-3.58112	500	1	2.10966	42.19328
3	-3.37167	500	1	1.86086	37.21714
4	-3.16223	500	1	1.63218	32.64362
5	-2.95279	500	1	1.42264	28.45283
6	-2.74335	502	1	1.23174	24.63484
7	-2.15160	514	1	0.79335	15.86703
8	-1.80104	521	1	0.60443	12.08863
9	-1.54917	526	1	0.50001	10.00021
10	-1.34993	530	1	0.43479	8.69582
11	-1.18271	533	1	0.39108	7.82169
12	-1.03658	536	1	0.36052	7.21031
13	-0.90499	539	1	0.33859	6.77179
14	-0.78376	541	1	0.32267	6.45341
15	-0.67000	543	2	0.31109	6.22185
16	-0.56167	546	2	0.30274	6.05475
17	-0.45725	548	2	0.29681	5.93614
18	-0.35563	550	2	0.29270	5.85396
19	-0.25594	552	2	0.28993	5.79861
20	-0.15753	554	2	0.28810	5.76209
21	-0.05993	556	2	0.28688	5.73759
22	0.03725	558	2	0.28597	5.71931
23	0.13427	559	2	0.28513	5.70266
24	0.23136	561	3	0.28423	5.68460
25	0.32872	563	3	0.28321	5.66416
26	0.42657	565	3	0.28214	5.64284
27	0.52520	567	3	0.28124	5.62477
28	0.62498	569	3	0.28082	5.61630
29	0.72645	571	3	0.28127	5.62537
30	0.83031	572	3	0.28303	5.66062
31	0.93746	576	4	0.28653	5.73063
32	1.04906	578	4	0.29216	5.84326
33	1.16654	580	4	0.30026	6.00528
34	1.29165	583	4	0.31111	6.22212
35	1.42659	585	4	0.32494	6.49880
36	1.57417	588	4	0.34217	6.84333
37	1.73837	589	4	0.36380	7.27592
38	1.92542	589	4	0.39240	7.84795
39	2.14648	589	4	0.43397	8.67947
40	2.42432	589	4	0.50248	10.04951
41	2.81425	589	4	0.63530	12.70595
42	3.50392	589	4	0.98016	19.60328
43	4.00000	590	4	1.31516	26.30327

Table L-38. Raw to Scaled Score Look-up Table—NM-MSSA SLA Grade 6 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	600	1	2.33683	46.73660
1	-3.80157	600	1	2.09547	41.90932
2	-3.60313	600	1	1.87377	37.47536
3	-3.40470	600	1	1.67137	33.42739
4	-3.20627	600	1	1.48768	29.75361
5	-3.00784	600	1	1.32175	26.43494
6	-2.80940	600	1	1.17225	23.44502

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
7	-2.23035	610	1	0.81674	16.33484
8	-1.85864	618	1	0.64398	12.87950
9	-1.58115	623	1	0.54082	10.81633
10	-1.35630	628	1	0.47210	9.44200
11	-1.16460	631	1	0.42293	8.45851
12	-0.99559	635	2	0.38560	7.71200
13	-0.84322	638	2	0.35570	7.11393
14	-0.70382	641	2	0.33074	6.61481
15	-0.57493	643	2	0.30960	6.19205
16	-0.45476	646	2	0.29199	5.83987
17	-0.34171	648	2	0.27799	5.55976
18	-0.23427	650	2	0.26766	5.35317
19	-0.13097	652	2	0.26089	5.21772
20	-0.03042	654	2	0.25729	5.14589
21	0.06858	656	2	0.25624	5.12489
22	0.16698	658	2	0.25686	5.13712
23	0.26542	659	2	0.25810	5.16204
24	0.36421	662	3	0.25902	5.18040
25	0.46341	664	3	0.25903	5.18056
26	0.56298	666	3	0.25816	5.16318
27	0.66296	668	3	0.25704	5.14080
28	0.76364	670	3	0.25660	5.13190
29	0.86558	672	3	0.25765	5.15294
30	0.96964	674	4	0.26057	5.21145
31	1.07680	676	4	0.26508	5.30151
32	1.18798	678	4	0.27018	5.40363
33	1.30394	681	4	0.27475	5.49493
34	1.42540	683	4	0.27858	5.57157
35	1.55359	686	4	0.28344	5.66873
36	1.69124	689	4	0.29303	5.86057
37	1.84385	689	4	0.31259	6.25172
38	2.02150	689	4	0.34936	6.98724
39	2.24209	689	4	0.41443	8.28857
40	2.53811	689	4	0.52527	10.50539
41	2.97606	689	4	0.71408	14.28160
42	3.75528	689	4	1.10795	22.15896
43	4.00000	690	4	1.25473	25.09463

Table L-39. Raw to Scaled Score Look-up Table—NM-MSSA SLA Grade 7 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	700	1	2.51124	50.22479
1	-3.82903	700	1	2.28043	45.60857
2	-3.65805	700	1	2.06415	41.28310
3	-3.48708	700	1	1.86194	37.23884
4	-3.31611	700	1	1.67317	33.46336
5	-3.14514	700	1	1.49711	29.94230
6	-2.97416	700	1	1.33309	26.66172
7	-2.80319	700	1	1.18061	23.61217
8	-2.23337	711	1	0.76083	15.21670
9	-1.88377	718	1	0.58427	11.68535
10	-1.62305	723	1	0.49567	9.91337
11	-1.40804	727	1	0.44658	8.93152
12	-1.21973	730	1	0.41773	8.35457
13	-1.04832	735	2	0.40022	8.00447
14	-0.88821	738	2	0.38936	7.78716

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
15	-0.73601	741	2	0.38229	7.64572
16	-0.58958	744	2	0.37705	7.54097
17	-0.44758	747	2	0.37223	7.44460
18	-0.30917	749	2	0.36694	7.33876
19	-0.17376	752	2	0.36089	7.21775
20	-0.04089	755	2	0.35437	7.08748
21	0.08994	757	2	0.34802	6.96043
22	0.21922	760	3	0.34238	6.84767
23	0.34748	763	3	0.33762	6.75243
24	0.47516	765	3	0.33344	6.66872
25	0.60261	768	3	0.32934	6.58684
26	0.73009	770	3	0.32518	6.50353
27	0.85796	773	3	0.32149	6.42972
28	0.98680	774	3	0.31942	6.38834
29	1.11759	778	4	0.32018	6.40353
30	1.25169	781	4	0.32451	6.49026
31	1.39070	783	4	0.33256	6.65125
32	1.53643	786	4	0.34407	6.88133
33	1.69091	789	4	0.35876	7.17511
34	1.85649	789	4	0.37656	7.53128
35	2.03603	789	4	0.39747	7.94947
36	2.23301	789	4	0.42124	8.42471
37	2.45191	789	4	0.44784	8.95690
38	2.69957	789	4	0.47974	9.59490
39	2.98932	789	4	0.52611	10.52217
40	3.35242	789	4	0.61113	12.22269
41	3.87482	789	4	0.80463	16.09259
42	4.00000	790	4	0.86408	17.28159
43	4.00000	790	4	0.86408	17.28159

Table L-40. Raw to Scaled Score Look-up Table—NM-MSSA SLA Grade 8 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	800	1	2.08010	41.60199
1	-3.84657	800	1	1.92775	38.55510
2	-3.69315	800	1	1.78288	35.65755
3	-3.53972	800	1	1.64504	32.90072
4	-3.38629	800	1	1.51400	30.28000
5	-3.23286	800	1	1.38974	27.79474
6	-3.07944	800	1	1.27237	25.44744
7	-2.42667	809	1	0.85501	17.10011
8	-2.02683	817	1	0.66593	13.31857
9	-1.73596	822	1	0.55721	11.14430
10	-1.50525	827	1	0.48609	9.72181
11	-1.31231	831	1	0.43579	8.71579
12	-1.14499	834	1	0.39847	7.96939
13	-0.99589	837	1	0.37000	7.39997
14	-0.86015	839	1	0.34799	6.95986
15	-0.73432	842	2	0.33097	6.61942
16	-0.61586	845	2	0.31794	6.35889
17	-0.50283	847	2	0.30821	6.16424
18	-0.39367	849	2	0.30124	6.02488
19	-0.28710	851	2	0.29662	5.93230
20	-0.18208	853	2	0.29396	5.87924
21	-0.07771	856	2	0.29296	5.85914
22	0.02674	858	2	0.29330	5.86592

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
23	0.13190	859	2	0.29469	5.89389
24	0.23829	862	3	0.29689	5.93785
25	0.34635	864	3	0.29967	5.99337
26	0.45643	866	3	0.30286	6.05714
27	0.56888	868	3	0.30637	6.12744
28	0.68404	870	3	0.31023	6.20450
29	0.80228	873	4	0.31453	6.29060
30	0.92413	876	4	0.31949	6.38982
31	1.05026	878	4	0.32537	6.50736
32	1.18157	881	4	0.33243	6.64857
33	1.31926	884	4	0.34091	6.81830
34	1.46489	886	4	0.35106	7.02113
35	1.62057	889	4	0.36321	7.26414
36	1.78924	889	4	0.37821	7.56411
37	1.97546	889	4	0.39807	7.96144
38	2.18706	889	4	0.42728	8.54551
39	2.43901	889	4	0.47541	9.50815
40	2.76362	889	4	0.56481	11.29619
41	3.24396	889	4	0.75898	15.17951
42	4.00000	890	4	1.19471	23.89413
43	4.00000	890	4	1.19471	23.89413

Table L-41. Raw to Scaled Score Look-up Table—NM-MSSA Mathematics (Spanish) Grade 3 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	300	1	2.73878	47.92865
1	-3.88661	300	1	2.58572	45.25005
2	-3.77322	300	1	2.43497	42.61203
3	-3.65983	300	1	2.28660	40.01551
4	-3.54644	300	1	2.14075	37.46306
5	-3.43305	300	1	1.99768	34.95940
6	-3.31966	300	1	1.85781	32.51161
7	-3.20627	300	1	1.72167	30.12929
8	-2.28699	312	1	0.84476	14.78325
9	-1.86103	320	1	0.61086	10.69010
10	-1.57167	325	1	0.50376	8.81583
11	-1.34698	329	1	0.44162	7.72843
12	-1.16026	332	1	0.40059	7.01038
13	-0.99861	335	1	0.37131	6.49800
14	-0.85473	337	1	0.34937	6.11395
15	-0.72406	340	1	0.33236	5.81628
16	-0.60352	340	1	0.31884	5.57975
17	-0.49095	344	2	0.30788	5.38790
18	-0.38474	346	2	0.29882	5.22939
19	-0.28368	347	2	0.29121	5.09610
20	-0.18685	349	2	0.28470	4.98217
21	-0.09349	351	2	0.27905	4.88336
22	-0.00300	352	2	0.27409	4.79664
23	0.08512	354	2	0.26971	4.71997
24	0.17131	355	2	0.26583	4.65205
25	0.25594	357	2	0.26241	4.59218
26	0.33933	358	2	0.25943	4.54010
27	0.42182	359	2	0.25691	4.49587
28	0.50368	361	3	0.25484	4.45977
29	0.58522	363	3	0.25327	4.43222

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
30	0.66673	364	3	0.25221	4.41369
31	0.74850	365	3	0.25169	4.40464
32	0.83086	367	3	0.25174	4.40549
33	0.91413	368	3	0.25238	4.41660
34	0.99869	370	3	0.25362	4.43827
35	1.08491	371	3	0.25548	4.47082
36	1.17325	373	3	0.25798	4.51472
37	1.26419	374	3	0.26119	4.57075
38	1.35832	376	3	0.26516	4.64022
39	1.45632	376	3	0.27001	4.72519
40	1.55902	380	4	0.27592	4.82864
41	1.66744	382	4	0.28313	4.95472
42	1.78292	384	4	0.29196	5.10922
43	1.90725	386	4	0.30289	5.30064
44	2.04290	388	4	0.31671	5.54249
45	2.19357	389	4	0.33474	5.85791
46	2.36515	389	4	0.35937	6.28896
47	2.56791	389	4	0.39531	6.91797
48	2.82207	389	4	0.45298	7.92723
49	3.17590	389	4	0.56127	9.82219
50	3.79836	389	4	0.84810	14.84176
51	4.00000	390	4	0.97136	16.99874

Table L-42. Raw to Scaled Score Look-up Table—NM-MSSA Mathematics (Spanish) Grade 4 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	400	1	5.55944	80.00000
1	-3.73211	400	1	4.34682	76.06932
2	-3.46422	400	1	3.39626	59.43447
3	-3.19634	400	1	2.65279	46.42389
4	-2.92845	401	1	2.07283	36.27448
5	-2.66056	406	1	1.62187	28.38268
6	-2.39267	410	1	1.27271	22.27237
7	-2.12479	415	1	1.00394	17.56886
8	-1.52538	426	1	0.61322	10.73131
9	-1.19886	431	1	0.48694	8.52140
10	-0.96504	435	1	0.42250	7.39367
11	-0.77795	439	1	0.38283	6.69950
12	-0.61900	441	1	0.35554	6.22193
13	-0.47892	443	1	0.33520	5.86596
14	-0.35244	446	2	0.31904	5.58320
15	-0.23629	448	2	0.30554	5.34691
16	-0.12829	450	2	0.29383	5.14204
17	-0.02691	452	2	0.28345	4.96044
18	0.06899	453	2	0.27418	4.79810
19	0.16032	455	2	0.26591	4.65336
20	0.24780	457	2	0.25861	4.52562
21	0.33206	458	2	0.25226	4.41454
22	0.41362	459	2	0.24683	4.31955
23	0.49294	461	3	0.24227	4.23971
24	0.57044	462	3	0.23850	4.17370
25	0.64647	464	3	0.23542	4.11993
26	0.72134	465	3	0.23296	4.07672
27	0.79533	466	3	0.23100	4.04245
28	0.86869	467	3	0.22947	4.01568

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
29	0.94164	469	3	0.22830	3.99527
30	1.01440	470	3	0.22745	3.98042
31	1.08719	471	3	0.22690	3.97071
32	1.16022	473	3	0.22664	3.96613
33	1.23374	474	3	0.22669	3.96709
34	1.30801	475	3	0.22711	3.97437
35	1.38333	476	3	0.22796	3.98922
36	1.46006	478	3	0.22933	4.01330
37	1.53864	479	3	0.23136	4.04875
38	1.61958	481	4	0.23419	4.09827
39	1.70354	482	4	0.23801	4.16519
40	1.79133	484	4	0.24307	4.25368
41	1.88400	485	4	0.24966	4.36904
42	1.98294	487	4	0.25819	4.51826
43	2.08999	489	4	0.26920	4.71100
44	2.20776	489	4	0.28350	4.96132
45	2.34004	489	4	0.30237	5.29142
46	2.49276	489	4	0.32799	5.73974
47	2.67611	489	4	0.36465	6.38142
48	2.91001	489	4	0.42223	7.38907
49	3.24327	489	4	0.52992	9.27352
50	3.86151	489	4	0.83443	14.60244
51	4.00000	490	4	0.92526	16.19197

Table L-43. Raw to Scaled Score Look-up Table—NM-MSSA Mathematics (Spanish) Grade 5 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	500	1	6.52341	80.00000
1	-3.77155	500	1	5.18696	80.00000
2	-3.54309	500	1	4.12079	72.11388
3	-3.31464	500	1	3.27457	57.30496
4	-3.08618	501	1	2.60569	45.59957
5	-2.85773	505	1	2.07860	36.37542
6	-2.62927	509	1	1.66394	29.11903
7	-2.40082	513	1	1.33775	23.41055
8	-2.17236	517	1	1.08062	18.91082
9	-1.62817	527	1	0.66411	11.62201
10	-1.30719	532	1	0.50928	8.91241
11	-1.07573	536	1	0.42792	7.48863
12	-0.89251	539	1	0.37789	6.61314
13	-0.73929	542	1	0.34400	6.02004
14	-0.60641	544	1	0.31943	5.59004
15	-0.48812	547	1	0.30070	5.26225
16	-0.38072	548	2	0.28587	5.00272
17	-0.28169	550	2	0.27379	4.79130
18	-0.18921	552	2	0.26373	4.61533
19	-0.10194	553	2	0.25523	4.46655
20	-0.01885	555	2	0.24797	4.33939
21	0.06088	556	2	0.24171	4.22997
22	0.13790	558	2	0.23632	4.13552
23	0.21274	559	2	0.23166	4.05405
24	0.28585	560	3	0.22766	3.98404
25	0.35763	561	3	0.22425	3.92429
26	0.42839	563	3	0.22136	3.87375
27	0.49845	564	3	0.21894	3.83150

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
28	0.56806	565	3	0.21695	3.79671
29	0.63748	566	3	0.21535	3.76866
30	0.70694	567	3	0.21411	3.74685
31	0.77666	569	3	0.21320	3.73106
32	0.84686	570	3	0.21265	3.72143
33	0.91780	571	3	0.21248	3.71844
34	0.98970	572	3	0.21274	3.72294
35	1.06285	574	4	0.21348	3.73592
36	1.13757	575	4	0.21477	3.75846
37	1.21421	576	4	0.21666	3.79161
38	1.29317	578	4	0.21922	3.83641
39	1.37496	579	4	0.22252	3.89414
40	1.46019	581	4	0.22667	3.96671
41	1.54965	582	4	0.23184	4.05728
42	1.64440	584	4	0.23835	4.17117
43	1.74596	586	4	0.24669	4.31712
44	1.85655	588	4	0.25769	4.50949
45	1.97966	589	4	0.27272	4.77252
46	2.12103	589	4	0.29428	5.14983
47	2.29114	589	4	0.32734	5.72852
48	2.51176	589	4	0.38362	6.71338
49	2.84035	589	4	0.49960	8.74301
50	3.52940	589	4	0.89131	15.59791
51	4.00000	590	4	1.29574	22.67549

Table L-44. Raw to Scaled Score Look-up Table—NM-MSSA Mathematics (Spanish) Grade 6 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	600	1	4.18583	73.25201
1	-3.84329	600	1	3.74186	65.48253
2	-3.68658	600	1	3.33236	58.31629
3	-3.52987	600	1	2.95620	51.73350
4	-3.37316	603	1	2.61217	45.71305
5	-3.21645	605	1	2.29905	40.23336
6	-3.05974	608	1	2.01560	35.27301
7	-2.90303	611	1	1.76062	30.81093
8	-2.74632	614	1	1.53293	26.82621
9	-2.00417	627	1	0.78478	13.73366
10	-1.62163	633	1	0.57308	10.02899
11	-1.35387	638	1	0.47529	8.31750
12	-1.14202	642	1	0.42007	7.35130
13	-0.96296	645	1	0.38517	6.74052
14	-0.80538	648	2	0.36131	6.32293
15	-0.66301	650	2	0.34389	6.01800
16	-0.53205	652	2	0.33034	5.78088
17	-0.41010	654	2	0.31912	5.58457
18	-0.29552	656	2	0.30926	5.41212
19	-0.18718	658	2	0.30016	5.25276
20	-0.08425	660	3	0.29142	5.09988
21	0.01387	662	3	0.28285	4.94985
22	0.10767	664	3	0.27436	4.80124
23	0.19752	665	3	0.26595	4.65420
24	0.28376	667	3	0.25771	4.50994
25	0.36668	668	3	0.24972	4.37014
26	0.44659	669	3	0.24209	4.23651

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
27	0.52379	671	3	0.23488	4.11048
28	0.59857	672	3	0.22818	3.99311
29	0.67126	673	3	0.22201	3.88516
30	0.74216	675	3	0.21642	3.78731
31	0.81161	676	3	0.21145	3.70034
32	0.87992	677	3	0.20716	3.62528
33	0.94744	678	3	0.20362	3.56338
34	1.01455	679	4	0.20092	3.51616
35	1.08164	681	4	0.19916	3.48527
36	1.14916	682	4	0.19843	3.47246
37	1.21760	683	4	0.19883	3.47952
38	1.28752	684	4	0.20048	3.50832
39	1.35956	685	4	0.20347	3.56079
40	1.43443	687	4	0.20794	3.63901
41	1.51299	688	4	0.21402	3.74538
42	1.59624	689	4	0.22187	3.88280
43	1.68537	689	4	0.23172	4.05516
44	1.78190	689	4	0.24389	4.26803
45	1.88778	689	4	0.25885	4.52994
46	2.00566	689	4	0.27740	4.85459
47	2.13938	689	4	0.30085	5.26480
48	2.29474	689	4	0.33145	5.80031
49	2.48127	689	4	0.37340	6.53447
50	2.71605	689	4	0.43516	7.61535
51	3.03372	689	4	0.53613	9.38233
52	3.51825	689	4	0.72949	12.76611
53	4.00000	690	4	0.96454	16.87952
54	4.00000	690	4	0.96454	16.87952

Table L-45. Raw to Scaled Score Look-up Table—NM-MSSA Mathematics (Spanish) Grade 7 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	700	1	6.55322	80.00000
1	-3.80127	700	1	5.36837	80.00000
2	-3.60254	700	1	4.39788	76.96295
3	-3.40380	700	1	3.60486	63.08507
4	-3.20507	703	1	2.95796	51.76431
5	-3.00634	707	1	2.43075	42.53808
6	-2.80761	710	1	2.00112	35.01966
7	-2.60887	714	1	1.65080	28.88907
8	-2.41014	717	1	1.36483	23.88452
9	-2.21141	721	1	1.13110	19.79416
10	-1.63301	731	1	0.66603	11.65554
11	-1.30865	736	1	0.50317	8.80546
12	-1.08019	740	1	0.41754	7.30688
13	-0.90187	743	1	0.36391	6.36849
14	-0.75408	746	1	0.32695	5.72161
15	-0.62663	747	1	0.29997	5.24956
16	-0.51356	750	2	0.27960	4.89302
17	-0.41106	752	2	0.26389	4.61803
18	-0.31656	754	2	0.25160	4.40309
19	-0.22823	755	2	0.24191	4.23350
20	-0.14474	757	2	0.23420	4.09857
21	-0.06509	758	2	0.22801	3.99022
22	0.01147	759	2	0.22298	3.90212

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
23	0.08553	761	3	0.21882	3.82928
24	0.15754	762	3	0.21530	3.76774
25	0.22785	763	3	0.21225	3.71435
26	0.29676	764	3	0.20952	3.66666
27	0.36449	766	3	0.20702	3.62286
28	0.43122	767	3	0.20466	3.58163
29	0.49710	768	3	0.20241	3.54217
30	0.56228	769	3	0.20023	3.50407
31	0.62687	769	3	0.19813	3.46731
32	0.69102	771	4	0.19612	3.43212
33	0.75483	772	4	0.19423	3.39900
34	0.81847	774	4	0.19250	3.36870
35	0.88209	775	4	0.19098	3.34213
36	0.94588	776	4	0.18974	3.32045
37	1.01005	777	4	0.18886	3.30497
38	1.07488	778	4	0.18840	3.29707
39	1.14066	779	4	0.18847	3.29820
40	1.20775	780	4	0.18913	3.30971
41	1.27656	782	4	0.19045	3.33286
42	1.34758	783	4	0.19251	3.36887
43	1.42139	784	4	0.19538	3.41917
44	1.49868	785	4	0.19919	3.48582
45	1.58035	787	4	0.20413	3.57232
46	1.66758	788	4	0.21056	3.68476
47	1.76205	789	4	0.21907	3.83369
48	1.86629	789	4	0.23069	4.03715
49	1.98432	789	4	0.24724	4.32671
50	2.12311	789	4	0.27207	4.76123
51	2.29610	789	4	0.31227	5.46470
52	2.53425	789	4	0.38636	6.76129
53	2.93931	789	4	0.56993	9.97384
54	4.00000	790	4	1.56018	27.30309

Table L-46. Raw to Scaled Score Look-up Table—NM-MSSA Mathematics (Spanish) Grade 8 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.00000	800	1	3.57676	62.59329
1	-3.83165	800	1	3.17486	55.56005
2	-3.66331	800	1	2.81287	49.22526
3	-3.49496	800	1	2.48694	43.52146
4	-3.32661	800	1	2.19392	38.39366
5	-3.15827	800	1	1.93139	33.79935
6	-2.98992	800	1	1.69755	29.70714
7	-2.82158	803	1	1.49107	26.09364
8	-2.65323	806	1	1.31078	22.93872
9	-1.88454	820	1	0.77589	13.57810
10	-1.44458	827	1	0.61947	10.84076
11	-1.12353	833	1	0.53585	9.37742
12	-0.86882	837	1	0.47564	8.32365
13	-0.65763	840	1	0.42839	7.49688
14	-0.47694	844	2	0.39174	6.85537
15	-0.31856	847	2	0.36359	6.36276
16	-0.17718	850	2	0.34152	5.97661
17	-0.04927	852	2	0.32349	5.66104
18	0.06762	854	2	0.30812	5.39207

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
19	0.17532	856	2	0.29463	5.15600
20	0.27527	857	2	0.28261	4.94564
21	0.36867	859	2	0.27184	4.75719
22	0.45652	861	3	0.26219	4.58836
23	0.53966	862	3	0.25358	4.43757
24	0.61883	864	3	0.24591	4.30348
25	0.69465	865	3	0.23914	4.18489
26	0.76767	866	3	0.23318	4.08069
27	0.83835	867	3	0.22799	3.98977
28	0.90709	869	3	0.22349	3.91108
29	0.97427	870	3	0.21963	3.84359
30	1.04020	871	3	0.21636	3.78636
31	1.10516	872	3	0.21363	3.73857
32	1.16943	873	3	0.21140	3.69953
33	1.23325	874	3	0.20964	3.66874
34	1.29686	875	3	0.20834	3.64593
35	1.36051	876	3	0.20749	3.63106
36	1.42444	877	3	0.20711	3.62435
37	1.48892	879	4	0.20722	3.62633
38	1.55422	880	4	0.20787	3.63779
39	1.62069	881	4	0.20914	3.65987
40	1.68868	882	4	0.21109	3.69403
41	1.75865	883	4	0.21384	3.74213
42	1.83112	885	4	0.21752	3.80656
43	1.90674	886	4	0.22231	3.89034
44	1.98633	887	4	0.22843	3.99748
45	2.07094	889	4	0.23619	4.13337
46	2.16195	889	4	0.24604	4.30566
47	2.26126	889	4	0.25860	4.52555
48	2.37157	889	4	0.27487	4.81019
49	2.49690	889	4	0.29641	5.18721
50	2.64363	889	4	0.32593	5.70379
51	2.82264	889	4	0.36842	6.44732
52	3.05477	889	4	0.43430	7.60029
53	3.38756	889	4	0.55029	9.63011
54	3.97490	889	4	0.82208	14.38643
55	4.00000	890	4	0.83587	14.62778

Table L-47. Raw to Scaled Score Look-up Table—NM-MSSA Science (Spanish) Grade 5 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-4.32534	500	1	1.25946	15.74320
1	-4.29742	500	1	1.24237	15.52964
2	-4.26949	500	1	1.22550	15.31877
3	-3.42403	511	1	0.80780	10.09747
4	-2.95027	517	1	0.64224	8.02795
5	-2.61701	521	1	0.55040	6.88002
6	-2.35716	524	1	0.49127	6.14087
7	-2.14215	527	1	0.44982	5.62281
8	-1.95720	529	1	0.41918	5.23974
9	-1.79367	531	1	0.39571	4.94634
10	-1.64607	533	1	0.37730	4.71629
11	-1.51070	535	1	0.36264	4.53306
12	-1.38492	536	1	0.35085	4.38562
13	-1.26680	538	1	0.34130	4.26619

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
14	-1.15486	539	1	0.33352	4.16900
15	-1.04800	540	1	0.32717	4.08963
16	-0.94530	542	1	0.32197	4.02463
17	-0.84605	543	1	0.31770	3.97121
18	-0.74968	544	2	0.31417	3.92714
19	-0.65571	545	2	0.31125	3.89065
20	-0.56373	547	2	0.30883	3.86033
21	-0.47342	548	2	0.30681	3.83510
22	-0.38448	549	2	0.30513	3.81416
23	-0.29667	550	2	0.30376	3.79698
24	-0.20974	551	2	0.30266	3.78321
25	-0.12349	552	2	0.30182	3.77271
26	-0.03774	553	2	0.30124	3.76546
27	0.04772	554	2	0.30093	3.76157
28	0.13305	555	2	0.30090	3.76120
29	0.21845	556	2	0.30117	3.76461
30	0.30411	557	2	0.30177	3.77206
31	0.39019	558	2	0.30271	3.78386
32	0.47691	559	2	0.30402	3.80030
33	0.56447	561	3	0.30574	3.82170
34	0.65306	562	3	0.30787	3.84833
35	0.74293	563	3	0.31044	3.88046
36	0.83430	564	3	0.31346	3.91829
37	0.92743	565	3	0.31696	3.96201
38	1.02257	566	3	0.32094	4.01169
39	1.12000	568	3	0.32539	4.06737
40	1.22003	569	3	0.33032	4.12901
41	1.32297	570	3	0.33572	4.19653
42	1.42915	571	3	0.34159	4.26983
43	1.53893	573	3	0.34791	4.34891
44	1.65272	573	3	0.35471	4.43390
45	1.77096	576	4	0.36202	4.52526
46	1.89413	577	4	0.36990	4.62378
47	2.02280	579	4	0.37846	4.73075
48	2.15765	581	4	0.38782	4.84781
49	2.29945	582	4	0.39814	4.97675
50	2.44910	584	4	0.40951	5.11892
51	2.60762	586	4	0.42195	5.27439
52	2.77612	588	4	0.43528	5.44095
53	2.95466	590	4	0.44902	5.61271
54	2.95466	590	4	0.44902	5.61271
55	2.95466	590	4	0.44902	5.61271
56	2.95466	590	4	0.44902	5.61271
57	2.95466	590	4	0.44902	5.61271
58	2.95466	590	4	0.44902	5.61271
59	2.95466	590	4	0.44902	5.61271
60	2.95466	590	4	0.44902	5.61271
61	2.95466	590	4	0.44902	5.61271
62	2.95466	590	4	0.44902	5.61271
63	2.95466	590	4	0.44902	5.61271
64	2.95466	590	4	0.44902	5.61271

Table L-48. Raw to Scaled Score Look-up Table—NM-MSSA Science (Spanish) Grade 8 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-5.56012	800	1	2.11037	21.10368
1	-5.25684	803	1	1.87240	18.72403
2	-4.95357	806	1	1.65813	16.58128
3	-3.74556	818	1	1.00791	10.07910
4	-3.12420	824	1	0.78011	7.80114
5	-2.70091	828	1	0.65942	6.59423
6	-2.37645	831	1	0.58377	5.83769
7	-2.11070	834	1	0.53174	5.31736
8	-1.88355	836	1	0.49377	4.93775
9	-1.68354	838	1	0.46492	4.64918
10	-1.50352	840	1	0.44230	4.42297
11	-1.33879	842	1	0.42412	4.24120
12	-1.18607	843	1	0.40921	4.09205
13	-1.04302	844	1	0.39674	3.96736
14	-0.90789	846	2	0.38613	3.86131
15	-0.77937	847	2	0.37697	3.76971
16	-0.65644	849	2	0.36895	3.68950
17	-0.53828	850	2	0.36185	3.61850
18	-0.42425	851	2	0.35552	3.55516
19	-0.31378	852	2	0.34984	3.49844
20	-0.20642	853	2	0.34476	3.44759
21	-0.10175	854	2	0.34021	3.40210
22	0.00057	855	2	0.33616	3.36156
23	0.10089	856	2	0.33256	3.32562
24	0.19950	857	2	0.32939	3.29391
25	0.29667	858	2	0.32661	3.26610
26	0.39266	859	2	0.32418	3.24182
27	0.48769	859	2	0.32208	3.22080
28	0.58199	861	3	0.32028	3.20282
29	0.67576	862	3	0.31878	3.18776
30	0.76920	863	3	0.31756	3.17563
31	0.86251	864	3	0.31666	3.16658
32	0.95588	865	3	0.31608	3.16085
33	1.04953	866	3	0.31588	3.15880
34	1.14366	867	3	0.31609	3.16088
35	1.23850	867	3	0.31676	3.16760
36	1.33429	868	3	0.31795	3.17948
37	1.43130	869	3	0.31971	3.19706
38	1.52981	870	3	0.32208	3.22080
39	1.63014	871	3	0.32510	3.25104
40	1.73260	872	3	0.32879	3.28795
41	1.83756	873	3	0.33315	3.33150
42	1.94537	875	3	0.33815	3.38146
43	2.05642	876	3	0.34374	3.43742
44	2.17112	877	3	0.34990	3.49898
45	2.28988	878	3	0.35660	3.56597
46	2.41317	879	3	0.36387	3.63875
47	2.54152	881	3	0.37184	3.71843
48	2.67554	881	3	0.38069	3.80693
49	2.81599	883	4	0.39066	3.90657
50	2.96376	885	4	0.40192	4.01919
51	3.11986	886	4	0.41447	4.14467
52	3.28539	888	4	0.42795	4.27947
53	3.46144	889	4	0.44165	4.41645
54	3.53988	890	4	0.44735	4.47350

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
55	3.53988	890	4	0.44735	4.47350
56	3.53988	890	4	0.44735	4.47350
57	3.53988	890	4	0.44735	4.47350
58	3.53988	890	4	0.44735	4.47350
59	3.53988	890	4	0.44735	4.47350
60	3.53988	890	4	0.44735	4.47350
61	3.53988	890	4	0.44735	4.47350
62	3.53988	890	4	0.44735	4.47350
63	3.53988	890	4	0.44735	4.47350
64	3.53988	890	4	0.44735	4.47350

Table L-49. Raw to Scaled Score Look-up Table—NM-MSSA Science (Spanish) Grade 11 Operational Form 1

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
0	-8.02951	1100	1	5.10105	38.25786
1	-6.92351	1108	1	3.44362	25.82716
2	-5.81752	1116	1	2.26694	17.00205
3	-4.71152	1124	1	1.46338	10.97538
4	-3.69428	1132	1	0.96814	7.26105
5	-3.12552	1136	1	0.76769	5.75767
6	-2.72773	1139	1	0.65386	4.90397
7	-2.42001	1142	1	0.57909	4.34321
8	-2.16760	1143	1	0.52586	3.94398
9	-1.95238	1145	1	0.48600	3.64502
10	-1.76370	1146	1	0.45511	3.41334
11	-1.59478	1148	1	0.43058	3.22933
12	-1.44104	1149	1	0.41073	3.08049
13	-1.29925	1150	1	0.39445	2.95838
14	-1.16704	1151	1	0.38095	2.85711
15	-1.04263	1152	1	0.36965	2.77239
16	-0.92464	1153	1	0.36014	2.70106
17	-0.81200	1153	1	0.35209	2.64071
18	-0.70382	1154	2	0.34527	2.58951
19	-0.59940	1155	2	0.33947	2.54603
20	-0.49816	1156	2	0.33455	2.50913
21	-0.39960	1157	2	0.33039	2.47791
22	-0.30330	1157	2	0.32689	2.45164
23	-0.20890	1158	2	0.32396	2.42971
24	-0.11607	1159	2	0.32155	2.41159
25	-0.02455	1159	2	0.31958	2.39686
26	0.06594	1160	3	0.31802	2.38515
27	0.15560	1161	3	0.31682	2.37616
28	0.24464	1162	3	0.31596	2.36966
29	0.33327	1162	3	0.31540	2.36550
30	0.42165	1163	3	0.31514	2.36357
31	0.50995	1164	3	0.31518	2.36383
32	0.59835	1164	3	0.31550	2.36628
33	0.68700	1165	3	0.31612	2.37092
34	0.77605	1166	3	0.31704	2.37777
35	0.86567	1166	3	0.31824	2.38681
36	0.95601	1167	3	0.31973	2.39801
37	1.04722	1168	3	0.32150	2.41126
38	1.13946	1168	3	0.32353	2.42646
39	1.23288	1169	3	0.32580	2.44347
40	1.32764	1170	3	0.32830	2.46222

continued

Raw Score	Theta	Scale Score	Performance Level	CSEM	Scaled CSEM
41	1.42391	1170	3	0.33103	2.48269
42	1.52185	1171	3	0.33400	2.50500
43	1.62167	1172	3	0.33726	2.52945
44	1.72359	1173	3	0.34086	2.55648
45	1.82786	1173	3	0.34489	2.58666
46	1.93480	1174	3	0.34943	2.62069
47	2.04476	1175	3	0.35457	2.65926
48	2.15816	1176	3	0.36041	2.70305
49	2.27549	1177	3	0.36703	2.75270
50	2.39733	1178	3	0.37452	2.80887
51	2.52434	1179	3	0.38298	2.87236
52	2.65732	1180	3	0.39258	2.94436
53	2.79723	1180	3	0.40354	3.02656
54	2.94524	1182	4	0.41618	3.12132
55	3.10277	1183	4	0.43085	3.23139
56	3.27159	1184	4	0.44793	3.35950
57	3.45379	1186	4	0.46770	3.50772
58	3.65191	1187	4	0.49037	3.67776
59	3.86908	1189	4	0.51646	3.87341
60	4.10383	1190	4	0.54679	4.10091
61	4.10383	1190	4	0.54679	4.10091
62	4.10383	1190	4	0.54679	4.10091
63	4.10383	1190	4	0.54679	4.10091
64	4.10383	1190	4	0.54679	4.10091
65	4.10383	1190	4	0.54679	4.10091
66	4.10383	1190	4	0.54679	4.10091
67	4.10383	1190	4	0.54679	4.10091
68	4.10383	1190	4	0.54679	4.10091

APPENDIX—M
iMSSA 2022-23 TECHNICAL REPORT ADDENDUM

New Mexico Interim Measure of Student Success and Achievement

2022–23 Technical Report Addendum

1. Introduction

The Interim Measure of Student Success and Achievement (iMSSA) includes assessments in mathematics, reading, and language usage that are administered online to students in New Mexico in grades 3–8. Schools can administer up to three different, equivalent test forms, one per administration window, during the school year. The iMSSA is designed to measure student achievement against college- and career-readiness standards, such as the Common Core State Standards or similar frameworks, in the assessed content areas. These academic content and process standards express what students should know and be able to do in each grade level and content area.

The iMSSA provides point-in-time information about student academic achievement and progress. Student results are reported according to academic achievement descriptors using cut scores established in standard setting for each of three achievement levels: Needs Support, Near Target, and On Target. The results from these assessments provide information about students' progress to guide the creation and modification of future educational practices to meet the needs of students, their families, and educators.

The iMSSA assessments are not required in New Mexico, except for K5 Plus schools required to take them.

This addendum builds upon the information provided in the Cognia Interim Assessments technical report. The intent of this document is to provide information specific to the administration of the iMSSA assessments in New Mexico in the 2022–23 school year.

2. Administration and Participation

The 2022–23 iMSSA assessments were administered during three testing windows:

- Beginning-of-Year (BOY): August 15, 2022 – October 14, 2022
- Middle-of-Year (MOY): December 5, 2022 – January 20, 2023
- End-of-Year (EOY): April 3, 2023 – May 26, 2023

Each administration window is preceded by a time for schools and districts to upload student rosters to the iTester platform and schedule tests for administration. Reports are then delivered to educators in eMetric's Data Interaction Platform and to families using the Parent Portal.

The iMSSA assessments are computer-based only, with certain accommodations embedded into the platform. A list of available accommodations can be found in the NM-ASR and NM-MSSA Spring 2022 Accessibility Features and Accommodations Manual.

2.1 Summary of 2022–23 iMSSA Administration

Table 1 provides a summary of the 2022–23 iMSSA administration overall and by administration window (i.e., BOY, MOY, EOY). Appendix 1 provides the counts of students participating in the iMSSA by school and district for each of the administration windows.

During the 2022–23 administration, valid responses to iMSSA were provided by a total of 51,424 students in grades 3–8 from 303 schools in 71 districts across New Mexico. Generally, the number of participating students, schools, and districts increased across the year; at BOY, there were approximately 6,700 to 7,800 students per grade level, whereas at EOY there were 7,000 to 8,400 students per grade level.

Table 1. Summary of 2022–23 iMSSA Administration

		Overall	BOY	MOY	EOY
Counts	Students	46739	44431	44322	37903
	Tests	358836	128820	128479	101537
	Schools	239	238	239	225
	Districts	56	56	56	52
Grade	3	7542	7166	7089	6682
	4	7554	7177	7187	6600
	5	7612	7290	7281	6824
	6	7917	7519	7560	6167
	7	7888	7474	7476	5713
	8	8246	7805	7731	5918

3. Scale Scores

Scale scores are computed as linear transformations of student-ability estimates resulting from responses to items on each of the mathematics, reading, and language usage test forms. Calculated separately for each grade level and content area, these continuous scales are defined according to common properties and specifications, which allow for convenient interpretation of student performance and aggregation at the classroom, school, district, or state level. These scale scores are constructed similarly for each administration period, facilitating longitudinal examination and comparison of student performance.

Three-digit scale scores are presented for each content area and specified according to parameters that facilitate interpretation of student performance within the current grade level:

- On Target cut points are located at the scale score where grade level is in the hundreds place followed by 60;
- Lowest obtainable scale score (LOSS) is defined as 100 points below the On Target cut point; and
- Highest obtainable scale score (HOSS) is defined as 80 points above the On Target cut point.

3.1 Summary of 2022–23 iMSSA Scale Scores

Scale scores for the 2022–23 administration of iMSSA are summarized by content area, grade level, and administration window in Table 2; box-and-whisker plots of corresponding student performance are presented in Figure 1.

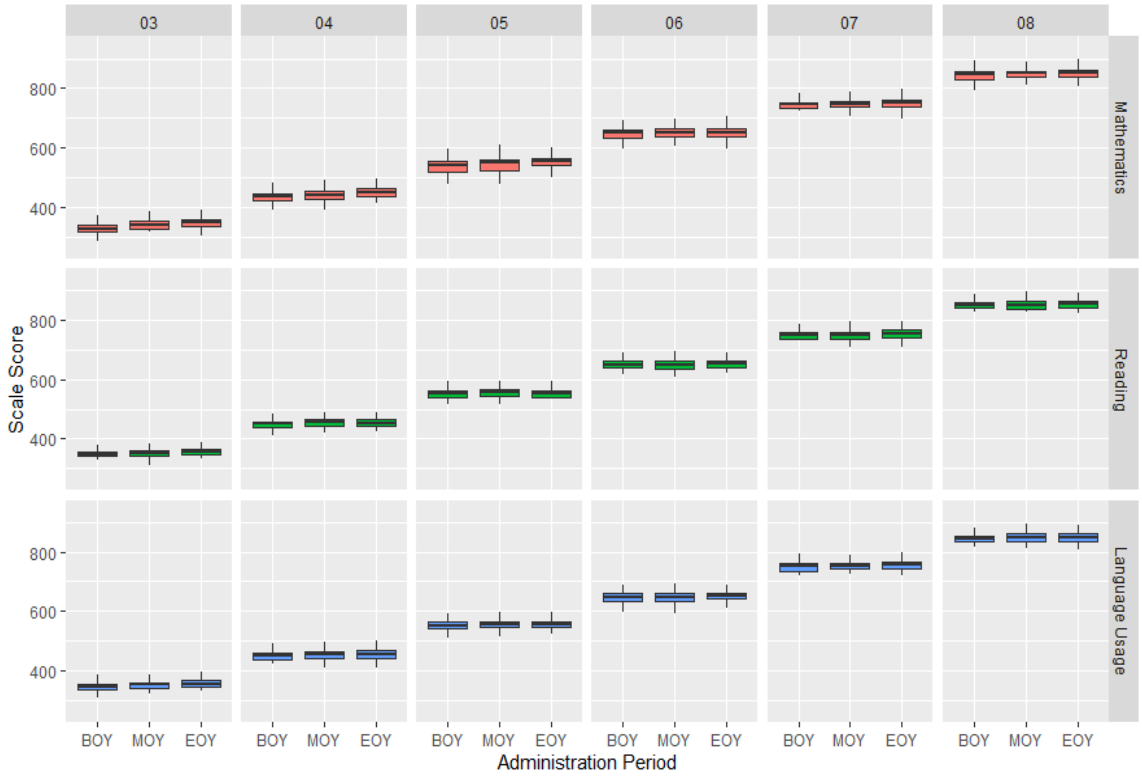
Scale scores follow the monotonically increasing pattern defined by grade level with averages typically near but below the On Target cut point (e.g., 360 for grade 3 test forms). In general, the 2022–23 EOY averages demonstrate an increase over BOY averages; between those administration windows, however, certain grade level/content area combinations demonstrate slight decreases from the prior administration window. For example, slight decreases between MOY and EOY scale scores are observed for reading in grade 5.

Table 2. Summary of Scale Scores* for 2022–23 iMSSA

Content Area	Grade	BOY	MOY	EOY
Mathematics	3	324.926 (19.91) *	333.235 (27.61)	346.597 (20.11)
	4	429.083 (24.19)	439.822 (19.89)	445.352 (26.75)
	5	537.739 (20.44)	542.181 (23.26)	549.318 (22.05)
	6	644.246 (19.47)	648.865 (19.5)	649.909 (22.13)
	7	736.939 (27.2)	743.392 (20.68)	746.796 (21.38)
	8	843.384 (17.92)	848.4 (16.99)	846.975 (22.35)
Reading	3	344.003 (19.96)	349.396 (19.16)	352.287 (21.24)
	4	446.928 (20.76)	451.087 (22.04)	452.038 (22.59)
	5	550.031 (20.9)	554.251 (20.22)	549.922 (21.71)
	6	647.929 (21.21)	648.488 (21.79)	649.98 (22.06)
	7	746.448 (23.6)	747.857 (20.67)	751.579 (22.68)
	8	848.614 (19.71)	847.363 (25.19)	850.81 (21.99)
Language Usage	3	340.513 (20.68)	349.147 (17.31)	353.786 (19.63)
	4	446.4 (23.35)	451.968 (21.51)	455.607 (21.16)
	5	549.522 (21.19)	553.624 (19.86)	554.571 (20.94)
	6	647.584 (21.26)	645.172 (22.09)	650.818 (23.07)
	7	747.51 (20.6)	753.403 (19.55)	753.139 (20.07)
	8	843.602 (18.11)	848.646 (18.98)	848.51 (20.98)

** Means; standard deviations in parentheses.*

Figure 1. Scale Score Distributions for 2022–23 iMSSA



4. Achievement Levels

Overall achievement levels are ordered categories labeled as Needs Support, Near Target, and On Target. These categories indicate the degree to which students can demonstrate knowledge and skills based on end-of-grade expectations in each content area. The specific boundaries of each of these achievement levels are based on cut points that were established during standard setting; the On Target cut points are always located at the scale score beginning with the numeric grade value followed by 60 (e.g., 360 for grade 3), whereas the other two cut points were independently determined for each content area and grade level.

4.1 Summary of iMSSA 2022–23 Achievement Levels

Table 3 summarizes the distributions of students across achievement levels, and Figure 2 provides a graphical summary.

Generally, the percentage of students categorized as Needs Support decreased across administration windows of the 2022–23 iMSSA. These decreases were absorbed across the Near Target and On Target achievement levels, showing increased percentages in one or both.

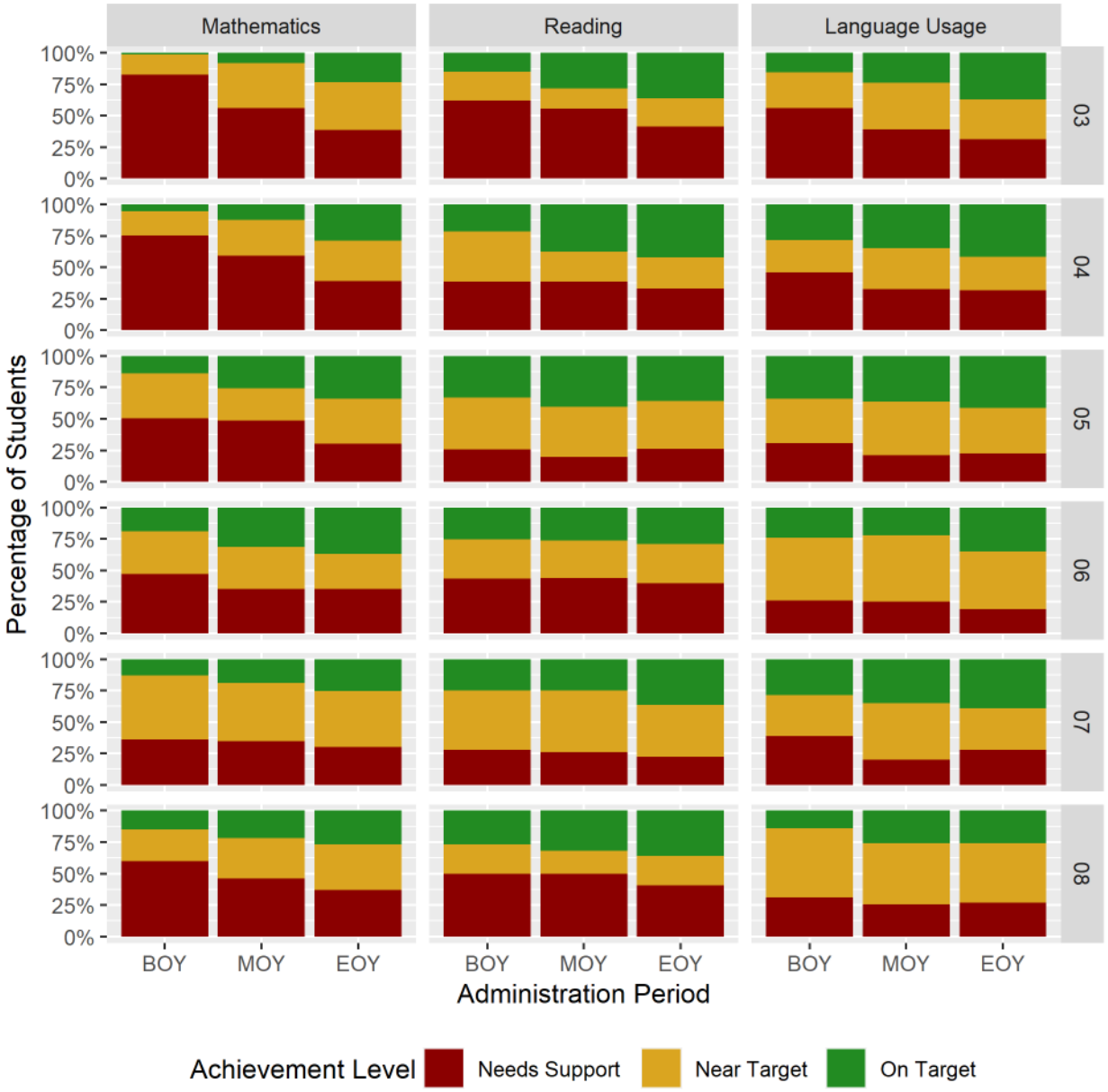
- For mathematics, 36%–83% of all students were categorized as Needs Support at BOY while less than 20% were categorized as On Target. By EOY, Needs Support decreased to less than 40% of students while On Target increased to 23%–37% of students.
- Distributions of achievement levels were similar for reading and language usage. From 26%–62% of students were categorized as Needs Support at BOY, decreasing to 19%–41% at EOY. The

distribution of On Target students increased from BOY and EOY, from less than 34% increased to 26%–42% of students.

Table 3. Summary of Achievement-Level Distributions for the 2022–23 iMSSA Administration

		Needs Support			Near Target			On Target		
Content Area	**Grade**	**BOY**	**MOY**	**EOY**	**BOY**	**MOY**	**EOY**	**BOY**	**MOY**	**EOY**
Mathematics	3	83%	56%	39%	16%	36%	38%	1%	8%	23%
	4	75%	59%	39%	19%	28%	32%	5%	12%	29%
	5	50%	48%	31%	36%	26%	36%	14%	26%	34%
	6	47%	35%	35%	34%	33%	28%	19%	31%	37%
	7	36%	35%	30%	50%	46%	44%	13%	19%	25%
	8	60%	46%	37%	25%	32%	36%	15%	22%	27%
Reading	3	62%	56%	41%	23%	16%	23%	15%	28%	36%
	4	39%	39%	33%	40%	24%	25%	22%	38%	42%
	5	26%	20%	26%	41%	40%	38%	33%	41%	36%
	6	44%	44%	40%	31%	30%	31%	25%	26%	29%
	7	28%	26%	22%	47%	49%	41%	25%	25%	36%
	8	50%	50%	41%	23%	19%	23%	27%	32%	36%
Language Usage	3	56%	39%	31%	29%	37%	32%	15%	24%	37%
	4	46%	33%	32%	25%	33%	27%	29%	35%	42%
	5	31%	21%	22%	35%	42%	37%	34%	36%	41%
	6	26%	25%	19%	50%	52%	46%	24%	22%	35%
	7	39%	20%	28%	33%	45%	33%	28%	35%	39%
	8	31%	25%	27%	55%	49%	47%	14%	26%	26%

Figure 2. Distribution of Achievement Levels for the 2022–23 iMSSA Administration



5. Differential Validity

For an interim testing program, it is important to examine differences in student performance that may result from construct-irrelevant factors (see *Standards for Educational and Psychological Testing*[1]). In addition to item and test design activities intended to limit the bias of any specific test content, statistical

[1] *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.*

analyses of the results are conducted to evaluate potential impact of such factors. The degree to which student performance differs as a function of identification of student subgroup is referred to as differential validity.

5.1 Summary of iMSSA 2022–23 Differential Validity

Student subgroup identification provided by the New Mexico Public Education Department is considered in the differential validity analysis as follows:

- Gender: Female, Male, Unknown
- Hispanic: Yes, No
- Race: Asian, Black, Caucasian/White, Native Hawaiian / Other Pacific Islander, American Indian / Alaska Native, Multi-race
- English Learner status: Initially Fluent English Proficient – Student never EL; Current EL Student; Reclassified Fluent English Proficient – exited Year 1; Reclassified Fluent English Proficient – exited Year 2; Reclassified Fluent English Proficient – exited Year 3; Reclassified Fluent English Proficient – exited Year 4; Reclassified Fluent English Proficient – exited Year 5
- Bilingual Education, Economically Disadvantaged, Foster Care, Homeless, Homeschool, Special Education / Individualized Education Plan, Migrant, Military, 504 Plan or status: Yes, No, Unknown for all.

Student distribution in each of the identified subgroups is summarized in Table 4.

Table 4. Summary of Student Subgroups for the 2022–23 iMSSA Administration

Subgroup	Description	Overall	BOY	MOY	EOY
Overall		46739	44431	44322	37903
Gender	Female	49%	49%	49%	49%
	Male	51%	51%	51%	51%
	Unknown	0%	0%	0%	0%
Ethnicity / Race	Hispanic	56%	56%	56%	56%
	American Indian / Alaska Native	15%	15%	15%	15%
	Asian	2%	2%	2%	2%
	Black / African American	3%	3%	3%	3%
	Native Hawaiian / Other Pacific Islander	1%	1%	1%	1%
	Caucasian / White	78%	78%	78%	78%
EL	Initially Fluent English Proficient – Student never EL	82%	83%	82%	82%
	Current EL Student	16%	15%	16%	16%
	Reclassified Fluent English Proficient – exited Year 1	0%	0%	0%	0%
	Reclassified Fluent English Proficient – exited Year 2	1%	1%	1%	1%
	Reclassified Fluent English Proficient – exited Year 3	0%	0%	0%	0%
	Reclassified Fluent English Proficient – exited Year 4	0%	0%	0%	0%
	Reclassified Fluent English Proficient – exited Year 5	0%	0%	0%	0%
Demographics	Bilingual Education	14%	13%	13%	14%
	Economically Disadvantaged	40%	36%	41%	43%
	Foster Care	*	*	*	*
	Homeless	1%	1%	1%	1%
	Homeschool	*	*	*	*
	Special Education / Individualized Education Plan	16%	15%	16%	16%
	Migrant	0%	0%	0%	0%
	Military	0%	0%	0%	0%
	504 Plan	1%	1%	1%	1%

** Results suppressed due to failure to meet minimum reporting threshold n > 20 students.*

For each of the 54 iMSSA test forms (i.e., three content areas, six grade levels, three administration windows) and the 18 student subgroups evaluated for each test, there is the very likely potential for inflation of Type I error; significant effects of subgroup on student performance may be spuriously identified given the large number of calculations conducted. Meaningful statistical results are therefore presented according to effect size calculations produced from regressing student scale scores on each subgroup. These effect sizes are calculated as η^2 and indicate the variability in student scale scores that may be attributed to a student subgroup. Guidelines exist to facilitate the interpretation of effect sizes[2]:

- Very small effect size: $\eta^2 < 0.02$;
- Small effect size: $0.02 \leq \eta^2 < 0.13$;
- Medium effect size: $0.13 \leq \eta^2 < 0.26$; and
- Large effect size: $\eta^2 \geq 0.26$.

For the 2022–23 iMSSA, evaluation of differential validity yielded no medium or large effect sizes ($\eta^2 \geq 0.13$) for any of the student subgroups participating in the assessments, which would have suggested closer inspection of specific results and test content. Small effect sizes ($0.02 \leq \eta^2 < 0.13$) are demonstrated only for the Hispanic, American Indian / Alaska Native, Special Education / Individualized Education Plan, Economically Disadvantaged, English Learner, and Gender subgroups across some of the content areas, grade levels, and administration windows, with η^2 ranges from 0.02004 to 0.059.

Scale scores are presented in Appendix 2 to demonstrate differential validity results of subgroups with small effect sizes for mathematics, reading, and language usage assessments by grade level and administration window. For example, for grade 8 mathematics administered at EOY, students identified as Hispanic demonstrate lower average scale scores (843.824) compared to non-Hispanic students (850.462).

Some trends that appear in these results:

- Students identified as Hispanic demonstrate lower average scale scores than non-Hispanic students for grade 8 in all three content areas.
- Students identified as American Indian / Alaska Native demonstrate lower average scale scores than other students. One or more administrations in all grade levels and content areas show evidence of differential validity for this student subgroup.
- In all instances where English Learners demonstrate small effect sizes, Current English Learners demonstrate lower average scale scores than all other students in Reading and Language Usage in grades 4–8 in one or more administrations.
- In one or more administrations, students identified as Special Education / Individualized Education Plans demonstrate lower average scale scores than all other students for Reading across all grade levels except grade 7, for Language Usage across all grade levels, and for Mathematics in grades 3–6.
- Generally, the effect sizes increase slightly across administration windows, from BOY to EOY, as differences between average scale scores increase between student subgroups.

2 Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155.

Appendix 1

Table 1.1. School and District Participation Results for the 2022–23 iMSSA Administration*

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
001	047	Albuquerque Public Schools	Public Academy For Performing Arts (PAPA)	212	205	205	129
001	051	Albuquerque Public Schools	Robert F. Kennedy Charter School	85	74	76	*
001	118	Albuquerque Public Schools	Christine Duncan Heritage Academy	241	230	228	122
001	333	Albuquerque Public Schools	Pajarito Elementary	138	130	135	*
001	781	Albuquerque Public Schools	The International School at Mesa Del Sol	183	173	172	95
002	135	Reserve Public Schools	Reserve Elementary	26	24	23	23
002	136	Reserve Public Schools	Reserve High	10	10	10	9
004	009	Roswell Independent Schools	Sidney Gutierrez Middle	135	130	132	130
004	024	Roswell Independent Schools	Berrendo Elementary	192	181	181	181
004	025	Roswell Independent Schools	Berrendo Middle	699	656	653	639
004	036	Roswell Independent Schools	Mountain View Middle	554	518	504	496
004	041	Roswell Independent Schools	Del Norte Elementary	274	258	254	256
004	042	Roswell Independent Schools	Mesa Middle	436	403	394	371
004	044	Roswell Independent Schools	East Grand Plains Elementary	122	114	112	109
004	050	Roswell Independent Schools	El Capitan Elementary	198	186	185	182
004	052	Roswell Independent Schools	Nancy Lopez Elementary	95	81	88	86
004	095	Roswell Independent Schools	Military Heights Elementary	192	182	177	176
004	100	Roswell Independent Schools	Missouri Avenue Elementary	137	133	129	127
004	105	Roswell Independent Schools	Monterrey Elementary	216	205	205	202
004	120	Roswell Independent Schools	Sunset Elementary	120	112	110	110
004	125	Roswell Independent Schools	Sierra Middle	664	615	601	603
004	126	Roswell Independent Schools	Pecos Elementary	143	135	137	130
004	161	Roswell Independent Schools	Valley View Elementary	251	237	242	242
005	054	Hagerman Municipal Schools	Hagerman Elementary	78	76	75	*
006	043	Dexter Consolidated Schools	Dexter Elementary	170	164	160	160
006	048	Dexter Consolidated Schools	Dexter Middle	190	*	190	*
007	073	Lake Arthur Municipal Schools	Lake Arthur Elementary	23	20	23	*
007	077	Lake Arthur Municipal Schools	Lake Arthur Middle	36	32	33	*
010	058	Springer Municipal Schools	Wilferth Elementary	28	26	25	27
010	150	Springer Municipal Schools	Springer High	20	20	20	20
012	040	Clovis Municipal Schools	Barry Elementary	158	134	137	137
012	042	Clovis Municipal Schools	Arts Academy at Bella Vista	157	152	152	137
012	058	Clovis Municipal Schools	Highland Elementary	151	135	145	134
012	066	Clovis Municipal Schools	James Bickley Elementary	152	140	137	123

continued

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
012	068	Clovis Municipal Schools	Cameo Elementary	130	116	119	117
012	072	Clovis Municipal Schools	La Casita Elementary	121	109	113	110
012	081	Clovis Municipal Schools	CMS iAcademy at Lincoln Jackson	84	72	74	33
012	084	Clovis Municipal Schools	Lockwood Elementary	162	149	144	144
012	091	Clovis Municipal Schools	Marshall Middle	530	498	490	459
012	095	Clovis Municipal Schools	Mesa Elementary	221	201	201	204
012	098	Clovis Municipal Schools	Yucca Middle	603	573	562	467
012	122	Clovis Municipal Schools	Parkview Elementary	211	200	192	196
012	145	Clovis Municipal Schools	Sandia Elementary	164	155	158	156
012	155	Clovis Municipal Schools	Zia Elementary	173	164	157	162
012	156	Clovis Municipal Schools	W.D. Gattis Middle	584	521	522	514
013	161	Texico Municipal Schools	Texico Elementary	114	111	110	105
013	163	Texico Municipal Schools	Texico Middle	126	121	122	124
014	093	Melrose Public Schools	Melrose Junior	49	47	44	43
014	094	Melrose Public Schools	Melrose Elementary	85	79	79	80
015	055	Grady Municipal Schools	Grady Elementary	44	38	33	39
015	056	Grady Municipal Schools	Grady Middle	49	44	44	46
016	051	Fort Sumner Municipal Schools	Fort Sumner Elementary	55	50	48	38
016	060	Fort Sumner Municipal Schools	Fort Sumner Middle	84	77	79	76
018	001	Hatch Valley Public Schools	Rio Grande Elementary	239	231	225	232
018	050	Hatch Valley Public Schools	Hatch Valley Middle	268	254	253	256
019	009	Gadsden Independent Schools	Sunrise Elementary	276	253	241	259
019	017	Gadsden Independent Schools	Gadsden Elementary	254	250	238	244
019	025	Gadsden Independent Schools	Yucca Heights Elementary	289	274	246	269
019	120	Gadsden Independent Schools	North Valley Elementary	183	172	174	171
022	001	Artesia Public Schools	Yeso Elementary	296	277	286	283
022	032	Artesia Public Schools	Central Elementary	91	80	78	85
022	056	Artesia Public Schools	Hermosa Elementary	180	174	171	167
022	128	Artesia Public Schools	Penasco Elementary	11	11	11	11
022	139	Artesia Public Schools	Roselawn Elementary	100	98	94	93
022	183	Artesia Public Schools	Yucca Elementary	196	184	189	69
022	187	Artesia Public Schools	Artesia Park Junior High	316	303	306	93
022	189	Artesia Public Schools	Artesia Zia Intermediate	558	532	527	521
024	023	Cobre Consolidated Schools	Bayard Elementary	70	64	64	63
024	033	Cobre Consolidated Schools	Central Elementary	68	65	60	56
024	059	Cobre Consolidated Schools	Hurley Elementary	64	59	57	61
024	143	Cobre Consolidated Schools	San Lorenzo Elementary	31	28	28	28
025	015	Santa Rosa Consolidated Schools	Rita A. Marquez Elementary	30	30	30	29
025	020	Santa Rosa Consolidated Schools	Anton Chico Middle	34	33	34	32
025	144	Santa Rosa Consolidated Schools	Santa Rosa Elementary	85	80	83	82

continued

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
025	150	Santa Rosa Consolidated Schools	Santa Rosa Middle	113	107	110	108
030	016	Animas Public Schools	Animas Elementary	23	22	20	21
030	017	Animas Public Schools	Animas 7–12 School	28	24	27	28
030	020	Animas Public Schools	Animas Middle	24	23	24	23
032	049	Eunice Municipal Schools	Caton Middle	164	160	159	152
033	008	Hobbs Municipal Schools	Heizer Middle	604	574	570	*
033	028	Hobbs Municipal Schools	Broadmoor Elementary	174	147	166	165
033	030	Hobbs Municipal Schools	College Lane Elementary	301	286	289	292
033	032	Hobbs Municipal Schools	Coronado Elementary	195	169	170	177
033	046	Hobbs Municipal Schools	Edison Elementary	122	110	95	114
033	057	Hobbs Municipal Schools	Highland Middle	832	773	802	*
033	059	Hobbs Municipal Schools	Houston Middle	846	815	798	*
033	066	Hobbs Municipal Schools	Jefferson Elementary	173	157	160	162
033	072	Hobbs Municipal Schools	Mills Elementary	195	177	182	184
033	075	Hobbs Municipal Schools	Stone Elementary	201	183	174	182
033	144	Hobbs Municipal Schools	Sanger Elementary	207	195	196	192
033	156	Hobbs Municipal Schools	Southern Heights Elementary	183	166	171	174
033	164	Hobbs Municipal Schools	Taylor Elementary	230	216	218	221
033	172	Hobbs Municipal Schools	B.T. Washington Elementary	43	40	41	38
033	176	Hobbs Municipal Schools	Will Rogers Elementary	133	112	111	120
033	177	Hobbs Municipal Schools	Murray Elementary	209	190	195	196
035	090	Tatum Municipal Schools	Tatum Junior High	55	52	51	18
035	162	Tatum Municipal Schools	Tatum Elementary	86	84	84	81
036	130	Ruidoso Municipal Schools	Ruidoso Middle	415	383	373	334
036	160	Ruidoso Municipal Schools	White Mountain Elementary	404	370	380	367
037	035	Carrizozo Municipal Schools	Carrizozo Elementary	32	31	31	32
037	157	Carrizozo Municipal Schools	Carrizozo Middle	34	30	32	15
038	038	Corona Municipal Schools	Corona Elementary	19	19	19	19
038	039	Corona Municipal Schools	Corona High	16	15	15	16
041	017	Los Alamos Public Schools	Aspen Elementary	243	237	237	232
041	021	Los Alamos Public Schools	Barranca Mesa Elementary	218	207	207	125
041	040	Los Alamos Public Schools	Chamisa Elementary	180	175	176	133
041	101	Los Alamos Public Schools	Mountain Elementary	238	224	232	163
041	124	Los Alamos Public Schools	Los Alamos Middle	617	598	590	589
041	127	Los Alamos Public Schools	Pinon Elementary	167	161	165	136
043	034	Gallup-McKinley Cty Schools	Catherine A. Miller Elementary	178	156	160	151
043	089	Gallup-McKinley Cty Schools	Tse'Yi'Gai High	52	21	48	19
043	130	Gallup-McKinley Cty Schools	Ramah High	118	109	109	91
043	174	Gallup-McKinley Cty Schools	Del Norte Elementary	255	237	240	214
044	104	Mora Independent Schools	Mora Elementary	106	101	91	92
							continued

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
046	003	Alamogordo Public Schools	Mountain View Middle	517	481	473	413
046	028	Alamogordo Public Schools	Buena Vista Elementary	114	107	106	110
046	033	Alamogordo Public Schools	Chaparral Middle	650	614	607	507
046	037	Alamogordo Public Schools	Holloman Middle	187	174	177	155
046	056	Alamogordo Public Schools	Sunset Hills Elementary	234	221	219	161
046	057	Alamogordo Public Schools	High Rolls Mountain Park Elementary	11	9	10	9
046	058	Alamogordo Public Schools	Holloman Elementary	231	207	203	201
046	072	Alamogordo Public Schools	La Luz Elementary	127	117	121	113
046	114	Alamogordo Public Schools	North Elementary	98	91	86	81
046	144	Alamogordo Public Schools	Desert Star Elementary	249	229	238	182
046	150	Alamogordo Public Schools	Sierra Elementary	156	147	145	130
046	181	Alamogordo Public Schools	Yucca Elementary	131	124	124	118
047	160	Tularosa Municipal Schools	Tularosa Intermediate	277	266	261	258
047	164	Tularosa Municipal Schools	Tularosa Middle	147	141	141	137
048	038	Cloudcroft Municipal Schools	Cloudcroft Elementary	89	81	81	80
048	042	Cloudcroft Municipal Schools	Cloudcroft Middle	110	100	99	96
049	163	Tucumcari Public Schools	Tucumcari Middle	219	210	199	195
052	001	San Jon Municipal Schools	San Jon Middle	33	30	27	29
052	144	San Jon Municipal Schools	San Jon Elementary	32	30	31	31
054	044	Dulce Independent Schools	Dulce Elementary	124	115	116	112
054	050	Dulce Independent Schools	Dulce Middle	135	116	119	121
058	047	Elida Municipal Schools	Elida Elementary	45	45	44	41
058	048	Elida Municipal Schools	Elida High	29	27	28	27
061	016	Bernalillo Public Schools	Algodones Elementary	71	70	70	66
061	020	Bernalillo Public Schools	Cochiti Elementary	82	78	77	77
061	024	Bernalillo Public Schools	Cochiti Middle	80	77	78	67
061	026	Bernalillo Public Schools	Bernalillo Middle	420	406	407	386
061	028	Bernalillo Public Schools	Santo Domingo Middle	95	92	90	87
061	127	Bernalillo Public Schools	Placitas Elementary	59	57	57	55
061	136	Bernalillo Public Schools	Bernalillo Elementary	306	296	297	291
061	151	Bernalillo Public Schools	Santo Domingo Elementary	127	125	124	119
063	170	Jemez Valley Public Schools	Jemez Valley Middle	61	56	56	41
064	001	Aztec Municipal Schools	Mosaic Academy	135	121	118	118
064	017	Aztec Municipal Schools	C.V. Koogler Middle	537	513	507	488
064	099	Aztec Municipal Schools	McCoy Avenue Elementary	91	88	84	86
064	123	Aztec Municipal Schools	Park Avenue Elementary	349	333	325	327
064	136	Aztec Municipal Schools	Lydia Rippey Elementary	78	74	76	76
065	015	Farmington Municipal Schools	Animas Elementary	203	182	182	194
065	017	Farmington Municipal Schools	Apache Elementary	243	227	227	228
065	019	Farmington Municipal Schools	Bluffview Elementary	189	180	176	172
							continued

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
065	037	Farmington Municipal Schools	Country Club Elementary	292	269	277	279
065	038	Farmington Municipal Schools	Esperanza Elementary	262	248	250	250
065	058	Farmington Municipal Schools	Hermosa Middle	654	618	594	588
065	059	Farmington Municipal Schools	Heights Middle	732	694	692	682
065	073	Farmington Municipal Schools	Ladera Del Norte Elementary	291	283	284	276
065	095	Farmington Municipal Schools	McCormick Elementary	225	195	202	204
065	100	Farmington Municipal Schools	McKinley Elementary	251	242	238	232
065	106	Farmington Municipal Schools	Mesa Verde Elementary	221	211	210	214
065	108	Farmington Municipal Schools	Mesa View Middle	508	480	470	473
065	118	Farmington Municipal Schools	Northeast Elementary	286	272	268	261
065	162	Farmington Municipal Schools	Tibbetts Middle	707	657	647	648
067	026	Central Consolidated Schools	Eva B. Stokely Elementary	118	113	117	101
067	034	Central Consolidated Schools	Kirtland Middle	436	421	414	*
067	038	Central Consolidated Schools	Kirtland Elementary	228	215	220	175
067	060	Central Consolidated Schools	Judy Nelson Elementary	300	291	288	211
067	075	Central Consolidated Schools	Ojo Amarillo Elementary	173	159	164	156
067	110	Central Consolidated Schools	Mesa Elementary	156	148	148	146
067	114	Central Consolidated Schools	Naschitti Elementary	31	31	31	25
067	116	Central Consolidated Schools	Newcomb Elementary	101	95	96	7
067	126	Central Consolidated Schools	Newcomb Middle	182	172	172	163
067	152	Central Consolidated Schools	Nizhoni Elementary	149	141	141	38
067	160	Central Consolidated Schools	Tse'Bit'Ai Middle	427	396	392	117
068	004	West Las Vegas Public Schools	Rio Gallinas School for Ecology & the Arts	49	46	46	43
068	050	West Las Vegas Public Schools	Valley Elementary	35	33	35	35
068	068	West Las Vegas Public Schools	Valley Middle	27	25	25	25
068	112	West Las Vegas Public Schools	Don Cecilio Martinez Elementary	88	82	85	80
068	125	West Las Vegas Public Schools	Tony Serna Jr. Elementary	111	109	106	105
068	157	West Las Vegas Public Schools	Union Elementary	115	108	111	105
068	172	West Las Vegas Public Schools	West Las Vegas Middle	301	296	291	283
071	005	Santa Fe Public Schools	Cesar Chavez Elementary	159	154	151	153
071	008	Santa Fe Public Schools	Acequia Madre Elementary	96	94	88	89
071	011	Santa Fe Public Schools	El Camino Real Academy Community	523	503	459	489
071	012	Santa Fe Public Schools	Desert Sage Academy	125	90	114	53
071	022	Santa Fe Public Schools	Carlos Gilbert Elementary	188	184	185	184
071	023	Santa Fe Public Schools	Ramirez Thomas Elementary	203	202	191	190
071	024	Santa Fe Public Schools	Academy For Tech & The Classics-ATC	158	153	152	*
071	033	Santa Fe Public Schools	Atalaya Elementary	151	150	145	145
071	054	Santa Fe Public Schools	Aspen Community School	235	233	218	202
071	057	Santa Fe Public Schools	Gonzales Elementary	240	225	231	226
071	070	Santa Fe Public Schools	Kearny Elementary	198	193	188	189
							continued

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
071	099	Santa Fe Public Schools	E.J. Martinez Elementary	101	94	90	91
071	100	Santa Fe Public Schools	Pinon Elementary	304	298	300	300
071	110	Santa Fe Public Schools	Edward Ortiz Middle	466	435	423	435
071	130	Santa Fe Public Schools	R.M. Sweeney Elementary	143	138	137	138
071	135	Santa Fe Public Schools	El Dorado Community School	293	279	273	279
071	141	Santa Fe Public Schools	Amy Biehl at Rancho Viejo Community School	208	199	200	201
071	143	Santa Fe Public Schools	Salazar Elementary	120	107	105	112
071	145	Santa Fe Public Schools	Francis X. Nava Elementary	99	91	94	90
071	146	Santa Fe Public Schools	Chaparral Elementary	127	116	124	123
071	160	Santa Fe Public Schools	Tesuque Elementary	43	40	41	43
071	170	Santa Fe Public Schools	Nina Otero	462	448	444	437
071	173	Santa Fe Public Schools	Mandela International Magnet	125	125	122	3
071	176	Santa Fe Public Schools	Wood-Gormley Elementary	183	177	174	176
071	189	Santa Fe Public Schools	Milagro Middle	403	375	356	362
073	016	T Or C Municipal Schools	Arrey Elementary	43	42	43	38
073	060	T Or C Municipal Schools	Sierra Elementary	149	136	144	133
073	063	T Or C Municipal Schools	T Or C Middle	293	274	263	250
073	162	T Or C Municipal Schools	T Or C Elementary	89	84	85	80
075	100	Magdalena Municipal Schools	Magdalena Middle	78	74	74	74
075	133	Magdalena Municipal Schools	Magdalena Elementary	64	59	57	58
079	001	Questa Independent Schools	Alta Vista Elementary	20	20	17	*
080	001	Estancia Municipal Schools	Upper Elementary	162	138	144	156
080	172	Estancia Municipal Schools	Estancia Middle	99	86	89	89
081	001	Moriarty-Edgewood Municipal Schools	Route 66 Elementary	169	158	155	128
081	003	Moriarty-Edgewood Municipal Schools	Edgewood Middle	304	293	288	179
081	100	Moriarty-Edgewood Municipal Schools	Moriarty Elementary	203	184	187	140
081	102	Moriarty-Edgewood Municipal Schools	Moriarty Middle	251	232	234	174
081	120	Moriarty-Edgewood Municipal Schools	South Mountain Elementary	169	166	159	*
088	038	Grants-Cibola County Schools	Cubero Elementary	122	114	112	113
088	056	Grants-Cibola County Schools	Los Alamitos Middle	431	415	411	402
088	058	Grants-Cibola County Schools	Laguna-Acoma Middle	35	30	29	32
088	099	Grants-Cibola County Schools	Mesa View Elementary	242	228	237	230
088	104	Grants-Cibola County Schools	Milan Elementary	251	238	233	233
088	106	Grants-Cibola County Schools	Mount Taylor Elementary	269	245	238	246
088	152	Grants-Cibola County Schools	San Rafael Elementary	29	27	28	28
088	155	Grants-Cibola County Schools	Seboyeta Elementary	21	21	19	18
088	915	Grants-Cibola County Schools	Bluewater Elementary	51	45	42	44
505	001	School of Dreams Academy	School of Dreams Academy	248	226	225	222
520	001	The ASK Academy	The ASK Academy	363	358	350	347
539	000	Red River Valley Charter School	Red River Valley Charter School DO	50	48	44	46
							continued

District Code	School Code	Parent Organization Name	Organization Name	N	BOY	MOY	EOY
550	000	Estancia Valley Classical Academy	Estancia Valley Classical Academy DO	315	303	303	296
557	001	Explore Academy Charter School	Explore Academy Charter School	514	497	493	481
574	001	Albuquerque Collegiate Charter School	Albuquerque Collegiate Charter School	74	72	70	*
579	001	Aces Technical Charter School	Aces Technical Charter School	125	114	105	99

** No school data available or no test records found.*

Appendix 2

Table 2.1. Differential Validity for the Hispanic Subgroup on the 2022–23 iMSSA Administration

Content Area	Grade	Window	No	Yes
Mathematics	8	EOY	850.462	843.824
Reading	8	EOY	854.1	847.738
Language Usage	8	EOY	851.871	845.398

Table 2.2. Differential Validity for the American Indian / Alaska Native Subgroup on the 2022–23 iMSSA Administration

Content Area	Grade	Window	Unknown	No	Yes
Mathematics	3	MOY	334.493	334.867	326.633
		EOY	347.598	345.198	341.137
	4	BOY	430.533	431.682	421.5
		MOY	440.788	441.672	434.701
		EOY	446.447	445.371	438.454
	5	MOY	543.652	540	534.762
		EOY	550.114	551.805	544.328
	6	EOY	650.982	645.541	644.131
	7	MOY	744.327	741.34	738.445
		EOY	748.36	747.3	737.564
	8	BOY	844.135	841.55	839.442
		EOY	847.973	844.466	842.262
Reading	3	EOY	353.117	355.378	347.983
	4	BOY	447.963	446.782	442.093
		MOY	452.139	453.281	445.836
		EOY	453.311	452.888	445.561
	5	BOY	551.044	549.982	545.063
		MOY	555.481	552.967	548.381
		EOY	551.039	550.98	544.207
	6	BOY	648.964	648.85	641.685
		MOY	649.236	650.4	643.775
	7	BOY	747.596	742.469	740.454
		MOY	748.957	744.806	742.221
		EOY	753.044	750.735	743.249
	8	BOY	849.706	849.232	842.606
		MOY	848.547	848.923	840.781
		EOY	851.942	850.39	845.546
Language Usage	3	BOY	341.477	343.832	335.484
		EOY	354.857	353.716	348.694
	4	BOY	447.606	446.321	440.752
		MOY	453.308	452.883	445.584
		EOY	456.66	456.158	449.592
	5	BOY	550.612	549.096	544.131
		MOY	554.799	553.639	547.983
		EOY	555.645	553.781	548.853

continued

Content Area	Grade	Window	Unknown	No	Yes
Language Usage	6	BOY	648.272	651.894	643.116
		MOY	646.135	647.566	639.075
		EOY	651.725	650.202	645.49
	7	BOY	748.878	744.913	740.331
		MOY	754.361	753.484	748.208
		EOY	754.334	751.286	746.403
	8	BOY	844.464	845.205	838.798
		MOY	849.27	851.681	844.93
		EOY	849.639	849.28	842.977

Table 2.3. Differential Validity for the English Learner Subgroup on the 2022–23 iMSSA Administration

Content Area	Grade	Window	0**	1	2	3	4	5	6
Reading	4	EOY	453.825	442.865	*	439.933	*	*	*
	5	BOY	551.852	541.15	562.237	535.326	*	*	*
		MOY	556.103	545.483	564.308	541.957	*	*	*
	6	EOY	651.46	640.458	657.2	645.833	*	*	*
	8	BOY	849.874	839.433	854.635	841.362	856.571	856.604	854.481
		MOY	849.346	835.004	853.078	846	853.573	854.875	854.691
		EOY	852.547	840.812	854.021	846.865	854.763	854.974	855.745
Language Usage	4	EOY	457.436	446.132	*	443.386	*	*	*
	5	BOY	551.241	540.948	561.026	539.087	*	*	*
	6	MOY	646.695	635.781	654.763	641.043	*	*	*
		EOY	652.242	641.47	658.192	644.946	*	*	*
	7	MOY	754.734	744.909	758.216	750.484	762.265	*	761.864
		EOY	754.471	743.848	760.25	744.118	761.519	*	761.7
	8	BOY	844.838	835.038	846.94	839.234	846.426	850.065	849.333
		MOY	850.123	839.298	851.275	847.487	853.906	856.021	856.054
		EOY	850.273	838.316	849.34	840.378	854.238	854.256	854.875

** Results suppressed due to failure to meet minimum reporting threshold n > 20 students.*

*** English Learner status: Unknown, 0 = Initially Fluent English Proficient – Student never EL, 1 = Current EL Student, 2 through 6 = Reclassified Fluent English Proficient – exited Year 1 through 5, respectively.*

Table 2.4. Differential Validity for the Special Education / Individualized Education Plan Subgroup on the 2022–23 iMSSA Administration

Content Area	Grade	Window	Unknown / Blank	No / 0	Yes / 1
Mathematics	3	MOY	337.176	334.972	322.111
		EOY	348.206	348.052	337.954
	4	EOY	442.789	447.591	434.228
	5	EOY	553.614	550.564	541.325
	6	EOY	651.559	651.276	642.874
Reading	3	MOY	352.207	350.564	341.771
		EOY	355.037	353.662	343.705
	4	MOY	452.628	452.757	442.139
		EOY	452.398	453.797	442.474
	5	BOY	552.577	551.832	540.473
		MOY	558.135	555.812	545.571
		EOY	553.782	551.378	541.138
	6	BOY	648.188	649.468	640.682
		MOY	649.48	649.96	641.343
		EOY	650.354	651.506	642.715
	8	BOY	845.553	850.345	840.883
		MOY	845.781	849.309	837.945
Language Usage	3	EOY	356.931	355.226	344.515
	4	BOY	445.932	447.998	438.615
		MOY	452.43	453.918	442.19
		EOY	453.659	457.553	446.021
	5	BOY	551.774	551.298	540.231
		MOY	557.066	555.46	543.77
		EOY	558.018	556.422	543.904
	6	BOY	646.989	649.091	640.64
		MOY	645.261	647.186	635.749
	7	EOY	751.676	754.765	746.199
	8	BOY	843.002	844.986	836.645
		MOY	848.419	850.329	839.945
		EOY	845.009	850.175	841.066

Table 2.5. Differential Validity for the Economically Disadvantaged Subgroup on the 2022–23 iMSSA Administration

Content Area	Grade	Window	Unknown	No	Yes
Mathematics	5	EOY	554.642	549.464	545.074
	6	MOY	652.457	648.224	644.976
		EOY	654.307	647.292	645.544
	7	MOY	747.087	743.496	739.064
		EOY	751.3	745.857	741.888
	8	BOY	846.779	840.781	840.614
Reading	6	MOY	652.448	647.377	644.415
	7	MOY	751.315	748.106	743.811
Language Usage	5	MOY	557.781	550.715	550.72
		EOY	559.405	552.392	551.087
	6	BOY	651.365	643.513	645.23
	7	MOY	756.795	753.798	749.313

Table 2.6. Differential Validity for the Gender Subgroup on the 2022–23 iMSSA Administration

Content Area	Grade	Window	Female	Male	Unknown
Language Usage	8	MOY	851.482	845.946	*

APPENDIX N—RELIABILITY

Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.

Table N-1. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 3, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		7,392	0.900	2.983	0.795
	Gender	Female	3,543	0.901	2.997	0.814
		Male	3,842	0.899	2.966	0.774
	Ethnicity	African American or Black	266	0.904	2.957	0.766
		American Indian or Alaska Native	900	0.863	2.946	0.707
		Asian	115	0.917	2.965	0.851
		Caucasian	5,969	0.901	2.987	0.800
		Hawaiian Native or Other Pacific Islander	25	--	--	--
		Multi	105	0.913	2.929	0.838
	Hispanic	Yes	4,603	0.891	2.982	0.765
		No	2,777	0.910	2.977	0.830
	Econ. Dis.	Yes	3,645	0.888	2.968	0.747
		No	2,925	0.903	2.996	0.837
	English Learners	Yes	1,138	0.862	2.942	0.663
		No	6,254	0.901	2.986	0.807
	Special Ed	Yes	1,460	0.839	2.863	0.529
		No	5,752	0.897	2.999	0.821
2	**Overall**		6,756	0.898	2.991	0.822
	Gender	Female	3,346	0.898	2.992	0.827
		Male	3,410	0.898	2.987	0.818
	Ethnicity	African American or Black	209	0.888	2.984	0.823
		American Indian or Alaska Native	808	0.852	2.994	0.730
		Asian	118	0.887	2.917	0.891
		Caucasian	5,466	0.900	2.991	0.827
		Hawaiian Native or Other Pacific Islander	25	--	--	--
		Multi	130	0.916	2.916	0.839
	Hispanic	Yes	4,215	0.890	3.001	0.802
		No	2,541	0.906	2.970	0.846
	Econ. Dis.	Yes	3,338	0.887	2.993	0.784
		No	2,636	0.899	2.977	0.856
	English Learners	Yes	999	0.861	2.997	0.730
		No	5,757	0.899	2.986	0.830
	Special Ed	Yes	1,042	0.870	2.896	0.641
		No	5,557	0.894	2.997	0.838
3	**Overall**		6,729	0.900	2.967	0.804
	Gender	Female	3,330	0.899	2.979	0.813
		Male	3,398	0.900	2.950	0.794
	Ethnicity	African American or Black	225	0.891	2.978	0.808
		American Indian or Alaska Native	845	0.873	2.944	0.702
		Asian	117	0.909	2.906	0.835
		Caucasian	5,418	0.900	2.968	0.809
		Hawaiian Native or Other Pacific Islander	18	--	--	--
		Multi	106	0.915	2.963	0.886
	Hispanic	Yes	4,209	0.891	2.980	0.787
		No	2,520	0.910	2.942	0.824
	Econ. Dis.	Yes	3,248	0.888	2.969	0.754
		No	2,727	0.902	2.956	0.848
	English Learners	Yes	937	0.865	2.935	0.681
		No	5,792	0.900	2.969	0.814
	Special Ed	Yes	1,034	0.874	2.840	0.573
		No	5,536	0.895	2.982	0.828

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-2. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 4, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		7,632	0.885	3.018	0.804
	Gender	Female	3,693	0.880	3.048	0.820
		Male	3,929	0.889	2.985	0.788
	Ethnicity	African American or Black	252	0.883	3.008	0.778
		American Indian or Alaska Native	939	0.860	2.960	0.706
		Asian	161	0.886	3.024	0.883
		Caucasian	6,112	0.885	3.023	0.811
		Hawaiian Native or Other Pacific Islander	31	--	--	--
		Multi	122	0.899	3.008	0.813
	Hispanic	Yes	4,692	0.875	3.023	0.786
		No	2,925	0.895	3.006	0.823
	Econ. Dis.	Yes	3,918	0.877	3.005	0.770
		No	3,109	0.885	3.026	0.837
	English Learners	Yes	1,349	0.852	2.968	0.676
		No	6,283	0.885	3.023	0.819
	Special Ed	Yes	1,587	0.843	2.872	0.576
		No	5,861	0.874	3.042	0.831
2	**Overall**		6,851	0.881	2.999	0.811
	Gender	Female	3,394	0.876	3.029	0.825
		Male	3,455	0.884	2.963	0.794
	Ethnicity	African American or Black	186	0.888	2.954	0.827
		American Indian or Alaska Native	846	0.844	3.000	0.720
		Asian	121	0.893	2.957	0.857
		Caucasian	5,565	0.882	2.999	0.816
		Hawaiian Native or Other Pacific Islander	25	--	--	--
		Multi	108	0.904	2.975	0.815
	Hispanic	Yes	4,240	0.874	3.007	0.795
		No	2,611	0.889	2.980	0.828
	Econ. Dis.	Yes	3,517	0.873	2.998	0.776
		No	2,779	0.880	2.995	0.841
	English Learners	Yes	1,113	0.853	2.978	0.696
		No	5,738	0.880	2.998	0.823
	Special Ed	Yes	1,078	0.865	2.854	0.621
		No	5,623	0.870	3.015	0.827
3	**Overall**		6,841	0.886	3.012	0.809
	Gender	Female	3,437	0.886	3.037	0.824
		Male	3,404	0.885	2.980	0.793
	Ethnicity	African American or Black	209	0.881	3.010	0.794
		American Indian or Alaska Native	836	0.860	2.986	0.729
		Asian	113	0.871	2.946	0.889
		Caucasian	5,536	0.885	3.014	0.812
		Hawaiian Native or Other Pacific Islander	19	--	--	--
		Multi	128	0.887	3.014	0.869
	Hispanic	Yes	4,243	0.876	3.012	0.784
		No	2,598	0.894	3.006	0.840
	Econ. Dis.	Yes	3,480	0.872	3.006	0.767
		No	2,800	0.887	3.017	0.846
	English Learners	Yes	1,145	0.853	2.977	0.681
		No	5,696	0.884	3.014	0.823
	Special Ed	Yes	1,106	0.879	2.858	0.636
		No	5,582	0.874	3.033	0.829

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-3. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 5, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		7,849	0.876	3.111	0.816
	Gender	Female	3,768	0.871	3.133	0.828
		Male	4,079	0.879	3.078	0.803
	Ethnicity	African American or Black	259	0.873	3.122	0.821
		American Indian or Alaska Native	925	0.847	3.081	0.729
		Asian	157	0.853	3.160	0.857
		Caucasian	6,342	0.877	3.112	0.822
		Hawaiian Native or Other Pacific Islander	29	--	--	--
		Multi	129	0.868	3.042	0.824
	Hispanic	Yes	4,954	0.865	3.111	0.798
		No	2,887	0.887	3.107	0.837
	Econ. Dis.	Yes	4,068	0.860	3.093	0.780
		No	3,138	0.879	3.122	0.843
	English Learners	Yes	1,537	0.815	3.059	0.691
		No	6,312	0.878	3.118	0.829
	Special Ed	Yes	1,688	0.832	2.921	0.605
		No	5,998	0.863	3.139	0.838
2	**Overall**		7,032	0.876	3.078	0.801
	Gender	Female	3,494	0.873	3.099	0.816
		Male	3,537	0.878	3.052	0.786
	Ethnicity	African American or Black	208	0.875	3.055	0.783
		American Indian or Alaska Native	823	0.841	3.054	0.719
		Asian	133	0.896	3.070	0.850
		Caucasian	5,733	0.876	3.081	0.806
		Hawaiian Native or Other Pacific Islander	21	--	--	--
		Multi	114	0.890	3.056	0.797
	Hispanic	Yes	4,429	0.864	3.082	0.782
		No	2,603	0.890	3.067	0.825
	Econ. Dis.	Yes	3,609	0.858	3.064	0.757
		No	2,871	0.882	3.081	0.835
	English Learners	Yes	1,267	0.826	3.019	0.670
		No	5,765	0.877	3.085	0.815
	Special Ed	Yes	1,124	0.861	2.885	0.593
		No	5,763	0.865	3.103	0.822
3	**Overall**		6,989	0.875	3.124	0.799
	Gender	Female	3,428	0.871	3.157	0.814
		Male	3,561	0.876	3.087	0.781
	Ethnicity	African American or Black	201	0.895	3.078	0.778
		American Indian or Alaska Native	809	0.846	3.056	0.691
		Asian	129	0.885	3.204	0.880
		Caucasian	5,719	0.873	3.130	0.805
		Hawaiian Native or Other Pacific Islander	17	--	--	--
		Multi	114	0.885	3.117	0.840
	Hispanic	Yes	4,368	0.862	3.113	0.774
		No	2,621	0.886	3.135	0.827
	Econ. Dis.	Yes	3,597	0.857	3.086	0.748
		No	2,838	0.877	3.162	0.839
	English Learners	Yes	1,298	0.831	3.024	0.650
		No	5,691	0.874	3.140	0.816
	Special Ed	Yes	1,124	0.863	2.907	0.613
		No	5,723	0.865	3.156	0.816

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-4. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 6, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		7,913	0.867	3.088	0.850
	Gender	Female	3,843	0.863	3.115	0.859
		Male	4,064	0.871	3.055	0.840
	Ethnicity	African American or Black	234	0.872	3.085	0.838
		American Indian or Alaska Native	916	0.844	3.025	0.794
		Asian	132	0.894	3.117	0.901
		Caucasian	6,464	0.867	3.094	0.854
		Hawaiian Native or Other Pacific Islander	39	--	--	--
		Multi	113	0.878	3.077	0.856
	Hispanic	Yes	5,006	0.856	3.081	0.833
		No	2,892	0.880	3.093	0.869
	Econ. Dis.	Yes	3,888	0.853	3.060	0.816
		No	3,472	0.867	3.114	0.872
	English Learners	Yes	1,467	0.809	3.008	0.747
		No	6,446	0.868	3.100	0.860
	Special Ed	Yes	1,732	0.846	2.900	0.713
		No	6,008	0.849	3.123	0.864
2	**Overall**		7,063	0.858	3.055	0.846
	Gender	Female	3,553	0.850	3.075	0.851
		Male	3,510	0.866	3.027	0.842
	Ethnicity	African American or Black	247	0.871	3.024	0.833
		American Indian or Alaska Native	817	0.838	3.036	0.810
		Asian	113	0.865	3.022	0.868
		Caucasian	5,768	0.857	3.056	0.849
		Hawaiian Native or Other Pacific Islander	23	--	--	--
		Multi	95	0.866	3.057	0.878
	Hispanic	Yes	4,454	0.846	3.053	0.832
		No	2,609	0.869	3.053	0.863
	Econ. Dis.	Yes	3,453	0.840	3.048	0.819
		No	3,100	0.859	3.058	0.863
	English Learners	Yes	1,220	0.794	3.000	0.746
		No	5,843	0.857	3.061	0.854
	Special Ed	Yes	1,160	0.851	2.912	0.742
		No	5,762	0.843	3.075	0.853
3	**Overall**		7,061	0.860	3.035	0.852
	Gender	Female	3,483	0.855	3.052	0.857
		Male	3,577	0.865	3.010	0.846
	Ethnicity	African American or Black	216	0.856	3.057	0.835
		American Indian or Alaska Native	813	0.833	3.005	0.810
		Asian	132	0.864	3.023	0.868
		Caucasian	5,760	0.860	3.036	0.854
		Hawaiian Native or Other Pacific Islander	32	--	--	--
		Multi	108	0.872	3.031	0.874
	Hispanic	Yes	4,432	0.848	3.032	0.837
		No	2,629	0.872	3.036	0.868
	Econ. Dis.	Yes	3,469	0.842	3.030	0.823
		No	3,060	0.861	3.034	0.868
	English Learners	Yes	1,159	0.797	2.984	0.756
		No	5,902	0.860	3.041	0.859
	Special Ed	Yes	1,107	0.873	2.888	0.769
		No	5,818	0.844	3.054	0.856

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-5. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 7, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		7,900	0.835	2.994	0.827
	Gender	Female	3,791	0.827	3.001	0.834
		Male	4,091	0.840	2.977	0.818
	Ethnicity	African American or Black	237	0.842	2.960	0.805
		American Indian or Alaska Native	962	0.809	2.999	0.793
		Asian	130	0.841	3.017	0.867
		Caucasian	6,411	0.836	2.991	0.829
		Hawaiian Native or Other Pacific Islander	37	--	--	--
		Multi	96	0.810	3.020	0.817
	Hispanic	Yes	5,007	0.824	2.989	0.812
		No	2,866	0.848	2.996	0.847
	Econ. Dis.	Yes	3,791	0.820	2.981	0.796
		No	3,505	0.836	2.999	0.847
	English Learners	Yes	1,503	0.780	2.988	0.744
		No	6,397	0.836	2.990	0.835
	Special Ed	Yes	1,582	0.804	2.942	0.697
		No	6,150	0.814	2.989	0.833
2	**Overall**		7,219	0.835	3.031	0.831
	Gender	Female	3,593	0.828	3.045	0.834
		Male	3,625	0.840	3.005	0.825
	Ethnicity	African American or Black	211	0.851	3.017	0.812
		American Indian or Alaska Native	897	0.801	3.031	0.782
		Asian	144	0.856	3.037	0.867
		Caucasian	5,833	0.834	3.030	0.834
		Hawaiian Native or Other Pacific Islander	40	--	--	--
		Multi	94	0.872	3.006	0.863
	Hispanic	Yes	4,420	0.818	3.020	0.810
		No	2,799	0.850	3.044	0.854
	Econ. Dis.	Yes	3,394	0.810	3.016	0.791
		No	3,278	0.837	3.034	0.853
	English Learners	Yes	1,360	0.774	2.999	0.734
		No	5,859	0.834	3.034	0.839
	Special Ed	Yes	1,117	0.844	2.951	0.737
		No	5,959	0.816	3.034	0.834
3	**Overall**		7,222	0.835	3.058	0.807
	Gender	Female	3,563	0.823	3.091	0.815
		Male	3,658	0.845	3.012	0.797
	Ethnicity	African American or Black	204	0.822	3.010	0.778
		American Indian or Alaska Native	881	0.804	3.059	0.746
		Asian	102	0.830	3.126	0.855
		Caucasian	5,901	0.837	3.054	0.812
		Hawaiian Native or Other Pacific Islander	40	--	--	--
		Multi	94	0.810	3.116	0.836
	Hispanic	Yes	4,525	0.821	3.043	0.786
		No	2,697	0.847	3.074	0.830
	Econ. Dis.	Yes	3,438	0.814	3.027	0.762
		No	3,256	0.839	3.078	0.836
	English Learners	Yes	1,308	0.764	3.001	0.686
		No	5,914	0.835	3.065	0.818
	Special Ed	Yes	1,097	0.841	2.925	0.678
		No	5,983	0.817	3.072	0.815

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-6. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 8, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		8,318	0.864	3.026	0.841
	Gender	Female	4,044	0.850	3.057	0.851
		Male	4,261	0.872	2.979	0.827
	Ethnicity	African American or Black	244	0.885	2.991	0.839
		American Indian or Alaska Native	947	0.834	3.027	0.796
		Asian	131	0.852	3.032	0.874
		Caucasian	6,845	0.864	3.025	0.843
		Hawaiian Native or Other Pacific Islander	40	--	--	--
		Multi	98	0.893	3.032	0.895
	Hispanic	Yes	5,310	0.855	3.020	0.826
		No	2,995	0.873	3.033	0.859
	Econ. Dis.	Yes	3,956	0.851	3.013	0.815
		No	3,761	0.868	3.030	0.857
	English Learners	Yes	1,508	0.813	2.970	0.742
		No	6,810	0.862	3.034	0.849
	Special Ed	Yes	1,691	0.842	2.902	0.727
		No	6,454	0.847	3.043	0.847
2	**Overall**		7,653	0.865	3.025	0.837
	Gender	Female	3,777	0.859	3.057	0.851
		Male	3,875	0.867	2.985	0.820
	Ethnicity	African American or Black	232	0.857	3.017	0.844
		American Indian or Alaska Native	925	0.842	3.017	0.799
		Asian	129	0.822	3.076	0.864
		Caucasian	6,244	0.865	3.023	0.838
		Hawaiian Native or Other Pacific Islander	25	--	--	--
		Multi	98	0.856	3.031	0.823
	Hispanic	Yes	4,797	0.852	3.013	0.818
		No	2,856	0.877	3.040	0.860
	Econ. Dis.	Yes	3,597	0.847	3.000	0.801
		No	3,477	0.868	3.041	0.860
	English Learners	Yes	1,331	0.807	2.940	0.717
		No	6,322	0.862	3.038	0.848
	Special Ed	Yes	1,190	0.854	2.863	0.723
		No	6,305	0.853	3.048	0.845
3	**Overall**		7,623	0.862	3.079	0.839
	Gender	Female	3,737	0.855	3.120	0.849
		Male	3,883	0.865	3.026	0.826
	Ethnicity	African American or Black	238	0.862	3.086	0.832
		American Indian or Alaska Native	890	0.836	3.087	0.802
		Asian	134	0.901	3.034	0.882
		Caucasian	6,228	0.862	3.077	0.840
		Hawaiian Native or Other Pacific Islander	42	--	--	--
		Multi	91	0.876	3.095	0.872
	Hispanic	Yes	4,821	0.851	3.066	0.821
		No	2,802	0.871	3.096	0.860
	Econ. Dis.	Yes	3,593	0.849	3.054	0.810
		No	3,483	0.866	3.093	0.858
	English Learners	Yes	1,330	0.793	2.980	0.719
		No	6,293	0.860	3.093	0.848
	Special Ed	Yes	1,199	0.850	2.922	0.731
		No	6,279	0.849	3.101	0.847

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-7. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 3, as a Function of Subgroup*

Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
Overall		20,702	0.899	3.163	0.756
Gender	Female	10,144	0.893	3.143	0.745
	Male	10,557	0.905	3.172	0.764
Ethnicity	African American or Black	687	0.893	3.068	0.700
	American Indian or Alaska Native	2,550	0.862	2.998	0.610
	Asian	353	0.916	3.278	0.896
	Caucasian	16,703	0.899	3.178	0.766
	Hawaiian Native or Other Pacific Islander	68	0.902	3.087	0.763
	Multi	341	0.920	3.245	0.796
Hispanic	Yes	12,880	0.885	3.130	0.720
	No	7,822	0.912	3.204	0.797
Econ. Dis.	Yes	10,152	0.881	3.100	0.684
	No	8,209	0.906	3.228	0.817
English Learners	Yes	3,033	0.867	3.018	0.617
	No	17,669	0.901	3.181	0.771
Special Ed	Yes	3,519	0.869	2.900	0.487
	No	16,699	0.897	3.201	0.791

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-8. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 4, as a Function of Subgroup*

Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
Overall		21,365	0.890	3.180	0.725
Gender	Female	10,540	0.874	3.144	0.705
	Male	10,813	0.901	3.208	0.740
Ethnicity	African American or Black	647	0.875	3.115	0.695
	American Indian or Alaska Native	2,619	0.841	3.011	0.562
	Asian	400	0.928	3.338	0.872
	Caucasian	17,248	0.890	3.195	0.734
	Hawaiian Native or Other Pacific Islander	75	0.848	3.174	0.705
	Multi	361	0.912	3.179	0.738
Hispanic	Yes	13,203	0.870	3.130	0.674
	No	8,147	0.907	3.246	0.782
Econ. Dis.	Yes	10,925	0.865	3.088	0.633
	No	8,718	0.901	3.276	0.801
English Learners	Yes	3,626	0.844	2.992	0.537
	No	17,739	0.893	3.209	0.750
Special Ed	Yes	3,776	0.878	2.877	0.445
	No	17,102	0.886	3.223	0.765

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-9. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 5, as a Function of Subgroup*

Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
Overall		21,872	0.900	3.363	0.747
Gender	Female	10,685	0.891	3.376	0.739
	Male	11,184	0.907	3.346	0.753
Ethnicity	African American or Black	668	0.890	3.324	0.710
	American Indian or Alaska Native	2,558	0.862	3.259	0.612
	Asian	423	0.924	3.402	0.889
	Caucasian	17,792	0.900	3.370	0.754
	Hawaiian Native or Other Pacific Islander	67	0.918	3.379	0.821
	Multi	356	0.918	3.346	0.777
Hispanic	Yes	13,751	0.886	3.331	0.703
	No	8,113	0.912	3.400	0.797
Econ. Dis.	Yes	11,278	0.876	3.293	0.655
	No	8,842	0.909	3.414	0.817
English Learners	Yes	4,108	0.853	3.216	0.569
	No	17,764	0.903	3.384	0.771
Special Ed	Yes	3,936	0.880	3.068	0.452
	No	17,487	0.896	3.391	0.786

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-10. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 6, as a Function of Subgroup*

Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
Overall		22,054	0.878	3.179	0.617
Gender	Female	10,885	0.873	3.182	0.605
	Male	11,162	0.883	3.171	0.628
Ethnicity	African American or Black	702	0.869	3.115	0.545
	American Indian or Alaska Native	2,547	0.812	3.025	0.399
	Asian	381	0.921	3.422	0.840
	Caucasian	17,998	0.879	3.190	0.629
	Hawaiian Native or Other Pacific Islander	94	0.827	3.132	0.526
	Multi	317	0.894	3.228	0.686
Hispanic	Yes	13,893	0.856	3.111	0.548
	No	8,146	0.897	3.279	0.698
Econ. Dis.	Yes	10,815	0.845	3.067	0.482
	No	9,644	0.892	3.295	0.725
English Learners	Yes	3,853	0.788	2.943	0.287
	No	18,201	0.883	3.219	0.656
Special Ed	Yes	3,993	0.845	2.899	0.196
	No	17,611	0.876	3.224	0.674

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-11. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 7, as a Function of Subgroup*

Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
Overall		22,310	0.888	3.243	0.537
Gender	Female	10,930	0.877	3.234	0.505
	Male	11,360	0.897	3.247	0.563
Ethnicity	African American or Black	653	0.871	3.133	0.444
	American Indian or Alaska Native	2,731	0.820	3.070	0.244
	Asian	378	0.926	3.505	0.817
	Caucasian	18,121	0.889	3.259	0.556
	Hawaiian Native or Other Pacific Islander	116	0.891	3.206	0.571
	Multi	284	0.914	3.323	0.571
Hispanic	Yes	13,921	0.865	3.177	0.442
	No	8,362	0.906	3.338	0.644
Econ. Dis.	Yes	10,612	0.848	3.125	0.344
	No	10,023	0.904	3.347	0.668
English Learners	Yes	4,174	0.764	2.975	0.062
	No	18,136	0.894	3.290	0.591
Special Ed	Yes	3,791	0.873	2.937	0.084
	No	18,066	0.886	3.288	0.597

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-12. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 8, as a Function of Subgroup*

Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
Overall		23,601	0.849	3.058	0.668
Gender	Female	11,548	0.833	3.036	0.660
	Male	12,035	0.862	3.075	0.674
Ethnicity	African American or Black	715	0.840	3.018	0.652
	American Indian or Alaska Native	2,764	0.754	2.913	0.535
	Asian	396	0.911	3.419	0.838
	Caucasian	19,319	0.849	3.065	0.672
	Hawaiian Native or Other Pacific Islander	109	0.849	3.173	0.703
	Multi	285	0.888	3.202	0.749
Hispanic	Yes	14,917	0.811	2.987	0.608
	No	8,671	0.878	3.169	0.738
Econ. Dis.	Yes	11,144	0.801	2.952	0.570
	No	10,741	0.869	3.160	0.733
English Learners	Yes	4,165	0.676	2.805	0.373
	No	19,436	0.856	3.104	0.700
Special Ed	Yes	4,089	0.815	2.808	0.399
	No	19,034	0.848	3.101	0.699

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-13. Scaled Score Descriptive Statistics for NM-ASR Science Grade 5, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		11,923	0.896	3.622	0.906
	Gender	Female	5,700	0.888	3.661	0.901
		Male	6,223	0.903	3.579	0.910
	Ethnicity	African American or Black	343	0.893	3.639	0.904
		American Indian or Alaska Native	1,630	0.850	3.597	0.863
		Asian	229	0.911	3.643	0.920
		Caucasian	9,462	0.899	3.622	0.908
		Hawaiian Native or Other Pacific Islander	31	--	--	--
		Multi	228	0.899	3.607	0.908
	Hispanic	Yes	7,309	0.882	3.620	0.893
		No	4,614	0.908	3.622	0.917
	Econ. Dis.	Yes	6,141	0.878	3.598	0.887
		No	4,553	0.904	3.634	0.914
	English Learners	Yes	2,442	0.833	3.552	0.844
		No	9,481	0.899	3.632	0.909
	Special Ed	Yes	2,911	0.869	3.437	0.866
		No	8,808	0.889	3.653	0.901
2	**Overall**		7,325	0.895	3.603	0.907
	Gender	Female	3,667	0.886	3.624	0.901
		Male	3,657	0.902	3.574	0.913
	Ethnicity	African American or Black	230	0.892	3.567	0.903
		American Indian or Alaska Native	690	0.862	3.547	0.874
		Asian	132	0.915	3.556	0.923
		Caucasian	6,148	0.894	3.608	0.907
		Hawaiian Native or Other Pacific Islander	25	--	--	--
		Multi	100	0.913	3.549	0.919
	Hispanic	Yes	4,760	0.878	3.611	0.893
		No	2,565	0.911	3.579	0.920
	Econ. Dis.	Yes	3,796	0.877	3.603	0.892
		No	3,142	0.901	3.588	0.912
	English Learners	Yes	1,221	0.833	3.543	0.847
		No	6,104	0.896	3.606	0.909
	Special Ed	Yes	750	0.915	3.476	0.916
		No	6,396	0.890	3.610	0.904
3	**Overall**		2,621	0.892	3.647	0.903
	Gender	Female	1,309	0.885	3.675	0.898
		Male	1,310	0.900	3.610	0.907
	Ethnicity	African American or Black	96	0.865	3.665	0.876
		American Indian or Alaska Native	225	0.869	3.569	0.877
		Asian	60	0.891	3.695	0.907
		Caucasian	2,190	0.892	3.651	0.903
		Hawaiian Native or Other Pacific Islander	11	--	--	--
		Multi	31	--	--	--
	Hispanic	Yes	1,685	0.880	3.647	0.891
		No	928	0.902	3.641	0.912
	Econ. Dis.	Yes	1,338	0.870	3.645	0.882
		No	1,148	0.899	3.639	0.908
	English Learners	Yes	450	0.859	3.558	0.865
		No	2,171	0.892	3.660	0.904
	Special Ed	Yes	270	0.905	3.523	0.903
		No	2,285	0.889	3.657	0.901

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-14. Scaled Score Descriptive Statistics for NM-ASR Science Grade 8, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		12,046	0.867	3.477	0.872
	Gender	Female	5,773	0.856	3.492	0.865
		Male	6,271	0.875	3.457	0.878
	Ethnicity	African American or Black	349	0.853	3.426	0.856
		American Indian or Alaska Native	1,664	0.805	3.447	0.817
		Asian	188	0.892	3.575	0.908
		Caucasian	9,641	0.870	3.478	0.876
		Hawaiian Native or Other Pacific Islander	50	0.855	3.551	0.869
		Multi	154	0.869	3.559	0.883
	Hispanic	Yes	7,496	0.845	3.450	0.850
		No	4,550	0.884	3.515	0.892
	Econ. Dis.	Yes	5,667	0.836	3.419	0.839
		No	5,155	0.881	3.525	0.890
	English Learners	Yes	2,229	0.745	3.328	0.747
		No	9,817	0.871	3.503	0.880
	Special Ed	Yes	2,978	0.815	3.288	0.800
		No	8,864	0.863	3.523	0.874
2	**Overall**		8,521	0.880	3.515	0.891
	Gender	Female	4,283	0.871	3.525	0.883
		Male	4,236	0.889	3.497	0.897
	Ethnicity	African American or Black	255	0.891	3.490	0.899
		American Indian or Alaska Native	828	0.825	3.437	0.833
		Asian	149	0.914	3.521	0.924
		Caucasian	7,157	0.880	3.520	0.891
		Hawaiian Native or Other Pacific Islander	42	--	--	--
		Multi	90	0.894	3.634	0.914
	Hispanic	Yes	5,450	0.858	3.495	0.869
		No	3,071	0.899	3.541	0.910
	Econ. Dis.	Yes	4,070	0.858	3.472	0.867
		No	4,079	0.890	3.549	0.902
	English Learners	Yes	1,422	0.758	3.349	0.772
		No	7,099	0.882	3.537	0.895
	Special Ed	Yes	811	0.891	3.346	0.882
		No	7,511	0.877	3.528	0.889
3	**Overall**		3,019	0.872	3.563	0.881
	Gender	Female	1,491	0.860	3.578	0.873
		Male	1,516	0.882	3.543	0.888
	Ethnicity	African American or Black	110	0.869	3.520	0.877
		American Indian or Alaska Native	271	0.826	3.517	0.831
		Asian	57	0.887	3.735	0.910
		Caucasian	2,512	0.873	3.564	0.883
		Hawaiian Native or Other Pacific Islander	15	--	--	--
		Multi	42	--	--	--
	Hispanic	Yes	1,978	0.854	3.554	0.865
		No	1,029	0.891	3.570	0.900
	Econ. Dis.	Yes	1,404	0.845	3.524	0.853
		No	1,479	0.881	3.594	0.893
	English Learners	Yes	524	0.760	3.463	0.785
		No	2,495	0.875	3.577	0.887
	Special Ed	Yes	294	0.874	3.502	0.872
		No	2,650	0.870	3.571	0.881

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

Table N-15. Scaled Score Descriptive Statistics for NM-ASR Science Grade 11, as a Function of OP Form and Subgroup *

OP Form	Grouping	Subgroup	Number of Students	Coefficient α	Classical SEM	IRT Marginal Reliability
1	**Overall**		9,984	0.876	3.722	0.885
	Gender	Female	4,963	0.855	3.717	0.867
		Male	5,017	0.891	3.716	0.898
	Ethnicity	African American or Black	271	0.852	3.689	0.858
		American Indian or Alaska Native	1,552	0.809	3.665	0.823
		Asian	190	0.896	3.911	0.916
		Caucasian	7,845	0.881	3.724	0.890
		Hawaiian Native or Other Pacific Islander	39	--	--	--
		Multi	87	0.905	3.817	0.864
	Hispanic	Yes	5,863	0.854	3.687	0.863
		No	4,121	0.891	3.760	0.901
	Econ. Dis.	Yes	3,639	0.825	3.635	0.830
		No	5,452	0.889	3.767	0.901
	English Learners	Yes	1,433	0.698	3.453	0.699
		No	8,551	0.878	3.754	0.890
	Special Ed	Yes	1,631	0.779	3.443	0.761
		No	8,352	0.875	3.757	0.889
2	**Overall**		8,290	0.872	3.804	0.886
	Gender	Female	4,249	0.848	3.811	0.868
		Male	4,038	0.889	3.787	0.899
	Ethnicity	African American or Black	226	0.841	3.772	0.857
		American Indian or Alaska Native	853	0.809	3.774	0.833
		Asian	158	0.897	3.901	0.915
		Caucasian	6,941	0.875	3.804	0.889
		Hawaiian Native or Other Pacific Islander	36	--	--	--
		Multi	76	0.886	3.829	0.901
	Hispanic	Yes	5,303	0.849	3.761	0.863
		No	2,987	0.888	3.862	0.905
	Econ. Dis.	Yes	3,259	0.830	3.734	0.844
		No	5,000	0.882	3.839	0.898
	English Learners	Yes	1,059	0.667	3.505	0.686
		No	7,231	0.871	3.834	0.889
	Special Ed	Yes	863	0.812	3.508	0.800
		No	7,427	0.870	3.827	0.887
3	**Overall**		2,883	0.875	3.753	0.886
	Gender	Female	1,468	0.861	3.744	0.874
		Male	1,413	0.887	3.749	0.896
	Ethnicity	African American or Black	84	0.876	3.699	0.890
		American Indian or Alaska Native	288	0.819	3.757	0.833
		Asian	63	0.887	3.878	0.906
		Caucasian	2,410	0.877	3.745	0.887
		Hawaiian Native or Other Pacific Islander	12	--	--	--
		Multi	25	--	--	--
	Hispanic	Yes	1,817	0.848	3.701	0.857
		No	1,065	0.892	3.820	0.907
	Econ. Dis.	Yes	1,133	0.845	3.695	0.852
		No	1,725	0.885	3.781	0.897
	English Learners	Yes	377	0.665	3.481	0.679
		No	2,506	0.876	3.778	0.889
	Special Ed	Yes	324	0.777	3.474	0.758
		No	2,558	0.874	3.776	0.887

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistic values are suppressed for those Content Areas/grades with fewer than 50 students.*

APPENDIX O
DECISION ACCURACY AND CONSISTENCY RESULTS

Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistic values are suppressed for those content areas/grades with fewer than 50 students.

Table O-1. Decision Accuracy for NM-MSSA ELA Forms, as a Function of OP Form, Grade, Performance Level, and Cut Score*

OP Form	Grade	Number of Students	Overall	PL 1	PL 2	PL 3	PL 4	Cut 1	Cut 2	Cut 3
1	3	7,392	0.751	0.863	0.758	0.448	0.817	0.919	0.903	0.919
	4	7,632	0.718	0.866	0.682	0.518	0.773	0.915	0.895	0.901
	5	7,849	0.704	0.854	0.668	0.539	0.685	0.901	0.883	0.914
	6	7,913	0.720	0.820	0.807	0.477	0.720	0.926	0.885	0.903
	7	7,900	0.676	0.833	0.728	0.571	0.533	0.926	0.865	0.880
	8	8,318	0.693	0.840	0.699	0.474	0.712	0.908	0.876	0.899
2	3	6,756	0.739	0.844	0.759	0.502	0.782	0.927	0.893	0.910
	4	6,851	0.706	0.859	0.674	0.525	0.744	0.915	0.890	0.894
	5	7,032	0.686	0.874	0.634	0.527	0.621	0.909	0.882	0.886
	6	7,063	0.709	0.808	0.775	0.555	0.667	0.929	0.874	0.900
	7	7,219	0.687	0.804	0.775	0.513	0.644	0.926	0.868	0.888
	8	7,653	0.688	0.830	0.718	0.444	0.727	0.915	0.875	0.886
3	3	6,729	0.749	0.851	0.773	0.420	0.829	0.922	0.901	0.916
	4	6,841	0.702	0.831	0.729	0.496	0.745	0.920	0.885	0.890
	5	6,989	0.697	0.881	0.593	0.545	0.684	0.908	0.889	0.891
	6	7,061	0.714	0.803	0.774	0.526	0.746	0.928	0.879	0.900
	7	7,222	0.679	0.803	0.766	0.547	0.563	0.931	0.861	0.881
	8	7,623	0.678	0.822	0.709	0.487	0.673	0.915	0.867	0.883

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-2. Decision Consistency for NM-MSSA Mathematics Forms, as a Function of Grade, Performance Level, and Cut Score*

Grade	Number of Students	Overall	PL 1	PL 2	PL 3	PL 4	Cut 1	Cut 2	Cut 3
3	20,702	0.782	0.910	0.709	0.620	0.000	0.909	0.912	0.960
4	21,365	0.771	0.888	0.682	0.650	0.753	0.895	0.915	0.960
5	21,872	0.743	0.880	0.601	0.613	0.787	0.902	0.900	0.937
6	22,054	0.742	0.878	0.598	0.653	0.741	0.892	0.900	0.946
7	22,310	0.760	0.898	0.572	0.498	0.788	0.892	0.912	0.947
8	23,601	0.739	0.899	0.657	0.508	0.000	0.891	0.872	0.972

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

.

Table O-3. Decision Accuracy for ASR Science Forms, as a Function of OP Form, Grade, Performance Level, and Cut Score*

OP Form	Grade	Number of Students	Overall	PL 1	PL 2	PL 3	PL 4	Cut 1	Cut 2	Cut 3
1	5	11,923	0.793	0.822	0.771	0.793	0.806	0.898	0.922	0.972
	8	12,046	0.801	0.708	0.811	0.847	0.855	0.881	0.924	0.995
	11	9,984	0.780	0.813	0.627	0.876	0.806	0.874	0.907	0.997
2	5	7,325	0.782	0.783	0.767	0.782	0.843	0.900	0.918	0.964
	8	8,521	0.816	0.721	0.811	0.864	0.849	0.909	0.913	0.994
	11	8,290	0.769	0.773	0.639	0.868	0.842	0.875	0.896	0.995
3	5	2,621	0.788	0.819	0.764	0.799	0.791	0.906	0.913	0.969
	8	3,019	0.813	0.743	0.816	0.844	0.828	0.910	0.908	0.995
	11	2,883	0.773	0.789	0.630	0.875	0.830	0.872	0.903	0.996

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-4. Decision Accuracy for NM-MSSA SLA Forms, as a Function of OP Form, Grade, Performance Level, and Cut Score*

OP Form	Grade	Number of Students	Overall	PL 1	PL 2	PL 3	PL 4	Cut 1	Cut 2	Cut 3
1	3	7,392	0.666	0.791	0.672	0.341	0.714	0.885	0.864	0.884
	4	7,632	0.625	0.796	0.577	0.408	0.657	0.881	0.853	0.861
	5	7,849	0.609	0.785	0.562	0.432	0.526	0.860	0.838	0.877
	6	7,913	0.625	0.715	0.733	0.377	0.564	0.894	0.838	0.866
	7	7,900	0.575	0.726	0.629	0.496	0.374	0.895	0.811	0.838
	8	8,318	0.596	0.756	0.599	0.370	0.576	0.869	0.828	0.855
2	3	6,756	0.650	0.762	0.676	0.391	0.677	0.895	0.852	0.871
	4	6,851	0.610	0.784	0.567	0.415	0.625	0.880	0.846	0.851
	5	7,032	0.594	0.810	0.517	0.439	0.472	0.872	0.835	0.848
	6	7,063	0.607	0.690	0.698	0.450	0.511	0.898	0.825	0.858
	7	7,219	0.586	0.682	0.690	0.424	0.475	0.893	0.815	0.849
	8	7,653	0.590	0.741	0.615	0.345	0.603	0.878	0.826	0.838
3	3	6,729	0.663	0.772	0.689	0.316	0.730	0.888	0.861	0.879
	4	6,841	0.603	0.746	0.626	0.388	0.627	0.886	0.840	0.844
	5	6,989	0.605	0.819	0.479	0.441	0.549	0.871	0.844	0.849
	6	7,061	0.614	0.684	0.696	0.419	0.606	0.896	0.831	0.859
	7	7,222	0.577	0.682	0.673	0.468	0.399	0.900	0.806	0.841
	8	7,623	0.577	0.728	0.607	0.386	0.542	0.878	0.817	0.835

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-5. Decision Consistency for NM-MSSA Spanish Mathematics Forms, as a Function of Grade, Performance Level, and Cut Score*

Grade	Number of Students	Overall	PL 1	PL 2	PL 3	PL 4	Cut 1	Cut 2	Cut 3
3	20,702	0.707	0.867	0.604	0.542	0.282	0.872	0.875	0.950
4	21,365	0.696	0.840	0.578	0.569	0.477	0.854	0.878	0.955
5	21,872	0.659	0.829	0.490	0.502	0.647	0.862	0.861	0.910
6	22,054	0.660	0.826	0.488	0.562	0.516	0.849	0.859	0.930
7	22,310	0.686	0.861	0.462	0.380	0.636	0.849	0.876	0.923
8	23,601	0.661	0.844	0.553	0.418	0.136	0.848	0.824	0.965

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-6. Decision Accuracy for ASR Spanish Science Forms, as a Function of OP Form, Grade, Performance Level, and Cut Score*

OP Form	Grade	Number of Students	Overall	PL 1	PL 2	PL 3	PL 4	Cut 1	Cut 2	Cut 3
1	5	11,923	0.709	0.750	0.691	0.699	0.671	0.857	0.891	0.959
	8	12,046	0.722	0.584	0.756	0.762	0.667	0.836	0.891	0.994
	11	9,984	0.701	0.745	0.514	0.807	0.615	0.824	0.869	0.996
2	5	7,325	0.694	0.692	0.688	0.690	0.733	0.860	0.885	0.948
	8	8,521	0.743	0.579	0.760	0.795	0.663	0.874	0.878	0.992
	11	8,290	0.689	0.689	0.531	0.801	0.669	0.827	0.854	0.994
3	5	2,621	0.702	0.734	0.688	0.709	0.647	0.868	0.879	0.954
	8	3,019	0.736	0.613	0.762	0.766	0.600	0.873	0.869	0.993
	11	2,883	0.692	0.712	0.519	0.807	0.659	0.821	0.863	0.994

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-7. Overall Kappa, Cut Score False Positive Rates, and Cut Score False Negative Rates for NM-MSSA ELA Forms, as a Function of OP Form, Subject, and Grade

OP Form	Grade	Number of Students	Kappa	FP Cut 1	FP Cut 2	FP Cut 3	FN Cut 1	FN Cut 2	FN Cut 3
1	3	7,392	0.540	0.039	0.046	0.044	0.042	0.051	0.036
	4	7,632	0.494	0.040	0.049	0.055	0.045	0.056	0.044
	5	7,849	0.459	0.051	0.052	0.049	0.048	0.064	0.037
	6	7,913	0.464	0.035	0.047	0.061	0.039	0.068	0.036
	7	7,900	0.406	0.031	0.063	0.080	0.043	0.072	0.040
	8	8,318	0.450	0.045	0.055	0.052	0.048	0.069	0.049
2	3	6,756	0.521	0.037	0.054	0.044	0.037	0.053	0.046
	4	6,851	0.476	0.040	0.050	0.055	0.045	0.060	0.051
	5	7,032	0.444	0.043	0.051	0.067	0.048	0.067	0.047
	6	7,063	0.442	0.032	0.060	0.054	0.039	0.066	0.046
	7	7,219	0.413	0.032	0.054	0.072	0.042	0.079	0.039
	8	7,653	0.446	0.043	0.051	0.059	0.042	0.074	0.056
3	3	6,729	0.536	0.039	0.043	0.047	0.040	0.055	0.037
	4	6,841	0.467	0.042	0.046	0.057	0.038	0.069	0.053
	5	6,989	0.460	0.042	0.054	0.058	0.050	0.057	0.051
	6	7,061	0.458	0.032	0.058	0.058	0.040	0.063	0.042
	7	7,222	0.400	0.031	0.058	0.079	0.037	0.081	0.040
	8	7,623	0.427	0.043	0.057	0.055	0.042	0.076	0.062

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-8. Overall Kappa, Cut Score False Positive Rates, and Cut Score False Negative Rates for NM-MSSA Mathematics Forms, as a Function of OP Form, Subject, and Grade

Grade	Number of Students	Kappa	FP Cut 1	FP Cut 2	FP Cut 3	FN Cut 1	FN Cut 2	FN Cut 3
3	20,702	0.548	0.042	0.038	0.040	0.049	0.050	0.000
4	21,365	0.537	0.052	0.040	0.038	0.053	0.046	0.002
5	21,872	0.516	0.051	0.050	0.039	0.047	0.050	0.024
6	22,054	0.499	0.054	0.049	0.045	0.054	0.051	0.010
7	22,310	0.502	0.055	0.046	0.033	0.053	0.041	0.020
8	23,601	0.465	0.047	0.055	0.028	0.062	0.074	0.000

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table O-9. Overall Kappa, Cut Score False Positive Rates, and Cut Score False Negative Rates for NM-MSSA Science Forms, as a Function of OP Form, Subject, and Grade

OP Form	Grade	Number of Students	Kappa	FP Cut 1	FP Cut 2	FP Cut 3	FN Cut 1	FN Cut 2	FN Cut 3
1	**5**	11,923	0.578	0.052	0.045	0.016	0.050	0.032	0.012
	8	12,046	0.540	0.056	0.043	0.004	0.063	0.033	0.001
	11	9,984	0.552	0.065	0.051	0.002	0.061	0.041	0.001
2	**5**	7,325	0.564	0.049	0.047	0.022	0.051	0.035	0.014
	8	8,521	0.578	0.038	0.048	0.005	0.053	0.038	0.001
	11	8,290	0.532	0.067	0.055	0.004	0.058	0.048	0.001
3	**5**	2,621	0.569	0.043	0.051	0.018	0.051	0.036	0.013
	8	3,019	0.569	0.041	0.050	0.004	0.049	0.043	0.001
	11	2,883	0.538	0.067	0.053	0.003	0.062	0.044	0.001

** Calculations based on those students attempting 5 or more items on the given NM-MSSA assessment. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

APPENDIX P
PROCESSING & REPORTING BUSINESS REQUIREMENTS

New Mexico Measures of Student Success and Achievement Summative Assessment in ELA/SLA and Mathematics (NM MSSA) and Assessment in Science Readiness (ASR)

New Mexico Public Education Department			
120350: NM MSSA Spring 2023 **130750: NM ASR Spring 2023**			
Version Number	**Date**	**Updated Content Description**	**Updated By / Name**
0.1	2/24/2023	Initial update to content	W. Bogle
0.11	3/10/2023	Updates from CD and PgM completed	W. Bogle
0.12	3/15/2023	Updates from internal review completed	W. Bogle
0.13	3/16/2023	Edits from PED incorporated	W. Bogle
0.14	3/23/2023	Addendum A added	W. Bogle
1.0	4/4/2023	Addendum B added	W. Bogle
1.1	5/10/2023	F was added as a valid value for Math and Science CR items	W. Bogle
2.0	5/12/23	Adjustment made to 4 pt Writing trait to be 3pt	W. Bogle

Glossary	
PM	Program Management
CBT	Computer Based Test
PBT	Paper Based Test
PED	Public Education Department
MC	Multiple Choice
SRB	Student Response Booklet
EL	English Learner
OE	Open Ended also called Open Response items
FT	Field Test

Approval			
Version	**Printed Name**	**Title**	**Date Approved**

Table of Contents

I. Overview

Testing for New Mexico assessments were done on paper and online. The iTester testing platform was used for online testing. The Spanish Language Arts (SLA) assessment is a transadapted version of the ELA test. Testing is done in grades 3-8 in ELA/SLA and Mathematics. Students in grades 5,8 and 11 take the Assessment for Science Readiness (ASR).

A. Points of Contact

Title	Name	Contact Email
Client Services Program Manager	Larry Ehret	Larry.Ehret@cognia.org
Client Services Program Manager	Sarah Owens	Sarah.Owens@cognia.org
Client Services Program Manager	Mara Allaire	Mara.Allaire@cognia.org

B. Changes from 21-22

1. The writing prompt is operational beginning in 22-23.
2. We are reporting **Lexile and Quantile** on the Student Report.
3. An ***item level grid*** is reported on the Student Report.
4. ***Ways to Support*** is removed from the Student Report.
5. The following demographics have been removed from the student results data file layout:
    a. Bilingual
    b. Plan504
6. There is a new 10-day preliminary reporting timeline.
7. Test Report Codes have been updated.
8. Writing Analysis and Language Conventions will be summarized in 22-23 as a reporting category on the Student Report.
9. Report sequence added to the item analysis data to correspond to the report sequence on the student report item level grid in #3 above.
10. Spanish items will be added to the Item Analysis report data file.
11. BIE students will be excluded from NM state aggregations.
12.

C. Assessments

1) All NM assessments were administered to students beginning March 27, 2023, and ending May 12, 2023.
2) Students were tested online (CBT) and on Paper (PBT).
3) Tests were administered in grades 03-08 for ELA/SLA and Mathematics. Science is administered in grades 5, 8 and 11.
4) A student should take either ELA or SLA, not both.

D. Reporting Deliverables

1) Prior to test administration
    - Outbound Roster
        - Produced prior to the beginning of the test administration.
        - Printed and shipped with the labels for the student answer documents for paper testers.
        - There is a roster for each subject.
        - The source of data on the roster is a list of students scheduled for interim MOY iMSSA testing, amended by state or districts pre-id with the rest of students for summative testing.
2) 10 Day Preliminary Reporting
    - Student Results data file and Item level file to PED
        - Full results provided after equating is complete.

    - Item level data file containing information on all items on the test.
    - Files follow NM2223StudentLevelDatafilelayout_Final.xlsx and NM2223ItemlevelDataFileLayout.xlsx
    - Files are posted to the ftp site.
    - The data files will exclude students attending a BIE school.
- eMetric will receive the student results file that PED will receive.
    - File follows the NM2223StudentLevelDatafilelayout_Final.xlsx
    - Contains all student and test level information needed for reporting in eMetric
    - The file contains students who tested at a BIE school.
    - The file is posted to the ftp site for eMetric to access.
- eMetric will receive from Cognia a summary file containing data summaries to aid in quality assurance of Data Interaction calculations. This data does not get loaded to DI.

3) Final Reporting
    - Student Results Labels
        - See Student Results Labels Specifications for more information.
        - Student Results Labels are only produced in English.
        - One label per student is produced.
    - Student Report
        - Contains the student performance on the test.
        - There is one report per student. The report includes all assessments tested at the student's grade.
        - The printed Student Report is produced in Spanish only if the student has HomeLanguage=" Spanish"
        - Cognia will provide Student Report PDFs to eMetric for the Download Hub. This will give schools access to download and print copies of the student report. Only Student Reports in English will be available to download.
        - *See Student Report Specifications for more information.*
    - Student Results data file and Item level file to PED
        - Student Results data file containing demographics, accommodations, overall and reporting category performance.
        - Item level data file containing information on all items on the test.
        - Files follow NM2223StudentLevelDatafilelayout_Final.xlsx and NM2223ItemlevelDataFileLayout.xlsx
        - Files are posted to the ftp site.
        - The files do not include BIE students.
    - eMetric will receive a student results file.
        - File follows the NM2223StudentLevelDatafilelayout_Final.xlsx
        - Contains student and test level information needed for reporting in eMetric (DI)
        - The file is posted to the ftp site for eMetric to access.
        - The file to eMetric contains the BIE students.
    - eMetric will receive from Cognia a summary file containing data summaries to aid in quality assurance of Data Interaction calculations. This data does not get loaded to DI.
    - eMetric will receive from Cognia a data file containing all data necessary to produce a summative Item Analysis Report in Data Interaction
        - Data included is defined in the ItemAnalysisReportFileLayout2223.xlsx

D. Delivery of Reports

- 1 copy of the Student Report is printed and shipped.
- 1 set of Student Results Labels is printed and shipped.
- Online reports are available to the school and district, in eMetric's Download Hub, where the student tested. Students who test at different schools are reported to the last school where they tested.

- Online reports are run by grade and school.
- Paper reports are shipped to the district associated with their tested school. The report is shipped to the district associated with the last school the student tested if the student tests at different schools.

II. Pre-Test Administration Processes

This section describes the data preparation for student records pre-test administration:

1) The Pre-ID data file is used to provide answer booklet labels for students in the Pre-ID data file.
   i) A total record count will be provided with the final label data to iCore Distribution.
   ii) Each student label has a unique Barcode associated with a Student ID
   iii) One student label is printed for each booklet being administered.
2) The Pre-ID data is used to produce the Outbound Rosters that accompany the answer booklet labels.

B. ELA Test Design

Each MSSA test is administered in 2 sessions. Session 1 is Reading and Session 2 is Writing & Language. The form contains core operational items, matrix field test items and a matrix operational writing prompt. The core operational items are seen by all students and count toward the student's overall ELA score. There are 3 operational Writing Prompts. There will be one writing prompt per form. The prompts differ by form. There is a Narrative writing prompt, Informational/Explanatory writing prompt and Opinion/Argument writing prompt. Beginning in 2023, the student's score on the writing prompt will count toward the student's overall ELA score.

ELAGrades 3-8 (Spring 2023)	Passage-Based Items				Total Items	Total Points	
	Passage Sets	MS-1	MS-2	WP		Min	Max
Core Operational Items	6	27	5	0	32	37	37
Matrix Operational Items	0	0	0	1	1	7	7
Matrix Field Test Items	2	10	2	0	12	14	14
Total Student Experience	**8**	**37**	**7**	**1**	**45**	**58**	**58**

ELA Item Types

Type	Description	Points
MS-1	Machine Scored-Multiple Choice or Multi-Select	1
MS-2	Machine Scored-Evidence based Selected Response (EBSR)	2
WP	Writing Prompt	7

EBSRs are 2-part items. Students can earn a score of 0, 1, or 2 on EBSR items.

C. Math Test Design

Each MSSA test is administered in 2 sessions. The form contains core operational items and matrix field test items. The core operational items are seen by all students and count toward the student's score.

Mathematics Grade 3,4,5	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	33	2	2	37	51	51
Matrix Operational	0	0	0	0	0	0
Matrix Field Test Items	5	1		6	8	11
Total Student Experience	**38**	**5**		**43**	**59**	**62**

Mathematics Grade 6,7	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	36	2	2	40	54	54
Matrix Operational	0	0	0	0	0	0
Matrix Field Test Items	5	1		6	8	11
Total Student Experience	**41**	**5**		**46**	**62**	**65**

Mathematics Grade 8	Discrete Items			Total Items	Total Points	
	MS-1	CR-3	CR-6		Min	Max
Core Operational Items	37	2	2	41	55	55
Matrix Operational Items	0	0	0	0	0	0
Matrix Field Test Items	5	1		6	8	11
Total Student Experience	**42**	**5**		**47**	**63**	**66**

Math Item Types

Type	Description	Points
MS-1	Machine Scored -Multiple Choice or Multi-Select	1
CR-3	Hand scored-Constructed Response-may be a single prompt or multi-part item	3
CR-6	Hand scored-Constructed Response-may be a single prompt or multi-part item	6

Partial credit allowed for multi-part items.

D. Science Test Design

Student Testing Experience							
Grade 5	Cluster/Passage Items*			Standalone Items*		Total Items	Total Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	8	3	35	64
Matrix Field Test Items	2	4	4	4	1	13	24
Total Student Experience	8	16	16	12	4	48	88
				Estimated Testing Time (min)			150

Student Testing Experience							
Grade 8	Cluster/Passage Items*			Standalone Items*		Total Items	Total Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	8	3	35	64
Matrix Field Test Items	2	4	4	4	1	13	24
Total Student Experience	8	16	16	12	4	48	88
				Estimated Testing Time (min)			150

Student Testing Experience							
Grade 11	Cluster/Passage Items*			Standalone Items*		Total Items	Total Points
	Stim/Psg	MS-1	MS-2	MS-2	OE		
Core Operational Items	6	12	12	10	3	37	68
Matrix Field Test Items	2	4	4	5	1	14	26
Total Student Experience	8	16	16	15	4	51	94
				Estimated Testing Time (min)			158

Science Item Types

Item Type	Definition	Valid Point Values	Scoring Method and Scoring Rules
MS-1	machine-scored item that may be multiple choice, multi-select, or TEI interaction	0,1	Machine scored; all or nothing scoring for the interaction, no partial credit
MS-2	machine-scored item with part a, part b; interactions may be any combination of multiple choice, multi-select, or TEI	0,1,2	Machine scored; part a and part b each worth one point; each part scored all or nothing (0,1); sum scoring for parts for total score of (0, 1, 2); each part scored independently
OE	hand-scored extended text interaction (traditional open response/constructed response item)	0,1,2,3,4	Hand scored; holistically scored; one rubric/one dimensional scoring

E. Scoring Data

Scoring division provides Reporting Data Processing with the open response scores for all tests.

1. Every score record will contain valid scores for all items.
   a) A validation of score values will be performed against the scoring specifications.
   b) If a score value is found to be invalid, resolution will be done by the Scoring Division
2. Each score record is associated with a Booklet ID or a Test ID
   a) If a score record is received without an associated Test or Booklet ID, resolution will be done with the Scoring Division
3. All unresolved scoring records will be included in a report to the Scoring Division, for research and resolution.
4. The following values will be received from Scoring for hand scored writing items:
   B=Blank with code number 21
   U=Unreadable with code number 51
   F=non-English with code number 53
   O=Off Topic with code number 54
   R=Refusal with code number 58
   I=Illegible with code number 55
   P=Repeats the prompt with code number 59
   A=Insufficient Amount to Score with code number 57
5. The following values will be received from Scoring for hand scored math and science items:
   B=Blank
   O=Off Topic
   U=Unreadable
   F=non-English
6. Score values of U are not valid for operational items.
7. Any item receiving a non-scorable conditions code will be given a score of 0 for analysis purposes.
8. Writing prompts are administered in all grades.
9. The writing prompt is scored on 2 dimensions:
   a) Production of Writing (PW)
   b) Use of Conventions (UC)
10. Valid scores for PW are 1-4
11. Valid scores for UC are 1-3
12. The total score for the writing prompt is the sum of the scores for PW and UC for a possible for total score range of 2-7.

F. Flawed Item

During the key verification process an item may be identified for various reasons to be “flawed”. If an item is identified as flawed it will be listed here. A decision may be made to not count the item in the student’s overall score for the subject. If it should not count in the student’s score, the item will be marked flawed, and the students will not be disadvantaged for their response on the item. An ‘X’ will be placed in the item attempt column for the item for all students.

G. Item Attemptedness

Item Type	Value to meet Attemptedness
MS-1	Non-blank response to the item,*=paper only for single select items
MS-2	Non-blank response to the item
CR	Numerical score given
WP	Non-B and Non-R condition code or Numerical score given.

H. Forms

1. There is a Spanish version of the Math test.
2. There is a Spanish version of the Math breach test.
3. The accommodated form is form 1 for each subject and grade.
4. SLA forms are the transadapted version of ELA form 1 English.
5. A breach form is available for each grade and subject. No breach forms were administered in 2023.
6. Science forms

Language	Grade	Number of forms	Additional accommodated form
English CBT	05	7 (1-6, Form AX) + TTS Form A1 {Form A1, A2, A3, AX; B4, B5, B6}	Yes (Text-to-Speech (TTS) & Form X)
	08	7 (1-6, Form AX) + TTS Form A1 {Form A1, A2, A3, AX; B4, B5, B6}	Yes (Text-to-Speech (TTS) & Form X)
	11	7 (1-6, Form AX) + TTS Form A1 {Form A1, A2, A3, AX; B4, B5, B6}	Yes (Text-to-Speech (TTS) & Form X)
English PBT	05	1 (Form AX)	Yes
	08	1 (Form AX)	Yes
	11	1 (Form AX)	Yes
Spanish CBT	05	2 Forms (Form A1, Text-To-Speech Form A1)	Yes
	08	2 Forms (Form A1, Text-To-Speech Form A1)	Yes
	11	2 Forms (Form A1, Text-To-Speech Form A1)	Yes
Spanish PBT	05	1 (Form AX)	Yes
	08	1 (Form AX)	Yes
	11	1 (Form AX)	Yes

Note: For spring 2023, there are 2 main Core Forms: Core Form A and Core Form B; For each core Form, there are A1, A2, A3, AX, B4, B5, B6 for all grades. The form AX is the accommodated form used for paper replacement, Large Print, Braille. Form A1 is used for both Spanish Form, and Spanish TTS plus English TTS

a. Spanish Form A1 CBT is the translated version of English Form A1 CBT for each grade.
b. English and Spanish accommodated computer-based test forms are Form 1 Text-To-Speech for each grade.
c. Spanish Form AX PBT is the translated version of English Form AX PBT for each grade. The Spanish form is seen as an accommodated form.
d. English and Spanish Paper accommodated forms (represented by Form AX which is the paper replacement of the online form) are provided for students with the following accommodations:

e. Large Print
f. Braille
g. Paper accommodated forms contain replacement items for the TEIs on the online Form A1 test form that becomes Form AX.
h. Spanish accommodated forms are translated version of the respective English accommodated forms per grade.
i. The online accommodated form has the TTS-accessibility feature and is CBT Form A1.

III. Post-Test Assessment Administration

The Test Assessment Administration window was defined and closed prior to processing and reporting for Student Assessment Reporting.

A. Student Data Processing

1. Student Names will have all periods, commas and apostrophes removed.
   a) Middle Name is the First Initial of the Middle Name or blank if not available.
   b) Special characters (any non-letter characters) shall be set to blank.
2. Records are suppressed from processing if all Name fields, Student ID, and Test Items are blank.

B. Scan Paper Delivery and Data Denotation

1. Each Paper test is scanned and delivered immediately to the Reporting Data Processing team. At the time of receipt, Data Processing will perform procedures to accurately identify inaccuracies in the data. The data will be formatted as specified in the Scan Delivery Layout Format.
2. All discrepancies with the Scan File will be resolved accordingly.
3. Any Student Response Booklet where VOID is bubbled and there is at least one item that is attempted shall be researched via Webdesk system. See Data Processing Specifications for the resolution of Void bubbles.
4. Paper tests are not included in 10-day reporting.

C. Discrepancy Processing

1. There is no discrepancy processing for 10-day reporting.
2. Duplicates may exist where there is more than one data record with the same Student ID, be the record online or paper.
3. Duplicate Test records with the same Student ID/Grade/Language will be combined or otherwise suppressed. See *Data Processing Specifications* for resolution of duplicate tests.
4. If there is a duplicate where the student takes one session in one test instance and another session in another test instance, the 2 sessions will be combined/merged to created one complete test.
   a) If the schools differ between session 1 and session 2, the school from where the last session taken will be used for reporting (if it can be determined by the session updated dates for online tests). This school is the transfer school.
   b) The record will be flagged in the data file as being a merged record.
   c) If an online session is merged with a paper session, test mode flag is set to “both”.
5. Duplicate Test Cross Language: a student has taken both a Spanish and an English form of the same test. If the forms have at least 1 item attempted, send a report to PED for research and resolution. PED will resolve by indicating which form/test will be invalidated.
6. Duplicate Cross Grade tests are identified as more than one test taken with two different grades from the same student.

a) Should the Student have no work in the off-grade test, or the test is void and there is work in the matching grade test, suppress the off-grade test.
b) If both tests have responses, send a report to Program Management for research and resolution.

7. Braille Validation-Paper tests only
   a) Send PM a report for confirmation of booklets with Student with the Braille Accommodation bubbled for any subject.
   b) Should PM determine student is not Braille, clear the Braille Accommodation bubbled.
   c) Program Management will provide a list of any items that could not be Brailled. There are no such items in 2023.
8. If a test has sessions split between paper and online, the sessions will be merged to create one test.
9. SpecialEd and Plan504 cannot both be marked for test record.

D. Scaling and Equating

1. After equating is approved, Psychometrics applies a scale score and achievement level for each test receiving a valid score.
2. The scale score range begins with the grade as follows:
3. Grade 3 - 300-390
4. Grade 4 - 400-490
5. Grade 5 - 500-590
6. Grade 6 - 600-690
7. Grade 7 - 700-790
8. Grade 8 - 800-890
9. Grade 11 - 1100-1190

E. Score calculations

1. Hand scored items scored on multiple dimensions will have the dimension scores summed for the final reported score for the item.
2. Only Core Operational items are included in a student's overall raw score.
3. The overall raw score is used to determine the student's scaled score.
4. The scaled score determines the achievement level the student has attained.
5. Flawed items will not count toward a student's overall score.

F. Reporting Categories

1. The PassageType column in NTS provides the Reporting Categories for the Reading items.

Grade(s)	Subject	Reporting Category	Abbreviations	Reporting Order
All	ELA	Text Type-Literary Text	LT	1
		Text Type-Informational Text	IT	2
		Reading Strategy-Comprehension	CM	3
		Reading Strategy-Analysis & Interpretation	AI	4
		Writing Analysis and Language Conventions	WL	5
3-5	Math	Operations & Algebraic Thinking	OA	1
		Number & Operations in Base Ten/Number & Operations-Fractions	NO	2
		Measurement & Data/Geometry	MG	3
		Problem Solving/Reasoning & Argument	PR	4
		Modeling/Structure & Repeated Reasoning	MS	5
6-7		Ratios & Proportional Relationships	RP	1
		The Number System/Expressions & Equations	NS	2
		Geometry/Statistics & Probability	GS	3
		Problem Solving/Reasoning & Argument	PR	4
		Modeling/Structure & Repeated Reasoning	MS	5
8		Functions	FN	1
		The Number System/Expressions & Equations	NS	2
		Geometry/Statistics & Probability	GS	3
		Problem Solving/Reasoning & Argument	PR	4
		Modeling/Structure & Repeated Reasoning	MS	5
5,8,11	Science	Physical Sciences	PS	1
		Life Sciences	LS	2
		Earth and Space Sciences	ES	3

2. Subdomain indicators provided by Psychometrics are reported for the reporting categories. Values:1=Below Standard, 2=At/Near Standard, 3=Above Standard
3. A Reading scale score, provided by Psychometrics, is reported on the overall ELA scale.
4. A Writing & Language scale score, provided by Psychometrics, is reported on the overall ELA scale.

G. Test Attemptedness

Attemptedness is based on attempts to operational items only. See item attempt rules above.

1) If a session is voided, any items attempted in that session will be blanked out and will not count toward test attemptedness.
2) Students with Parental Refusal will have their items blanked out and no item will count toward test attemptedness.
3) Blanking of items as referenced in #1 and #2 above is done prior to determining test attemptedness. Therefore, no students with Parental Refusal marked will meet test attemptedness. Students with all sessions voided will also not meet test attemptedness
4) Only field test items can have a null score meaning that the item was not scored.
5) Field test items do not count toward attemptedness
6) The writing prompt is treated as one item when determining test attemptedness
7) A student is classified into 2 possible attempt groups of Attempt Status:
   a) Attempt Status 0 is assigned to the test if the student did not provide a valid response to at least 5 operational items.
   b) Attempt Status 1 is assigned to the test if the student provided a valid response to at least 5 operational items on the test.

H. Not Tested Reasons

The following not tested reasons are applicable:

Subject	Code	Not Tested Reason
ELA/SLA	01	Withdrew Before Test Completion
	03	Language Exempt for Reading Only
	04	PED-Approved Medical Exemption
	05	Parental Refusal/Non-compliance
	07	PED-Approved Test Invalidation
	08	Absent During Window/Chronically Absent
	10	Other
Math and Science	01	Withdrew Before Test Completion
	04	PED-Approved Medical Exemption
	05	Parental Refusal/Non-compliance
	07	PED-Approved Test Invalidation
	08	Absent During Window/Chronically Absent
	10	Other

1. Not tested reasons can be marked by the test administrator in iTester during the testing window or marked on the scannable for paper testers.
2. Currently a not tested reason can be assigned by an LEA for a student only at a subject-level, not a session-level.
3. The following hierarchy is applied if more than one not tested reason is marked. Priority is listed from highest to lowest.

ELA/SLA:

Invalidated test
Void test
Language Exempt
Medical Exemption
Parental Refusal
Absent
Withdrew
Other

Math and Science:

Invalidated test
Void test
Medical Exemption
Parental Refusal
Absent
Withdrew
Other

4. Not tested reasons are applied if a test does not meet the test attemptedness rule above. If the test meets attemptedness the not tested reason is ignored and the test receives a score. See section G above to see how Parental Refusal and all sessions Voided are dealt with.
5. If a test does not meet attemptedness and no Not Tested Reason is marked, the student is classified as "Did Not Reach Minimum Attempt".
6. Not tested reasons are applied per subject.
7. Void Tests
   - Paper tests can be voided by filling in the void bubble.
   - Online tests can be voided by test administrator in the testing platform.
   - Voids online are applied by session. While Voids for paper test are applied to a whole subject.
   - Void sessions will have any attempted items blanked out.
   - Test report code will be set to 99 for PED approved void tests.

1) Test Invalidations:

PED's decision to invalidate a test or session is determined by the testing irregularity that is reported by the LEA. The invalidations are classified as impactful or non-impactful. Impactful irregularities will be invalidated.

- PED may choose to submit Invalidations during data discrepancy period.
- Due to testing irregularities such as cheating, a test can be invalidated.
- Before a test can be invalidated by the district, the invalidation must be approved by the state.

I. Participation Status

Based on the above rules a student is assigned a participation status for each subject.

1) Participation status is determined using both the test "Attempt Status" value and the "Not Tested Reason."
   a) If Attempt Status is 0 (the test has no operational items with valid attempts or less than 5 items with valid attempts), and
      the test has a "Not Tested Reason", then the Not Tested Reason is reported, otherwise, the test is reported as "Did Not Reach Minimum Attempt".
   b) If Attempt Status is 1 (the test has at least 5 operational items with valid attempts),
      i) The student is classified as Tested and will receive a scaled score for a test based on non-blanked items as described in section G above and in the table below.

ii) If the student has a Not Tested Reason, the Not Tested Reason is ignored. Exception is parental refusal.

2) Regardless of the test attempt status, if a student is on the test invalidation list from PED, their test will be marked as Invalidated.
3) Only "Tested" students, that is, students who meet attemptedness will be included in analyses.
4) The following table summarizes participation.

Participation Status	Code	Assigned a Scaled Score and Achievement Level	Included in the State Results Data File	Reported in DI*	Included in Aggregations
Tested	Z	Yes	Yes	Yes	Yes
Absent	J	No	Yes	Yes	No
Medical Exemption	F	No	Yes	Yes	No
Parent Refusal	G	No	Yes	Yes	No
Withdrew	C	No	Yes	Yes	No
Other	P	No	Yes	Yes	No
Language Exempt	E	No	Yes	Yes	No
Test Invalidation	L	No	Yes	Yes	No
Did Not Reach Minimum Attempt	B	No	Yes	Yes	No
Void test	K	No	Yes	Yes	No

**In DI, if a student meets attemptedness and has a test report code, both the score and the test report code will be displayed. If the student does not meet attemptedness the participation status will be displayed.*

IV. Aggregations

1) All aggregations for Spanish tests will only include Spanish tests and all aggregations for English tests will only include English tests.
2) School, district, and state scale score averages rounded to the nearest whole number.
3) Aggregations are reported only if the entity has at least 10 included students. Only students with participation status='Z' are included in aggregations.
4) In the item analysis report, school, district, and state mean scores are rounded to 2 decimal places.
5) BIE schools are excluded from NM state results.
6) See list in Addenda section of BIE schools.
7) BIE schools are only aggregated with other BIE schools.

V. Rules pertaining to Calculating Classical Stats

A. Ethnicity:

Race and Ethnicity will stay the same as entered by state or overwritten by district during pre-id window. To perform DIF stats, the following process will be followed to collapse the Hispanic and Race fields into one variable:

- If a student has a Yes for Hispanic, the Ethnic value for the student will be H=Hispanic
- Otherwise, the Ethnic value will be equal to the Race value.
  B=Black
  P= Native Hawaiian/Other Pacific Islander
  A=Asian
  I=American Indian/Alaskan Native
  C=Caucasian/White
  M=Multirace
- For the purposes of DIF stats the Asian and Native Hawaiian/Other Pacific Islander categories are combined

B. DIF Stats Definition table:

DifVariable	DemoVariable	RefValue	FocValue	RefText	FocText
MF	Gender	M	F	Male	Female
WB	Ethnic	C	B	White	Black
WH	Ethnic	C	H	White	Hispanic
1	Ethnic	C	I	White	Native American
2	Ethnic	C	A	White	Asian/Native Hawaiian/Other Pacific Islander
6	Ethnic	C	M	White	Multirace
3	SpecialEd	N	Y	Non-Sped	Sped
4	EconDis	N	Y	Non EconDis	EconDis
5	EL	*0,2,3,4,5,6	1	Non-EL	EL

**EL values to be combined to create the non-EL Reference group*

IV. Data Deliverables Specific Rules

A. Student Results data file delivered to the PED and eMetric

a. PED receives a student level results data file and an item level results data file.
b. eMetric receives a student level results data file.
c. The data file contains tests with a Tested status and tests with a Not Tested reason.
d. Naming convention of the data files: NM2223StudentResultsPED.csv and NM2223StudentResultsAll.csv and NM2223ItemLevelResults.csv
e. NM2223StudentResultsPED.csv will not include students who test at a BIE school. This file will be delivered to the PED via the sftp site.
f. NM2223StudentResultsAll.csv will include all students including the BIE students. This file will be delivered to eMetric via the sftp site.
g. If a student's test was merged to create one test, then the *mergedtest* flag is set to 1, otherwise it is set to 0.
h. If the *mergedtest* flag is set to 1 and the student tested at 2 different locations, the last school (where the last session was attempted) is reported as *Discode, Schcode*. The first school (different from the last) is reported as *TransferDiscode, TransferSchcode*.
i. If the *mergedtest* flag is not 1 then *TransferDiscode* and *TransferSchcode* are blank.
j. The files are stacked by subject so that a student appears as many times as they have tests in the student results file.
k. Students with Homeschool flag set will be reported back to the district where they tested.
l. Each file contains all grades.
m. NumAttempted is the number of operational items in the test that met the item attemptedness rules described above. NumAttempted does not include Field Test items.
n. Students with a not tested reason and meet attemptedness will be reported in the file with their assigned scaled score and achievement level. In DI, they will be reported with both their score and their not tested reason.
o. All items are included in the item level data file.

B. Item Analysis Report data to eMetric

a. The file contains all data needed for eMetric to produce the Item Analysis report in DI.
b. The file follows the layout NM2223ItemAnalysis.xlsx
c. The file is posted to the sftp site for eMetric to access.
d. The file is produced as a csv file.
e. The file contains the relevant data for all subjects: ELA, Math, and Science.

f. For the calculation of the item mean scores, if the number of included items is less than 50 the mean score is suppressed and not reported.
g. Psychometrics provides the values for Difficulty Order and the Achievement Level for the data file.
h. Reportsequence is added to the file. This column corresponds to the report sequence on the student report item level grid. Items are grouped by reporting category and standard.
i. English and Spanish items will be included in the data file.

V. Report Specific Rules

A. Student Report

1. A student receives a student report if at least one subject has partstatus='Z'. That is, the student is classified as Tested for at least one subject.
2. Student Reports are sent back to where the student tested last across all subjects.
3. The report combines the results for all subjects, ELA/SLA, Math, and Science.
4. Students with HomeLanguage=" Spanish" will receive their student report in Spanish.
5. In 2023, lexile and quantile scores are reported on the front page of the student report. A lookup table is provided by Psychometrics that provides the correspondence of the student's scale score on ELA and Math to the lexile and quantile score respectively.
6. For each subject reported an item level grid is added to the subject page. The grid provides all operational items. It is ordered by the report sequence provided by Cognia's Content Development team. The items are grouped by reporting category and standard. The grid reports the reporting categories, content domains, and practices where available. The writing prompt dimensions are reported in separated labeled columns.
7. The points earned on each item is formatted as points earned/total possible points.
8. Student reports for students at a BIE school will have school and district aggregations reported but state aggregations will be blank.
9. If a student is receiving a report and has 1 or 2 subjects with a not tested reason, the not tested subject(s) is reported in the following manner:
   i. On the front page "Student did not test in this area" appears under the subject title
   ii. The rest of the subject section on the front page is left blank.
   iii. On the inside pages (ELA or Math) or the back (Science) if applicable, the reporting category names are printed. The rest of the table is left blank.
   iv. The comparison school, district and state bars are printed unless the bars are suppressed due to N-size suppression rules.
   v. There is no student score vertical bar printed.
   vi. In the item level grid, all rows except for the points earned are populated. The points earned row is left blank for all items.
10. See *Student Report Specifications* document for further details on the student report.

B. Student Results Labels

a. A Student receives a student results label if at least one subject has partstsatus='Z'. That is, the student is classified as Tested for at least one subject.
b. Student Labels are sent back to where the student tested last across all subjects.
c. The label combines the results for all subjects: ELA/SLA, Math, and Science according to the student's tested grade.
d. See *Student Results Label Specifications* document for further details on the student labels.

VI. Cognia Use Only

A. Reporting Products

Contract Code	Description	Report Type	Report For	Grade(s)	Report Subtype	Content Code	Qty
120350	Student Labels	07	1	03-08,11	03	00	1
120350	Student Report-Parent copy	07	1	03-08,11	02	00	1

B. Details for Item Analysis Report data file

1. Values of 2 and 7 in Process field in NTS indicate Math Practices
2. Both English and Spanish items will be aggregated.
3. Report Sequence is derived from an algorithm provided by CD. The algorithm uses the reporting category and standard for the item to group and order items.

C. eMetric Metadata file for Student Report PDFs

1. The column headings for the file are: ProgramName,ReportName,Year,Grade, Org_Num, PDF_name
2. The file is a csv file.
3. The naming convention for the file is NM2223_PDFmetadata.csv
4. The file is posted to the ftp site for eMetric to access.
5. Org_Num=Districtcode-Schoolcode
6. Year=2023
7. ProgramName=MSSA and ASR
8. ReportName=Individual Student Report
9. Web file naming convention: NM2223StudentReport_Gr[GG]_<districtcode||schoolcode>.pdf

VII. Appendix

VIII. Addenda

A. (3/24/23) PED's contract with MetaMetrics does not allow for Lexile and Quantile scores to be applied to non-English tests. Therefore, Lexile and Quantile scores are only reported for students who took their test in English.

B. (4/4/23) In 2023, the following steps will be taken to determine the students who will receive a Spanish printed report.
   i. For each student in the final reporting dataset (data after applying all data processing rules), look across all tests submitted by the student.
   ii. If the student submitted at least one Spanish form (regardless of attemptedness on the test), the student receives a printed Spanish report.

C. (4/25/23) During testing an item was found to have an error in the key resulting in no correct answer to the item. The item was in Grade 3 Math on the paper form. The Asset ID for the affected item is 619084. The solution is that for all students who took the form, the item will be treated as a flawed item. An "X" will be placed in the item attempt column of their test. The students will be on a separate scaleform

D. List of BIE schools taking the NM tests. These schools should have rules applied as described in the aggregations and Deliverables sections:

District Code	District Name	School Code	School Name
615	Laguna Middle DOE		
615	Laguna Middle DOE	615	Laguna Middle School
614	Laguna Elementary DOE		
614	Laguna Elementary DOE	614	Laguna Elementary School
623	Ohkay Owingeh Community School		
623	Ohkay Owingeh Community School	694	Ohkay Owingeh Community School
320	Santa Fe Indian School		
320	Santa Fe Indian School	625	Santa Fe Indian School
631	Haak'u Community Academy		
631	Haak'u Community Academy	031	Haak'u Community Academy
692	Navajo Preparatory School		
692	Navajo Preparatory School	020	Navajo Preparatory School

E. (5/12/23) Due to the lack of students at the highest score point in the Production of Writing (PW) writing trait, the scores will be adjusted so that students who receive a score point of 4 will have their score adjusted to 3. 3 will be treated as the maximum score. The maximum sum of the writing prompt traits will be 6 rather than 7.

APPENDIX Q
CUMULATIVE SCALED-SCORE DISTRIBUTIONS

Figure Q-1. Cumulative Scaled-Score Distribution for ELA—Grade 3

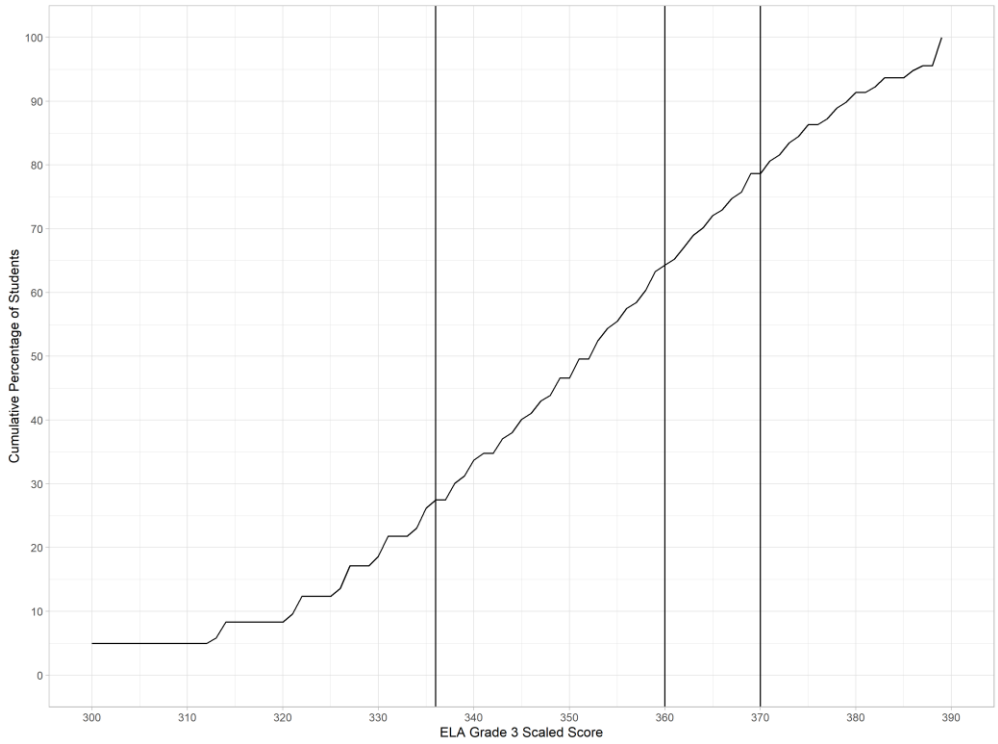


Figure Q-2. Cumulative Scaled-Score Distribution for ELA—Grade 4

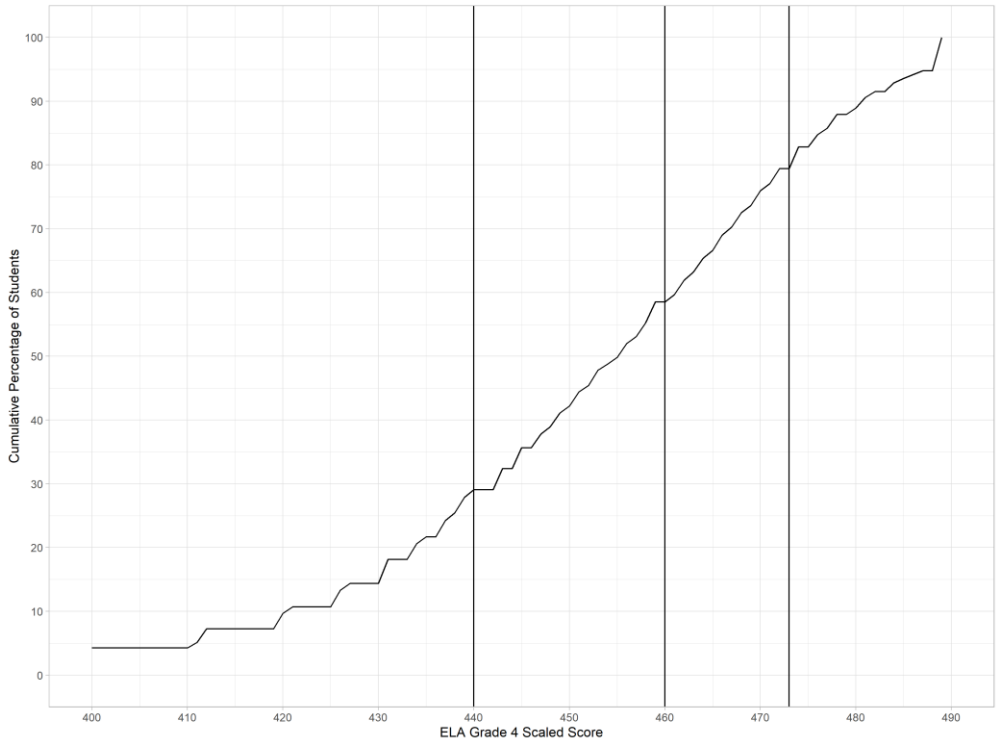


Figure Q-3. Cumulative Scaled-Score Distribution for ELA—Grade 5

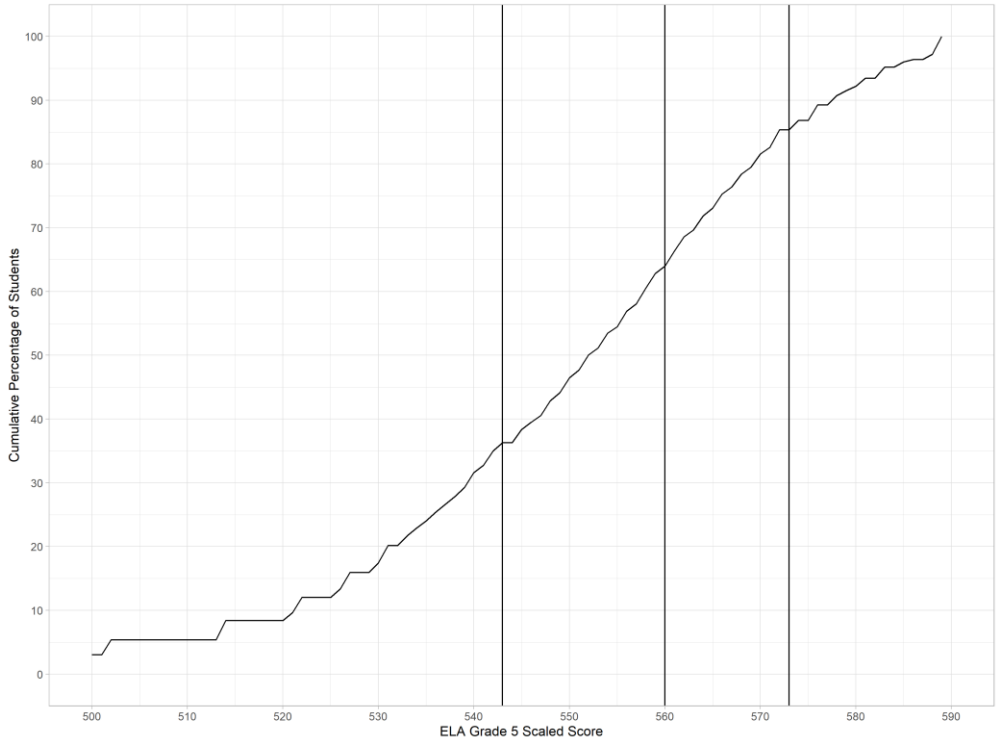


Figure Q-4. Cumulative Scaled-Score Distribution for ELA—Grade 6

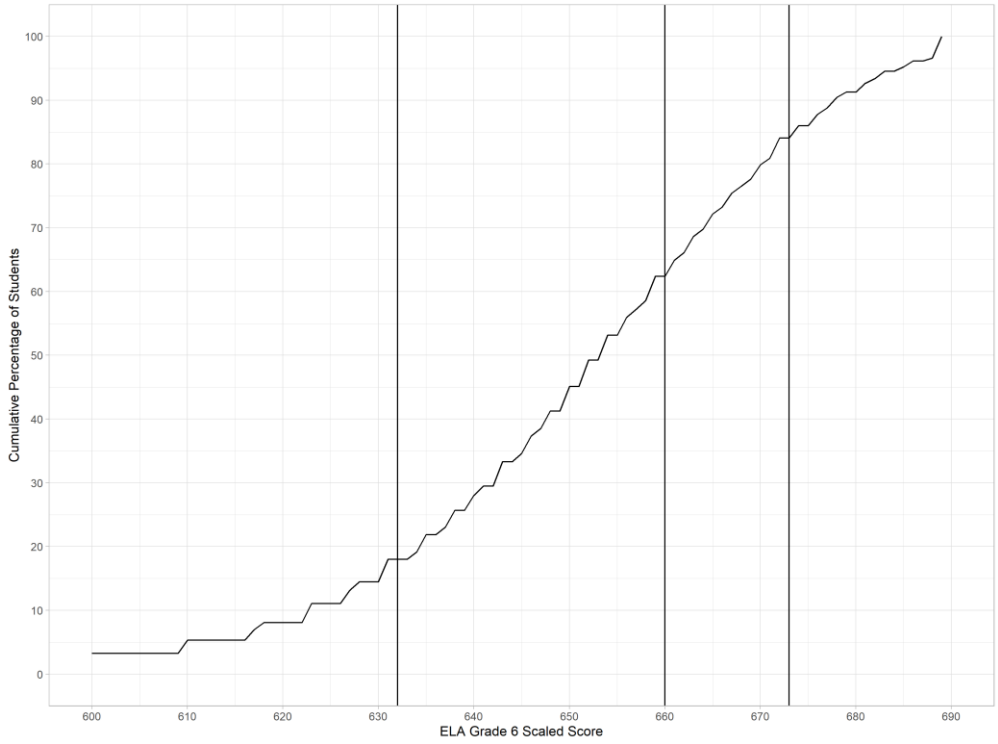


Figure Q-5. Cumulative Scaled-Score Distribution for ELA—Grade 7

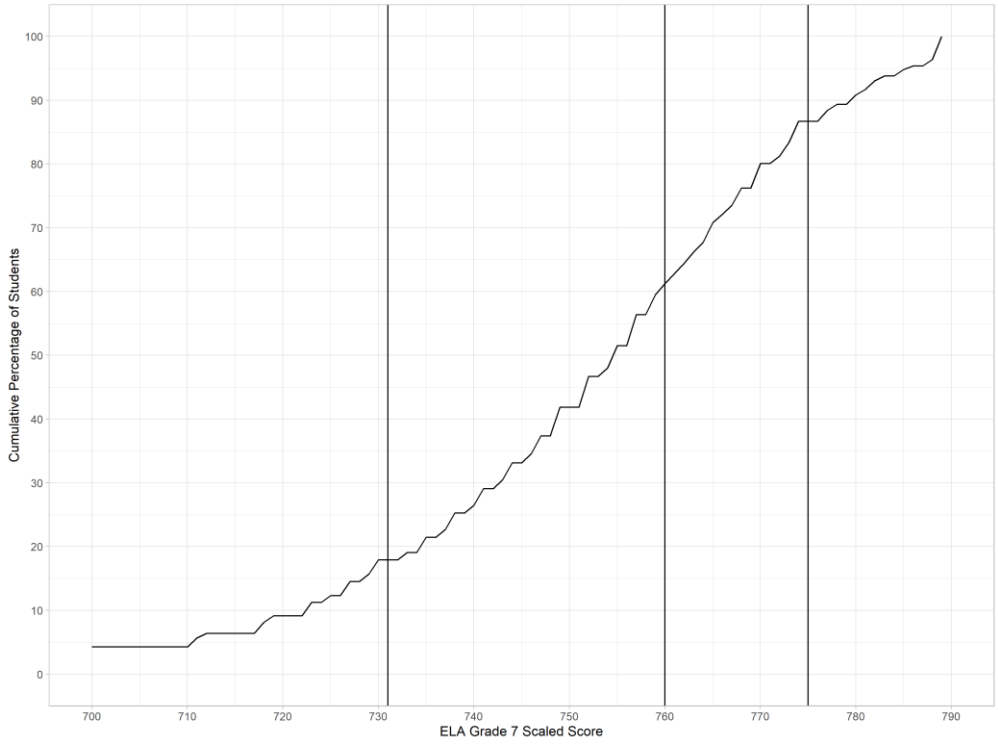


Figure Q-6. Cumulative Scaled-Score Distribution for ELA—Grade 8

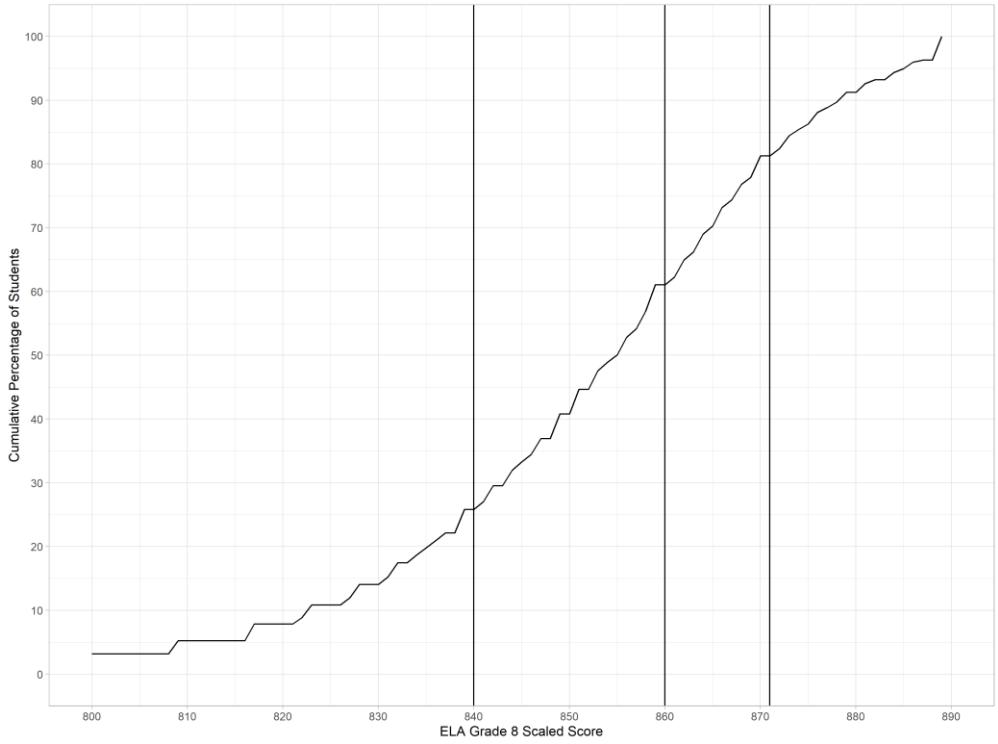


Figure Q-7. Cumulative Scaled-Score Distribution for Mathematics—Grade 3

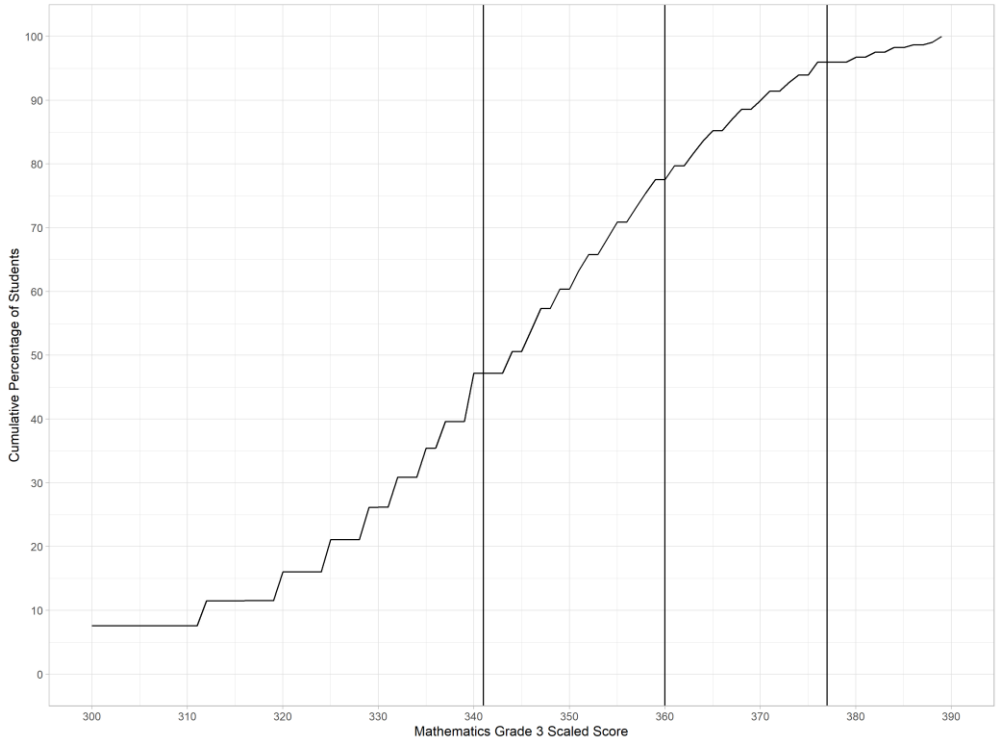


Figure Q-8. Cumulative Scaled-Score Distribution for Mathematics—Grade 4

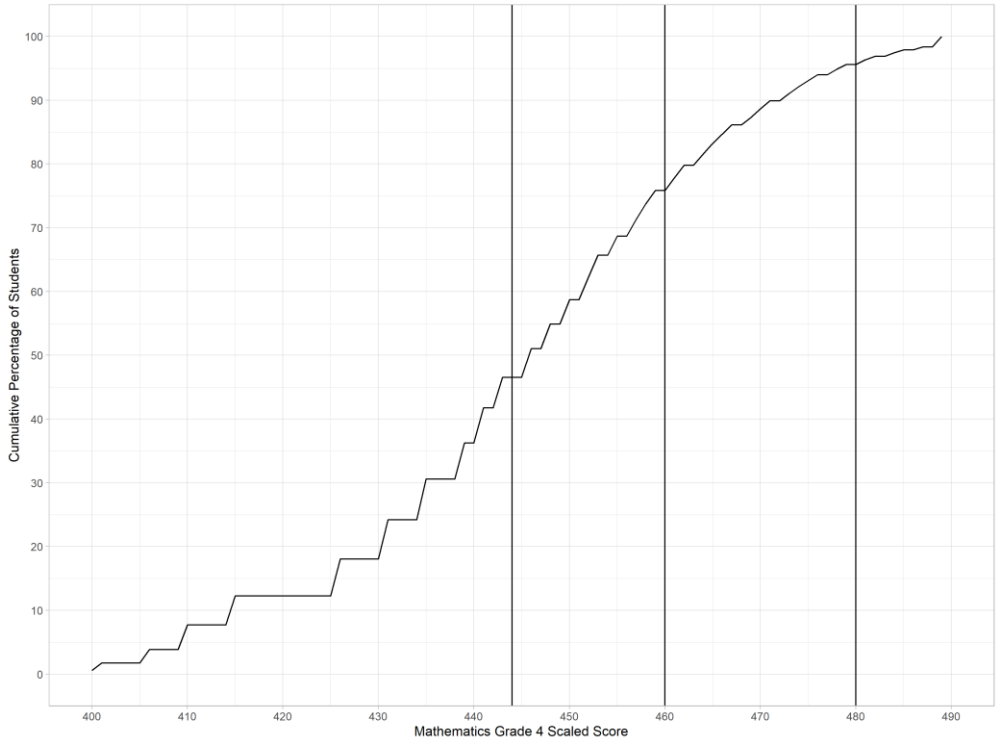


Figure Q-9. Cumulative Scaled-Score Distribution for Mathematics—Grade 5

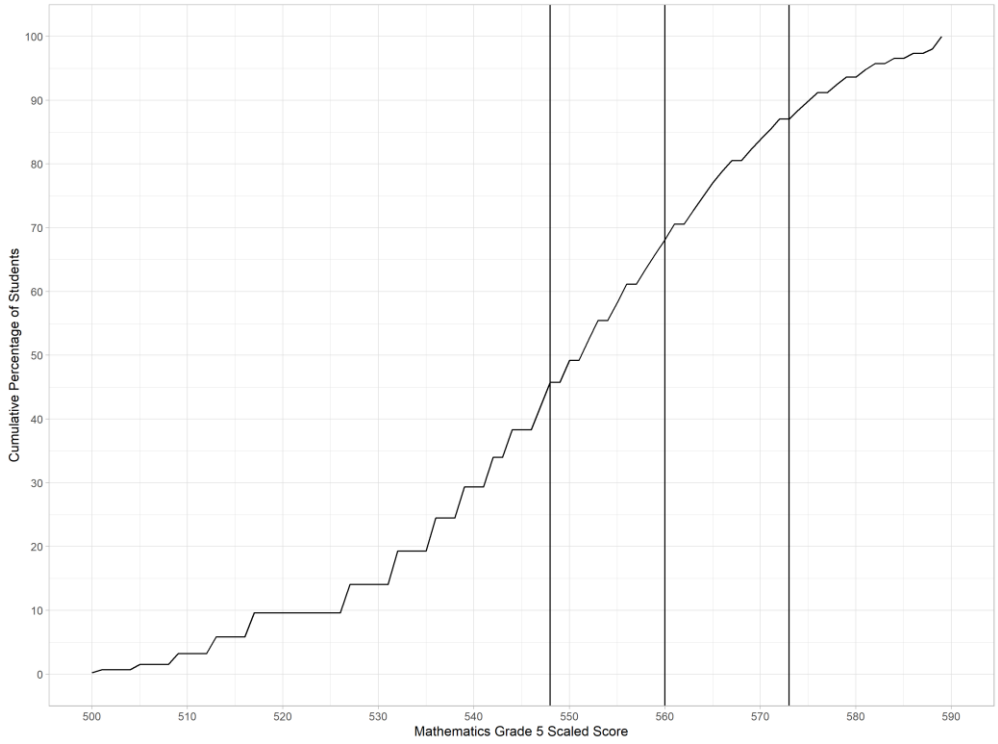


Figure Q-10. Cumulative Scaled-Score Distribution for Mathematics—Grade 6

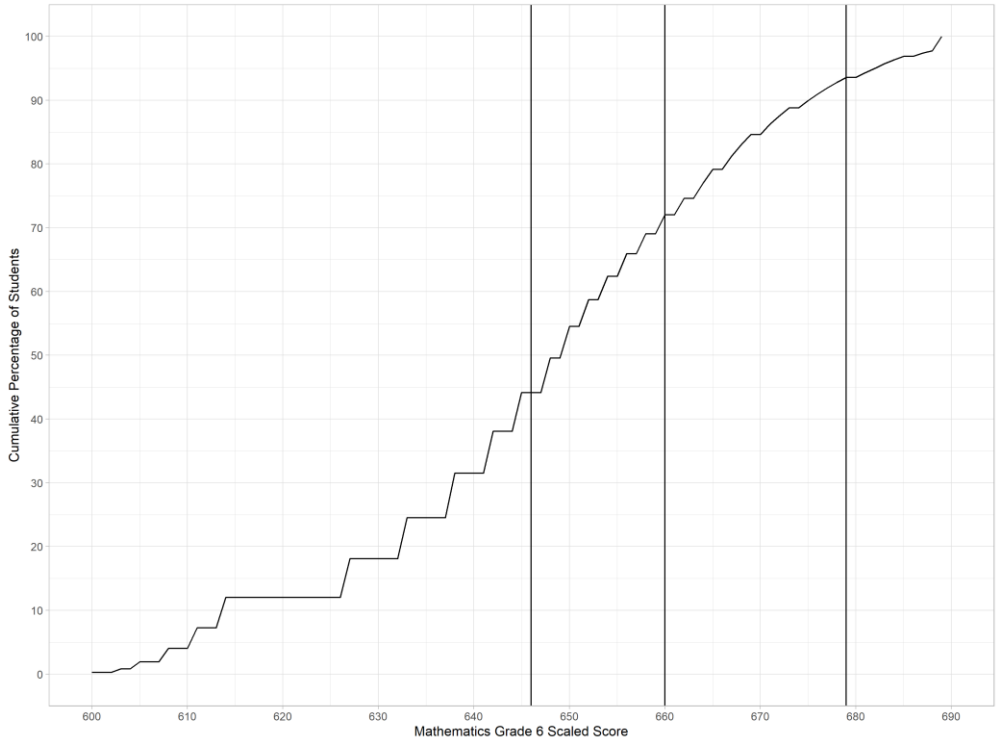


Figure Q-11. Cumulative Scaled-Score Distribution for Mathematics—Grade 7

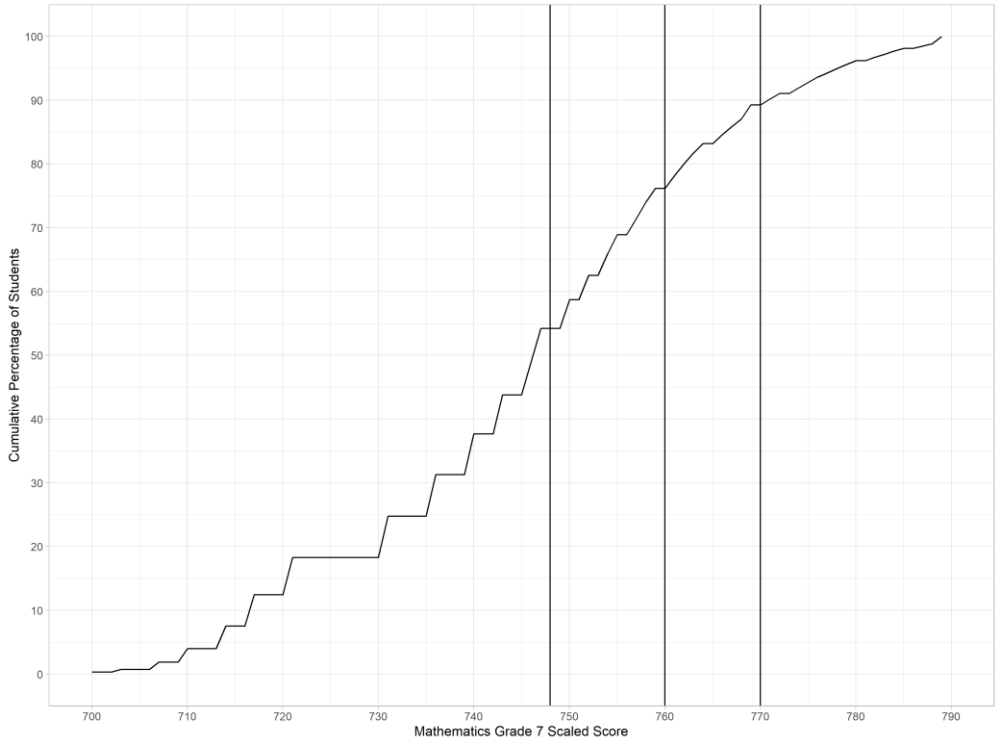


Figure Q-12. Cumulative Scaled-Score Distribution for Mathematics—Grade 8

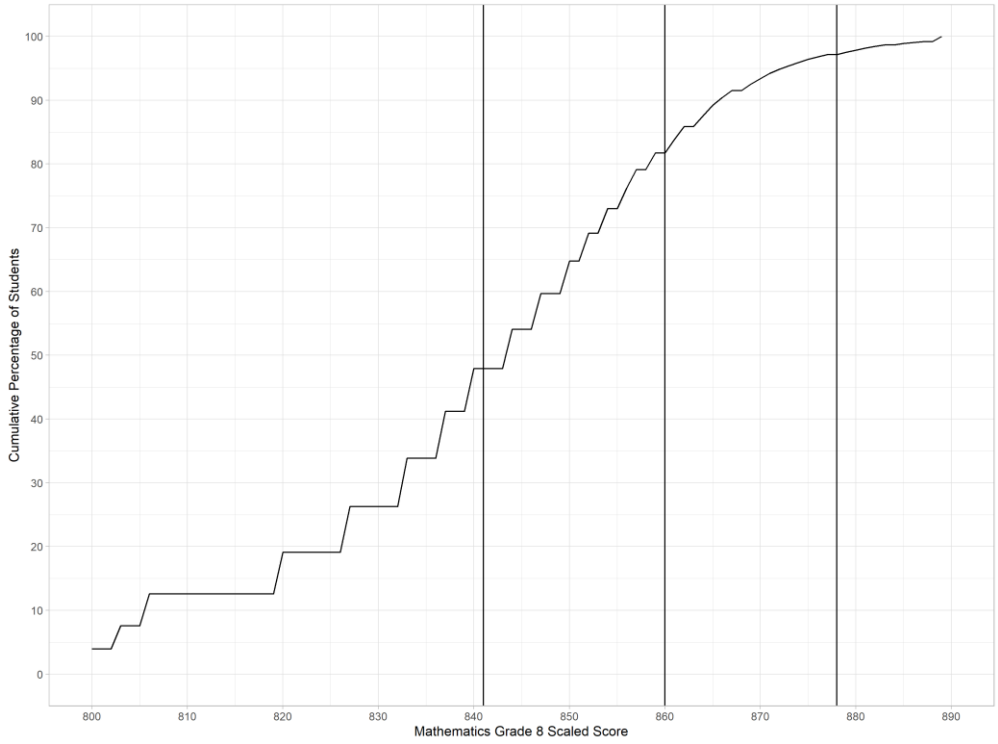


Figure Q-13. Cumulative Scaled-Score Distribution for Science—Grade 5

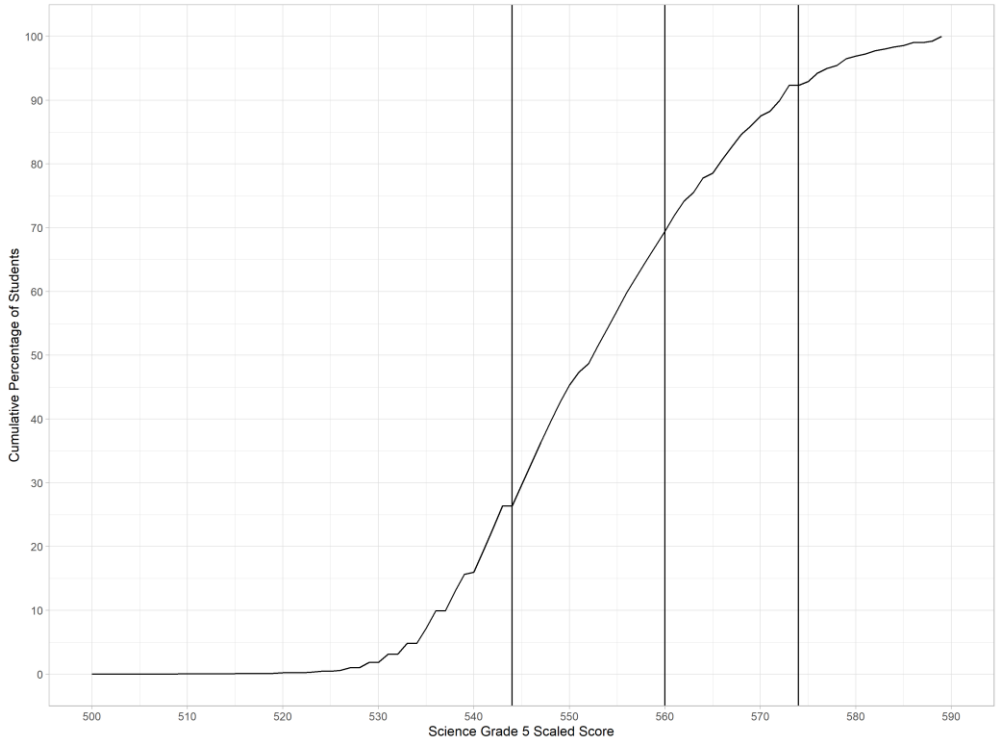


Figure Q-14. Cumulative Scaled-Score Distribution for Science—Grade 8

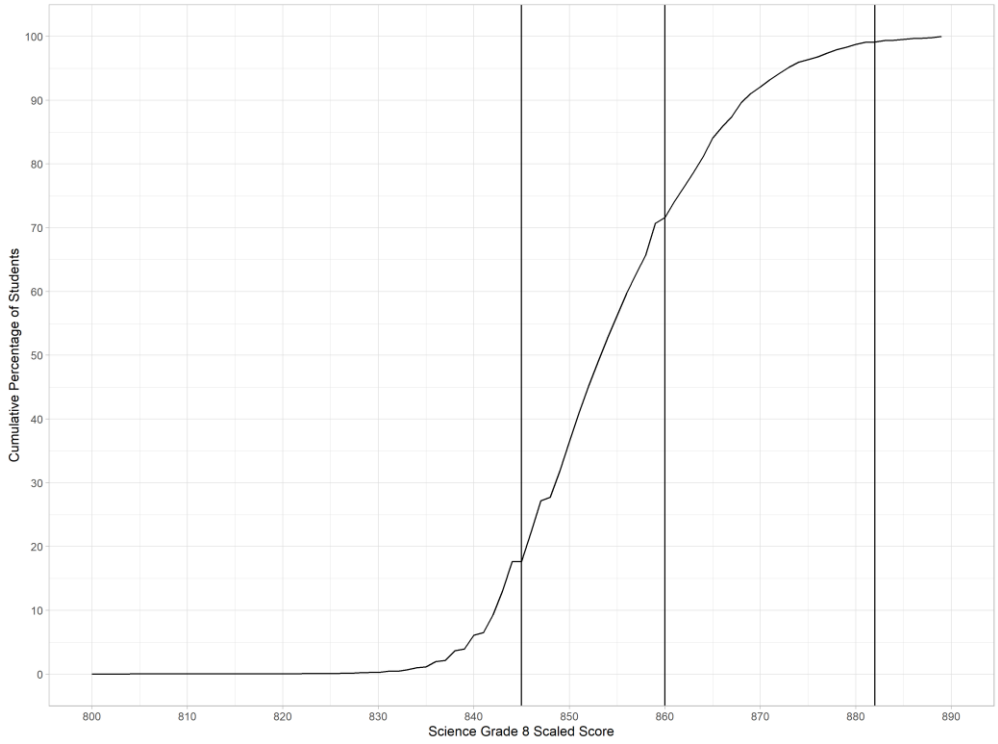


Figure Q-15. Cumulative Scaled-Score Distribution for Science—Grade 11

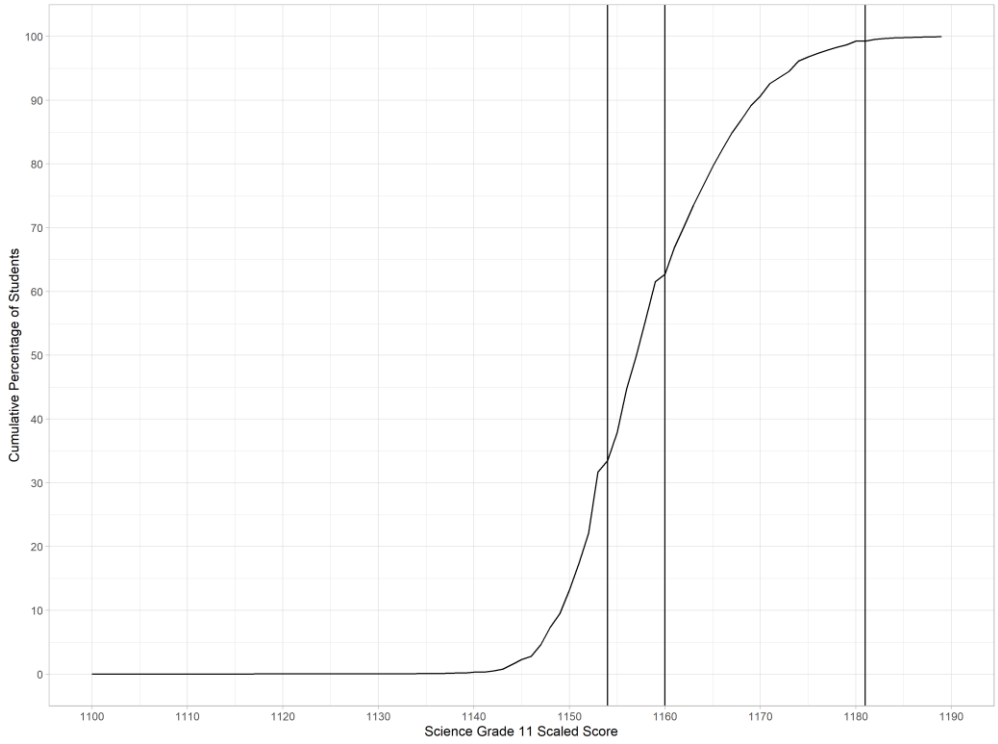


Figure Q-16. Cumulative Scaled-Score Distribution for SLA—Grade 3

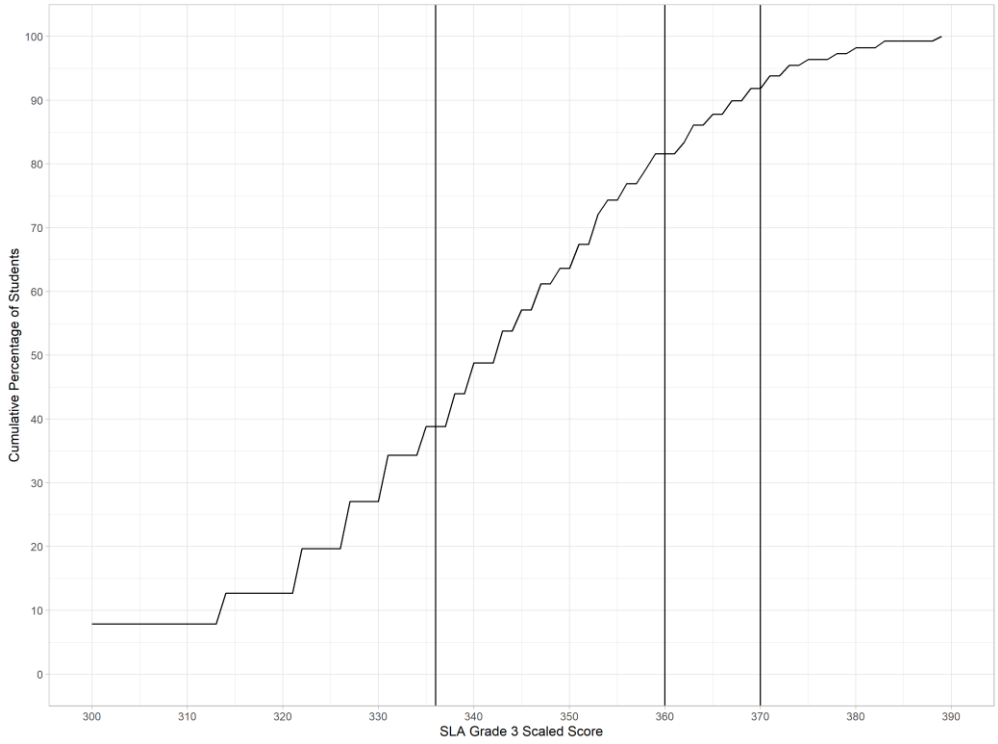


Figure Q-17. Cumulative Scaled-Score Distribution for SLA—Grade 4

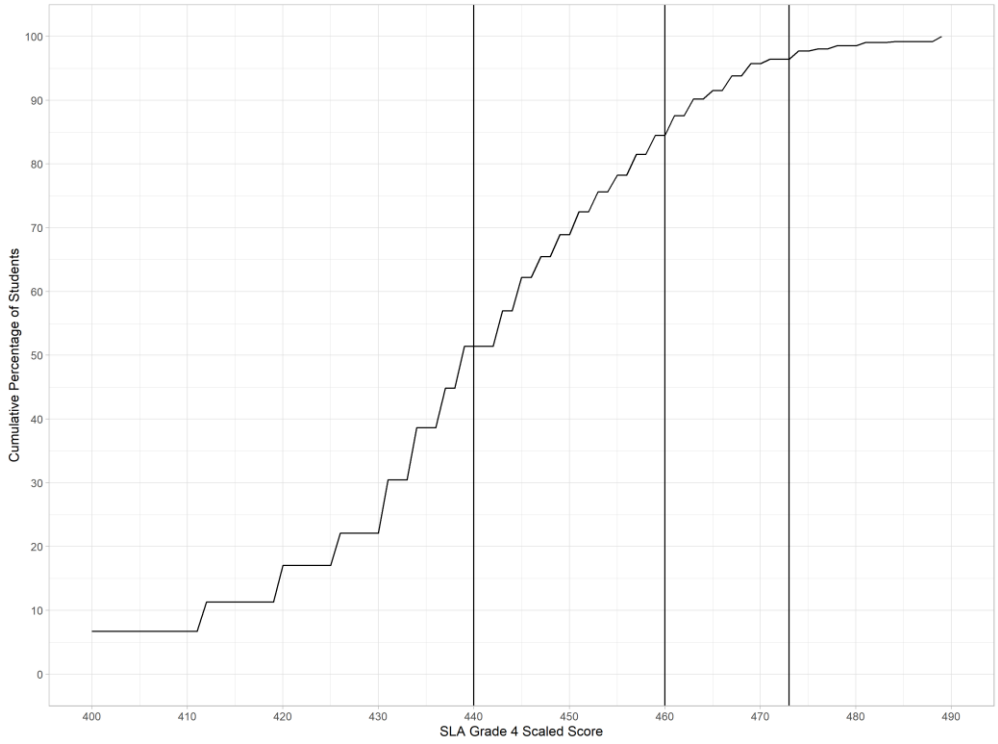


Figure Q-18. Cumulative Scaled-Score Distribution for SLA—Grade 5

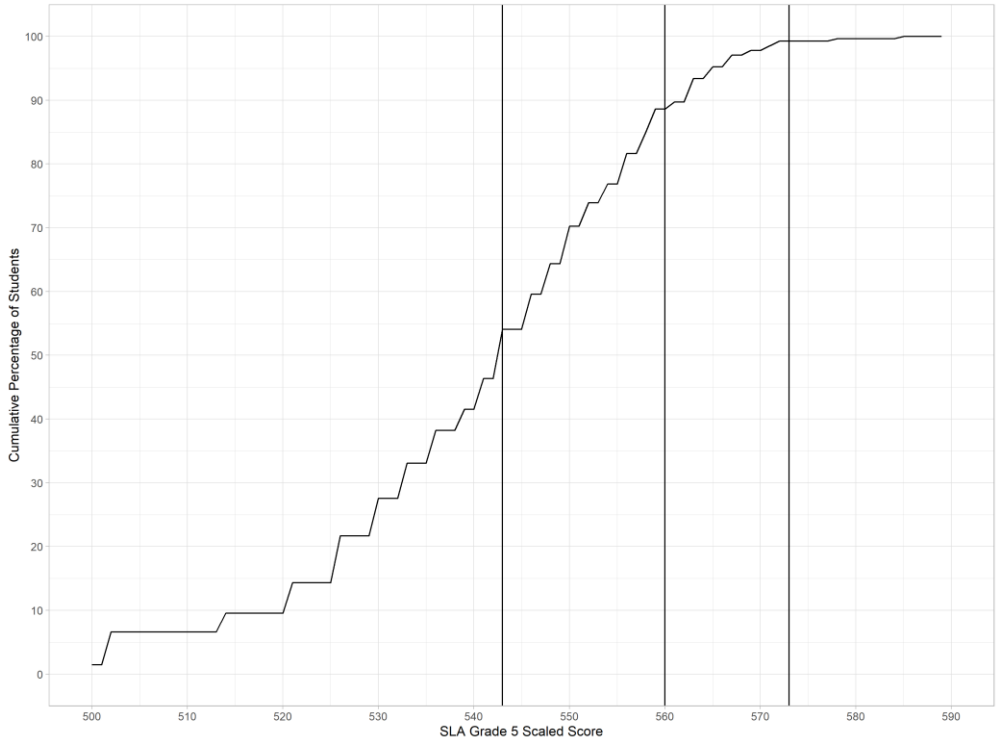


Figure Q-19. Cumulative Scaled-Score Distribution for SLA—Grade 6

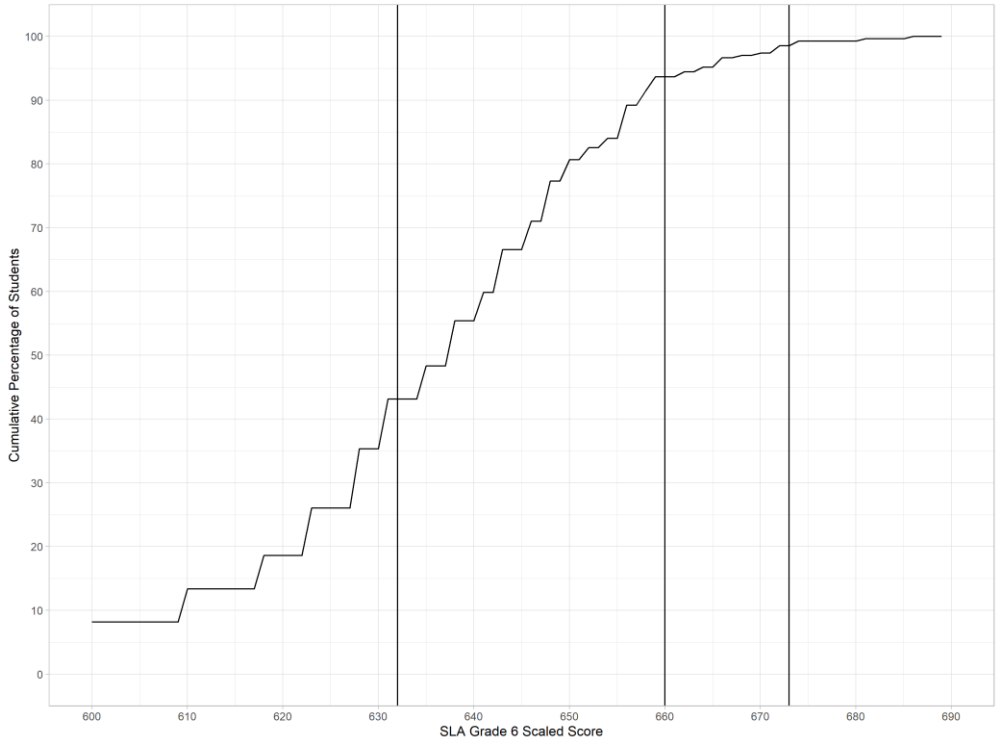


Figure Q-20. Cumulative Scaled-Score Distribution for SLA—Grade 7

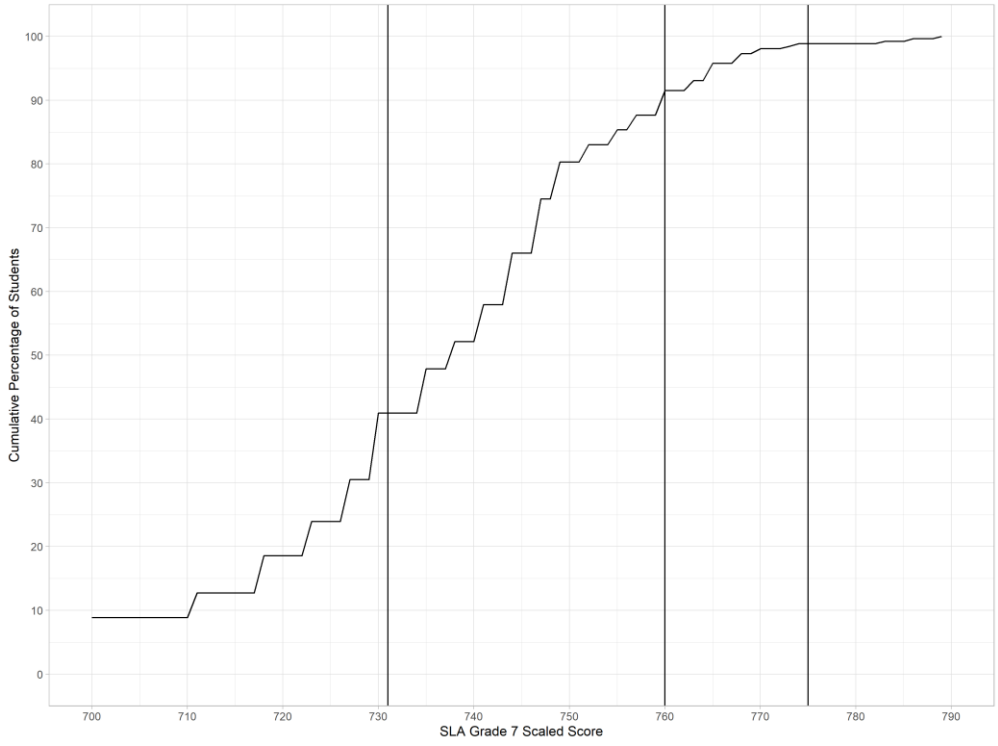


Figure Q-21. Cumulative Scaled-Score Distribution for SLA—Grade 8

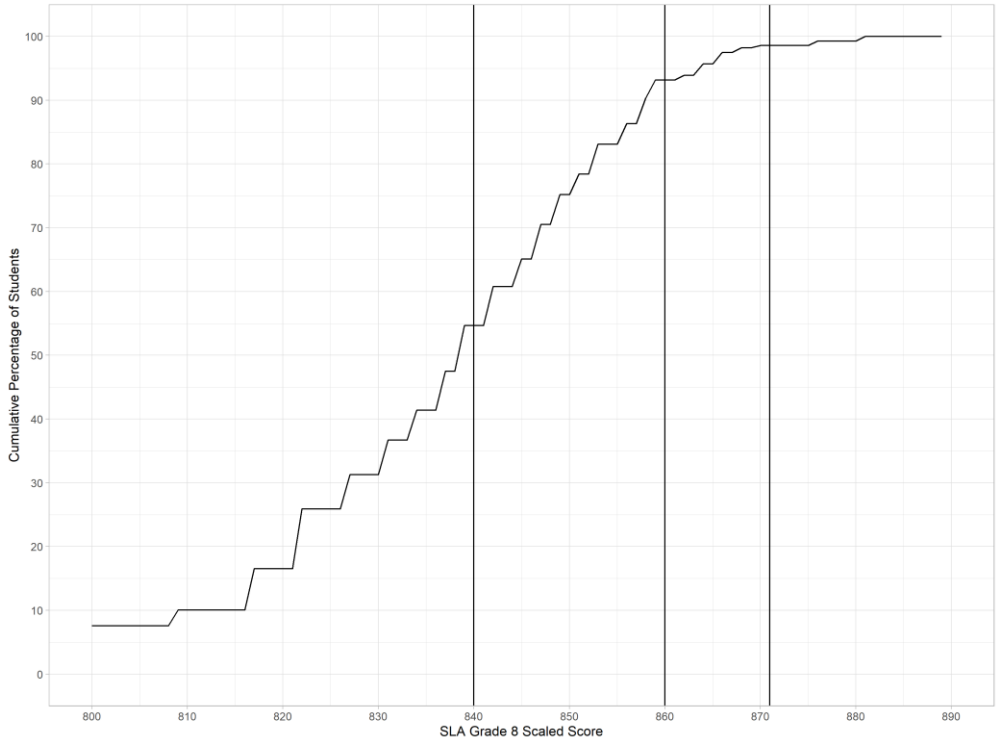


Figure Q-22. Cumulative Scaled-Score Distribution for Mathematics (Spanish Transadapted) Grade 3

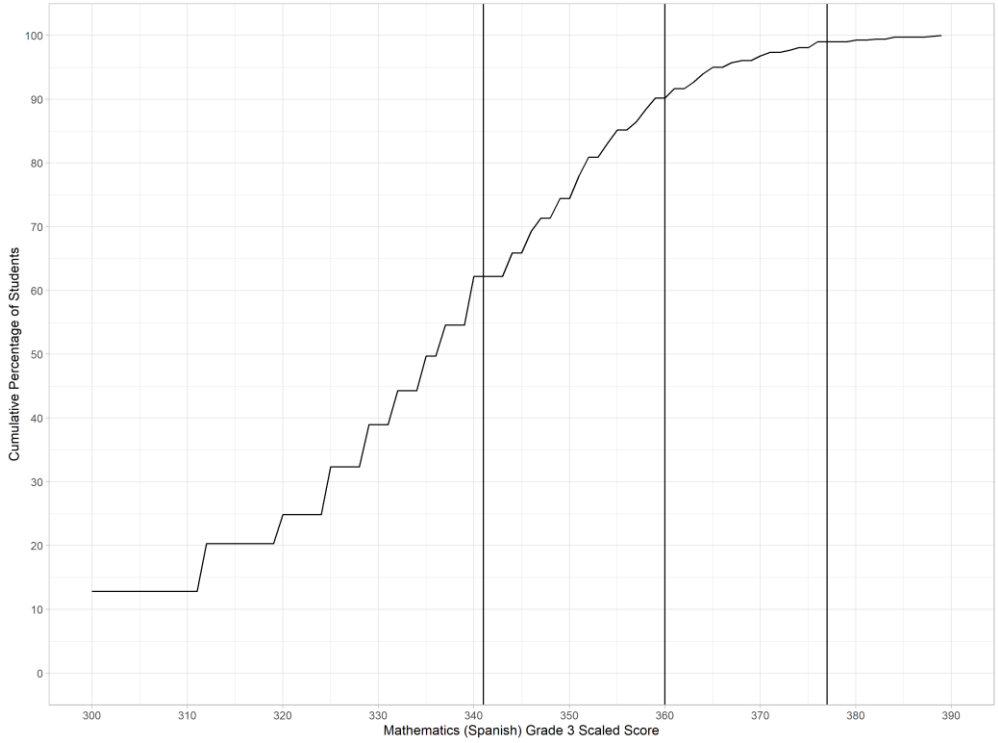


Figure Q-23. Cumulative Scaled-Score Distribution for Mathematics (Spanish Transadapted) Grade 4

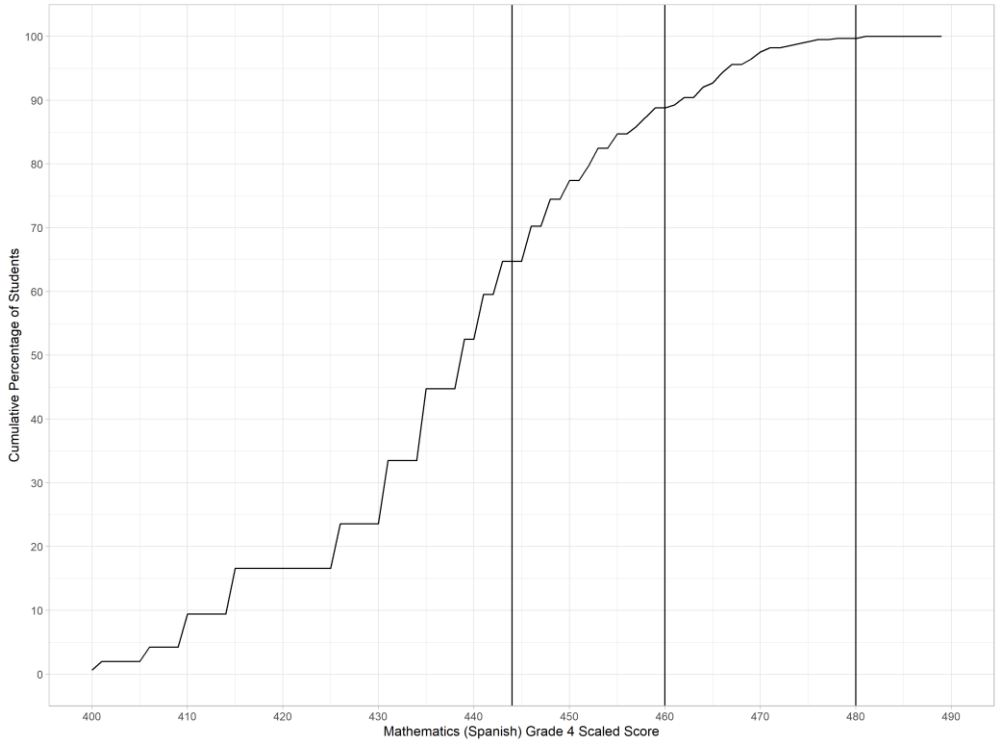


Figure Q-24. Cumulative Scaled-Score Distribution for Mathematics (Spanish Transadapted) Grade 5

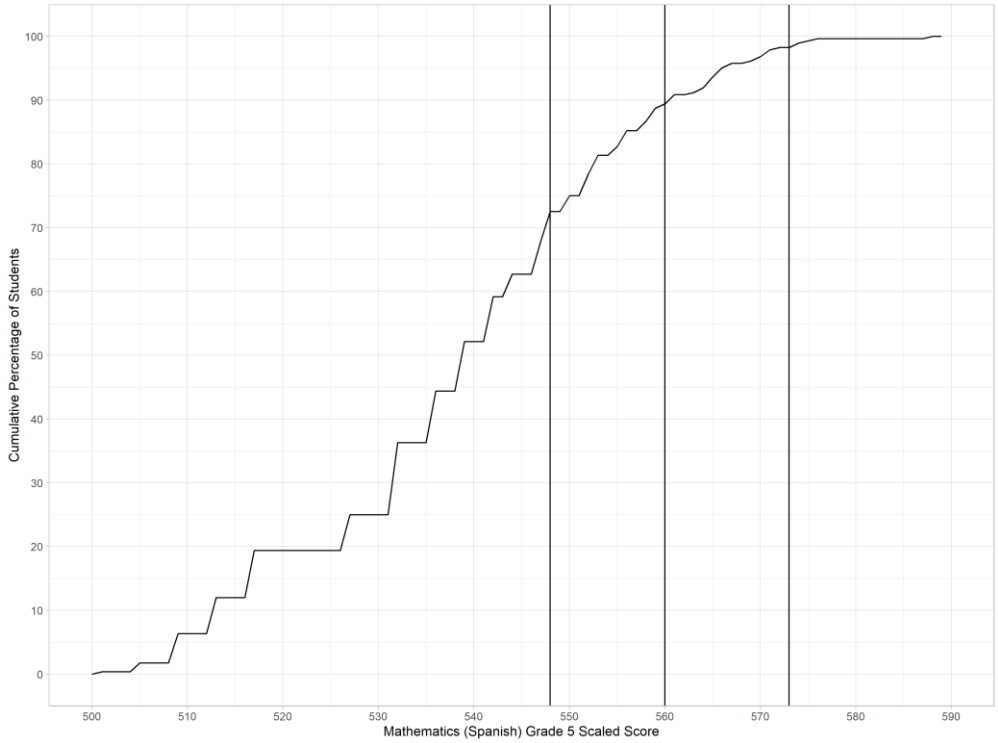


Figure Q-25. Cumulative Scaled-Score Distribution for Mathematics (Spanish Transadapted) Grade 6

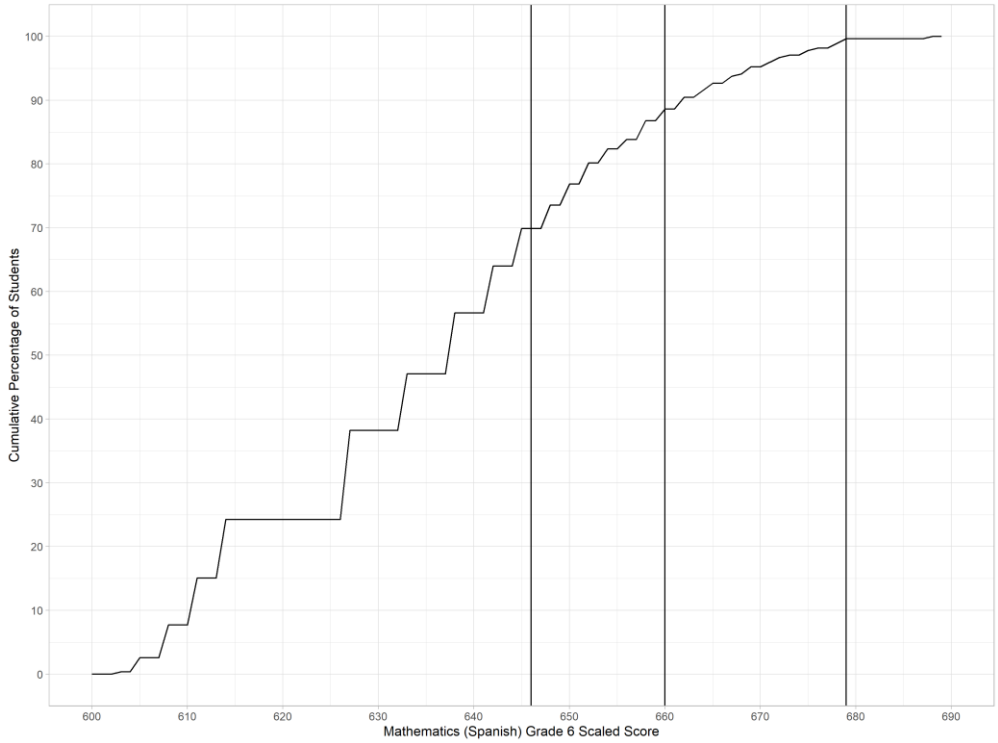


Figure Q-26. Cumulative Scaled-Score Distribution for Mathematics (Spanish Transadapted) Grade 7

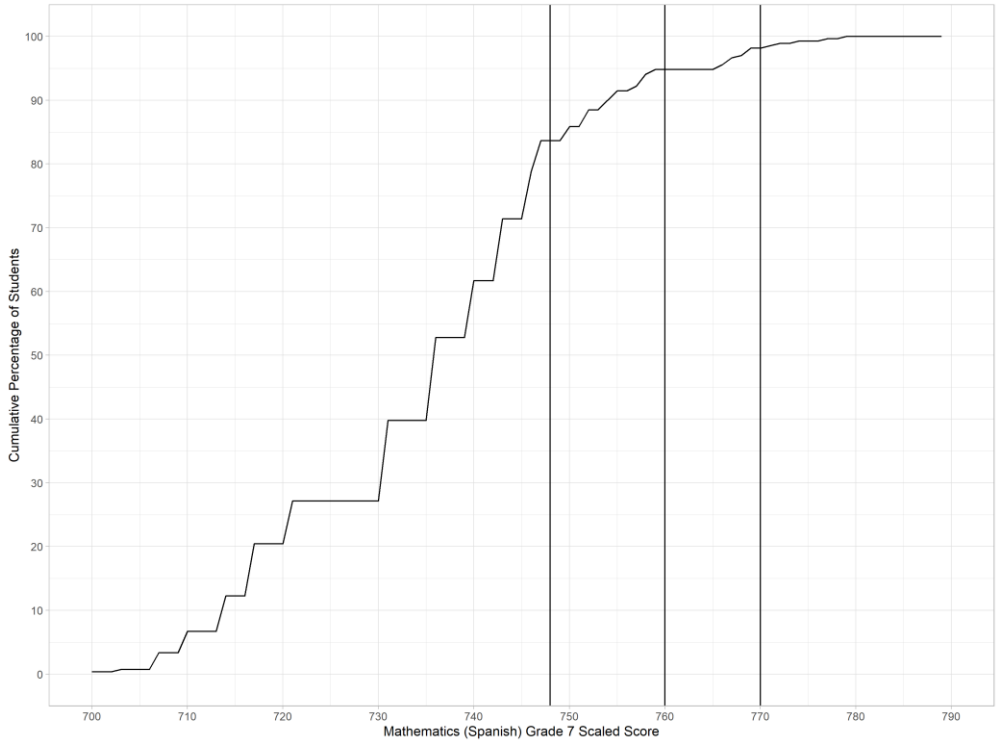


Figure Q-27. Cumulative Scaled-Score Distribution for Mathematics (Spanish Transadapted) Grade 8

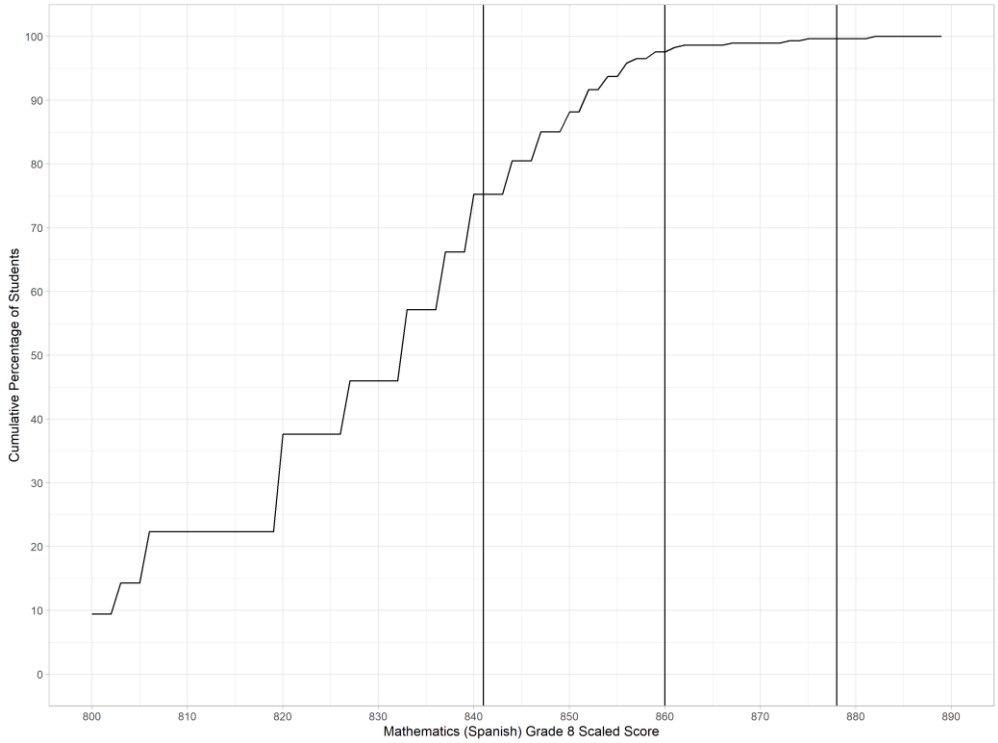


Figure Q-28. Cumulative Scaled-Score Distribution for Science (Spanish Transadapted) Grade 5

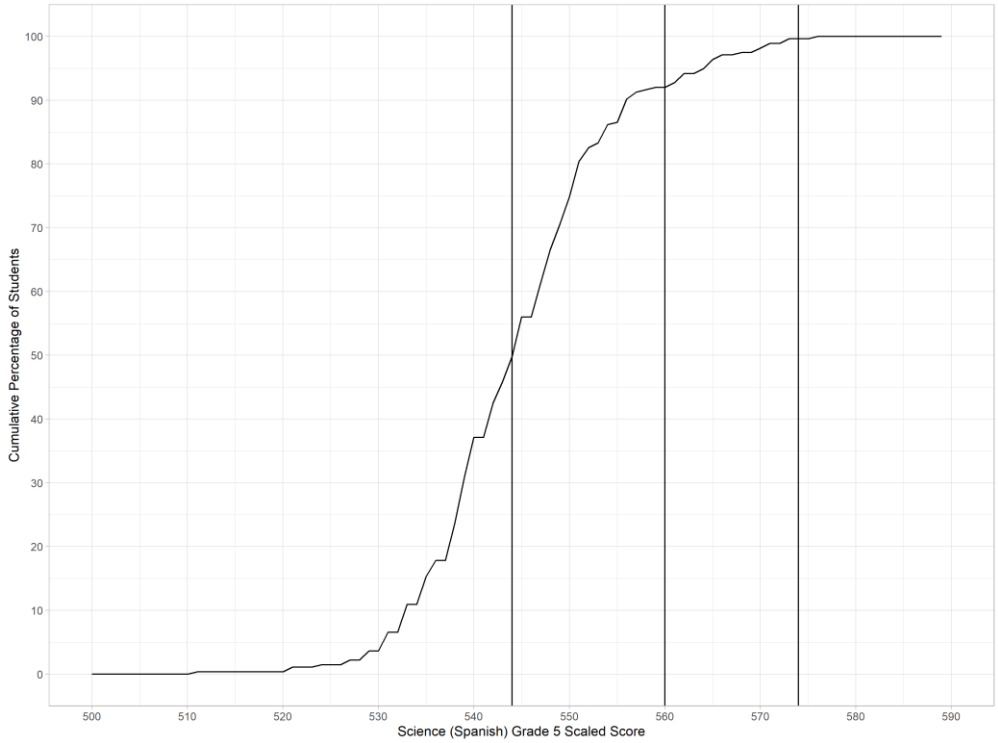


Figure Q-29. Cumulative Scaled-Score Distribution for Science (Spanish Transadapted) Grade 8

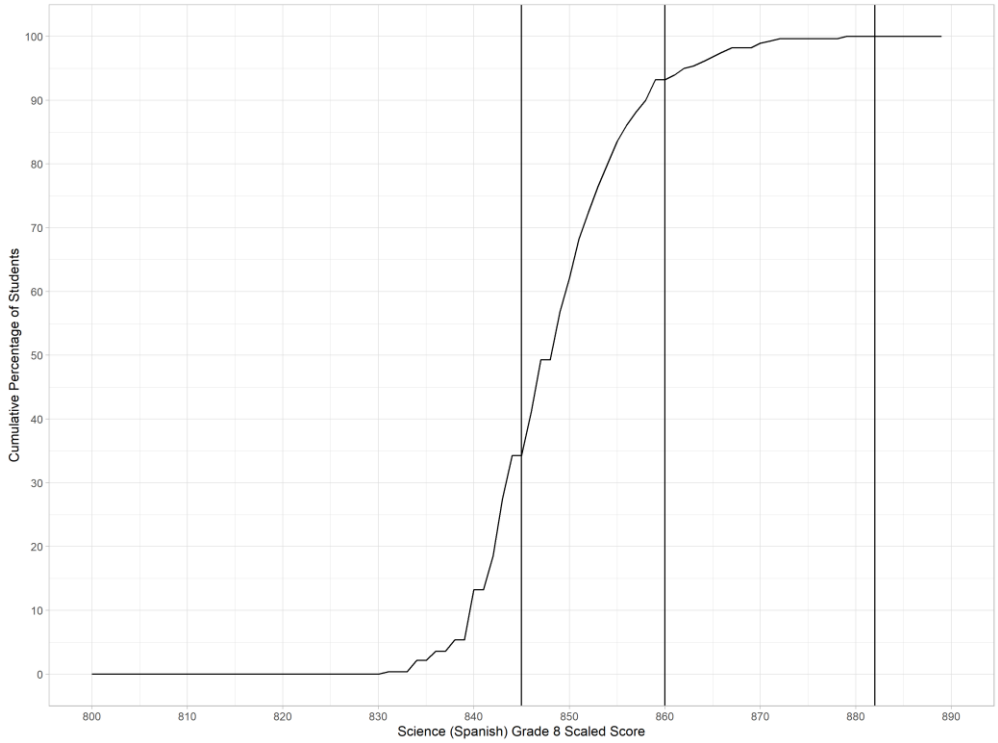
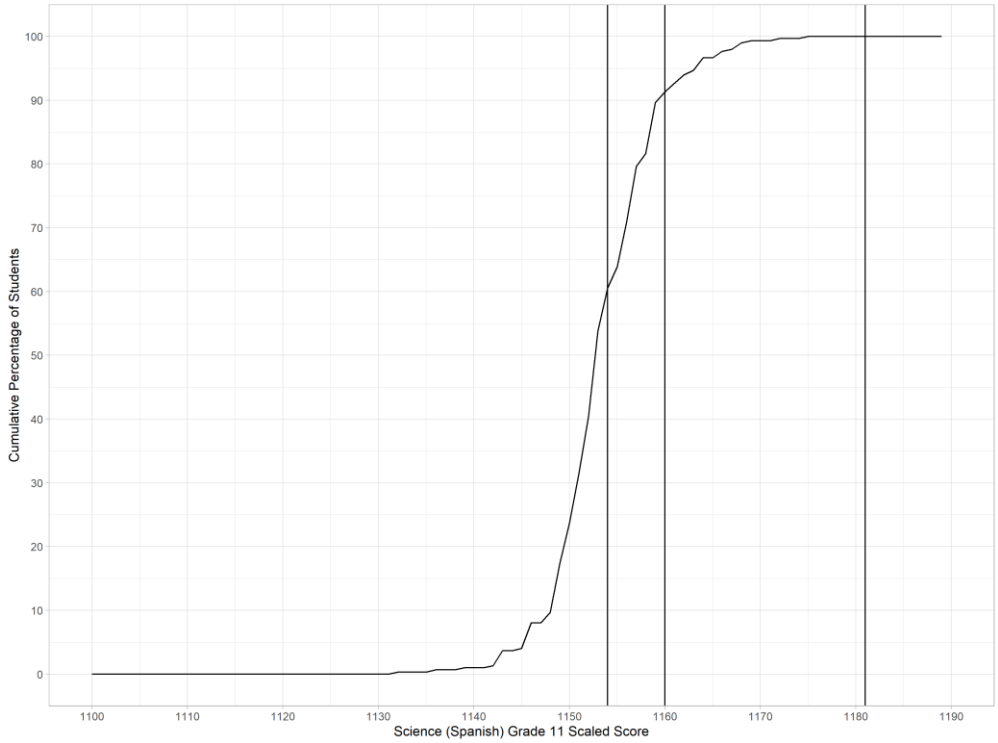


Figure Q-30. Cumulative Scaled-Score Distribution for Science (Spanish Transadapted) Grade 11



APPENDIX R
SCALED SCORE DESCRIPTIVE STATISTICS

Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.

Table R-1. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 3, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		20,877	350.6	353.0	22.5	-0.327	-0.462
Gender	Female	10,219	352.0	354.0	22.3	-0.380	-0.381
	Male	10,650	349.3	351.0	22.6	-0.276	-0.523
Ethnicity	African American or Black	700	348.0	349.0	22.2	-0.203	-0.517
	American Indian or Alaska Native	2,553	342.6	343.0	20.5	-0.130	-0.297
	Asian	350	362.3	367.0	21.2	-0.822	0.224
	Caucasian	16,853	351.6	353.0	22.5	-0.377	-0.431
	Hawaiian Native or Other Pacific Islander	68	348.9	350.0	22.3	-0.411	-0.300
	Multi	341	355.7	357.0	23.2	-0.479	-0.354
Hispanic	Yes	13,027	348.7	351.0	22.0	-0.299	-0.415
	No	7,838	353.7	356.0	23.1	-0.420	-0.482
Econ. Dis.	Yes	10,231	346.4	347.0	22.0	-0.213	-0.451
	No	8,288	356.6	359.0	21.8	-0.544	-0.225
English Learners	Yes	3,074	341.1	340.0	20.8	-0.061	-0.353
	No	17,803	352.2	354.0	22.4	-0.399	-0.395
Special Ed	Yes	3,536	335.5	335.0	21.1	0.241	-0.251
	No	16,845	353.7	355.0	21.6	-0.446	-0.236

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-2. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 4, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		21,324	453.0	456.0	22.5	-0.473	-0.340
Gender	Female	10,524	454.9	457.0	21.9	-0.522	-0.212
	Male	10,788	451.2	453.0	22.9	-0.419	-0.449
Ethnicity	African American or Black	647	452.8	456.0	22.8	-0.535	-0.369
	American Indian or Alaska Native	2,621	445.1	447.0	21.3	-0.271	-0.386
	Asian	395	465.5	468.0	20.0	-0.836	0.269
	Caucasian	17,213	454.0	456.0	22.4	-0.515	-0.285
	Hawaiian Native or Other Pacific Islander	75	449.8	455.0	20.8	-0.675	0.021
	Multi	358	454.8	458.0	23.2	-0.458	-0.496
Hispanic	Yes	13,175	451.2	453.0	22.0	-0.435	-0.331
	No	8,134	456.1	459.0	22.9	-0.576	-0.283
Econ. Dis.	Yes	10,915	449.3	451.0	22.2	-0.381	-0.424
	No	8,688	458.5	462.0	21.7	-0.675	0.012
English Learners	Yes	3,607	442.7	443.0	21.7	-0.191	-0.482
	No	17,717	455.1	458.0	22.1	-0.556	-0.203
Special Ed	Yes	3,771	436.4	435.0	22.4	0.230	-0.428
	No	17,066	456.7	459.0	20.8	-0.604	0.049

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-3. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 5, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		21,870	550.5	552.0	21.8	-0.418	-0.323
Gender	Female	10,690	552.6	555.0	21.2	-0.488	-0.142
	Male	11,177	548.5	550.0	22.2	-0.345	-0.458
Ethnicity	African American or Black	668	548.5	550.0	22.7	-0.323	-0.546
	American Indian or Alaska Native	2,557	543.2	545.0	20.9	-0.295	-0.392
	Asian	419	563.1	566.0	20.6	-0.956	0.917
	Caucasian	17,794	551.3	554.0	21.7	-0.448	-0.283
	Hawaiian Native or Other Pacific Islander	67	554.9	557.0	20.9	-0.801	0.537
	Multi	357	552.0	552.0	21.8	-0.381	-0.338
Hispanic	Yes	13,751	548.6	550.0	21.2	-0.390	-0.282
	No	8,111	553.7	557.0	22.5	-0.527	-0.307
Econ. Dis.	Yes	11,274	546.3	548.0	21.1	-0.324	-0.359
	No	8,847	556.4	559.0	21.3	-0.644	0.026
English Learners	Yes	4,102	540.5	541.0	20.0	-0.229	-0.412
	No	17,768	552.8	556.0	21.6	-0.514	-0.184
Special Ed	Yes	3,936	534.4	533.0	21.6	0.251	-0.364
	No	17,484	554.0	556.0	20.3	-0.526	0.028

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-4. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 6, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		22,037	652.0	654.0	21.1	-0.418	-0.228
Gender	Female	10,879	653.3	654.0	20.3	-0.399	-0.156
	Male	11,151	650.7	652.0	21.7	-0.414	-0.328
Ethnicity	African American or Black	697	650.2	652.0	21.9	-0.472	-0.215
	American Indian or Alaska Native	2,546	645.6	646.0	20.3	-0.248	-0.243
	Asian	377	662.3	665.0	20.9	-0.816	0.278
	Caucasian	17,992	652.7	654.0	20.9	-0.444	-0.196
	Hawaiian Native or Other Pacific Islander	94	648.3	656.0	22.0	-0.938	0.128
	Multi	316	654.7	656.0	21.7	-0.445	-0.297
Hispanic	Yes	13,892	650.1	652.0	20.5	-0.399	-0.185
	No	8,130	655.2	658.0	21.6	-0.514	-0.220
Econ. Dis.	Yes	10,810	647.8	650.0	20.6	-0.346	-0.256
	No	9,632	657.2	659.0	20.3	-0.590	0.035
English Learners	Yes	3,846	641.7	643.0	19.0	-0.289	-0.241
	No	18,191	654.2	656.0	20.8	-0.505	-0.112
Special Ed	Yes	3,999	636.7	635.0	22.1	0.215	-0.442
	No	17,588	655.4	656.0	19.2	-0.471	0.053

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-5. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 7, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		22,341	752.4	755.0	21.8	-0.510	-0.194
Gender	Female	10,947	754.7	757.0	20.7	-0.571	0.049
	Male	11,374	750.2	752.0	22.5	-0.434	-0.392
Ethnicity	African American or Black	652	750.3	754.0	22.3	-0.505	-0.211
	American Indian or Alaska Native	2,740	746.1	749.0	21.1	-0.328	-0.342
	Asian	376	762.9	765.0	20.8	-1.010	0.941
	Caucasian	18,145	753.2	757.0	21.6	-0.544	-0.144
	Hawaiian Native or Other Pacific Islander	117	752.0	755.0	21.8	-0.646	0.309
	Multi	284	755.7	757.0	21.1	-0.519	-0.062
Hispanic	Yes	13,952	750.4	752.0	21.3	-0.504	-0.188
	No	8,362	755.8	759.0	22.1	-0.583	-0.143
Econ. Dis.	Yes	10,623	747.9	749.0	21.3	-0.430	-0.296
	No	10,039	757.5	760.0	21.1	-0.673	0.106
English Learners	Yes	4,171	742.1	744.0	20.5	-0.345	-0.439
	No	18,170	754.8	757.0	21.3	-0.597	-0.027
Special Ed	Yes	3,796	736.3	735.0	22.9	0.212	-0.533
	No	18,092	755.9	757.0	19.9	-0.599	0.201

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-6. Scaled Score Descriptive Statistics for NM-MSSA ELA Grade 8, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		23,594	852.7	855.0	21.0	-0.498	-0.122
Gender	Female	11,558	855.6	858.0	19.8	-0.527	0.115
	Male	12,019	850.0	853.0	21.7	-0.436	-0.340
Ethnicity	African American or Black	714	852.4	855.0	21.5	-0.492	-0.203
	American Indian or Alaska Native	2,762	847.5	849.0	20.2	-0.374	-0.194
	Asian	394	865.1	868.0	19.1	-0.913	0.826
	Caucasian	19,317	853.2	856.0	21.0	-0.522	-0.088
	Hawaiian Native or Other Pacific Islander	107	857.5	858.0	19.2	-0.526	0.534
	Multi	287	857.4	861.0	21.6	-0.664	-0.047
Hispanic	Yes	14,928	850.7	853.0	20.6	-0.492	-0.118
	No	8,653	856.2	859.0	21.3	-0.573	-0.063
Econ. Dis.	Yes	11,146	848.8	851.0	20.7	-0.439	-0.194
	No	10,721	856.9	859.0	20.7	-0.623	0.080
English Learners	Yes	4,169	841.4	844.0	19.5	-0.346	-0.360
	No	19,425	855.2	858.0	20.5	-0.585	0.056
Special Ed	Yes	4,080	838.0	839.0	21.3	0.135	-0.357
	No	19,038	856.0	858.0	19.6	-0.614	0.278

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-7. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 3, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		20,702	343.1	344.0	21.3	-0.221	-0.414
Gender	Female	10,230	342.4	344.0	20.7	-0.202	-0.343
	Male	10,646	344.1	346.0	21.9	-0.251	-0.472
Ethnicity	African American or Black	702	339.8	340.0	21.3	-0.117	-0.478
	American Indian or Alaska Native	2,555	335.6	337.0	19.7	-0.129	-0.488
	Asian	355	359.3	361.0	19.9	-0.520	-0.121
	Caucasian	16,851	344.2	346.0	21.2	-0.265	-0.378
	Hawaiian Native or Other Pacific Islander	68	341.5	340.0	21.1	0.072	0.010
	Multi	341	346.5	347.0	23.4	-0.264	-0.514
Hispanic	Yes	13,022	341.2	340.0	20.5	-0.226	-0.383
	No	7,850	346.7	347.0	22.2	-0.297	-0.448
Econ. Dis.	Yes	10,240	339.0	340.0	20.6	-0.147	-0.431
	No	8,290	348.8	351.0	21.1	-0.382	-0.232
English Learners	Yes	3,079	335.6	337.0	20.1	-0.068	-0.491
	No	17,805	344.6	346.0	21.3	-0.265	-0.367
Special Ed	Yes	3,541	329.7	329.0	20.8	0.264	-0.529
	No	16,847	346.1	347.0	20.4	-0.288	-0.204

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-8. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 4, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		21,365	445.9	446.0	20.1	-0.157	-0.357
Gender	Female	10,540	444.8	446.0	19.1	-0.165	-0.251
	Male	10,813	447.0	448.0	20.9	-0.177	-0.459
Ethnicity	African American or Black	647	443.8	446.0	19.8	-0.180	-0.515
	American Indian or Alaska Native	2,619	439.5	441.0	18.5	-0.100	-0.338
	Asian	400	460.0	462.0	21.4	-0.641	-0.058
	Caucasian	17,248	446.7	448.0	20.0	-0.190	-0.335
	Hawaiian Native or Other Pacific Islander	75	444.9	446.0	17.6	-0.205	0.170
	Multi	361	446.6	446.0	21.8	-0.087	-0.531
Hispanic	Yes	13,203	444.0	443.0	19.2	-0.181	-0.311
	No	8,147	449.1	450.0	21.0	-0.211	-0.442
Econ. Dis.	Yes	10,925	442.1	443.0	19.3	-0.121	-0.344
	No	8,718	451.1	452.0	20.0	-0.302	-0.258
English Learners	Yes	3,626	438.6	439.0	18.9	-0.050	-0.435
	No	17,739	447.4	448.0	20.0	-0.199	-0.313
Special Ed	Yes	3,776	433.8	435.0	20.1	0.465	-0.018
	No	17,102	448.6	450.0	19.1	-0.246	-0.155

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-9. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 5, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		21,872	550.5	552.0	19.7	-0.262	-0.355
Gender	Female	10,685	550.3	552.0	19.0	-0.280	-0.245
	Male	11,184	550.7	552.0	20.3	-0.252	-0.453
Ethnicity	African American or Black	668	548.0	548.0	19.4	-0.184	-0.401
	American Indian or Alaska Native	2,558	545.6	547.0	18.1	-0.277	-0.252
	Asian	423	563.4	566.0	19.4	-0.648	-0.061
	Caucasian	17,792	551.0	552.0	19.7	-0.285	-0.344
	Hawaiian Native or Other Pacific Islander	67	554.4	555.0	21.2	-0.396	-0.626
	Multi	356	551.1	552.0	21.5	-0.200	-0.740
Hispanic	Yes	13,751	548.4	550.0	19.0	-0.251	-0.324
	No	8,113	554.2	555.0	20.2	-0.361	-0.348
Econ. Dis.	Yes	11,278	546.3	547.0	18.8	-0.219	-0.329
	No	8,842	555.5	558.0	19.6	-0.411	-0.261
English Learners	Yes	4,108	543.5	544.0	18.0	-0.201	-0.345
	No	17,764	552.1	553.0	19.7	-0.318	-0.319
Special Ed	Yes	3,936	537.7	536.0	19.4	0.371	-0.099
	No	17,487	553.3	555.0	18.6	-0.359	-0.107

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-10. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 6, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		22,054	648.4	650.0	20.7	-0.213	-0.453
Gender	Female	10,885	647.9	648.0	20.5	-0.204	-0.432
	Male	11,162	648.9	650.0	20.9	-0.224	-0.471
Ethnicity	African American or Black	702	646.2	648.0	20.1	-0.134	-0.329
	American Indian or Alaska Native	2,547	642.1	642.0	19.2	-0.122	-0.499
	Asian	381	661.5	664.0	21.2	-0.589	-0.315
	Caucasian	17,998	649.1	650.0	20.7	-0.248	-0.432
	Hawaiian Native or Other Pacific Islander	94	646.7	648.0	18.9	-0.336	-0.170
	Multi	317	649.9	650.0	20.8	-0.156	-0.465
Hispanic	Yes	13,893	646.3	648.0	20.0	-0.210	-0.408
	No	8,146	652.0	652.0	21.4	-0.298	-0.496
Econ. Dis.	Yes	10,815	644.2	645.0	19.8	-0.156	-0.462
	No	9,644	653.5	654.0	20.5	-0.358	-0.319
English Learners	Yes	3,853	639.9	642.0	18.8	-0.123	-0.527
	No	18,201	650.2	650.0	20.6	-0.271	-0.404
Special Ed	Yes	3,993	636.1	638.0	20.1	0.332	-0.282
	No	17,611	651.2	652.0	19.8	-0.309	-0.248

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-11. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 7, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		22,310	746.2	747.0	19.6	-0.090	-0.522
Gender	Female	10,930	745.3	746.0	19.0	-0.077	-0.448
	Male	11,360	747.0	747.0	20.1	-0.112	-0.588
Ethnicity	African American or Black	653	742.6	743.0	18.8	0.097	-0.474
	American Indian or Alaska Native	2,731	739.8	740.0	17.9	-0.024	-0.618
	Asian	378	759.1	759.0	20.5	-0.504	-0.399
	Caucasian	18,121	747.0	747.0	19.5	-0.128	-0.495
	Hawaiian Native or Other Pacific Islander	116	746.3	747.0	19.7	-0.130	-0.673
	Multi	284	747.5	747.0	22.0	-0.107	-0.742
Hispanic	Yes	13,921	743.9	746.0	18.8	-0.090	-0.502
	No	8,362	749.9	750.0	20.3	-0.183	-0.558
Econ. Dis.	Yes	10,612	741.9	743.0	18.3	-0.055	-0.508
	No	10,023	750.9	752.0	19.9	-0.238	-0.479
English Learners	Yes	4,174	737.5	740.0	16.8	-0.047	-0.669
	No	18,136	748.2	750.0	19.6	-0.165	-0.484
Special Ed	Yes	3,791	735.3	736.0	19.0	0.561	0.007
	No	18,066	748.4	750.0	18.9	-0.195	-0.366

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-12. Scaled Score Descriptive Statistics for NM-MSSA Mathematics Grade 8, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		23,601	841.5	844.0	20.7	-0.284	-0.420
Gender	Female	11,548	841.5	844.0	19.9	-0.328	-0.299
	Male	12,035	841.4	844.0	21.5	-0.248	-0.529
Ethnicity	African American or Black	715	840.0	840.0	20.4	-0.217	-0.437
	American Indian or Alaska Native	2,764	836.9	840.0	18.7	-0.297	-0.401
	Asian	396	856.9	860.0	21.1	-0.707	0.347
	Caucasian	19,319	841.8	844.0	20.8	-0.308	-0.425
	Hawaiian Native or Other Pacific Islander	109	843.9	850.0	22.0	-0.530	-0.532
	Multi	285	845.3	850.0	22.6	-0.392	-0.463
Hispanic	Yes	14,917	839.2	840.0	20.0	-0.302	-0.488
	No	8,671	845.5	847.0	21.3	-0.348	-0.327
Econ. Dis.	Yes	11,144	837.3	840.0	20.0	-0.224	-0.539
	No	10,741	845.9	847.0	20.7	-0.418	-0.211
English Learners	Yes	4,165	831.7	833.0	18.5	-0.169	-0.706
	No	19,436	843.6	847.0	20.6	-0.364	-0.306
Special Ed	Yes	4,089	830.0	833.0	20.2	0.282	-0.298
	No	19,034	844.0	847.0	20.0	-0.399	-0.189

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-13. Scaled Score Descriptive Statistics for NM-MSSA Science Grade 5, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		21,874	553.8	553.0	13.5	0.300	-0.378
Gender	Female	10,680	553.6	553.0	13.0	0.359	-0.288
	Male	11,191	554.0	553.0	14.0	0.248	-0.462
Ethnicity	African American or Black	669	551.9	550.0	13.2	0.322	-0.196
	American Indian or Alaska Native	2,547	549.0	548.0	11.4	0.408	0.143
	Asian	421	562.0	562.0	14.8	-0.078	-0.555
	Caucasian	17,803	554.3	554.0	13.6	0.263	-0.438
	Hawaiian Native or Other Pacific Islander	67	554.6	554.0	13.2	-0.034	-0.211
	Multi	359	555.2	554.0	14.4	0.269	-0.420
Hispanic	Yes	13,755	552.1	551.0	12.6	0.313	-0.264
	No	8,111	556.6	556.0	14.5	0.160	-0.610
Econ. Dis.	Yes	11,277	550.9	549.0	12.4	0.363	-0.127
	No	8,846	557.8	558.0	14.0	0.106	-0.596
English Learners	Yes	4,113	547.2	546.0	10.8	0.383	0.092
	No	17,761	555.3	555.0	13.6	0.215	-0.458
Special Ed	Yes	3,933	545.8	543.0	12.8	0.866	0.792
	No	17,492	555.6	555.0	13.1	0.258	-0.392

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-14. Scaled Score Descriptive Statistics for NM-MSSA Science Grade 8, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		23,595	854.9	854.0	10.4	0.391	0.164
Gender	Female	11,553	855.0	854.0	10.0	0.419	0.348
	Male	12,026	854.8	854.0	10.8	0.373	0.009
Ethnicity	African American or Black	714	853.9	853.0	10.5	0.433	0.000
	American Indian or Alaska Native	2,763	852.3	852.0	8.6	0.423	0.435
	Asian	394	862.5	863.0	11.8	-0.087	-0.070
	Caucasian	19,319	855.1	854.0	10.5	0.363	0.132
	Hawaiian Native or Other Pacific Islander	107	855.7	855.0	10.4	0.089	-0.762
	Multi	286	857.6	857.0	10.7	0.354	0.135
Hispanic	Yes	14,926	853.6	853.0	9.7	0.331	0.309
	No	8,657	857.2	856.0	11.2	0.335	-0.195
Econ. Dis.	Yes	11,143	852.8	852.0	9.5	0.369	0.419
	No	10,720	857.1	856.0	10.9	0.303	-0.083
English Learners	Yes	4,175	849.4	849.0	7.9	0.239	2.053
	No	19,420	856.1	855.0	10.5	0.324	-0.029
Special Ed	Yes	4,087	849.3	847.0	9.6	1.026	2.512
	No	19,030	856.2	855.0	10.2	0.318	0.086

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-15. Scaled Score Descriptive Statistics for NM-MSSA Science Grade 11, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		21,157	1158.7	1158.0	8.1	0.518	0.416
Gender	Female	10,680	1158.2	1157.0	7.5	0.584	0.480
	Male	10,468	1159.3	1158.0	8.7	0.425	0.261
Ethnicity	African American or Black	581	1157.7	1157.0	7.4	0.526	0.112
	American Indian or Alaska Native	2,693	1156.8	1156.0	6.6	0.529	0.544
	Asian	411	1164.0	1164.0	9.0	-0.001	-0.551
	Caucasian	17,196	1158.9	1158.0	8.2	0.506	0.340
	Hawaiian Native or Other Pacific Islander	87	1158.4	1158.0	8.1	0.715	0.700
	Multi	188	1161.3	1161.0	9.6	-0.710	3.832
Hispanic	Yes	12,983	1157.5	1156.0	7.4	0.503	0.820
	No	8,173	1160.7	1159.0	8.7	0.397	-0.117
Econ. Dis.	Yes	8,031	1156.6	1156.0	7.0	0.438	1.316
	No	12,177	1160.3	1159.0	8.5	0.430	-0.047
English Learners	Yes	2,869	1153.3	1153.0	5.2	0.472	1.519
	No	18,288	1159.6	1158.0	8.1	0.433	0.357
Special Ed	Yes	2,818	1153.5	1153.0	6.3	0.939	3.805
	No	18,337	1159.6	1158.0	8.0	0.463	0.324

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-16. Scaled Score Descriptive Statistics for NM-MSSA SLA Grade 3, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		662	341.4	343.0	20.6	-0.186	-0.390
Gender	Female	329	343.2	345.0	20.6	-0.252	-0.281
	Male	333	339.7	340.0	20.4	-0.128	-0.456
Ethnicity	African American or Black	8	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	3	--	--	--	--	--
	Caucasian	645	341.6	343.0	20.6	-0.193	-0.381
	Hawaiian Native or Other Pacific Islander	4	--	--	--	--	--
	Multi	2	--	--	--	--	--
Hispanic	Yes	651	341.5	343.0	20.6	-0.186	-0.387
	No	11	--	--	--	--	--
Econ. Dis.	Yes	484	341.5	343.0	21.5	-0.175	-0.490
	No	137	340.6	340.0	17.9	-0.169	-0.263
English Learners	Yes	605	341.4	343.0	20.8	-0.156	-0.426
	No	57	341.6	345.0	17.8	-0.701	0.110
Special Ed	Yes	70	327.2	329.0	18.8	0.063	-0.337
	No	590	343.2	343.0	20.1	-0.220	-0.324

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-17. Scaled Score Descriptive Statistics for NM-MSSA SLA Grade 4, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		611	440.4	439.0	19.0	-0.210	-0.119
Gender	Female	308	441.8	443.0	18.5	-0.139	0.073
	Male	303	438.9	439.0	19.5	-0.255	-0.319
Ethnicity	African American or Black	3	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	3	--	--	--	--	--
	Caucasian	599	440.3	439.0	19.1	-0.219	-0.147
	Hawaiian Native or Other Pacific Islander	3	--	--	--	--	--
	Multi	2	--	--	--	--	--
Hispanic	Yes	599	440.2	439.0	19.0	-0.227	-0.138
	No	11	--	--	--	--	--
Econ. Dis.	Yes	448	439.2	439.0	18.9	-0.203	-0.201
	No	138	443.3	443.0	20.0	-0.230	-0.039
English Learners	Yes	561	440.0	439.0	18.8	-0.225	-0.134
	No	50	444.1	443.0	21.5	-0.231	-0.031
Special Ed	Yes	56	429.7	431.0	15.9	-0.250	0.066
	No	545	441.4	443.0	19.1	-0.250	-0.101

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-18. Scaled Score Descriptive Statistics for NM-MSSA SLA Grade 5, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		272	541.0	543.0	17.3	-0.500	-0.046
Gender	Female	140	543.8	546.0	16.0	-0.471	0.229
	Male	132	538.2	541.0	18.3	-0.445	-0.354
Ethnicity	African American or Black	4	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	2	--	--	--	--	--
	Caucasian	261	541.0	543.0	17.5	-0.498	-0.091
	Hawaiian Native or Other Pacific Islander	1	--	--	--	--	--
	Multi	4	--	--	--	--	--
Hispanic	Yes	265	541.0	543.0	17.4	-0.493	-0.039
	No	7	--	--	--	--	--
Econ. Dis.	Yes	144	538.4	541.0	18.8	-0.338	-0.605
	No	114	543.4	544.5	15.6	-0.504	0.782
English Learners	Yes	229	541.3	543.0	17.5	-0.477	-0.056
	No	43	--	--	--	--	--
Special Ed	Yes	7	--	--	--	--	--
	No	260	541.1	543.0	17.5	-0.503	-0.055

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-19. Scaled Score Descriptive Statistics for NM-MSSA SLA Grade 6, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		269	635.9	638.0	18.2	-0.177	-0.277
Gender	Female	134	637.6	638.0	17.7	-0.326	0.041
	Male	135	634.1	631.0	18.6	-0.028	-0.427
Ethnicity	African American or Black	2	--	--	--	--	--
	American Indian or Alaska Native	3	--	--	--	--	--
	Asian	1	--	--	--	--	--
	Caucasian	258	635.4	638.0	18.3	-0.136	-0.279
	Hawaiian Native or Other Pacific Islander	2	--	--	--	--	--
	Multi	3	--	--	--	--	--
Hispanic	Yes	264	635.9	638.0	18.2	-0.180	-0.238
	No	5	--	--	--	--	--
Econ. Dis.	Yes	139	634.3	635.0	18.2	-0.137	-0.237
	No	113	636.9	638.0	18.8	-0.168	-0.327
English Learners	Yes	221	636.8	638.0	18.1	-0.166	-0.159
	No	48	--	--	--	--	--
Special Ed	Yes	12	--	--	--	--	--
	No	256	636.7	638.0	18.0	-0.191	-0.222

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-20. Scaled Score Descriptive Statistics for NM-MSSA SLA Grade 7, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		259	736.6	738.0	18.7	-0.152	-0.174
Gender	Female	130	739.2	741.0	18.0	0.031	0.325
	Male	129	734.0	735.0	19.2	-0.262	-0.723
Ethnicity	African American or Black	4	--	--	--	--	--
	American Indian or Alaska Native	3	--	--	--	--	--
	Asian	4	--	--	--	--	--
	Caucasian	243	736.9	738.0	18.4	-0.186	-0.167
	Hawaiian Native or Other Pacific Islander	2	--	--	--	--	--
	Multi	3	--	--	--	--	--
Hispanic	Yes	253	736.8	738.0	18.6	-0.138	-0.145
	No	6	--	--	--	--	--
Econ. Dis.	Yes	141	734.9	735.0	18.9	-0.327	-0.530
	No	103	737.7	738.0	19.2	0.224	0.015
English Learners	Yes	215	737.1	741.0	18.7	-0.161	-0.092
	No	44	--	--	--	--	--
Special Ed	Yes	3	--	--	--	--	--
	No	255	736.8	738.0	18.5	-0.132	-0.127

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-21. Scaled Score Descriptive Statistics for NM-MSSA SLA Grade 8, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		278	837.1	839.0	18.0	-0.301	-0.345
Gender	Female	130	840.3	842.0	18.4	-0.541	-0.108
	Male	148	834.3	837.0	17.3	-0.143	-0.301
Ethnicity	African American or Black	2	--	--	--	--	--
	American Indian or Alaska Native	2	--	--	--	--	--
	Asian	2	--	--	--	--	--
	Caucasian	268	837.0	839.0	18.2	-0.307	-0.354
	Hawaiian Native or Other Pacific Islander	4	--	--	--	--	--
	Multi	0	--	--	--	--	--
Hispanic	Yes	277	837.1	839.0	18.1	-0.296	-0.351
	No	1	--	--	--	--	--
Econ. Dis.	Yes	147	836.6	839.0	18.5	-0.279	-0.352
	No	119	838.0	839.0	17.5	-0.318	-0.279
English Learners	Yes	248	837.0	839.0	17.4	-0.390	-0.276
	No	30	--	--	--	--	--
Special Ed	Yes	3	--	--	--	--	--
	No	274	837.2	839.0	18.1	-0.315	-0.348

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-22. Scaled Score Descriptive Statistics for NM-MSSA Mathematics (Spanish Transadapted) Grade 3, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		680	335.0	337.0	20.3	-0.117	-0.585
Gender	Female	335	334.9	335.0	20.7	-0.059	-0.504
	Male	345	335.1	337.0	20.0	-0.179	-0.668
Ethnicity	African American or Black	8	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	3	--	--	--	--	--
	Caucasian	662	335.2	337.0	20.4	-0.130	-0.583
	Hawaiian Native or Other Pacific Islander	5	--	--	--	--	--
	Multi	2	--	--	--	--	--
Hispanic	Yes	666	335.1	337.0	20.2	-0.130	-0.587
	No	14	--	--	--	--	--
Econ. Dis.	Yes	493	335.2	337.0	21.4	-0.108	-0.684
	No	147	333.7	335.0	17.2	-0.164	-0.414
English Learners	Yes	619	334.8	335.0	20.5	-0.105	-0.619
	No	61	337.6	337.0	17.7	-0.161	-0.138
Special Ed	Yes	70	321.7	325.0	18.8	0.454	-0.614
	No	608	336.7	337.0	19.9	-0.175	-0.450

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-23. Scaled Score Descriptive Statistics for NM-MSSA Mathematics (Spanish Transadapted) Grade 4, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		615	438.6	439.0	17.1	-0.116	-0.331
Gender	Female	308	436.8	439.0	16.6	-0.123	-0.326
	Male	307	440.4	441.0	17.5	-0.145	-0.331
Ethnicity	African American or Black	3	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	3	--	--	--	--	--
	Caucasian	604	438.8	439.0	17.0	-0.120	-0.298
	Hawaiian Native or Other Pacific Islander	3	--	--	--	--	--
	Multi	1	--	--	--	--	--
Hispanic	Yes	602	438.6	439.0	17.1	-0.124	-0.302
	No	12	--	--	--	--	--
Econ. Dis.	Yes	447	438.6	439.0	17.7	-0.118	-0.461
	No	142	438.9	439.0	15.7	-0.023	0.139
English Learners	Yes	559	439.0	439.0	17.1	-0.118	-0.341
	No	56	434.9	437.0	17.1	-0.109	-0.152
Special Ed	Yes	55	433.1	435.0	16.1	-0.033	-0.276
	No	550	439.1	439.0	17.2	-0.131	-0.323

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-24. Scaled Score Descriptive Statistics for NM-MSSA Mathematics (Spanish Transadapted) Grade 5, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		284	539.1	539.0	17.2	-0.068	-0.481
Gender	Female	146	539.3	539.0	17.4	-0.002	-0.382
	Male	138	538.9	540.5	16.9	-0.148	-0.585
Ethnicity	African American or Black	4	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	2	--	--	--	--	--
	Caucasian	272	539.2	539.0	17.3	-0.070	-0.496
	Hawaiian Native or Other Pacific Islander	2	--	--	--	--	--
	Multi	4	--	--	--	--	--
Hispanic	Yes	277	539.1	539.0	17.2	-0.074	-0.476
	No	7	--	--	--	--	--
Econ. Dis.	Yes	152	537.6	539.0	17.8	0.169	-0.466
	No	118	540.4	540.5	16.1	-0.318	-0.211
English Learners	Yes	236	539.3	539.0	17.3	-0.070	-0.386
	No	48	--	--	--	--	--
Special Ed	Yes	7	--	--	--	--	--
	No	272	539.2	539.0	17.3	-0.076	-0.492

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-25. Scaled Score Descriptive Statistics for NM-MSSA Mathematics (Spanish Transadapted) Grade 6, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		272	636.6	638.0	19.2	0.173	-0.711
Gender	Female	136	634.5	633.0	18.0	0.151	-0.765
	Male	136	638.6	638.0	20.2	0.132	-0.746
Ethnicity	African American or Black	2	--	--	--	--	--
	American Indian or Alaska Native	3	--	--	--	--	--
	Asian	1	--	--	--	--	--
	Caucasian	261	636.5	638.0	19.4	0.188	-0.725
	Hawaiian Native or Other Pacific Islander	2	--	--	--	--	--
	Multi	3	--	--	--	--	--
Hispanic	Yes	267	636.5	638.0	19.1	0.187	-0.663
	No	5	--	--	--	--	--
Econ. Dis.	Yes	140	635.9	638.0	18.7	0.189	-0.681
	No	116	636.3	638.0	20.3	0.264	-0.745
English Learners	Yes	226	636.8	638.0	19.4	0.189	-0.724
	No	46	--	--	--	--	--
Special Ed	Yes	11	--	--	--	--	--
	No	260	637.2	638.0	19.2	0.131	-0.713

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-26. Scaled Score Descriptive Statistics for NM-MSSA Mathematics (Spanish Transadapted) Grade 7, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		269	735.8	736.0	15.7	0.067	-0.284
Gender	Female	134	735.9	736.0	15.6	0.027	-0.241
	Male	135	735.7	736.0	15.9	0.106	-0.289
Ethnicity	African American or Black	4	--	--	--	--	--
	American Indian or Alaska Native	4	--	--	--	--	--
	Asian	4	--	--	--	--	--
	Caucasian	252	735.6	736.0	15.6	0.069	-0.228
	Hawaiian Native or Other Pacific Islander	2	--	--	--	--	--
	Multi	3	--	--	--	--	--
Hispanic	Yes	263	735.8	736.0	15.5	0.025	-0.355
	No	6	--	--	--	--	--
Econ. Dis.	Yes	145	734.8	736.0	15.3	0.082	-0.178
	No	107	736.6	736.0	16.9	0.082	-0.477
English Learners	Yes	223	736.3	736.0	15.6	0.057	-0.424
	No	46	--	--	--	--	--
Special Ed	Yes	3	--	--	--	--	--
	No	265	735.8	736.0	15.7	0.066	-0.272

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-27. Scaled Score Descriptive Statistics for NM-MSSA Mathematics (Spanish Transadapted) Grade 8, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		287	829.4	833.0	18.1	-0.074	-0.710
Gender	Female	135	830.7	833.0	18.1	-0.160	-0.456
	Male	152	828.3	827.0	18.1	0.002	-0.887
Ethnicity	African American or Black	2	--	--	--	--	--
	American Indian or Alaska Native	2	--	--	--	--	--
	Asian	2	--	--	--	--	--
	Caucasian	277	829.5	833.0	18.3	-0.080	-0.731
	Hawaiian Native or Other Pacific Islander	4	--	--	--	--	--
	Multi	0	--	--	--	--	--
Hispanic	Yes	286	829.3	833.0	18.1	-0.067	-0.704
	No	1	--	--	--	--	--
Econ. Dis.	Yes	150	830.6	833.0	18.1	-0.144	-0.680
	No	123	827.8	833.0	18.3	0.026	-0.664
English Learners	Yes	255	829.6	833.0	17.9	-0.140	-0.687
	No	32	--	--	--	--	--
Special Ed	Yes	2	--	--	--	--	--
	No	284	829.6	833.0	18.0	-0.074	-0.690

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-28. Scaled Score Descriptive Statistics for NM-MSSA Science (Spanish Transadapted) Grade 5, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		275	545.1	545.0	9.8	0.394	0.839
Gender	Female	143	545.7	545.0	9.2	0.678	1.086
	Male	132	544.4	543.0	10.5	0.231	0.588
Ethnicity	African American or Black	4	--	--	--	--	--
	American Indian or Alaska Native	0	--	--	--	--	--
	Asian	2	--	--	--	--	--
	Caucasian	263	545.1	544.0	9.8	0.457	0.884
	Hawaiian Native or Other Pacific Islander	2	--	--	--	--	--
	Multi	4	--	--	--	--	--
Hispanic	Yes	268	545.0	544.5	9.7	0.421	0.876
	No	7	--	--	--	--	--
Econ. Dis.	Yes	147	544.7	544.0	10.5	0.352	0.644
	No	114	545.5	544.0	8.9	0.671	1.053
English Learners	Yes	229	545.1	545.0	9.3	0.645	0.666
	No	46	--	--	--	--	--
Special Ed	Yes	6	--	--	--	--	--
	No	265	545.3	545.0	9.8	0.389	0.845

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-29. Scaled Score Descriptive Statistics for NM-MSSA Science (Spanish Transadapted) Grade 8, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		280	848.8	849.0	7.5	0.688	0.943
Gender	Female	132	849.2	849.0	7.8	0.593	0.748
	Male	148	848.5	847.0	7.3	0.781	1.230
Ethnicity	African American or Black	2	--	--	--	--	--
	American Indian or Alaska Native	2	--	--	--	--	--
	Asian	2	--	--	--	--	--
	Caucasian	270	848.9	849.0	7.6	0.690	0.921
	Hawaiian Native or Other Pacific Islander	4	--	--	--	--	--
	Multi	0	--	--	--	--	--
Hispanic	Yes	279	848.8	849.0	7.5	0.697	0.962
	No	1	--	--	--	--	--
Econ. Dis.	Yes	150	848.7	849.0	7.6	0.550	0.442
	No	117	848.9	847.0	7.6	0.947	1.646
English Learners	Yes	247	848.5	847.0	7.5	0.696	1.107
	No	33	--	--	--	--	--
Special Ed	Yes	2	--	--	--	--	--
	No	277	848.8	849.0	7.6	0.685	0.915

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

Table R-30. Scaled Score Descriptive Statistics for NM-MSSA Science (Spanish Transadapted) Grade 11, as a Function of Subgroup*

Group	Subgroup	Number of Students	Mean	Median	SD	Skewness	Kurtosis
Overall		299	1153.8	1153.0	5.5	0.220	1.919
Gender	Female	162	1153.6	1153.0	5.3	0.097	2.328
	Male	137	1154.1	1153.0	5.6	0.325	1.574
Ethnicity	African American or Black	3	--	--	--	--	--
	American Indian or Alaska Native	5	--	--	--	--	--
	Asian	0	--	--	--	--	--
	Caucasian	287	1153.9	1153.0	5.5	0.196	1.880
	Hawaiian Native or Other Pacific Islander	4	--	--	--	--	--
	Multi	0	--	--	--	--	--
Hispanic	Yes	293	1153.9	1153.0	5.5	0.200	1.848
	No	6	--	--	--	--	--
Econ. Dis.	Yes	169	1154.0	1153.0	5.3	0.314	0.504
	No	130	1153.6	1153.0	5.7	0.141	3.385
English Learners	Yes	240	1153.8	1153.0	5.5	0.134	1.932
	No	59	1154.2	1154.0	5.4	0.617	2.020
Special Ed	Yes	2	--	--	--	--	--
	No	296	1153.8	1153.0	5.5	0.225	1.892

**Calculations based on those students attempting 5 or more items on the given NM-MSSA & ASR assessments. Statistical values are suppressed for those content areas/grades with fewer than 50 students.*

APPENDIX S
SCORE REPORT INTERPRETATION GUIDE

NEW MEXICO
Public Education Department

S I G

Score Report Interpretation Guide

for Computer-Based and Paper-Based Tests

Spring 2022

NM-MSSA Grades 3–8
NM-ASR Grades 5, 8, and 11

NM-MSSA

NEW MEXICO MEASURES OF STUDENT SUCCESS AND ACHIEVEMENT

NM-ASR

NEW MEXICO ASSESSMENT OF SCIENCE READINESS

PED Contact Information

General Administration Questions	Policy Questions
Cognia New Mexico Customer Care Center & Help Desk Team	New Mexico Public Education Department Assessment Bureau Helpdesk*
Telephone: 877-676-6722 **Email:** nmtechsupport@cognia.org	**Telephone:** 505-827-5861 **Email:** ped.assessment@state.nm.us

*The PED should only be contacted by the district test coordinator (DTC). Test administrators should contact their school test coordinator or DTC with any questions or concerns.

Note: This manual is available online at newmexico.onlinehelp.cognia.org/combined-manuals-summatives.

Table of Contents

1.0 General Information for Families and Educators

1.1 Background

The New Mexico Measures of Student Success and Achievement (NM-MSSA) is the summative assessment in Language Arts, and Mathematics for students in grades 3–8 aligned to the New Mexico Common Core State Standards (NMCCSS) for math and language arts. The assessment measures a student's grade level proficiency and progress toward college and/or career readiness.

The NM-MSSA Spanish Language Arts Assessment for students in grades 3–8 is aligned to the Common Core Español Standards for Language Arts. The assessment measures a student's grade level proficiency and progress toward college and/or career readiness.

The New Mexico Assessment of Science Readiness (NM-ASR) Is a summative assessment in Science for students in grades 5, 8, and 11 aligned to the New Mexico STEM Ready! Science Standards. The assessment measures whether students are on track to be ready for college and/or career.

1.2 NM-MSSA and NM-ASR Assessments

The NM-MSSA is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the NMCCSS. The NM-ASR is designed to measure whether students are on track to be ready for college or career, as defined by the State, by showing they have mastered the New Mexico STEM Ready! Science Standards.

The Spring 2022 NM-MSSA assessments were administered in either computer-based or paper-based format. The Reading assessment contained items that focused on understanding key ideas and details, analyzing elements of craft and structure, and integrating knowledge and ideas using informational and literary texts. The Writing and Language assessment contained items that focused on communicating clearly and effectively for a particular task and purpose, determining the meaning of grade-appropriate words, and applying conventions of standard English grammar, usage, and mechanics. The Mathematics assessment focused on understanding and applying skills and concepts, solving multi-step problems that require abstract reasoning, and modeling real-world problems with precision, perseverance, and strategic use of tools. The Science assessement focused on the integration of Science and Engineering Practices, Disciplinary Core Ideas, and Crosscutting Concepts to explain phenomena and solve problems. In each content area, students demonstrated their acquired skills and knowledge by answering selected-response items, multi-select selected response items, and extended response items.

1.3 Confidentiality of Reporting Results

Individual student performance results on NM-MSSA and NM-ASR assessments are confidential and may be released only in accordance with the Family Educational Rights and Privacy Act of 1974 (20 U.S.C. Section 1232g). Aggregated student performance data are made available to the public and do not contain the names of individual students or teachers.

1.4 Purpose of this Guide

This guide provides information on the individual student reports, school reports, and district reports provided for NM-MSSA and NM-ASR assessment results. Section 2.0, which outlines and explains elements of the individual student report, may be shared with families. This section will help families understand their child's test results. "3.0 Understanding the NM-MSSA and NM-ASR School and District Reports" on page 8 outlines and explains elements of the school and district reports. New Mexico state policies and calculations for accountability reporting may differ from the policies and calculations used for assessment reports.

Sample reports included in this guide are for illustration purposes only. They are provided to show the basic layout of the reports and the information they provide. Sample reports do not include actual data from any administration.

2.0 Understanding the NM-MSSA and NM-ASR Individual Student Report (ISR)

2.1 Types of Scores on the NM-MSSA and NM-ASR ISR

Student performance on NM-MSSA and NM-ASR assessments is described on the individual student report using the interim scale scores, performance levels, standard error, and subclaim performance indicators.

2.1.1 Scale Score

A scale score is a numerical value that summarizes student performance. Not all students respond to the same set of test items, so each student's scaled score accounts for the slight differences in difficulty among the various forms and administrations of the test. The resulting scale score allows for an appropriate comparison across test forms and administration years within a grade or course and content area. NM-MSSA and NM-ASR reports provide overall scale scores for Language Arts, Mathematics, and Science, each of which determines a student's performance level in the respective content area. Scale score ranges differ by grade for all tests.

For example, a student who earns an overall scale score of 800 on one form of the grade 8 Mathematics assessment would be expected to earn an overall scale score of 800 on any other form of the grade 8 Mathematics assessment. Furthermore, the student's overall scale score and level of mastery of concepts and skills would be comparable to a student who took the same assessment the previous year or following year.

2.1.2 Performance Level

Each performance level is a broad, categorical level defined by a student's overall scale score and is used to report overall student performance by describing how well students met the expectations for their grade level/course. Each performance level is defined by a range of overall scale scores for each subject. There are four performance levels for NM-MSSA assessments:

- Level 4: Advanced
- Level 3: Proficient
- Level 2: Nearing Proficiency
- Level 1: Novice

Students who are Proficient or Advanced display mastery of grade-level expectations. They display satisfactory or thorough understanding and use of college- and career-readinesses standards.

Performance Level Descriptors (PLDs) describe the knowledge, skills, and practices that students should know and be able to demonstrate at each performance level in each content area (Language Arts, Mathematics, and Science), and at each grade level/course.

Web links to the PLDs are listed in "Appendix B: Performance Level Descriptors" on page 14.

2.1.3 Reporting Category Performance Indicators

Reporting category performance for NM-MSSA and NM-ASR assessments is reported to indicate whether the student performed above standard, at/near standard, and below standard in a given reporting category.

2.2 Description of Individual Student Reports

The following pages show examples of student reports. The text below describes what the information represents.

General Information

A ***Identification Information***

The ISR lists the student's name, state student ID, date of birth, language in which the student tested, the grade level of the test, the grade level of the student when assessed, the district name, and the school name.

B ***Family Letter***

This letter, written by Secretary of Education Dr. Ryan Stewart, explains how this report was created and the special considerations of this school year. There is information here to guide families to more assessment literacy resources.

Overall Assessment Scores for Each Content Area

C ***Overall Scale Score and Performance Level***

This section of the report provides the student's overall scale score and performance level for each assessment taken (refer to Section 2.1). Students receive an overall scale score and, based on that score, are placed in one of four performance levels, with Level 3 indicating the student is on target and Level 1 indicating the student needs support.

Performance by Reporting Category

D ***Reporting Category***

Within NM-MSSA and NM-ASR, there are specific skill sets (reporting categories) students demonstrate on the assessments. Each reporting category includes the header identifying the reporting category, a raw score indicating the number of points earned out of the total points possible, and an explanation of whether the student has met the expectations of the reporting category.

E ***Reporting Category Performance Indicators***

A student's reporting category performance indicator represents how well the student performed in that category.

Reporting category performance indicators are:

- Above Standard
- At/Near Standard
- Below Standard

F ***Ways to Support***

For each reporting category additional resources are provided for supporting families in the development of these skills at home.

Comparison to the School, District, and State

G ***Achievement Levels***

This lists the four performance levels and provides a brief description of each.

H ***Scale Score Range***

Indicates the highest and lowest scale score for each performance level

I ***Peer Comparison***

This section of the report shows a side-by-side comparison of a student's overall scale score with the average scale score of their peers in their school, in their district, and in the state.

Sample Student Report page 1–Overview

New Mexico Measures of Student Success and Achievement and Assessment of Science Readiness

Spring 2022 Student Report

A

Student Name:	LASTNAME123, FIRST123 M.
SSID:	D00000123
Date of Birth:	07/29/2009
Tested Grade:	05
Student Grade:	05
District:	Demonstration District A
School:	Demonstration School 2

B

Dear Parents and Guardians,

Thank you for your continued support and partnership with the Public Education Department to ensure that all New Mexico students are healthy, secure in their identity, and holistically prepared for college, career, and life. I am especially grateful for your time and sacrifice on behalf of your student during the immense challenges of the last two years.

This Individual Score Report describes how your student performed on spring 2022 state assessments. These assessments are summative in nature. They were not designed to inform your student's teachers about short-term teaching strategies or potential interventions but to give them, and you, a snapshot of where your student finished the 2021-2022 school year relative to state-adopted content standards and instruction.

In particular, this year's assessments are important as a starting point, post-pandemic, for determining new baseline end-of-grade math and reading levels.

If you have specific questions about your student's performance on the assessment, I encourage you to reach out directly to your local school administration. The Family Report Interpretation Guide is available at https://newmexico.onlinehelp.cognia.org/. In addition, should you have specific questions about the assessment, please visit the PED assessment bureau's Parent Resource page at https://webnew.ped.state.nm.us/bureaus/assessment/parent-and-student-resources/.

The PED appreciates the opportunity to be a part of your student's educational success.

Kurt Steinhaus, Ed.D.
Secretary of Education, New Mexico Public Education Department

How did your student do on the New Mexico MSSA and ASR assessment? You can look at your student's scale scores, Achievement Level labels, and Achievement Level Descriptors to determine how well your student has done and whether additional support may be necessary. Your student's teacher can help you with interpreting this report and deciding on next steps for your student.

LANGUAGE ARTS

C

508

Proficient

Confidence band is 498 to 518

Your child's English Language Arts score is in the **Proficient** level, on track for college and career readiness.

MATHEMATICS

480

Nearing Proficiency

Confidence band is 470 to 490

Your child's Mathematics score is in the **Nearing Proficiency** level, not yet on track for college and career readiness.

SCIENCE

538

Advanced

Confidence band is 528 to 548

Your child's Science score is in the **Advanced** level, on track for college and career readiness.

Confidence bands: Your student's test score indicates performance on the day of the test. The confidence bands indicate the range of possible test scores your student would be expected to achieve on a different day.

Novice **Nearing Proficiency** **Proficient** **Advanced**

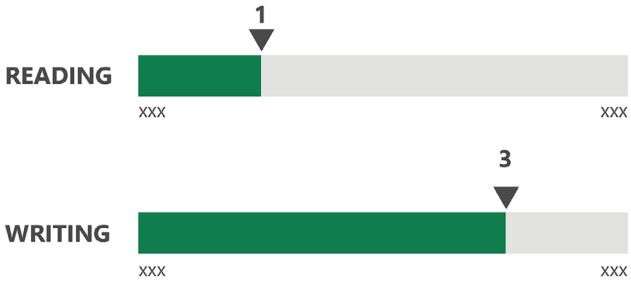
Sample Student Report page 2–English Language Arts

LANGUAGE ARTS

First123's Language Arts Performance by Reporting Category

	Points Earned / Points Possible	Subdomain Indicator	Ways to Support First123 F
Text type - Literary Text	D 13 / 16	E Above Standard	• Read stories with your student, allowing them to take the lead and read out loud as you listen and follow along. • After your student reads a story, ask them to summarize what happened in the story.
Text type - Informational Text	9 / 12	At/Near Standard	• Help your student choose materials on topics they are interested in, such as their favorite animals or famous people. Take turns reading with them. • Have your student (or yourself) point out interesting words while you are reading together.
Reading Strategy - Comprehension	3 / 6	Above Standard	• Read stories with your student, allowing them to take the lead and read out loud as you listen and follow along. • After your student reads a story, ask them to summarize what happened in the story.
Reading Strategy - Analysis and Interpretation	1 / 4	Below Standard	• Take turns reading with your student. Encourage and reassure them as they read. • After your student reads a story, ask them basic questions about the story. Ask them to try to use examples from the story to support their answers.

Key: x / y = x points earned out of y possible points

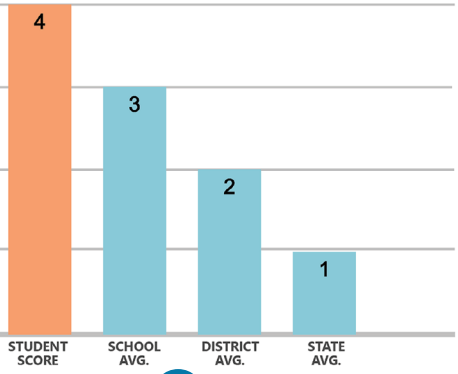


The Writing & Language scale score is based on student responses to the Writing & Language items. These items assess mastery of skills that require students to:

- analyze the use of introductions and conclusions,
- develop writing topics,
- use precise and effective language (grades 4-8),
- use transition words within pieces of writing,
- use correct conventions of standard English,
- use knowledge of language, and
- acquire and use vocabulary.

G COMPARISON to the SCHOOL, DISTRICT, and STATE

ACHIEVEMENT LEVELS	H SCALE SCORE RANGE
Advanced – Students show **thorough** understanding of how to interpret and analyze texts, use language and its conventions, and write for a task and purpose.	xxx – xxx
Proficient – Students show **satisfactory** understanding of how to interpret and analyze texts, use language and its conventions, and write for a task and purpose.	xxx – xxx
Nearing Proficiency – Students show **partial** understanding of how to interpret and analyze texts, use language and its conventions, and write for a task and purpose.	xxx – xxx
Novice – Students show **emerging** understanding of how to interpret and analyze texts, use language and its conventions, and write for a task and purpose.	xxx – xxx



I

Sample Student Report page 3–Mathematics

MATHEMATICS

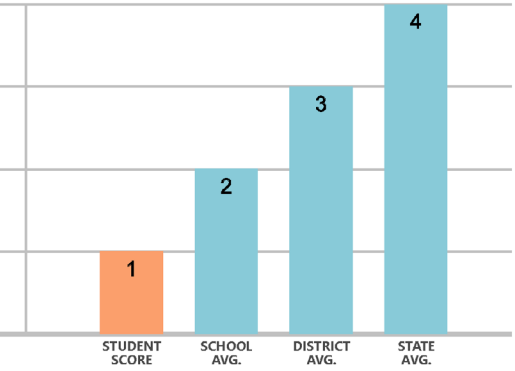
First123's Mathematics Performance by Reporting Category

	Points Earned / Points Possible	Subdomain Indicator	Ways to Support First123
Operations & Algebraic Thinking	15 / 23	Below Standard	• Solve multi-step word problems using addition and subtraction or multiplication and division with decimals. • Understand that multiplication and division can be used to compare quantities. For example, explain that a rubber band can stretch to three times its usual length.
Number & Operations in Base Ten/Number & Operations - Fractions	3 / 9	Below Standard	• Have your child visually model fractions, for example, drawing ½. Have them explain what ½ of certain shapes would look like. • Ask your child to multiply a number ending in zero by 10, 100 or 1000 and have them explain the place value of certain digits.
Measurement & Data/Geometry	9 / 12	At/Near Standard	• Pour liquids into two different sized cups. Ask your child if they have the same or different volume. Have them explain their reasoning. • Solve problems using the coordinate grid (graphs). For example, discuss a graph showing how temperature changes over the course of a year.
Problem Solving/Reasoning & Argument	5 / 6	At/Near Standard	• Describe, analyze, compare, and classify shapes using types of lines and angles. For example, compare the types of angles in two triangles. • Describe, analyze, compare, and classify shapes using types of lines and angles. For example, compare the types of angles in two triangles.
Modeling/Structure & Repeated Reasoning	3 / 8	Below Standard	• Encourage your child to experiment with representing problem situations in multiple ways, including writing numbers, creating math drawings, using objects, writing equations, and making a chart, list or graph. • Use rules (like add 3) to make patterns of numbers (like 2, 5, 8, 11...).

Key: x / y = x points earned out of y possible points

COMPARISON to the SCHOOL, DISTRICT, and STATE

ACHIEVEMENT LEVELS	SCALE SCORE RANGE
Advanced – Students show **thorough** understanding of mathematical concepts and strong procedural skill, fluency, and application to solve problems.	xxx – xxx
Proficient – Students show **satisfactory** understanding of mathematical concepts and adequate procedural skill, fluency, and application to solve problems.	xxx – xxx
Nearing Proficiency – Students show **partial** understanding of mathematical concepts and some procedural skill, fluency, and application to solve problems.	xxx – xxx
Novice – Students show **emerging** understanding of mathematical concepts and beginner procedural skill, fluency, and application to solve problems.	xxx – xxx



Page 3 of 4

Sample Student Report page 4–Science

SCIENCE

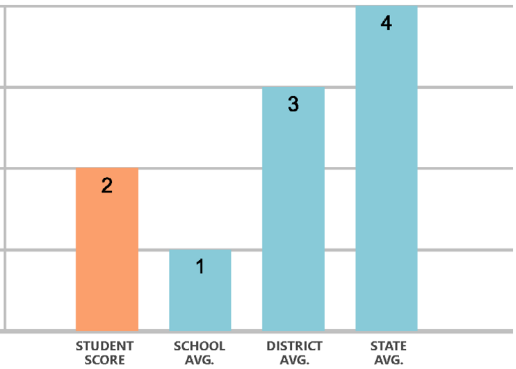
First123's Science Performance by Reporting Category

	Points Earned / Points Possible	Subdomain Indicator	Ways to Support First123
Practices and Crosscutting Concepts in Physical Sciences	12 / 15	Above Standard	• Ask your child to cook with you and discuss how they sometimes form a new substance when you mix two or more substances together. Develop a plan to investigate whether the mass of substances changes. • Ask your child to cook with you and discuss how they sometimes form a new substance when you mix two or more substances together. Develop a plan to investigate whether the mass of substances changes.
Practices and Crosscutting Concepts in Life Sciences	8 / 12	At/Near Standard	• Explore and explain how humans process information from our senses through the brain to keep us alive. • Explore and model how the different parts of plants (like seeds, leaves, roots, and fruit) and animals (like bones, legs, ears, and eyes) help them grow and survive.
Practices and Crosscutting Concepts in Earth and Space Sciences	15 / 18	Above Standard	• Understand the relationships between the Sun, Earth, Moon, and stars. For example, model how the Earth orbits the Sun and the Moon orbits the Earth and the effect of gravity on the orbits. • Work with your child to develop an model of a local ecosystem, showing how plants and animals get substances from energy and matter transfer. Discuss how the construction of a new building might affect the ecosystem.

Key: x / y = x points earned out of y possible points

COMPARISON to the SCHOOL, DISTRICT, and STATE

ACHIEVEMENT LEVELS	SCALE SCORE RANGE
Advanced – Students show **thorough** understanding of all 3 dimensions in making sense of phenomena and designing solutions in all 3 science domains.	xxx – xxx
Proficient – Students show **satisfactory** understanding of all 3 dimensions in making sense of phenomena and designing solutions in all 3 science domains.	xxx – xxx
Nearing Proficiency – Students show **partial** understanding of all 3 dimensions in making sense of phenomena and designing solutions in all 3 science domains.	xxx – xxx
Novice – Students show **emerging** understanding of all 3 dimensions in making sense of phenomena and designing solutions to problems in all 3 science domains.	xxx – xxx



3.0 Understanding the NM-MSSA and NM-ASR School and District Reports

3.1 Purpose and Use of NM-MSSA and NM-ASR Results

The NM-MSSA is New Mexico's statewide summative assessment for Language Arts and Mathematics, administered at the end of grades 3–8. The NM-ASR is New Mexico's statewide summative assessment for science, administered at the end of grades 5, 8, and 11. As the NM-MSSA and NM-ASR are singular measures at the end of a grade band, interpretations and uses of NM-MSSA and NM-ASR scores should be supplemented with additional measures, including information from classroom summative and formative assessments in Language Arts, Mathematics, and Science, as well as interim assessments.

3.2 NM-MSSA and NM-ASR School and District Reports

Districts and schools will have access to digital ISRs and a dynamic, customizable grade-level Student List in the Data Interaction reporting platform. The Student List can be customized by adding or removing data fields and by sorting and filtering selected data fields.

Data tools can be used to summarize scores and review score distributions for the whole group or disaggregate scores by subgroups. Bivariate analyses, both cross-tab and scatterplot, can be used to explore the relationship between scores.

ISRs and the Student List can be downloaded in a variety of formats for printing, presentations, or uploading into other analysis tools. Performance on NM-MSSA and NM-ASR assessments is described on the school and district reports using scale scores, performance levels, and reporting category performance levels.

3.3 Types of Scores on the NM-MSSA and NM-ASR School and District Reports

Performance on NM-MSSA and NM-ASR assessments is described on the school and district reports using scale scores, performance levels, and reporting category performance indicators. Information about state, district, and school average results is included in relevant sections of the report to help schools and districts understand how student and school performance compares to other students and schools. In some instances, a dash (–) will appear in place of average results for a school and/or district. This indicates that there are too few students to maintain student privacy and, therefore, results are not reported.

3.3.1 Scale Score

A scale score is a numerical value that summarizes student performance. Not all students respond to the same set of test items, so each student's scale score takes into account the slight differences in difficulty among the various forms of the test. The scale score allows for an appropriate comparison across test forms and administration years within a grade or course and content area. This year, NM-MSSA and NM-ASR reports provide overall scale scores for Language Arts, Mathematics, and Science, each of which determines a student's performance level in the respective content area. You can reference the NM-MSSA and NM-ASR scale score ranges in a table that appears on page 13.

For example, a student who earns an overall scale score of 800 on one form of the grade 8 Mathematics assessment would be expected to earn an overall scale score of 800 on any other form of the grade 8 Mathematics assessment. Furthermore, the student's overall scale score and level of mastery of concepts and skills would be comparable to a student who took the same assessment the previous year or following year.

3.3.2 Performance Level

Each performance level is a broad, categorical level defined by a student's overall scale score and is used to report overall student performance by describing how well students met the expectations for their grade level/course in the given content area. Each performance level is defined by a range of overall scale scores for the assessment. This year, there are four performance levels for NM-MSSA assessments:

- Level 4: Advanced
- Level 3: Proficient
- Level 2: Nearing Proficiency
- Level 1: Novice

Performance Level Descriptors (PLDs) describe the knowledge, skills, and practices that students should know and be able to demonstrate at each performance level in each content area (Language Arts, Mathematics, and Science), and at each grade level/course.

Web links to the PLDs are listed in "Appendix B: Performance Level Descriptors" on page 14.

3.3.3 Reporting Category Performance Indicators

Reporting category performance for NM-MSSA and NM-ASR assessments is reported to indicate whether the student performed above standard, at/near standard, and below standard in a given reporting category.

3.4 Description of The Student List and Summary Statistics

The following pages show examples of student reports. The text below describes what the information represents.

General Information

A ***Assessment Information***

The Student List displays the assessment, state, year, and the grade level.

B ***Identification Information***

The first column of the Student List displays the students in the school by last name. The students' first names and State Student IDs are shown in the next two columns.

Overall Assessment Scores

C ***Scale Score***

This column provides the student's overall scale score. Students receive a numerical score and, based on that score, are placed in one of three performance levels.

D ***Performance Levels***

This column provides the student's performance levels, with Level 3 indicating the student is on target, Level 2 indicating the student is near target, and Level 1 indicating the student needs support.

Report Functionality

E ***Options***

The options menu provides the capability to customize the Student List report. Student demographic fields and score data can be added or removed from the report. Additional scores can also be added or removed. This includes subject level scale scores and subclaim performance levels.

F ***Save***

Each report and the current selections can be saved in Data Interaction, allowing the user to conveniently retrieve the report at a later date. Saved reports can be retrieved by clicking on the Save icon.

G ***Download***

Tabular reports can be downloaded as an excel, CSV or PDF file by clicking on the Download icon. Charts can be downloaded as PDFs.

H ***Print***

Each report can be printed.

I ***Help***

A detailed user guide is available by selecting the Help icon.

J ***Student ISR***

The ISR for each student in the Student List Report can be viewed by clicking on the Student icon.

Performance by Reporting Category

K ***Reporting Category***

Within NM-MSSA and NM-ASR, there are specific skill sets (reporting categories) students demonstrate on the assessments. Each reporting category includes the header identifying the reporting category; state, district, and school averages; and an indicator of the student's performance.

Summary Statistics

Click on Scale Score > Summarize on the Student List page to view summary statistics for the selected organization.

L ***Population***

This count includes both valid and invalid students.

M ***Valid N***

This count includes only valid student records.

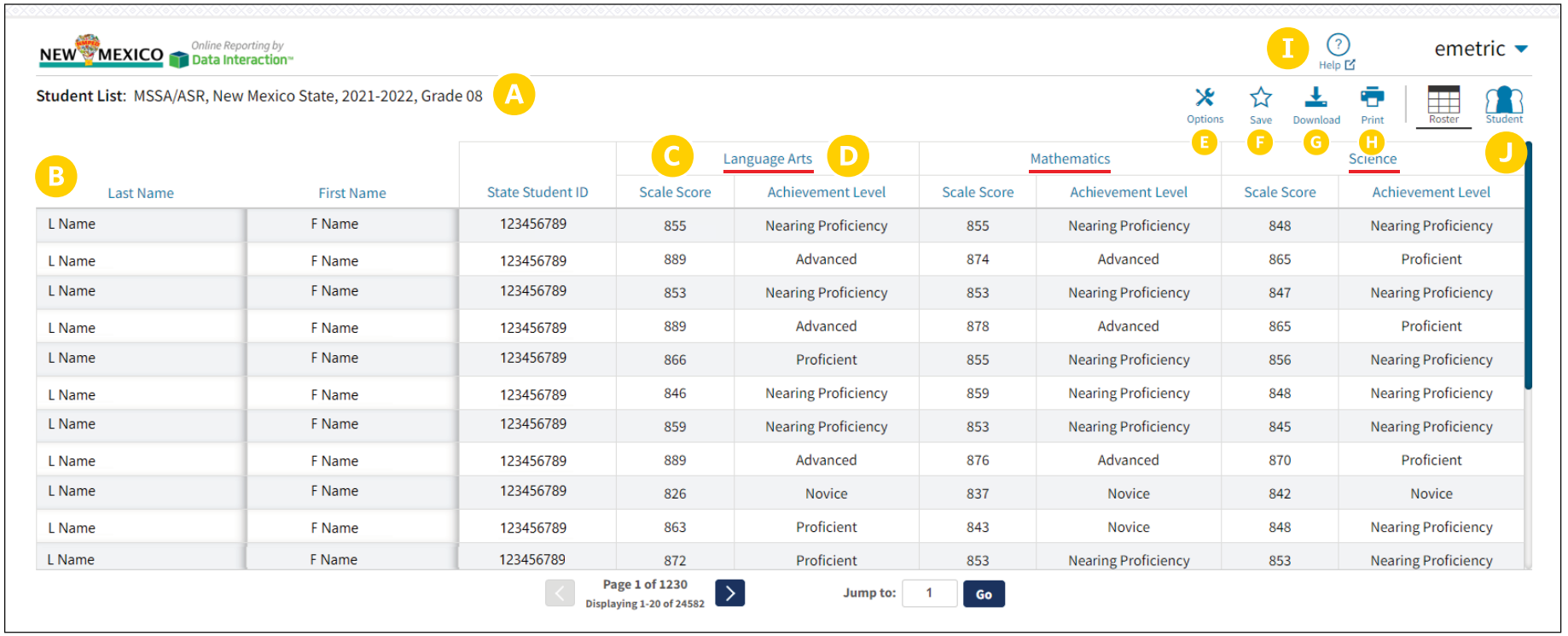
N ***Summary Statistics***

These statistics include mean, standard deviation and the minimum and maximum of the selected score.

O ***Recent Reports***

Recently generated reports can be viewed, within the session.

Sample Student List (Language Arts, Mathematics, and Science and Summary Statistics)



NEW MEXICO Online Reporting by Data Interaction

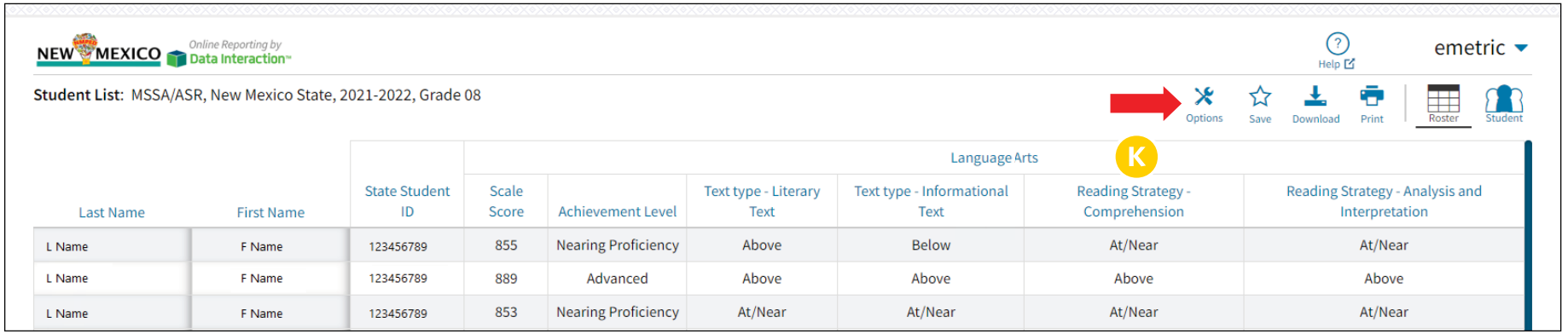
I Help emetric

Student List: MSSA/ASR, New Mexico State, 2021-2022, Grade 08 A

Options E Save F Download G Print H Roster Student J

B Last Name	First Name	State Student ID	C Language Arts Scale Score	D Language Arts Achievement Level	Mathematics Scale Score	Mathematics Achievement Level	Science Scale Score	Science Achievement Level
L Name	F Name	123456789	855	Nearing Proficiency	855	Nearing Proficiency	848	Nearing Proficiency
L Name	F Name	123456789	889	Advanced	874	Advanced	865	Proficient
L Name	F Name	123456789	853	Nearing Proficiency	853	Nearing Proficiency	847	Nearing Proficiency
L Name	F Name	123456789	889	Advanced	878	Advanced	865	Proficient
L Name	F Name	123456789	866	Proficient	855	Nearing Proficiency	856	Nearing Proficiency
L Name	F Name	123456789	846	Nearing Proficiency	859	Nearing Proficiency	848	Nearing Proficiency
L Name	F Name	123456789	859	Nearing Proficiency	853	Nearing Proficiency	845	Nearing Proficiency
L Name	F Name	123456789	889	Advanced	876	Advanced	870	Proficient
L Name	F Name	123456789	826	Novice	837	Novice	842	Novice
L Name	F Name	123456789	863	Proficient	843	Novice	848	Nearing Proficiency
L Name	F Name	123456789	872	Proficient	853	Nearing Proficiency	853	Nearing Proficiency

Page 1 of 1230 Displaying 1-20 of 24582 Jump to: 1 Go



NEW MEXICO Online Reporting by Data Interaction

Help emetric

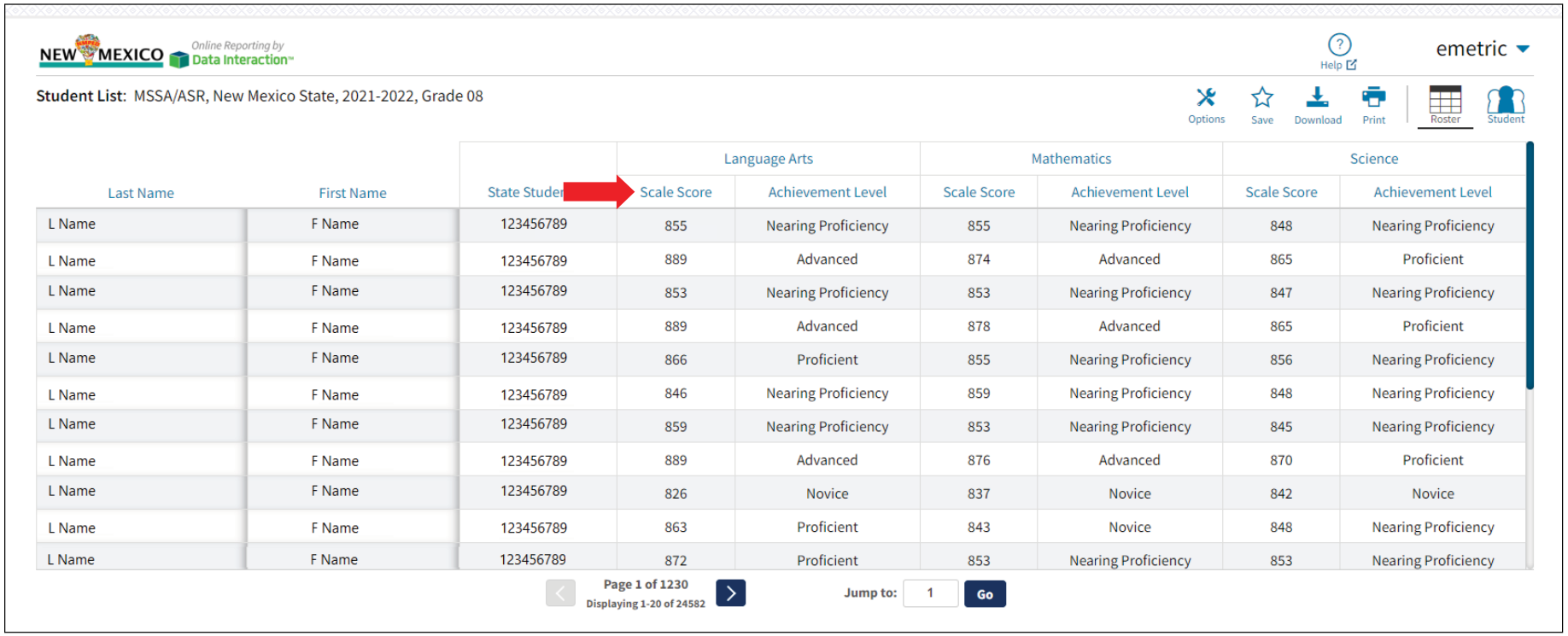
Student List: MSSA/ASR, New Mexico State, 2021-2022, Grade 08

Options Save Download Print Roster Student

Last Name	First Name	State Student ID	Language Arts K Scale Score	Achievement Level	Text type - Literary Text	Text type - Informational Text	Reading Strategy - Comprehension	Reading Strategy - Analysis and Interpretation
L Name	F Name	123456789	855	Nearing Proficiency	Above	Below	At/Near	At/Near
L Name	F Name	123456789	889	Advanced	Above	Above	Above	Above
L Name	F Name	123456789	853	Nearing Proficiency	At/Near	At/Near	At/Near	At/Near

Select *Options*-> Additional Scores can be added or removed. This includes subject level scale scores and reporting category achievement levels.

Sample Student List (Language Arts, Mathematics, and Science Summary Statistics) continued



NEW MEXICO Online Reporting by Data Interaction™

Help emetric

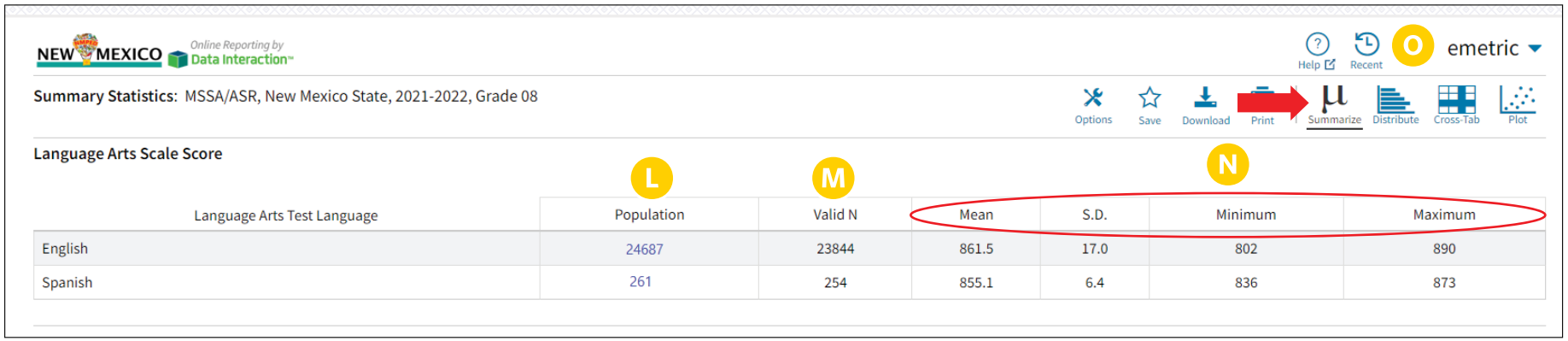
Student List: MSSA/ASR, New Mexico State, 2021-2022, Grade 08

Options Save Download Print Roster Student

			Language Arts		Mathematics		Science	
Last Name	First Name	State Studer	Scale Score	Achievement Level	Scale Score	Achievement Level	Scale Score	Achievement Level
L Name	F Name	123456789	855	Nearing Proficiency	855	Nearing Proficiency	848	Nearing Proficiency
L Name	F Name	123456789	889	Advanced	874	Advanced	865	Proficient
L Name	F Name	123456789	853	Nearing Proficiency	853	Nearing Proficiency	847	Nearing Proficiency
L Name	F Name	123456789	889	Advanced	878	Advanced	865	Proficient
L Name	F Name	123456789	866	Proficient	855	Nearing Proficiency	856	Nearing Proficiency
L Name	F Name	123456789	846	Nearing Proficiency	859	Nearing Proficiency	848	Nearing Proficiency
L Name	F Name	123456789	859	Nearing Proficiency	853	Nearing Proficiency	845	Nearing Proficiency
L Name	F Name	123456789	889	Advanced	876	Advanced	870	Proficient
L Name	F Name	123456789	826	Novice	837	Novice	842	Novice
L Name	F Name	123456789	863	Proficient	843	Novice	848	Nearing Proficiency
L Name	F Name	123456789	872	Proficient	853	Nearing Proficiency	853	Nearing Proficiency

Page 1 of 1230
Displaying 1-20 of 24582

Jump to: 1 Go



NEW MEXICO Online Reporting by Data Interaction™

Help Recent O emetric

Summary Statistics: MSSA/ASR, New Mexico State, 2021-2022, Grade 08

Options Save Download Print Summarize Distribute Cross-Tab Plot

Language Arts Scale Score

L M N

Language Arts Test Language	Population	Valid N	Mean	S.D.	Minimum	Maximum
English	24687	23844	861.5	17.0	802	890
Spanish	261	254	855.1	6.4	836	873

Click on *Scale Score->Summarize* on the Student List page to view summary statistics such as mean, standard deviation, and the minimum and maximum score, for the selected organization.

Appendix A: Scale Score Ranges

		Scale Score Range			
Subject	**Grade**	**Novice**	**Nearing Proficiency**	**Proficient**	**Advanced**
Language Arts	3	300–335	336–359	360–369	370–390
	4	400–439	440–459	460–472	473–490
	5	500–542	543–559	560–572	573–590
	6	600–631	632–659	660–672	673–690
	7	700–730	731–759	760–774	775–790
	8	800–839	840–859	860–870	871–890
Mathematics	3	300–340	341–359	360–376	377–390
	4	400–443	444–459	460–479	480–490
	5	500–547	548–559	560–572	573–590
	6	600–645	646–659	660–678	679–690
	7	700–747	748–759	760–769	770–790
	8	800–840	841–859	860–877	878–890
Science	5	500–543	544–559	560–574	575–590
	8	800–843	845–859	860–881	882–890
	11	1100–1153	1154–1159	1160–1181	1182–1190

Appendix B: Performance Level Descriptors

Grades 3–8 Language Arts and Mathematics

PLDs for grades 3-8 Language Arts and Mathematics are available at webnew.ped.state.nm.us/bureaus/assessment/state-assessments/#assessment-nmmssa

Grades 5, 8, 11 Science

PLDs for grades 5, 8, 11 Science are available at webnew.ped.state.nm.us/bureaus/assessment/state-assessments/#assessment-nmasr

NEW MEXICO
Public Education Department

NM-ASR

NEW MEXICO ASSESSMENT
OF SCIENCE READINESS

NM-MSSA

NEW MEXICO MEASURES
OF STUDENT SUCCESS AND
ACHIEVEMENT